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)	Productivity loss due to menstruation-related symptoms: a nationwide cross-sectional survey among 32,748 women
AUTHORS	Schoep, Mark; Adang, Eddy; Maas, Jacques; De Bie, Bianca; Aarts, Johanna; Nieboer, Theodoor

VERSION 1 - REVIEW

REVIEWER	Kristy Sanderson University of East Anglia UK
REVIEW RETURNED	29-Aug-2018

GENERAL COMMENTS	This is an important topic which is under-researched. The survey is large-scale and had good patient and public input into the study. I had some queries around the measures and analysis. Measures The absenteeism and presenteeism questions are critical as the two outcomes for this paper. An existing measure was adapted, and a brief description was provided in the survey summary (Appendix), but I was still unclear on: - were these modifications validated? - was the non-MRS abs/pres questions about other health-related productivity loss or due to any reason/unspecified? Why was a 6-month recall period used when the other measures used 3 months? - how was the presenteeism % productivity loss question presented? For a typical day after the number of presenteeism days question? It would be helpful to provide the full text of the absenteeism and presenteeism questions in the Methods section. Menstrual cycle assessment - was a validated measure used, if so please provide a reference, and if not, provide further detail on what items were used and their origin. Data analysis It seems strange to have an online survey which did not control valid response options in the survey set-up, thus requiring censoring in the analysis. It should not be possible to report

anything beyond the maximum number of possible days of presenteeism in a 3-month period, for example.
Why were number of hours worked/studied censored at 40 hours when it is entirely plausible to work/study more than this?
The analysis is described as a binary logistic regression, which I am assuming was 0 vs 1+ days. This is not consistent with the text in Table 4 which states that "ORs >1 correlate with more absenteeism or presenteeism". If the outcome is binary the OR above 1 is interpreted as odds of having any days, it doesn't say anything about the number of days. Please clarify.
All measures are extrapolated to a 12 month time frame, whether they were a 3-month or 6-month recall period. I appreciate wanting to make the measures comparable post-hoc, but did the authors consider extrapolating just the 3-month measures to 6 months, and having a 6-month time frame? It would require re-doing the analysis, but it seems less messy and more valid than extrapolating all measures.
Related to this, please comment on the validity of extrapolating from a 3-month recall period to a 12-month recall period for absenteeism and presenteeism measures.

Discussion
Various other studies are mentioned that also assessed productivity loss, but nothing is mentioned of the likely variability in absenteeism and presenteeism assessment which will influence ability to compare across studies.

REVIEWER	Klara Hasselrot
	Dpt of Gynecology and Obstetrics, Danderyds Hospital, Stockholm Sweden Affiliation Karolinska Institutet, Stockholm, Sweden
REVIEW RETURNED	24-Sep-2018

GENERAL COMMENTS	Dear Authors,
	thank you for your contribution in the research field of menstruation-related symtoms.
	Please find the remarks in the attached file.
	Review
	" Productivity loss due to menstruation-related symptoms: a nationwide cross-sectional survey among 32.748 women" Schoep M et al
	Recommendation: Major revision, and additional specialist statistical review needed.
	This manuscript is easy to read, and are contributing to the important knowledge about women´s limitations due to menstrual-related symptoms. The objective is straight-forward and answered clearly.
	The remarks come in chronological order:
	1 Abstract: please rephrase "…and women´s preferences regarding conditions at work and school" since this is a too wide definition and does not reflect the questions asked (conditions related to menstrual-related symptoms)

2 Methods, page 6: "A group of women…. was involved in the conduct of this study at several stages" is to me a strange piece of information. Only being a woman clearly not qualifies to design scientific studies. Who are these people, and what are their qualifications? Furthermore, in the acknowledgement section, 3 out of 6 mentioned contributors seem to be male?

3 Methods, page 9: how can 1 full working/studying account for 8 hours, but a full week 36 hours?

4 Methods, page 10: the recall period seem to vary between different questions asked (from 1-6months). Please elaborate on this, and the possible effects on the results. Non menstrual-related symptoms had significantly longer recall period.

5 Methods, page 10: the data regarding "extreme outliers" needs to be clarified in order to publish these data. Clearly, one could suggest for the future to pose less open questions (i.e fixed numbers/intervals of days of presenteeism or number of sanitary pads/day). But as of today; on what grounds have the authors considered an answer "an extreme outlier"? Where is the cut-off, and is it made with any scientific methodology?

