

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)	A registry-based analysis of participator representativeness: A source of concern for sickness absence research?
AUTHORS	Knapstad, Marit; Love, Jesper; Holmgren, Kristina; Hensing, Gunnel; Øverland, Simon

VERSION 1 - REVIEW

REVIEWER	Juha Karvanen Department of Mathematics and Statistics University of Jyväskylä Finland
REVIEW RETURNED	10-May-2016

GENERAL COMMENTS	<p>Non-participation is becoming more and more severe problem in surveys. Only few countries have high quality registries available to study non-participants. From this perspective, the problem considered in the manuscript is important and interesting.</p> <p>I have concerns related to the design, data, methods and conclusions.</p> <p>1) I had difficulties in understanding what kind of data are available for each group. On page 3, it is said "The use of registries yielded complete and unbiased data from participants and nonparticipants alike." but on page 9 it is told that only aggregate data are available for the non-participants. What do you mean by aggregate data? Do you have the joint distributions of the variables or only marginal distributions? The confusion about the data makes it difficult to evaluate the validity of the methods and conclusions.</p> <p>2) Sickness absence data is not one number but an event history that can be characterized, for instance, by recency, frequency and duration. I can imagine many ways how sickness absence could impact survey participation: a) individuals currently on the sick leave may not be able to participate b) individuals with recent sickness absence may have high/low probability to participate. c) individuals that are often on the sickness leave may have high/low probability to participate d) individuals who have had long sickness absences may have high/low probability to participate. As you mention, the distribution of annual sickness absence days is likely to be skewed and calculating the mean is not the best way to handle such data.</p> <p>3) Taking into account comments 1 and 2, I am not convinced that the question stated in the title can be answered reliably. The aggregate data may not be sufficient to find the way sickness absences impact the participation.</p>
-------------------------	---

	<p>4) On page 11, you mention that you accounted for the gender and age differences between the comparisons groups. Did you consider other potential confounders, especially country of birth, marital status and income mentioned on page 9?</p> <p>5) I believe the statistical methods could be improved but it is difficult to be specific when I do not fully understand the structure of the data. Maybe you could have a statistical model that combines all years.</p> <p>6) For me, the first two sentences in Discussion sound conflicting. The same sentences are also in the abstract.</p> <p>Minor comments:</p> <p>7) Page 4: There are more relevant references to selection bias than (1). See, for instance, Keiding & Louis (2016) JRSS A and references therein.</p> <p>8) How do you handle individuals who have moved in or out the region?</p> <p>9) Page 8: Why the inclusion condition is >14 days in one sample and >1 day in another?</p> <p>10) Page 10: Is co-habiting classified as "not married"?</p> <p>11) Page 11 footnote: Do you mean "negligible" when you write "trivial"?</p> <p>12) Page 15: I do not understand expression "statistically lower mean number" but I would understand it without word "statistically".</p> <p>13) Page 20: Could you explain how "including a population-based sample increases generalizability of the findings."?</p> <p>14) Maybe the flow chart could also tell about the type of available data (individual / aggregate)</p> <p>15) When you talk about follow-up you mean the years before the survey, which can be confusing for a reader.</p> <p>16) Page 3: What do you mean by "mechanism for survey participation"? I understand that this means roughly the same as the factors that affect survey participation.</p> <p>17) In reference 12 the name of the author should be Läärä, not Laara.</p>
--	---

REVIEWER	Marika Wenemark Department of Medicine and Health Sciences Linköping University, Sweden
REVIEW RETURNED	12-May-2016

