PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Prevalence and factors associated with fatigue in the Lausanne middle-aged population: A population based cross-sectional survey
AUTHORS	Galland, Coralie; Marques-Vidal, Pedro; Vollenweider, Peter

VERSION 1 - REVIEW

REVIEWER	James N. Baraniuk, MD Georgetown University Washington DC USA
REVIEW RETURNED	13-Oct-2018

GENERAL COMMENTS	Abstract: "Primary outcome measure: Prevalence of chronic fatigue, defined as a score ≥4 using the Fatigue severity scale (FSS)." page, line 28: “Fatigue severity during the last week….” In the abstract and conclusion you claim that 20% of the population has chronic fatigue, yet they were asked about fatigue in the past week. Do you define chronic fatigue as 1 week? Such a conclusion may have been valid if you had fatigue measurements to compare from the first and this, the second, samplings. However, fatigue was only assessed the second time. Please clarify. page 9, line 9: You use the original cut-off score ≥4. This should be repeated on page 9 lines 9-10 because your prevalence values will be considered important and should be qualified by the level of fatigue, particularly because you contrast your results to Lerdal et al. who used a higher cut-off (≥5). Your cut-off may also be mentioned in the legend to Table 2. The logistic regression and bivariate analysis would be strengthened by also showing the results with the higher cut-off (≥5) so that the level of fatigue, exclusion, and co-variates can be understood when stratified by fatigue severity. Your most significant features could be plotted as the % against increasing fatigue scores starting at 0 show the stratification of the most significant variables by fatigue level. This is relevant because fatigue can change over time. The stratified data would be of interest in general practice to emphasize the characteristics of each stratum of fatigued patients. Showing the prevalence of fatigue at each level of the FSS would be very useful for planning future epidemiological studies of fatigue. page 8, line 27: The propensity index is mentioned in the methods but I do not see it in the results or discussion.

	page 16, Line 33: “our results provide a first estimation of the prevalence of fatigue in the general population,” However, you report similar findings from other studies on page 12 lines 38 to 54. Please clarify. The high exclusion rate is a significant limitation. Conclusion: Comment on the almost 40% of the surveyed population who were excluded from the study because they may have biased the outcomes. Supplemental table 1: Add odds ratios for exclusion variables

REVIEWER	Solveig Dohrmann Centre of Martitime Health and Society, The University of Southern Denmark, Denmark
REVIEW RETURNED	29-Nov-2018

GENERAL COMMENTS	Thank you for presenting this manuscript, I have read it with great interest. This manuscript can contribute important knowledge on the prevalence of fatigue and on its determinants in the general population. Therefore, I believe that the manuscript should be published by BMJ Open. However, there are areas that need attention and therefore I recommend revision before publication. On the positive side the method (survey) suits the purpose. With respect to the overall structure, the manuscript is well organized with heading, subheadings and tables, which all contributes to making the manuscript really reader friendly. Further, the manuscript is well articulated, and almost always logical. As for the areas that need attention, from the beginning: Introduction Please consider/address: - You provide a definition, state that fatigue varies in duration and intensity and provide prevalence rates (line 5-15, page 4). Please elaborate – add a brief explanation why fatigue is a health concern (with a specific focus on the purpose of your study (exploring determinants of cardiovascular disease)), make a clear distinction between acute fatigue, fatigue of a more chronic nature and chronic fatigue syndrome, and based on the definition you provide specify what type of fatigue you focus on. - Line 11-13, page 4: You state that fatigue is a common symptom in the general population, and along line 36-37, page 4 you state that only two studies have assessed the prevalence of fatigue in the general population – please elaborate. - Line 13-15, page 4: I recommend that you move the last sentence ‘this ten-fold range….. assess fatigue’ to the discussion, and discuss this issue in greater details. - Line 37-39, page 4: I recommend that you underpin your punchline – do you address other potential determining factor than these studies? Have most of these studies been conducted in the US, meaning that we lack knowledge on determinants of fatigue in European countries? Or? - Line 46, page 4: Why do you focus on a sample aged 45-86 (I recommend you move this information to the paragraph ‘Study population’? Population and Methods Please consider/address: - Study population:

	o Line 8-23, page 5: I'm a little in doubt how your study/your sample relates to the CoLaus study - please elaborate (a flow diagram (including the 'supplementary figure 1: The reason for exclusion) could be very helpful as well as including the paragraph 'Exclusion criteria'). - Fatigue scale: o Please explain in a short sentence how the fatigue severity scale fits your definition of fatigue/suits the purpose of your study. o Please provide Cronbach alphas for FSS (all scale used) - Covariates: o Please make sure that the introduction underlines your decisions on potential confounders. □ Hereunder, why is depression is included as the only mental condition – other conditions are also relevant (e.g. stress, burnout, anxiety) – which you also indicate in your introduction? Why are lifestyle factors not mentioned in the introduction, like smoking which is mentioned under 'Covariates'? o Why are work-related factors not considered (e.g. shiftwork/working night time). And children (i.e. having small children)? I recommend you include these potential confounding factors into your analyses. If that is not possible please address this/these potential confounders in the discussion (under strength and limitations). o In the paragraph 'Results' and the two tables caffeinated drinks, cardiovascular disease, diabetes etc. appears, but these variables are not described in 'Covariates' – please make sure that 'Introduction', 'Covariates' and 'Results' (incl. tables and the footnote you provide for tables) agree. o Line 17-19, page 6: Please provide a reference for your statement that 'It has been used in other…..and cardiovascular risk factors' - Statistical analysis: o line 15-19, page 8: The sentence 'All categorical variables…..were included in the multivariable analyses' please specify on what level ($p<0.05$?) o Line 38, page 8: The sentence 'Statistical significance was…..test with $p<0.05$' please specify what this relates to. Results - Table 2 (line 2-56, page 10): Under 'Bivariate' it says 'No' and 'Yes' – I suppose that relates to fatigue. Please make sure that the categories used in table 1 and table 2 agree. Discussion - In general. Please provide study design and participants for the studies that you relate to (eg. line 40-44, page 12 (Loge et al and Lerdal et al) as well as line 11-20, page 13 (Engberg). It would be very helpful for the reader when comparing the results. Please also specify, argue why results from studies conducted in other settings, with different populations/participants and using different fatigue scales are relevant with respect to the results of your study (Line 50-54, page 12 (Pawlikowska et el and Bultmann et al (please notice the study by Bultmann et al is conducted in Maastricht, the Netherlands and not in Denmark). - Clinical and social determinants of fatigue: o Line 40-44, page 14: This sentence is redundant - please erase it - as this fact is addressed in 'Strengths and limitations' - Implications for clinical practice: You aim is to address prevalence of fatigue and its determinants in a population based sample – which means that many of your participants are healthy

	individual while others are faced with physical and/or mental conditions (which is also indicated in table 1 and 2). Therefore, I encourage you to also focus on prevention – and along those lines argue why preventing fatigue is at all needed in Lausanne (you find that the prevalence of fatigue is 22.1% - which is in accordance with the findings of others. From that one could speculate that 'this may just a condition of life – for instance due to children, gran children, work, spear time activities, illness….' - Strengths and limitations: o Line 9-11, page 16: The sentence 'Secondly, the age group……of interest for general practitioners and internists' – with respect to my comments on 'Implications for clinical practice' – I disagree that this is one of the main strength of your study. o Line 29-31, page 16: The sentence 'Hence, it is possible….or over-estimated'. Please also discuss how using the French version of FSS (non-validated) may have affected your results? And please address 'how likely is it is it to be problem that the French version of FSS has not been validated' – do you have reason to believe that this is a problem (if yes, why/in no, why). o As for strengths and limitations: I encourage you to also address the following matters: □ The definition of fatigue you have applied vs. the fact that fatigue tend to be hard to define (and therefore also to measure) • How does the definition you have chosen apply to the tool you have chosen (FSS) o How does this relate to your sample? o What about other (standardized) fatigue scales? o How does FSS relates to the other tools you apply (FSS asks respondents to 'refer to your usual way of life within the last week', while ISI asks respondents to refer to the last month) – how does that relate? □ Self-administered questionnaire - please discuss all bias' to be considered □ Confounders – the relevance of those you have considered and potential relevant once that you have not considered (such as other mental conditions than depression, work-related matters). □ Please discuss the findings from your sensitivity analysis against your results – what does it indicate that excluded participants were more frequently women, were older etc.? □ Validity – internal and external o As a concluding remark; your findings are – to a large extend – in accordance with the findings of others. Hence, I will encourage you to discuss this fact - including what you would suggest in terms of future research (what should future research aim at (Prevalence of fatigue? Determinants of fatigue? Planning, implementation and evaluation of fatigue preventive strategies? What study design to be used? What Settings/populations to focus on (general populations, workers, general practice attendees, others – and in this specific subgroup (e.g. older/younger age, gender male/female)? Other?). You may want to state that your findings are – in a large extend - in accordance with those of others. Then mention agreements (something like 'Thus for instance we found that obese subjects reported more fatigue (references) which we also found to be the case for those who also reported insomnia, depression and antidepressant medication (reference). Move on to disagreements/controversial findings (contrary to prior findings, we found that…./hypothyroidism is often cited…). And then present a discussion of 'why this is the case and what it implies' followed by strength and limitations of your study.