6 Table 1: please check the percentages in all columns, for example it does not sums up correctly regarding hormonal contraception (37,5% instead of 36,6%) and 63,7% instead of 63,4% in non-hormonal contraception. Also: why is PCOS per se considered a MRS diagnosis in your material? Is there any evidence of overrepresentation of HMB or dysmenorrhea in this group, like the other mentioned diagnosis? If so, please provide reference.

7 Results, page 12: how can the mean of total absenteeism not related to MRS be lower than in both of the subdivided groups?

8 Results, page 13: please explain why presenteeism due to non-MRS factors was not included in the questionnaire. This is a major concern, especially since the productivity loss due to presenteeism is so much higher than absenteeism in this study group.

9 Discussion: please define/exemplify what you mean by "cultural diversity". The problem with such a large difference regarding presenteeism in different countries is not addressed by using a reference discussing absenteeism.

10 Discussion: considering the age-dependent productivity loss is the main objective of this work, this is not addressed in the discussion. Please elaborate why younger women are more heavily affected. For example, older women have increasing problems of HMB, but this is not reflected in this data set?

REVIEWER	Fiona Cocker
	University of Tasmania, Australia
REVIEW RETURNED	04-Nov-2018

GENERAL COMMENTS	An excellent study which makes a significant contribution to the literature. However, three minor points I would like the authors to clarify before publication.
	1. Did the incidence of reporting menstrual symptoms as the main reason for absenteeism and/or presenteeism differ by age group? That is, are older more comfortable attributing their work attendance behaviours to MRS?

	2. Can you please include a few studies that have explored the impact of other health conditions on absenteeism and presenteeism in order to provide context for the current findings. I.e. how much more or less of an impact do MRS have on work attendance behaviours than other conditions. 3. Finally, have the results been adjusted for co-morbid health conditions?

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1

Reviewer Name: Kristy Sanderson

Institution and Country: University of East Anglia

UK

Please state any competing interests or state 'None declared': None declared

Please leave your comments for the authors below

This is an important topic which is under-researched. The survey is large-scale and had good patient and public input into the study.

I had some queries around the measures and analysis.

Measures

The absenteeism and presenteeism questions are critical as the two outcomes for this paper. An existing measure was adapted, and a brief description was provided in the survey summary (Appendix), but I was still unclear on:

Were these modifications validated?

The questions in itself were not adapted. However, as stated in the methods section we adapted the recall period. These modifications were obviously not validated. As stated in the discussion section (line 434 and further) we are aware of the fact that extending the recall period increases the risk on a recall bias.

Was the non-MRS abs/pres questions about other health-related productivity loss or due to any reason/unspecified?

We did not ask the participants the specific reason for the non-MRS related absenteeism of absenteeism, but merely mentioned that it had to be unrelated to the menstrual period in any way. We added a remark in the Methods section. (line 150 and further)

Why was a 6-month recall period used when the other measures used 3 months? We realize that this could be a fundamental drawback in our manuscript. We assumed the amount of presenteeism to be larger than the amount of absenteeism. Therefore, the recall period for absenteeism was extended to six months to maintain accuracy. These lines have been added to the Methods section. (line 143 and further)

How was the presenteeism % productivity loss question presented? For a typical day after the number of presenteeism days question?

We used a visual analog scale in which the left end said "I could not do anything" and the right end said "I could do just as much as I do normally". We also adapted the appendix to provide full-text of the absenteeism and presenteeism questions.

It would be helpful to provide the full text of the absenteeism and presenteeism questions in the Methods section.

The full-text questions on absenteeism and presenteeism have been added to the appendix.

Menstrual cycle assessment - was a validated measure used, if so please provide a reference, and if not, provide further detail on what items were used and their origin.

Participants were asked in a lay manner how long their menstrual cycle was and what the exact meaning of a menstrual cycle was. The duration of the cycle was divided in several categories. These are now provided in the Methods section. (line 132 and further)

Data analysis

It seems strange to have an online survey which did not control valid response options in the survey set-up, thus requiring censoring in the analysis. It should not be possible to report anything beyond the maximum number of possible days of presenteeism in a 3-month period, for example.