GENERAL COMMENTS	The manuscript address an interesting and important question: do we get representative participants in surveys with low response rates? Specifically the manuscript want to investigate whether
-------------------------	---

participants with sickness absence are more inclined to participate in a survey on sickness absence, health and work. My main worry is that there are (at least) two important factors that influence participation: interest in topic (relevance) and problems to participate due to health-problems. These two factors will probably influence participation in opposite directions. So the finding that participants and non-participants have similar sickness-absence history may for the sick-listed samples be due to more response due to interest as well as less responses due to health-problems. This is addressed by the authors on page 23 (line 2-10) but need to be discussed earlier and more substantial in the discussion. The question is if the analysis is optimal due to this problem. The data material is very interesting and could be used in analyses that better answer the research questions. For example the detailed presentations of year-by-year mean days could be presented as individual sickness absence history over the last 8 years (for example no history, short episodes, long recent episode, and long less recent episode). To be able to separate influence of interest from influence of health problems I would like to see for example separate analyses for people who are now employed and has no sickness absence during the last month).

Beside this main point I have some specific questions:
The descriptions of the samples could be clarified. It is not clear whether the employee and non-employee samples are random or not. On P9L3 the text "In the employee-sample, the total sampling frame was invited to participate..." is confusing. Is the sampling frame a sample? If every person in the frame is invited it is not a sample. The random population sample could be titled no-sickness absence sample. I would suggest to call them groups instead and describe that the employee group consists of all people that met the inclusion criteria... the non-employee group is based on a sample... and so on.

I would like to see response rates in the different groups in Table 1. Page 15: I would like to see a table of what is declared in the header "Differences in survey participation by registered sickness absence days..." instead of the mean number of days which is now showed. Table 2 is unnecessary complicated since it is not obvious that the two right-side columns is different for different parts of the table. I would rather see Participants and Nonparticipants for all three groups. I would like to see the estimates of gender, age, country of birth and so on together with mean absence days for participants in each of the three groups and the target population values (as the truth) for each group. That would show the representativeness of the different groups as compared to the true population values.

Page 17: Just as the header on page 15 this header declare something very interesting which is not shown in the table. I would like to see a table of response rates in different groups of sickness absence now and previous. Especially the employed groups 1) no recent sickness absence >14 days and no history of long absence periods 2) no recent sickness absence >14 days and history of long episode periods (or other similar groups) would be interesting to compare since they should have the same capabilities to participate with regards to their current health status (since they are employed and have no recent sickness absence).

This manuscript is based on what seems to be high quality data that can be used more efficient to answer the research questions raised.

VERSION 1 – AUTHOR RESPONSE

REVIEWER#1: Comments and responses

Non-participation is becoming more and more severe problem in surveys. Only few countries have high quality registries available to study non-participants. From this perspective, the problem considered in the manuscript is important and interesting.

I have concerns related to the design, data, methods and conclusions.

1) I had difficulties in understanding what kind of data are available for each group. On page 3, it is said "The use of registries yielded complete and unbiased data from participants and nonparticipants alike." but on page 9 it is told that only aggregate data are available for the non-participants. What do you mean by aggregate data? Do you have the joint distributions of the variables or only marginal distributions? The confusion about the data makes it difficult to evaluate the validity of the methods and conclusions.

RESPONSE COMM#1: Thank you for pointing out that the structure of the data available was not clear enough. We hope that the following description is clarifying. We have revised the manuscript to better describe the kind of data that were available for each group. By aggregate we mean that the data are merged into gender and age groups. As the "aggregate" term might be confusing, we have changed the description to "group-level" in the revised ms (which also better contrast individual-level). Also, we added information about the level of the data available for each group in the figure, as suggested, as well as in the abstract, to better highlight this information.

About the extraction of the registry data:

Firstly, the Health Assets Project (HAP) was set up as a survey study including linkages to registry data. As described in the ms, three cohorts were invited: two recent sick-listed cohorts and a random population cohort (not recent sick-listed). For all participants, in all three groups we have survey data and individual-level registry data on various demographic variables and sickness absence across nine years. Regarding sickness absence, we more specifically got the total number of sickness absence days and episodes each individual had got reimbursed by the Social Insurance Agency per year. As these data are registrations of actual monetary reimbursements they are regarded highly accurate, nearly complete (i.e. there is little missing data) and unbiased (as opposed to challenge of e.g. recall bias when using self-reported data).