REVIEWER	Anja Leppin University of Southern Denmark Faculty of Health Sciences Denmark
REVIEW RETURNED	11-Feb-2019

GENERAL COMMENTS	The present paper presents the results of a population-based cross-sectional survey investigating fatigue and factors potentially associated with fatigue experience. There are comparatively few studies which have investigated fatigue in general populations, so the present study could potentially make a contribution to the field. However there are a series of relevant issues which need clarification/change. Background 1. The introduction of the fatigue concept is a bit too condensed and unsystematic. Particularly for readers not entirely familiar with the topic it may, for instance, be confusing to first read that “fatigue is usually defined as an unpleasant…. symptom … not relieved by common strategies that restore energy” and then two sentences later “In healthy subjects, fatigue is a natural occurrence after physical or mental efforts, … and is usually relieved by rest.” Especially the differentiation of fatigue from tiredness/sleepiness and the specific role of fatigue as a consequenc of disease and/or treatment should be made clearer. 2. The “hypothesis” put forward is not really a hypothesis, i.e. such a statement would need more precision than “fatigue would be relatively prevalent and associated with several clinical, biological and socio-demographical characteristcs”. At least “the terms “relevant” and “several” would require specification as well as a theoretical/empirical rationale. Else, using “research questions” might be a better choice altogether, also given the cross-sectional research design. Population and methods 3. It does not become clear how the study population was recruited. The recruitment period named is quite long, i.e. May 2014 to April 2017 and needs explanation. Also, no information has been provided about whether this was an online survey, a mail survey etc., how respondents were selected/approached. It is stated that “detailed descriptions of the study design have been reported elsewhere”, but the very essentials should be repeated here, at least briefly. Measurement 4. The way fatigue was measured is crucial and should be described in a bit more detail than has been provided here. The FSS is one of the oldest fatigue scales but not every reader can be expected to be familiar with it. In particular, what exactly does the questionnaire measure in terms of item content? Example items should be provided. Further, in the background section the authors refer to a definition of fatigue as a multidimensional phenomenon but this is not reflected on the operational level. What has been the rationale for choosing a unidimensional instrument? 5. It is mentioned that the questionnaire has been validated in a Swiss setting. This might be misleading, since later, in the limitations section, the authors write that only a German version has been validated (in Switzerland?), not the French one, which presumably has been used in the present study given that the

	study was conducted in Lausanne. In any case, even if no validation study for the French version was available, to which extent are the populations involved here and in the validation study comparable (beyond the language difference)? Further, information is lacking about which translation procedures were used and in which ways equivalence was ensured. 6. Information should be added on whether the cut-off score of 4 for the fatigue scale is based on a clinical criterion. 7. The text mentions that “lifestyle variables were collected” but there is only information on measurement of smoking status, even though the results/tables suggest that other variables such as caffeine intake were assessed. 8. The description about the Insomnia Severity Index is missing information about the type of response scale, i.e. is it assessing frequency, intensity? 9. The frequency response scale of the CESD is unipolar, thus not a Likert scale. 10. Under exclusion criteria it is stated that those persons who lacked either one of the relevant variables, i.e. for instance the fatigue questionnaire score or clinical covariates etc. were excluded. (Supplementary figure 1 adds helpful information about this). At first glance, these missing participants might seem more like drop-outs but in any case, information is needed about the reasons for these quite large numbers of “excluded” participants. Results 11. The presentation of the tables is a bit confusing because the annotations (text under the tables) have been set in normal text font so that they appear as part of the “normal” text. 12. In general, the description of the statistical analyses should be under the respective paragraph in the methods section and not be repeated (or appear first) in the results section. 13. Table 1 may be expendable, since the relevant variables have also been included (in categorized form) in table 2, which contains the findings of the logistic regression model. 14. In table 1, why have age and BMI not been carried forward to the multivariable analysis? 15. Among the covariates tested, quality of life is listed, but does not appear in the results table and neither has any information on measurement been provided. Discussion 16. The comparison with other prevalence findings is difficult, as the authors themselves note, since different assessment instruments have been used. It would be very interesting if some information could be added on what these differences are and how they might have affected findings. 17. The initially found gender differences disappeared in the multivariable analysis but then only an explanation for the bivariate result is offered. Either only the multivariable results should be interpreted. Or - if the relevance of gender as an influence factor is to be debated on a larger scale - explanations about which of the covariates (for instance depression?) might have made the difference should have been provided. 18. The authors explain their finding that older people reported less fatigue than younger ones, which stands in contrast to prior studies from 1990/2000, with increased quality of life in the elderly. This is certainly interesting, but is there evidence for a significant change in Quol in the elderly within the last 15 years? Could there also be other explanations? To which extent could, for