We realize this as a drawback of our study. In retrospect, it would have been better that in the SurveyMonkey questionnaire, data limits were set at specific minimum and maximum levels. However, less than 0.3% of the participants were excluded based on impossible answers.

Why were number of hours worked/studied censored at 40 hours when it is entirely plausible to work/study more than this?

We agree with the reviewer it is possible to work more than 40 hours a week. For reasons of clarity in the calculations and comparability with the data of the Dutch Central Bureau of Statistics, we maximized the worked and studied hours at 40. We additionally provide data on the number of women that reported to work or study more than 40 hours a week (3.2% and 1.0%, respectively) in table 1.

The analysis is described as a binary logistic regression, which I am assuming was 0 vs 1+ days. This is not consistent with the text in Table 4 which states that "ORs >1 correlate with more absenteeism or presenteeism". If the outcome is binary the OR above 1 is interpreted as odds of having any days, it doesn't say anything about the number of days. Please clarify.

We adapted the text in table 4. Indeed, ORs > 1 correlate with a higher prevalence instead of a larger amount of absenteeism or presenteeism.

All measures are extrapolated to a 12 month time frame, whether they were a 3-month or 6-month recall period. I appreciate wanting to make the measures comparable post-hoc, but did the authors consider extrapolating just the 3-month measures to 6 months, and having a 6-month time frame? It would require re-doing the analysis, but it seems less messy and more valid than extrapolating all measures.

Related to this, please comment on the validity of extrapolating from a 3-month recall period to a 12-month recall period for absenteeism and presenteeism measures.

We did consider to present the data in a three or six month time-frame. In the tables we also present the data in their original format. However, mainly for reasons of clarity and comparability with other studies and CBS data, we decided to answer our main question based on a 12 month extrapolation. We of course agree with the reviewer that this intrinsically implies some sort of uncertainty and we have commented on this in the Discussion section. (line 440 and further)

Discussion

Various other studies are mentioned that also assessed productivity loss, but nothing is mentioned of the likely variability in absenteeism and presenteeism assessment which will influence ability to compare across studies.

We made a short remark on this in the discussion section. (line 352)

Reviewer: 2

Reviewer Name: Klara Hasselrot

Institution and Country: Dpt of Gynecology and Obstetrics, Danderyds Hospital, Stockholm Sweden

Affiliation Karolinska Institutet, Stockholm, Sweden

Please state any competing interests or state 'None declared': None declared

Please leave your comments for the authors below

Dear Authors,

thank you for your contribution in the research field of menstruation-related symtoms. Please find the remarks in the attached file.etflix

Review

" Productivity loss due to menstruation-related symptoms: a nationwide cross-sectional

survey among 32.748 women"

Schoep M et al

Recommendation: Major revision, and additional specialist statistical review needed.

This manuscript is easy to read, and are contributing to the important knowledge about women´s limitations due to menstrual-related symptoms. The objective is straight-forward and answered clearly.

The remarks come in chronological order:

1 Abstract: please rephrase "…and women´s preferences regarding conditions at work and school" since this is a too wide definition and does not reflect the questions asked (conditions related to menstrual-related symptoms).

We rephrased this sentence.

2 Methods, page 6: "A group of women…. was involved in the conduct of this study at several stages" is to me a strange piece of information. Only being a woman clearly not qualifies to design scientific

studies. Who are these people, and what are their qualifications? Furthermore, in the acknowledgement section, 3 out of 6 mentioned contributors seem to be male?

We have elucidated this in the Methods section. Note that not all contributors have been mentioned in the acknowledgement section because they did not meet the criteria for this.

3 Methods, page 9: how can 1 full working/studying account for 8 hours, but a full week 36 hours?

We understand the confusion. We used the Dutch guidelines for cost research in which a work-day accounts for 8 hours. For most sectors in the Netherlands however, a full-time work-week is 36 hours. We have tried to clarify this in the Methods section. (line 199 and further)

4 Methods, page 10: the recall period seem to vary between different questions asked (from 1-6months). Please elaborate on this, and the possible effects on the results. Non menstrual-related symptoms had significantly longer recall period.