In addition, as part of an overall nonresponse report for the project, Statistics Sweden extracted registry data on participants and invited nonparticipants on the following demographic characteristics (those displayed in table 1): gender, age group (19-30, 31-50, 51-64), country of birth (Nordic, others), marital status (married, not married), income (SEK 0- 149K, 150K-299K, 300K+). I.e. data merged into groups. This what we mean by aggregate data, that we do not have the specific value per individual, but rather the number of individuals per group for participants vs. nonparticipants.

Later, we wanted to examine the representativeness of the HAP participants concerning the main topic of HAP –sickness absence (i.e. the current study). To do this we contacted Statistics Sweden ones more to retrieve the registry data we needed to analyze this. Specifically, we asked for sickness absence data corresponding to what we had available from participants, for the target populations of the two recent sick-listed cohorts. In this round, we got data grouped (aggregated) by gender, age groups (19-30, 31-50, 51-64), and year. Again, the data is regarded complete, by including all sickness absence days reimbursed by the Social Insurance Agency. And by "aggregate data" we mean that the data were merged into the mentioned groups. For instance, in the recent sick-listed employed cohort, we got the information that among women aged 31-50, n=3979 individuals had in total 311526 number of reimbursed sickness absence days and 5303 number of sickness absence episodes in 2008.

Regarding the random population cohort, the statistics about the Social Insurance Agency's reimbursements in the Västra Götaland region (the target population) was already freely available on Statistics Sweden's web pages. These data are grouped/aggregated by gender and age groups.

Amendments in revised manuscript:

- Title: Changed to "A registry-based analysis of participator representativeness: A source of concern for sickness absence research?"
- Labelling throughout ms: Changed the labelling of the structure of the target population/nonparticipant data from "aggregate" to "group-level" data.
- Figure 1: Added information about the level of the data available for each group.
- Changed "bullet point" regarding the data available to increase precision of the statement regarding the registry data
- Revised text about data collection on the target populations, p. 9

COMM 2) Sickness absence data is not one number but an event history that can be characterized, for instance, by recency, frequency and duration. I can imagine many ways how sickness absence could impact survey participation: a) individuals currently on the sick leave may not be able to participate b) individuals with recent sickness absence may have high/low probability to participate. c) individuals that are often on the sickness leave may have high/low probability to participate d) individuals who have had long sickness absences may have high/low probability to participate. As you mention, the distribution of annual sickness absence days is likely to be skewed and calculating the mean is not the best way to handle such data.

RESPONSE COM#2: We fully agree that sickness absence in general is complex as is the way sickness absence could impact on survey participation. We therefore regard that the title of the study, which you comment on in comm#3, better should reflect our focus on representativeness, rather than whether/how sickness absence influence survey participation. The latter framing might lead to much focus to the mechanisms involved, which we did not have the design to address. We rather wanted to give a description of the overall "net effect" of sickness absence, to contribute to the question of whether (and to what extent) we are able to reach representative participants through survey-based studies.

We have tried various approaches to get the best out of the data available. Though not exhaustive, the analyses include duration (comparisons of mean number of days annually, Q#2), proxies of recency (through Q#1 and by including sickness absence from one to seven years before the survey in Q#2), and to some extent frequency (by comparing proportions with previous episodes of sickness absence, Q#3).

Secondly, we agree and acknowledge the challenge in doing analyses on sickness absence days, which indeed tend to be skewed. We therefore have considered whether there were alternative ways to analyze data and conferred with a biostatistician:

For nonparticipants and the target population, we only had group-level data. This did not therefore include data on standard deviations and median (or other measures of central tendency than the arithmetic mean), and the lack of this constrained the scope of possible analytic approaches: Poisson regression, negative binomial regression and other non-parametric alternatives, such as Wilcoxon rank-sum test, were not applicable as they require data on the median and skewness (standard deviation) from both comparison groups, and are also based on an assumption of equal distributions between the groups. For Poisson regression, even though suitable for a "count" variable as sickness absence days, we would need to use a zero-inflated model due to the many "zeros" which would challenge interpretations of the data.