	instance, the restricted age range in the present study (45 upwards) have contributed to the finding? 19. The association with depression – even though found before and seemingly "obvious" – deserves more discussion particularly regarding the "symptom overlap". 20. As for the positive association between depression medication and fatigue, it is probably not only – as suggested by the authors – that medication does not always remove fatigue symptoms but in fact a common side effect of some types of antidepressant medication is tiredness/sleepiness. 21. The suggestions for clinical practice are interesting but the rationale for the suggested sequence should be made explicit. If this is based on the ORs found, why for instance prioritize depression over health status (particularly if the latter one can be assessed with a single item)? Limitations 22. In general, the limitations of the study and their implications deserve a more throrough discussion. One issue not mentioned is the fact that the study was conducted in Lausanne only. But how representative is the population of one city in the French speaking part of Switzerland for Switzerland as a whole? 23. Also, despite the supplementary analysis for selection bias with propensity scores, implications (or eventually non-relevance) of the considerable exclusion/drop-out from the initial study population should be explained/discussed. 24. Further, there seems to have been a prior approximate 25% dropout from baseline to the 2nd follow-up, which provided the data used for the present study. This drop-out may also be suspected to have been correlated with fatigue levels, and/or depression and/or health status and should therefore be discussed. 25. Given that the study took place in Lausanne, I suppose it was the French version of the FSS which was used, which, however, was not validated yet. The possible implications may warrant a more in-depth discussion than just one sentence that this might have led to over- or underestimation. 26. Recruitment period was 3 years. Why was the period so long and might this have affected the findings? Conclusion 27. This is a mere summary/repeat of the findings and not a conclusion. General issues: 28. The title of the paper "A Swiss population study" might be somewhat misleading given that the survey was conducted only in Lausanne and did not include the entire age range of the adult population either but just the middle-aged to elderly. 29. In general, the manuscript needs editing regarding use of English language, quite many missing words, incomplete sentences, e.g. on page 6, 2nd para: "Depression was assessed… (by using?) the CES-D ….(which?) is a 20 item self-report instrument developed for research in the general population … (and which?) is used to assess the severity of depressive symptoms over the past week on a 4-point scale. 30. Given the cross-sectional character of the study and the conceptual overlap of some of the covariates with fatigue, the wording should (in parts) be more conservative, i.e. it would for

	instance be more appropriate to speak of associations and potential determinants rather than of “determinants”. Minor issues: 31. The headlines of the tables should describe the contents, i.e. the associations investigated rather than refer to the statistical analysis. 32. Table 2: Under multivariable “OR” is missing in the headline.

VERSION 1 – AUTHOR RESPONSE

Answers to reviewer: 1, James N. Baraniuk, MD

1) Abstract: "Primary outcome measure: Prevalence of chronic fatigue, defined as a score ≥4 using the Fatigue severity scale (FSS)."

We changed the text to: “Primary outcome measure: Prevalence of fatigue, defined as a score ≥4 using the Fatigue severity scale (FSS)."

2) page, line 28: “Fatigue severity during the last week….”

We changed the text to “Fatigue severity the previous week…”

3) In the abstract and conclusion you claim that 20% of the population has chronic fatigue, yet they were asked about fatigue in the past week. Do you define chronic fatigue as 1 week? Such a conclusion may have been valid if you had fatigue measurements to compare from the first and this, the second, samplings. However, fatigue was only assessed the second time. Please clarify.

We agree with the reviewer that we did not assess « chronic » fatigue, but only fatigue during the previous week. Hence, we removed all occurrences of “chronic” in the abstract and in the main text and indicated that only fatigue the previous week was assessed.

4) page 9, line 9: You use the original cut-off score ≥4. This should be repeated on page 9 lines 9-10 because your prevalence values will be considered important and should be qualified by the level of fatigue, particularly because you contrast your results to Lerdal et al. who used a higher cut-off (≥5). Your cut-off may also be mentioned in the legend to Table 2.

We added the statement “as defined by a FSS ≥4” as requested in the text and in the legend of table 2

5) The logistic regression and bivariate analysis would be strengthened by also showing the results with the higher cut-off (≥5) so that the level of fatigue, exclusion, and co-variates can be understood when stratified by fatigue severity. Your most significant features could be plotted as the % against increasing fatigue scores starting at 0 show the stratification of the most significant

variables by fatigue level. This is relevant because fatigue can change over time. The stratified data would be of interest in general practice to emphasize the characteristics of each stratum of fatigued patients. Showing the prevalence of fatigue at each level of the FSS would be very useful for planning future epidemiological studies of fatigue.

We now provide the results for the higher cut-off (Supplemental tables 3 and 4) and a histogram (supplemental figure 2) with the distribution of the FSS values. We also modified the discussion to take into account the findings using the higher cut-off (≥5)

We decided not to plot the results according to the cut-off for the following reasons

1) It was not the objective of the study to test the different cut-offs
2) All the literature used either the 4 or the 5 cut-off, so the rationale for using other cut-offs would be difficult to justify.
3) The high number of associations tested would raise the issue of multiple testing
4) The extra tables and text would increase considerably the manuscript size, making it no longer compliant with the journal's recommendations

Still, if the editor considers that this is an important issue to develop, we will perform the analysis.

6) page 8, line 27: The propensity index is mentioned in the methods but I do not see it in the results or discussion.

The propensity index was used for inverse probability weighting and the results are presented in supplementary table 2. We changed the sentence in the text to

"Sensitivity analysis using inverse probability weighting by the propensity score led to similar findings, except that anaemia and antidepressants were no longer associated with fatigue, while a positive association was found between low TSH levels and fatigue (Supplemental table 2)."