We comment on that in the discussion section; lines 404-409. The recall period for absenteeism due to MRSs was 6 months and for presenteeism due to MRSs 3 months. For absenteeism not related to MRSs, the recall period was 6 months. Moreover, we added a remark regarding the possible uncertainty initiated by extrapolating our results to a 12 months timeframe. (line 440 and further)

5 Methods, page 10: the data regarding "extreme outliers" needs to be clarified in order to publish these data. Clearly, one could suggest for the future to pose less open questions (i.e fixed numbers/intervals of days of presenteeism or number of sanitary pads/day). But as of today; on what grounds have the authors considered an answer "an extreme outlier"? Where is the cut-off, and is it made with any scientific methodology?

Indeed, "impossible answers" would be a better description, like we used in figure 1. We excluded answers that exceeded the maximum number of days possible. For example: more days with blood loss, abdominal pain or psychological complaints than the length of a menstrual cycle. Note that less than 0.3% of the participants were excluded based on impossible answers.

6 Table 1: please check the percentages in all columns, for example it does not sums up correctly regarding hormonal contraception (37,5% instead of 36,6%) and 63,7% instead of 63,4% in non-hormonal contraception.

Participants were able to give more than one answer, as mentioned in the legend of table 1. Therefore, only the 2 categories hormonal and no hormonal contraception sum up to 100%

Also: why is PCOS per se considered a MRS diagnosis in your material? Is there any evidence of overrepresentation of HMB or dysmenorrhea in this group, like the other mentioned diagnosis? If so, please provide reference.

We agree that PCOS is not a MRS diagnosis per se. However, PCOS is associated with a decreased quality of life, more pain and impaired mental health. See for example Li et al. Fertil Steril 2011;1996(2):452-8

7 Results, page 12: how can the mean of total absenteeism not related to MRS be lower than in both of the subdivided groups?

A total of 5,215 women (those not working or studying more than 5 hours a week) were not represented in one of the subdivided groups. Therefore the means can differ. In our opinion it makes sense that women who do not have study or work >5 hours a week have lower rates of absenteeism since they do not have to be at work or school for the majority of the week.

8 Results, page 13: please explain why presenteeism due to non-MRS factors was not included in the questionnaire. This is a major concern, especially since the productivity loss due to presenteeism is so much higher than absenteeism in this study group.

This is a fair point. However, including these questions could have made the questionnaire more difficult to fill in. We made a remark on this in the Discussion section. (line 426 and further)

9 Discussion: please define/exemplify what you mean by "cultural diversity". The problem with such a large difference regarding presenteeism in different countries is not addressed by using a reference discussing absenteeism.

We rephrased this sentence.

10 Discussion: considering the age-dependent productivity loss is the main objective of this work, this is not addressed in the discussion. Please elaborate why younger women are more heavily affected. For example, older women have increasing problems of HMB, but this is not reflected in this data set?

We have made an extra remark on this in the Discussion section (line 379 and further). Furthermore, women aged 46 and above were excluded in this study to prevent interference with perimenopausal complaints.

Reviewer: 3

Reviewer Name: Fiona Cocker

Institution and Country: University of Tasmania, Australia

Please state any competing interests or state 'None declared': None declared.

Please leave your comments for the authors below

An excellent study which makes a significant contribution to the literature. However, three minor points I would like the authors to clarify before publication.

1. Did the incidence of reporting menstrual symptoms as the main reason for absenteeism and/or presenteeism differ by age group? That is, are older more comfortable attributing their work attendance behaviours to MRS?

Initially we did not present these data. Younger women tend to be less open to report menstrual symptoms as being the reason for calling in sick (12.0% in women younger than 21 versus 27.0% in women aged 21 and above). We added this in the Results section. (line 316 and further)

2. Can you please include a few studies that have explored the impact of other health conditions on absenteeism and presenteeism in order to provide context for the current findings. I.e. how much more or less of an impact do MRS have on work attendance behaviours than other conditions.

This is an interesting but difficult point. We added additional data on this in the Discussion section. (line 383 and further)

3. Finally, have the results been adjusted for co-morbid health conditions?

No, we now added this in the Discussion section. (line 426 and further)

VERSION 2 – REVIEW

REVIEWER	Fiona Cocker
	University of Tasmania, Australia
REVIEW RETURNED	15-Jan-2019

GENERAL COMMENTS	My original queries have been satisfactorily addressed. Thank you.