We discussed this challenge with a biostatistician who considered the t-test to be a sound and valid approach based on the data available, for the following reasons: The t-test is developed to handle non-normally distributed data and is very robust for comparing means. According to “the central limit theorem”, the distribution of means will be normally distributed at samples >30-40 observations, of which our data is well beyond [1]. Additionally, “Slutsky’s theorem” shows that differences in sample variance have little effect on the distribution of t-test when the sample size is large.

That said, interpreting means in itself can be problematic when the distribution of the data is skewed. Even though means of sickness absence days probably is fairly meaningful, we therefore focus more on the differences between the comparison groups than the mean values in the interpretation of our results. To inform the reader about these considerations, we elaborate on this in the discussion (p. 22 (last paragraph)-23(first paragraph)).

Amendments in revised manuscript:

In sum, we aimed to describe central tendencies. To better convey this aim, we have revised the title and manuscript to better highlight our focus on representativeness. Regarding the various way that sickness absence could impact survey participation, we raise this discussion throughout the manuscript (abstract conclusion, introduction, interpretation part of the discussion). In the revised manuscript, we have put even more emphasis on this this issue in the discussion (p.23, line 4-5 and line 16-21).

COMM 3) Taking into account comments 1 and 2, I am not convinced that the question stated in the title can be answered reliably. The aggregate data may not be sufficient to find the way sickness absences impact the participation.

RESPONSE COM#3: As described above, we have altered the title to better get our aim of the study across, i.e. whether we reach representative participants concerning sickness absence.

COMM 4) On page 11, you mention that you accounted for the gender and age differences between the comparisons groups. Did you consider other potential confounders, especially country of birth, marital status and income mentioned on page 9?

RESPONSE COM #4:

Yes, we have considered the factors you mention as potential confounders for a relationship between sickness absence and survey participation. Unfortunately, the data we have on these variables were retrieved separately from the sickness absence data (as described above), are not directly linkable, and we are thus not able to control statistically for these factors. We therefore addressed this in the limitation section in the discussion. We chose to specifically mention socio-economic status, whereof income is a key indicator, as there is solid evidence for an association between socio-economic status and both sickness absence and survey participation. In the revised manuscript, we have also specified this country of birth and marital status as possible residual confounders in the discussion (p.22, last line – p.23 line 1-2).

COMMENT 5) I believe the statistical methods could be improved but it is difficult to be specific when I do not fully understand the structure of the data. Maybe you could have a statistical model that combines all years.

RESPONSE COM#5: We hope that the revisions of the manuscript improved the readability about the structure of our dataset (see amendments made under comm#1). As suggested, we have added a column in table 2, showing differences in mean days across all years combined. For the population-cohort, we did not calculate total values as we found the available statistics for the general population

too crude to make reliable sum scores. We are a bit unsure whether this is the exact model you meant. If you find it useful, we are fine with keeping it.

In addition, based on the comments from reviewer 2, we have also revised table 1 and 2, which we believe improved the statistical approach and presentation. Please review the amendments made. As mentioned, the group-level data of the non-participants/target population precluded some of the more sophisticated analyses, but suggestions for further improvements are of course welcomed.

COM 6) For me, the first two sentences in Discussion sound conflicting. The same sentences are also in the abstract.

RESPONSE COM#6: By reviewing these sentences concerning our main finding, we agree that they may sound conflicting. The message we wanted to get across was that even though we find some statistical significant differences between comparison groups in sickness absence history, we regard the size of these differences as being not substantial - especially seen in concert with the many statistically non-significant results and the possibility for residual confounding.

In the revised manuscript we have revised these sections (abstract and main results, p. 20) to improve clarity of the sum up and interpretations of our main findings.

Minor comments:

7) Page 4: There are more relevant references to selection bias than (1). See, for instance, Keiding & Louis (2016) JRSS A and references therein.