7) page 16, Line 33: "our results provide a first estimation of the prevalence of fatigue in the general population," However, you report similar findings from other studies on page 12 lines 38 to 54. Please clarify.

The reviewer is correct. We mean the "first estimation of the prevalence of fatigue in the Swiss general population". This has been corrected

8) The high exclusion rate is a significant limitation.

We agree. We added the following statement in the limitations chapter. "Thirdly, a sizable fraction of the sample was excluded, which might limit the generalizability of the findings. Still, an analysis using a propensity score weighting for the probability of being excluded led to similar findings."

9) Conclusion: Comment on the almost 40% of the surveyed population who were excluded from the study because they may have biased the outcomes.

We added the following sentence at the end of the conclusion: “The results should be interpreted taking into account the high exclusion rate.”

10) Supplemental table 1: Add odds ratios for exclusion variables

The results are now provided in supplemental table 2.

Answers to reviewer: 2, Solveig Dohrmann

Thank you for presenting this manuscript, I have read it with great interest.

This manuscript can contribute important knowledge on the prevalence of fatigue and on its determinants in the general population. Therefore, I believe that the manuscript should be published by BMJ Open. However, there are areas that need attention and therefore I recommend revision before publication.

On the positive side the method (survey) suits the purpose. With respect to the overall structure, the manuscript is well organized with heading, subheadings and tables, which all contributes to making the manuscript really reader friendly. Further, the manuscript is well articulated, and almost always logical.

Introduction

1. You provide a definition, state that fatigue varies in duration and intensity and provide prevalence rates (line 5-15, page 4). Please elaborate – add a brief explanation why fatigue is a health concern (with a specific focus on the purpose of your study (exploring determinants of cardiovascular disease)), make a clear distinction between acute fatigue, fatigue of a more chronic nature and chronic fatigue syndrome, and based on the definition you provide specify what type of fatigue you focus on.

We added the following text in the introduction:

“In healthy subjects, tiredness or sleepiness are a natural occurrence after physical or mental efforts, and are usually relieved by rest. [1 2] While fatigue is defined as extreme and persistent tiredness, weakness or exhaustion, of mental and/or physical origin [3] that is not relieved by rest. Fatigue is defined in duration as recent (<1 month) prolonged (1 to 6 months) and chronic (>6 months) [4]. When unexplained, chronic fatigue can be considered either as a syndrome (characterized by severe, disabling fatigue and other symptoms, including musculoskeletal pain, sleep disturbance, impaired concentration, and headaches) [5] or as idiopathic (absence of other symptoms).

Fatigue is one of the most common complaints reported in primary care [6] and is associated with a decreased quality of life and increased morbidity and mortality in the general population. [7]“

2. Line 11-13, page 4: You state that fatigue is a common symptom in the general population, and along line 36-37, page 4 you state that only two studies have assessed the prevalence of fatigue in the general population – please elaborate.

While several studies have assessed the prevalence of fatigue in different settings, only two actually assessed the prevalence of fatigue in the general population. We agree that the statement lines 11-13 is misleading and we changed it to “Indeed, fatigue is a common symptom, with prevalence rates varying between 4 and 45%.[2-4]This ten-fold range in prevalence rates is likely due to the different settings (i.e. general practice .[5] or workers [6]) or to the different methods used to assess fatigue.[7]

3. Line 13-15, page 4: I recommend that you move the last sentence ‘this ten-fold range….. assess fatigue’ to the discussion, and discuss this issue in greater details.

As we used this sentence to elaborate as requested in the previous query, we decided to keep it in the same place.

4. Line 37-39, page 4: I recommend that you underpin your punchline – do you address other potential determining factor than these studies? Have most of these studies been conducted in the US, meaning that we lack knowledge on determinants of fatigue in European countries? Or?

We added the following statement in the introduction: “Further, most studies focused on socio-economic and disease determinants of fatigue, while information regarding the biological determinants (i.e. anemia or thyroid pathology) [7]or the medications associated with fatigue is scarce.”

5. Line 46, page 4: Why do you focus on a sample aged 45-86 (I recommend you move this information to the paragraph ‘Study population’?

We deleted the age range in the introduction and added a sentence in the Methods, Study population, end of the paragraph: “At the second follow-up, participants were aged 45-86 years.”

Population and Methods

1) Study population:

    a) Line 8-23, page 5: I’m a little in doubt how your study/your sample relates to the CoLaus study - please elaborate (a flow diagram (including the ‘supplementary figure 1: The reason for exclusion) could be very helpful as well as including the paragraph ‘Exclusion criteria’).

