PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	How do women with chronic fatigue syndrome/myalgic encephalomyelitis rate quality and coordination of health care services? A cross-sectional study
AUTHORS	Hansen, Anne; Lian, Olaug S.

VERSION 1 - REVIEW

REVIEWER	Inger Johanne Bakken Norwegian Institute of Public Health
REVIEW RETURNED	06-Nov-2015

GENERAL COMMENTS	In this paper, the authors describe results from a cross-sectional study where CFS/ME patient's assessment of quality of primary health care, specialist health care, and coordination of care has been investigated. The aim was to test the association between self-rated health and self-rated degree of CFS/ME and these outcome measures. Overall, the paper is very well written. Aims and background are well described, the result section is organized and easy to follow and the discussion is interesting with comprehensive reference to the literature. My main concern with the current paper regards the analyses of the data, and the presentation of the results. Logistic regression analyses have been used which is appropriate in this context. However, it is stated in the text describing the results presented in Table 3 that interaction terms (health x education and GP frequency x self-rated health) are significant. These results are not presented in the table, and the information in the table is not sufficient to understand which analyses have been carried out. In the presence of interaction it is not possible to interpret the meaning of an odds ratio estimate in the same straightforward manner as in the absence of interaction (see for instance Hosmer and Lemeshow: Applied logistic regression, chapter 3 in the second edition, or http://www.ats.ucla.edu/stat/stata/seminars/interaction_sem/interaction_sem.htm). Often, stratified analyses can solve this problem, but in this case there are two interaction terms that are described in the text. Also, all variables are added as continuous variables in the models, even though most of these variables actually are categorized. Finally, is it really expected that the relationship between the outcomes and age should follow a linear relationship? Minor comments Table 3 includes data on three different models, each with different

	outcomes (poor primary health care, poor specialist health care, poor coordination). However, only the first model is mentioned in the text. If the authors do not find the results in the two latter columns worth commented in the result section, they should be removed from the text. Probably, the authors should rather comment upon these results as well in the text. Top of page 9: The paragraph starting on line 9 describes the same results as the previous paragraph (first model in table 3). This should not be a separate paragraph, but rather follow the sentence above (ending with "… coordination of care (Table 3)." Throughout the paper, "bad quality of care" is used several times, while "poor quality" is used otherwise. My suggestion is to use "poor quality" throughout, including in the table headings (i.e., replace "bad/very bad" with "poor / very poor" etc. In the Discussion sentence, the results for the quality scores for specialist care are only very briefly mentioned. The last paragraph on page 11 starts with the sentence "We found that quality scores for specialist health care were better than for primary care", but the rest of the paragraph only discusses GP scores. Are results for specialist health care satisfaction not worth mentioning further?

REVIEWER	Schwarz Charité-Universitätsmedizin Berlin Germany
REVIEW RETURNED	22-Nov-2015

GENERAL COMMENTS	The authors studied how women with CFS/ME rate quality and coordination of health care services. The study was cross-sectional. General queries: Content: In the reviewer´s opinion the disease might not be the adequat focus for a questionaire survey as the definition of CFS/ME is very difficult. This might be already a big influence factor. An additional influence factor are the communication skills of the doctors involved. No checklist is added. Language: Common speech should be corrected into technical (scintific claim). The authors should not repeat the words „in line with" as much as they did in the text. Language should be corrected by native speaker. Results Page: 7 Line 43: Mean age is missing in this line. Please change the location to this position. Table3: Numbers should not be highlightened (same for table4). And the author do not explain why the results are differently marked. Page 11: Paragraph from line 7-56:

	As this is one major aspect of this study you must discuss this point in more detail and must explain these findings clearly. Page 12: Line 10: Why is there a question inside the text without explaining that? This should be explained or deleted. Line 12-14: This sentence needs a literature. Line 17-19: This is not very scientific and more common speech. Page 13: Line 31: common speech again (correction by native speaker).

VERSION 1 – AUTHOR RESPONSE

Reviewer 1 - Inger Johanne Bakken

My main concern with the current paper regards the analyses of the data, and the presentation of the results. Logistic regression analyses have been used which is appropriate in this context. However, it is stated in the text describing the results presented in Table 3 that interaction terms (health x education and GP frequency x self-rated health) are significant. These results are not presented in the table, and the information in the table is not sufficient to understand which analyses have been carried out.

In the presence of interaction it is not possible to interpret the meaning of an odds ratio estimate in the same straightforward manner as in the absence of interaction (see for instance Hosmer and Lemeshow: Applied logistic regression, chapter 3 in the second edition, or http://www.ats.ucla.edu/stat/stata/seminars/interaction_sem/interaction_sem.htm). Often, stratified analyses can solve this problem, but in this case there are two interaction terms that are described in the text.