RESPONSE COM#7: Thank you for the suggestion.

We have added this reference (now ref 2) to refer to a more update and broader discussion of selection bias than ref 1. In the discussion part (p23, line 9) we have also added some more references to empirical examinations of selection bias [see refflist below: 2, 3].

8) How do you handle individuals who have moved in or out the region?

RESPONSE COM#8: The sickness absence register (LISA) cover all sickness absence reimbursed from the National Social Insurance Agency, which include those having Swedish citizenship the year of registration. For participants in all three groups and nonparticipants in the two recent sick-listed groups, we thus have data on registered sickness absence even though they have moved into (or out and into - must have returned to the Västra Götaland region by 2008 to be included in the HAP study) the region during the years 2001-2007.

In the random population group, we compared participants to the average level of registered sickness absence in the Västra Götaland region per year. Participants living in other regions some of the years analysed should thus strictly speaking have been compared to the average of their respective region the respective year(s). Yet, we regard both this group and the differences in region average in sickness absence to be too small to make any important difference for the analyses.

There is probably some missing in registered sickness absence due to migration in/out of Sweden for some of the years 2001-2007. As mentioned, the LISA register only include those having Swedish citizenship the year of registration. We know from a report from the Statistics Sweden from 2011 on the LISA register 1990-2009 [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. Missing due to migration (or registration errors) was treated using listwise deletion.

Migration might be more common among nonparticipants than participants, supported by the findings that country of birth outside the Nordic countries (whereof non-Swedish citizenship is more likely) was more common among nonparticipants than participants. Nonetheless, we regard such non-random missing to be unlikely to have an impact for the sickness absence – survey participation association.

9) Page 8: Why the inclusion condition is >14 days in one sample and >1 day in another?

RESPONSE COM#9:

The inclusion condition is having had a new episode of sickness absence granted by the Social Insurance Agency during the inclusion period. As described in the method section, the LISA register includes sickness absence days reimbursed by the Social Insurance Agency (SIA). For employees, the employer covers the first 14 days of a sickness absence episode and SIA covers sickness absence beyond 14 days, making >14 days of sickness absence the inclusion criteria for the employed group. For non-employed (including self-employed, unemployed and students) SIA covers sickness absence beyond 1 day (upon application from the non-employed), thus making >1 day the inclusion criteria for this group.

Thank you for making us aware that this was not made clear enough in the ms. We have revised the text regarding inclusion criteria in the method section (p.8 , line 12-14).

10) Page 10: Is co-habiting classified as "not married"?

RESPONSE COM#10:

Yes, in the registry data only formal marriage is classified as “married”. Other categories included are “single”, “divorced” and “widow/widower”, grouped as “not married”.

Some of those classified as not married is co-habiting. We do not have the exact number, but this is general currently common in Sweden. From survey participants we have self-report data whether they had a partner relationship (thus, not necessarily co-habiting). Across samples, 65,2 % of those registered as single, 55,9 % of those registered as divorced and 37,2 % of those registered as widow/widower in the registries responded to be in a partner relationship. A cautious assumption is that the “co-habiting group” is more similar to the “married” group than the “non-cohabiting” group, and that grouping the two former in theory would increase the strength of the association between civil status and survey participation. As this is a speculation only, we have not commented on this in the manuscript. We do nonetheless believe that the data available in the registry give a crude indication of the association between civil status and survey participation.

In the revised ms we have specified that “co-habiting” is classified as “not married” (p.10).

11) Page 11 footnote: Do you mean "negligible" when you write "trivial"?

RESPONSE COM#11: Yes, thank you for spotting this mistake. This is corrected in the revised ms (p. 10)

12) Page 15: I do not understand expression "statistically lower mean number" but I would understand it without word "statistically".

RESPONSE COM#12: Thank you for your thorough read of the manuscript. This was a misspelling; we did mean “statistically significant lower mean number”. This is corrected in the revised ms (p.16).