Our study is part of the CoLaus study, as it relies on the second follow-up of the CoLaus study. The reasons for exclusion are provided in the Supplementary figure 1, and we would be very grateful if the reviewer could provide some hints on how we can elaborate more on the flowchart. We increased the paragraph Exclusion criteria to “Participants were excluded if they lacked 1) any answer to the fatigue questionnaire; 2) clinical data such as age, body mass index, smoking, depression, insomnia or medications; 3) biological measures such as haemoglobin or thyroid hormones and 4) socioeconomic data such as educational level.“

2) Fatigue scale:

    a) Please explain in a short sentence how the fatigue severity scale fits your definition of fatigue/suits the purpose of your study.

We added the following text in the methods: “The FSS is one of the most commonly used fatigue questionnaires. It had been validated in a general population setting in German-speaking Switzerland

[8], Portugal [9] and Norway [10]. It is a simple, time-saving, self-administrated questionnaire allowing its use in large epidemiological studies and has a high test-retest reliability. [3]"

b) Please provide Cronbach alphas for FSS (all scale used)

We added the following statement in the methods, fatigue scale: “The Cronbach’s alpha for the questionnaire was 0.918, suggesting an excellent internal consistency.”

3) Covariates:

a) Please make sure that the introduction underlines your decisions on potential confounders.

Please see our response to query 3.d)

b) Hereunder, why is depression is included as the only mental condition – other conditions are also relevant (e.g. stress, burnout, anxiety) – which you also indicate in your introduction? Why are lifestyle factors not mentioned in the introduction, like smoking which is mentioned under ‘Covariates’?

Please see our response to query 3.d)

c) Why are work-related factors not considered (e.g. shiftwork/working night time). And children (i.e. having small children)? I recommend you include these potential confounding factors into your analyses. If that is not possible please address this/these potential confounders in the discussion (under strength and limitations).

We have no information regarding the shiftwork or the number of small children. Still, as the age range of the sample is 45-86 years, it can be estimated that 1) the number of participants with small children is small, and 2) a sizable fraction of the sample (>40%) is retired. We added the following statement in the limitations paragraph: “Fourthly, no information was available regarding shift work or the presence of very young children. Still, as a sizable fraction (almost 70%) of the sample was aged over 55 and over 36% of the sample was aged over 64, it is likely that the number of participants either on shift work or with very young children would be small. "

d) In the paragraph ‘Results’ and the two tables caffeinated drinks, cardiovascular disease, diabetes etc. appears, but these variables are not described in ‘Covariates’ – please make sure that ‘Introduction’, ‘Covariates’ and ‘Results’ (incl. tables and the footnote you provide for tables) agree.

We added the following statement in the Introduction: “Further, most studies focused on socio-economic and disease determinants of fatigue, while information regarding the biological determinants (i.e. anemia or thyroid pathology) [7] or the medications associated with fatigue is scarce.”

We added the following statement in the Covariates:

“Caffeinated drinks consumption was assessed by the question “How many cups or cans of drinks containing caffeine (coffee, tea, coke or similar) do you drink per day?”, with possible answers “None”, “1-3”, “4-6” and “7 or more”.

Participants were asked to report all medications (prescribed or bought over the counter) they took during the last 6 months. Medications were coded using the Anatomical, Therapeutic Chemical (ATC) classification of the world health organization (www.whocc.no/atc_ddd_index/). Antihistamics were defined as any ATC code beginning with “R06”; antidepressants were defined as an ATC code beginning with “N05BD" or N06AA" or “N06AB" or “N06AF" or “N06AG" or “N06AX" or “N06CA"; hypnotics were defined as any ATC code beginning with “N05C”. Antihypertensive drugs were defined by asking the participants if they were taking drugs for hypertension.

Diabetes was defined by a fasting plasma glucose ≥7 mmol/L and/or the presence of an antidiabetic drug treatment (oral or insulin). Personal history of CVD was assessed by asking the participant if he/she had sustained a coronary event (myocardial infarction or angina pectoris) or a stroke.”

e) Line 17-19, page 6: Please provide a reference for your statement that ‘It has been used in other…..and cardiovascular risk factors’

We now provide the following reference: Hamieh N, Meneton P, Wiernik E, Limosin F, Zins M, Goldberg M, et al. Depression, treatable cardiovascular risk factors and incident cardiac events in the Gazel cohort. Int J Cardiol. 2018.

4) Statistical analysis:

a) line 15-19, page 8: The sentence ‘All categorical variables…..were included in the multivariable analyses’ please specify on what level ($p<0.05$?)

We changed the text to “All categorical variables significantly ($p<0.05$) associated with fatigue in the bivariate analysis were included…”

b) Line 38, page 8: The sentence ‘Statistical significance was…..test with $p<0.05$’ please specify what this relates to.

This level relates to all tests performed. To our knowledge, this is a common statement among published papers. Still, we will very happy if the reviewer could provide some examples of different statements.

Results

1) Table 2 (line 2-56, page 10): Under ‘Bivariate’ it says ‘No’ and ‘Yes’ – I suppose that relates to fatigue. Please make sure that the categories used in table 1 and table 2 agree.