We find the reviewer's concern highly relevant, and agree that more information is needed. On the reviewer's request, we have put effort into clarifying the interactions by major improvements in tables and text. Two new tables (Table 3 and 4) replace the former Table 3. We have extended the explanation of the interactions in the text. We hope that the reviewer and the editor will find this satisfying.

Also, all variables are added as continuous variables in the models, even though most of these variables actually are categorized.

We agree that this is a highly relevant concern, since some information might get lost by performing trend analyses. Initially, we performed all the analyses as dummy-analyses as well as trend analyses. However, we decided to report the trend analyses exclusively since some of the groups were small, and we revealed no significant lack of linearity. We have added this information to the Methods section (Analyses).

Finally, is it really expected that the relationship between the outcomes and age should follow a linear relationship?

Since the reviewer is asking specifically about the linearity of age in relation to the outcome variables, we have created the following Excel charge to illustrate this relationship. Overall, we concluded that there is not a significant lack of linearity for the three outcome variables.

Minor comments
Table 3 includes data on three different models, each with different outcomes (poor primary health care, poor specialist health care, poor coordination). However, only the first model is mentioned in the text. If the authors do not find the results in the two latter columns worth commented in the result section, they should be removed from the text. Probably, the authors should rather comment upon these results as well in the text.

The results are mentioned in the text, however, they are not particularly emphasised since there were only one statistically significant finding regarding specialist care, and no statistically significant findings regarding coordination of care (the present Table 4). However, from our experience, the question whether findings should be reported in great detail in text or tables or both is a recurring topic of discussion where there will be different views. We will be happy to be guided by the editor in this regard.

Top of page 9: The paragraph starting on line 9 describes the same results as the previous paragraph (first model in table 3). This should not be a separate paragraph, but rather follow the sentence above (ending with "… coordination of care (Table 3)."

We agree, and this has been changed according to the reviewer's suggestion.

Throughout the paper, "bad quality of care" is used several times, while "poor quality" is used otherwise. My suggestion is to use "poor quality" throughout, including in the table headings (i.e., replace "bad/very bad" with "poor / very poor" etc.

We are grateful for this observation, and the text has been changed according to the reviewer's suggestion.

In the Discussion sentence, the results for the quality scores for specialist care are only very briefly mentioned. The last paragraph on page 11 starts with the sentence "We found that quality scores for specialist health care were better than for primary care", but the rest of the paragraph only discusses GP scores. Are results for specialist health care satisfaction not worth mentioning further?

We agree with the reviewer's suggestion. We have made efforts in addressing the quality scores for specialist health care services in the new version of the discussion. Please, see the sixth paragraph of the Discussion section.

Reviewer 2 – Carsten Schwartz

General queries:
Content:
In the reviewer´s opinion the disease might not be the adequat focus for a questionaire survey as the definition of CFS/ME is very difficult. This might be already a big influence factor. An additional influence factor are the communication skills of the doctors involved. No checklist is added.

This concern regards the very conduct of this study. In our approach, we have related to the fact that people with the diagnosis chronic fatigue syndrome/myalgic encephalomyelitis (CFS/ME) have established a patient organization and that they are users of health care services. This study is not an evaluation of the diagnosis CFS/ME as such, but a study of how women with CFS/ME assess the

quality of the care they have received.

We agree that the communication skills of the doctors involved is a central topic in the assessment of quality of care. We consider the communication skills to be an integrated part of the patients' experiences of health care quality. Unfortunately, we have not been able to evaluate communication skills more specifically in this study.

A STROBE checklist has been followed, and is included as a supplementary information file, as requested by the Editor and this reviewer.

Language:
Common speech should be corrected into technical (scientific claim).
The authors should not repeat the words „in line with"as much as they did in the text. Language should be corrected by native speaker.

There will often be different opinions whether one should use common speech or a technical scientific language in a paper like the current one. We consider our text a mixture of both. However, since the reviewer does not mention any particular paragraphs in this regard, we would be happy to be guided by the Editor's opinion whether changes should be made. We appreciate the observation that "in line with" is repeated several times in the text, and have rewritten this according to the reviewers concern.

Results
Page: 7
Line 43: Mean age is missing in this line. Please change the location to this position.

We agree. The text has been changed according to the reviewer's suggestion.

Table3: Numbers should not be highlightened (same for table4). And the author do not explain why the results are differently marked.