13) Page 20: Could you explain how "including a population-based sample increases generalizability

of the findings."?

RESPONSE COM#13:

As referred to on p .20, most studies on sickness absence as a predictor for survey participation have employed specific occupational or diagnostic groups. These groups may for instance have a different distribution of sickness absence and demographic characteristics etc. than the general population, and the observed associations between sickness absence and survey participation in these studies not necessarily applicable to other groups. As the current study examined the association between sickness absence history and survey participation in a population-based survey, the observed results may to a greater extent be regarded as general tendencies.

We acknowledge that this was not conveyed clearly enough in the manuscript and have amended the text on the issue of generalizability p. 20 as well as in the "strength and limitation bullet points" (p.3).

14) Maybe the flow chart could also tell about the type of available data (individual / aggregate)

RESPONSE COM#14:

Type of available data (individual/group level) is now included in the flowchart, as suggested. We thank the reviewer for the suggestion, which we believe can make it easier to for readers to understand what kind of data we have available for each group (re comment 1 from the reviewer).

15) When you talk about follow-up you mean the years before the survey, which can be confusing for a reader.

RESPONSE COM#15:

We agree that "follow-up" can be confusing when we are "following" people back in time and not prospectively. We have changed this term to a more direct description in the revised manuscript.

16) Page 3: What do you mean by "mechanism for survey participation"? I understand that this means roughly the same as the factors that affect survey participation.

RESPONSE COM#16:

As noted above, the aim of our study was to examine representativeness of the participating sample regarding sickness absence history.

When referring to mechanisms for survey participation, we mean the obstacles or motives driving the participating/not participating, i.e. the mediators driving the relationship between various factors, like sickness absence history, and survey participation. As presented in the introduction and elaborated on in the discussion, topic interest/relevance might be one such mechanism possibly increasing likelihood for participation, whereas health-related obstacles may reduce likelihood of participation. We agree that this was not clearly conveyed in the short bullet point, and have revised the text to improve clarity (p.21 line 10-13, bullet point p. 3=)

17) In reference 12 the name of the author should be Läärä, not Laara.

RESPONSE COM#17: Thank you for spotting this mistake, we have corrected this error in the revised manuscript.

REVIEWER#2: Comments and responses

COMMENT#1: The manuscript address an interesting and important question: do we get representative participants in surveys with low response rates? Specifically the manuscript want to

investigate whether participants with sickness absence are more inclined to participate in a survey on sickness absence, health and work. My main worry is that there are (at least) two important factors that influence participation: interest in topic (relevance) and problems to participate due to health-problems. These two factors will probably influence participation in opposite directions. So the finding that participants and non-participants have similar sickness-absence history may for the sick-listed samples be due to more response due to interest as well as less responses due to health-problems. This is addressed by the authors on page 23 (line 2-10) but need to be discussed earlier and more substantial in the discussion. The question is if the analysis is optimal due to this problem.

RESPONSE COM#1:

We agree with the reviewer that mechanisms for survey participation are important AND complex to study. As highlighted by the reviewer and addressed in the discussion, it is possible that the two factors a) interest in topic/relevance and b) problems to participate due to ill health influence participation in opposite directions. We have been somewhat hesitant to put too much weight on mechanisms in the manuscript, as we did not have the design to analyze this directly. Thank you for helping us to see the need to give this issue more focus. In the revised manuscript, we have lifted and elaborated on the issue in the discussion.

Now, over and above the question about mechanisms is the question of representativity, which we also fully agree with the reviewer is important to address. In the HAP study we had the data to quantify representativeness regarding history of sickness absence, of which few previous studies have been able to do. We see this study as an important backdrop for further studies to answer whether and how various obstacles or motivations impact on survey participation.

Amendments in revised manuscript:

- To better convey that the aim of the current study is to examine representativeness, we have altered the title of the study.
- Regarding the discussion on mechanisms for survey participation that may have influenced our overall results, we have revised and elaborated on this issue in the discussion (p.23, line 4-5 and line 16-21). Additionally, in the bullet point (p.3) we have revised the text to better convey that we were not able to examine these mechanisms.