We changed the column headings of table 2 to “No fatigue” and “Fatigue”.

Discussion

1) In general. Please provide study design and participants for the studies that you relate to (eg. line 40-44, page 12 (Loge et al and Lerdal et al) as well as line 11-20, page 13 (Engberg). It would be very helpful for the reader when comparing the results. Please also specify, argue why results from studies conducted in other settings, with different populations/participants and using different fatigue scales are relevant with respect to the results of your study (Line 50-54, page 12 (Pawlikowska et el and Bultmann et al (please notice the study by Bultmann et al is conducted in Maastricht, the Netherlands and not in Denmark).

We changed the first paragraph of chapter “Prevalence of fatigue” to

“Using the cut-off of ≥4, fatigue was present in one out of five participants (22.1%), a finding in agreement with the cross-sectional study by Loge et al. [1], which reported a prevalence of 22% among 2323 participants using the Chalder fatigue scale. Conversely, while the cross-sectional study by Lerdal et al. [10], which used the FSS in a sample of 1893 participants, reported a prevalence of fatigue of 46.7% and 23.1% using a cut-off of ≥4 and ≥5, respectively (to compare with 22.1% and 10.9% in our study). A study conducted in general practice attendees reported a prevalence of fatigue of 38% using the Chalder fatigue scale,[11] and a study conducted in the Dutch working population reported a prevalence of fatigue of 22% using other fatigue measures. [12] Overall, our results suggest that the prevalence of fatigue in the Lausanne population is similar or lower than reported previously, although the use of different scales to assess fatigue complicates comparison between studies.”

Reviewer #3 also commented on this.

2) -Clinical and social determinants of fatigue:

   a) Line 40-44, page 14: This sentence is redundant - please erase it - as this fact is addressed in ‘Strengths and limitations’

The sentence has been deleted.

3) -Implications for clinical practice: You aim is to address prevalence of fatigue and its determinants in a population based sample – which means that many of your participants are healthy individual while others are faced with physical and/or mental conditions (which is also indicated in table 1 and 2). Therefore, I encourage you to also focus on prevention – and along those lines argue why preventing fatigue is at all needed in Lausanne (you find that the prevalence of fatigue is 22.1% - which is in accordance with the findings of others. From that one could speculate that ‘this may just a condition of life – for instance due to children, gran children, work, spear time activities, illness….’

We completely rewrote the chapter “implications for clinical practice”, which is now “implications of the study”. The chapter now reads:

“Based on our study findings, we propose to focus on specific clinical and biological factors amenable to treatment at the individual level. Regarding clinical factors, sleep disturbances such as insomnia and sleep apnea (namely in presence of a patient with obesity) and the presence of depression should be assessed. Hence, lifestyle measures to improve sleep quality and quantity should be preferred to medication.[13] In case of depression, it will be important to warn patients that antidepressor medication might not necessarily lead into rapid relief of fatigue. Regarding biological factors, anemia should be ruled out, while screening for hypothyroidism is not recommended as a first step.

At the population level, preventive measures such as stress management and health promotion like relaxation, time management and cognitive reframing (for example within the work environment) could improve sleep quality, increase self-rated health {Hasson, 2005 #615} and so reduce fatigue.“

Reviewer #3 also commented on this topic.

4) Strengths and limitations:

a) Line 9-11, page 16: The sentence ‘Secondly, the age group……of interest for general practitioners and internists’ – with respect to my comments on ‘Implications for clinical practice’ – I disagree that this is one of the main strength of your study.

The statement has been deleted.

b) Line 29-31, page 16: The sentence ‘Hence, it is possible….or over-estimated’. Please also discuss how using the French version of FSS (non-validated) may have affected your results? And please address ‘how likely is it is it to be problem that the French version of FSS has not been validated’ – do you have reason to believe that this is a problem (if yes, why/in no, why).

Please see our reply to query 4) c) ix)

c) As for strengths and limitations: I encourage you to also address the following matters:

i) The definition of fatigue you have applied vs. the fact that fatigue tend to be hard to define (and therefore also to measure)

Please see our reply to query 4) c) ix)

ii) How does the definition you have chosen apply to the tool you have chosen (FSS)

Please see our reply to query 4) c) ix)

iii) How does this relate to your sample?

We failed to understand what the reviewer means by this question. Could you please elaborate?

iv) What about other (standardized) fatigue scales?

To our knowledge, there is no fatigue scale standardized or validated for the French-speaking Swiss population. We would be very glad if the reviewer could provide us some fatigue scales validated in French-speaking Switzerland.

v) How does FSS relates to the other tools you apply (FSS asks respondents to 'refer to your usual way of life within the last week', while ISI asks respondents to refer to the last month) – how does that relate?

Please see our reply to query 4) c) ix)

vi) Self-administered questionnaire - please discuss all bias' to be considered.

Please see our reply to query 4) c) ix)

vii) Confounders – the relevance of those you have considered and potential relevant once that you have not considered (such as other mental conditions than depression, work-related matters).