We have explained in the notes under the tables that statistically significant findings are marked in bold. There might be different opinions regarding highlighting of statistically significant numbers, and we will be happy to be guided by the Editor in this regard.

Page 11:
Paragraph from line 7-56: As this is one major aspect of this study you must discuss this point in more detail and must explain these findings clearly.

We agree that this is a major aspect of this study. We have clarified and extended the discussion regarding the assessment of specialist health care services, as was also requested by the other reviewer. We hope that the editor and the reviewers will find the extended version of the discussion satisfying in meeting their concerns.

Page 12:
Line 10: Why is there a question inside the text without explaining that? This should be explained or deleted.

We must admit that at this point we do not fully understand this concern. We have used a rhetoric question, and reflected upon its content in the following paragraph. However, if changes are needed according to this point, we hope that we can get a further explanation of this concern.

Line 12-14: This sentence needs a literature.

Line 17-19: This is not very scientific and more common speech.

We agree regarding the sentence in line 12-14, and a reference has been added. Regarding line 17-19, we refer to our comments above.

Page 13:
Line 31: common speech again (correction by native speaker).

We must admit that at this point we do not understand the concern raised. We will be happy to be guided by the editor according to this point.

VERSION 2 – REVIEW

REVIEWER	Inger Johanne Bakken Norwegian Institute of Public Health
REVIEW RETURNED	05-Jan-2016

GENERAL COMMENTS	The main concern regarding the previous version was the statistical analyses and the presentation of results from these. It was stated that interaction terms (health x education and GP frequency x self-rated health) were significant without presentation of findings. In the new version, stratified analyses have been presented (stratified on age, education and GP frequency). It is not clear to me that stratifying on for instance age removes the interaction between self-rated health and education. Is interaction no longer a “problem” in these “new” models? Also, in the table heading the word “probability” is used while odds ratios are presented in the table. The authors should probably consider discussing their analyses with a senior statistician. Minor comments Page 13, line 44: Here, the authors describe their questionnaire as “well-designed systemically tested”. However, from the methods section it seems that the questionnaire has been designed for this particular study, and there is no description of any “systematic testing”. Additional information on the preparation and testing of the questionnaire should be added to the methods section, or alternatively, the sentences on page 13 should be removed or changed. Line 19, page 14: “While,” should probably be removed (i.e. start sentence with “Some studies…”

VERSION 2 – AUTHOR RESPONSE

Reviewer Inger Johanne Bakken

The main concern regarding the previous version was the statistical analyses and the presentation of results from these. It was stated that interaction terms (health x education and GP frequency x self-rated health) were significant without presentation of findings. In the new version, stratified analyses have been presented (stratified on age, education and GP frequency). It is not clear to me that stratifying on for instance age removes the interaction between self-rated health and education. Is interaction no longer a “problem” in these “new” models?

In the first version of the paper (submitted 20. October 2015) we mentioned three statistically significant interactions. In addition to the two interactions mentioned by the reviewer above, we described interaction between self-rated health and age. This third interaction is the reason why we have also stratified on age. We have consulted senior statisticians, particularly regarding the interactions, and we think that the interactions are now properly taken care of.

However, we would like to remove from the text the statement just ahead of Table 3 ("Also, quality in primary care was more likely reported poor/very poor by women with a shorter GP relation if they had university education or less frequent GP visits the previous year"). The first reason is that we did not find statistically significant interaction terms between GP duration and education, or between GP duration and GP frequency. In addition, the odds ratios in the different strata are not very different from each other (even though some of them differ from 1.0).

Also, in the table heading the word "probability" is used while odds ratios are presented in the table. The authors should probably consider discussing their analyses with a senior statistician.

We are fully aware that odds and probabilities are not exactly the same. However, there is a tradition that these terms are used interchangeably, and we have allowed ourselves (not only in the heading of Table 3, but throughout the paper) to confirm with the simplification that high odds correspond to high probabilities, and low odds to low probabilities.

A senior statistician has evaluated our analyses and presentation of results, particularly regarding the interactions.

Page 13, line 44: Here, the authors describe their questionnaire as "well-designed systemically tested". However, from the methods section it seems that the questionnaire has been designed for this particular study, and there is no description of any "systematic testing". Additional information on the preparation and testing of the questionnaire should be added to the methods section, or alternatively, the sentences on page 13 should be removed or changed.

Thank you for this observation. We have added information on the preparation and testing of the questionnaire, and a new reference. Please, see the Methods section.

Line 19, page 14: "While," should probably be removed (i.e. start sentence with "Some studies…"

We agree. This has been changed; please see the text in the Discussion section.