COMMENT#2: The data material is very interesting and could be used in analyses that better answer the research questions. For example the detailed presentations of year-by-year mean days could be presented as individual sickness absence history over the last 8 years (for example no history, short episodes, long recent episode, and long less recent episode). To be able to separate influence of interest from influence of health problems I would like to see for example separate analyses for people who are now employed and has no sickness absence during the last month).

RESPONSE COM#2:

Thank you for all suggestions on ways to analyze history of sickness absence and to separate the separate influence from the influence of health problems. We have also tried to get our heads around to best illuminate this. The aggregate level of the nonparticipant/target population does however put constraint on which operationalizations that are possible to make. Regarding history of sickness absence the last 8 years, we have done the exact operationalization that you suggest (no history, short episodes, long recent episodes, long less recent episodes) in a previous paper based on the HAP-participant data [5]. For nonparticipants/target population, we do however not have individual-level data which is needed to examine the mentioned variations in history of sickness absence. For these, we only have annual information on n individuals, having in total x days and z episodes of sickness absence (by gender and age groups).

Furthermore, as the LISA register only include number of sickness absence days and episodes per

year, without information on timing of events, we are neither able to examine differences in participation by sickness absence status last month.

What we are able to do, and have done in the revised manuscript, is to show differences in total number of sickness absence days across the 8 years between participants and nonparticipants/target population within the recent sick-listed cohorts.

COMMENT #3: Beside this main point I have some specific questions:

The descriptions of the samples could be clarified. It is not clear whether the employee and non-employee samples are random or not. On P9L3 the text "In the employee-sample, the total sampling frame was invited to participate..." is confusing. Is the sampling frame a sample? If every person in the frame is invited it is not a sample. The random population sample could be titled no-sickness absence sample. I would suggest to call them groups instead and describe that the employee group consists of all people that met the inclusion criteria... the non-employee group is based on a sample... and so on.

RESPONSE COM#3:

Thank you for pointing out that the descriptions of the samples were not clear enough and for suggestions for improvement. It is correct that the total sampling frame were invited in the "employee-sample", and only a random sample of the sampling frame of the "non-employee-sample". We initially chose the "sample" term to obtain a uniform set of labels of the groups, but we now see and agree that the term "sample" can be misleading concerning the "employee group". It could work to call the "groups", but we will suggest calling them "cohorts" instead to separate them from the variable categories (which we currently label groups: "participant group" vs "nonparticipant group", age groups).

In the revised manuscript we therefore have chosen the terms "employee-cohort", "non-employee-cohort" and "population cohort". If the reviewers and editor find it more appropriate to use the "group" term we will amend the labels accordingly.

Regarding the random population cohort: We agree that it per se would communicate well to call this group a "no-sickness absence group" (or "no-sickness absence cohort"). It is true that this group did not start a new registered of sickness absence episode during inclusion. However, due to the fluctuating nature of sickness absence, some had ongoing sickness absence (starting before the inclusion period) and some probably started a new episode after inclusion (between the inclusion period and the time they received the questionnaire by mail). Hence, we find it somewhat misleading to use the label "no-sickness absence" label and would suggest to keep the "random population" label.

COMMENT# 4: I would like to see response rates in the different groups in Table 1.

RESPONSE COM#4: Table 1 revised as requested.

COMMENT #5: Page 15: I would like to see a table of what is declared in the header "Differences in survey participation by registered sickness absence days..." instead of the mean number of days which is now showed.

RESPONSE COM#5: As we understand the comment, the reviewer suggests an additional table or figure showing participation rate by registered sickness absence days. We agree that this would be interesting to analyze, but are again restricted by the aggregate-level data of the nonparticipants/target population. We apologize for the misleading subtitle and have specified this according to what we have shown – i.e. "Differences in mean days of registered sickness absence days between participants and comparison groups, within cohorts" (p.16).

COMMENT# 6: Table 2 is unnecessary complicated since it is not obvious that the two right-side columns is different for different parts of the table. I would rather see Participants and Nonparticipants for all three groups. I would like to see the estimates of gender, age, country of birth and so on together with mean absence days for participants in each of the three groups and the target population values (as the truth) for each group. That would show the representativeness of the different groups as compared to the true population values.

RESPONSE COM#6:

Thank you for your thoughtful suggestions.

In the revised table 2, we have added a column to first show non-participants and secondly the target populations. Note that we do not have information about the nonparticipants for the population-cohort, and therefore only compare participants to the overall population (true) values.

As for separate sickness absence estimates by gender, age etc., we have previously considered whether to show all stratified estimates in a table or perhaps as forest plots (that is for gender and age groups where we have available data). In table 3 we show gender stratified estimates. The reason we have chosen not to show stratify by all subgroups, was that we would get a high number of data point (years x participation status x age groups x gender x 3 cohorts). We found that such solution would be too difficult to grasp/complex to display for the reader weighted against the added information we gained. Thus, we instead have chosen to only show the sum estimates, and instead weighted these estimates for the gender and age distribution among participants. Please tell if you rather regard that we should show both gender and age-stratified estimates.

COMMENT# 7: Page 17: Just as the header on page 15 this header declare something very interesting which is not shown in the table. I would like to see a table of response rates in different groups of sickness absence now and previous. Especially the employed groups 1) no recent sickness absence >14 days and no history of long absence periods 2) no recent sickness absence >14 days and history of long episode periods (or other similar groups) would be interesting to compare since they should have the same capabilities to participate with regards to their current health status (since they are employed and have no recent sickness absence).

RESPONSE COM# 7:

Again, we agree that a table on participation rates by various sickness absence histories would be interesting to show, but are again restricted by the data available. We do nonetheless believe that the comparisons of sickness absence history between survey participants and nonparticipants/target population is a valuable and important contribution as empirical evidence for whether we reach representative participants trough surveys regarding sickness absence.

Again, we apologize for the misleading subtitle, and have revised the title to better convey what is analyzed (p.18)

COMMENT#8: This manuscript is based on what seems to be high quality data that can be used more efficient to answer the research questions raised.

RESPONSE COM#8:

We thank you acknowledging the quality of the data. We are sorry that we are not able to meet all suggestions for analyses. We have also tried ourselves to get the most out of the data available. Without individual-level data for both comparison groups some analyses are not possible to perform.

We do nonetheless believe that the study and data available is of high value at a general level, to contribute with empirical data on participator representativeness regarding sickness absence. In sickness absence research, survey-methodology is important to gain more detailed information about for instance mechanisms driving sickness absence processes, yielding information not available in registries only. Thus, it is key to know whether the participants of such surveys are representative of

the population we want to learn something about.

References:

1. Pedhazur EJ, Schmelkin LP. *Measurement, Design, and Analysis: An Integrated Approach*. Hillsdale, New Jersey: Laurence Erlbaum associates, 1991.
2. Langhammer A, Krokstad S, Romundstad P, et al. The HUNT study: participation is associated with survival and depends on socioeconomic status, diseases and symptoms. *BMC Medical Research Methodology* 2012 Sep; 12.
3. Nilsen RM, Vollset SE, Gjessing HK, et al. Self-selection and bias in a large prospective pregnancy cohort in Norway. *Paediatric and Perinatal Epidemiology* 2009; 23(6): 597-608.
4. StatisticsSweden. *Integrated database for labour market research [Longitudinell integrationsdatabas för Sjukförsäkrings- och Arbetsmarknadsstudier (LISA) 1990–2009]*, 2011.
5. Knapstad M, Holmgren K, Hensing G, et al. Previous sickness absence and current low perceived social support at work among employees in the general population: a historical cohort study. *BMJ open* 2014; 4(10): e005963.