Please see our reply to query 4) c) ix)

viii) Please discuss the findings from your sensitivity analysis against your results – what does it indicate that excluded participants were more frequently women, were older etc.?

Please see our reply to query 4) c) ix)

ix) Validity – internal and external

We modified the text in the limitations paragraph as follows:

"This study has also several limitations. Firstly, its cross sectional setting precludes the identification of the causes of fatigue, as reverse causality is possible (i.e. fatigue leading to depression and vice-versa).[14] All participants of the CoLaus study are currently being re-contacted and re-examined, so that a prospective analysis of the causes of fatigue will be feasible within two years. Secondly, there is no gold standard for the evaluation of fatigue and no official definition of fatigue. Hence, results might vary according to the scale applied or how participants interpret the term "fatigue". In this study, we chose to use a scale that was previously applied by other authors to facilitate comparisons. Thirdly, only the German version of the FSS has been validated in Switzerland; the French version used in this study has not yet been validated. Hence, it is possible that the true prevalence levels of fatigue might be under- or over-estimated, or that some items of the questionnaire might not be informative. Still, the Cronbach's alpha for the questionnaire was 0.918, suggesting an excellent internal consistency. Further, our results provide a first estimation of the prevalence of fatigue in the Swiss French-speaking general population, which could serve as a reference for further studies. Fourthly, a sizable fraction of the sample was excluded, both between the baseline and the second follow-up, and within the current study, which might limit the generalizability of the findings. For instance, excluded participants were more frequently women; as women reported more frequently fatigue, this might lead to an underestimation of prevalence rates or a decrease in the strength of the associations. Still, an analysis using a propensity score weighting for the probability of being excluded led to similar findings. Conversely, it was not possible to assess the reasons why participants did not complete the questionnaire. Fifthly, no information was available regarding shift work or the presence of very young children. Still, as a sizable fraction (almost 70%) of the sample was aged over 55 and over 36% of the sample was aged over 64, it is likely that the number of participants either on shift work or with very young children would be small. Sixthly, the FSS explored fatigue during the previous

week while the ISI score explored the sleep during the previous month. Hence, it is possible that the time association between the two variables might not be optimal. Still, as the FSS lies within the period encompassed by the ISI, we believe that the associations obtained are clinically relevant. Seventhly, the study is limited to the population of Lausanne aged 45 to 86, and its generalizability remains to be assessed. For instance, no information was collected regarding other confounders among younger subjects, where prevalence of fatigue might be higher due to parental and professional duties.[15]. Finally, possible biases related to the self-reporting of fatigue could not be avoided, such as over- or under-estimation of symptoms or misunderstanding of what the term "fatigue" meant; still, this dilution bias would lead to a decrease in the strength of the associations, and it would be too restrictive in our opinion to provide a definition of the term "fatigue" to the participants, as different interpretations of the definition itself could also occur."

d) As a concluding remark; your findings are – to a large extend – in accordance with the findings of others. Hence, I will encourage you to discuss this fact - including what you would suggest in terms of future research (what should future research aim at (Prevalence of fatigue? Determinants of fatigue? Planning, implementation and evaluation of fatigue preventive strategies? What study design to be used? What Settings/populations to focus on (general populations, workers, general practice attendees, others – and in this specific subgroup (e.g. older/younger age, gender male/female)? Other?). You may want to state that your findings are – in a large extend - in accordance with those of others. Then mention agreements (something like 'Thus for instance we found that obese subjects reported more fatigue (references) which we also found to be the case for those who also reported insomnia, depression and antidepressant medication (reference). Move on to disagreements /controversial findings (contrary to prior findings, we found that…./hypothyroidism is often cited…). And then present a discussion of 'why this is the case and what it implies' followed by strength and limitations of your study.

We added a new chapter "Recommendations for future studies" with the following text:

"Future studies on the prevalence of fatigue in the general population should focus on the following topics: 1) validate the questionnaires in the population of interest; 2) whenever possible, use a standard questionnaire to allow comparison between studies.

While some factors such as obesity [7 16], depression [7 17-19] and antidepressor medications [14] were consistently associated with fatigue in our study and in the literature, controversial findings such as the (absence of) association between fatigue and gender, age groups and anemia should be further explored."

Answers to reviewer: 3, Anja Leppin

The present paper presents the results of a population-based cross-sectional survey investigating fatigue and factors potentially associated with fatigue experience. There are comparatively few studies which have investigated fatigue in general populations, so the present study could potentially make a contribution to the field. However there are a series of relevant issues which need clarification/change.

Background

1. The introduction of the fatigue concept is a bit too condensed and unsystematic. Particularly for readers not entirely familiar with the topic it may, for instance, be confusing to first read that "fatigue is usually defined as an unpleasant…. symptom … not relieved by common strategies that restore energy" and then two sentences later "In healthy subjects, fatigue is a natural occurrence after physical or mental efforts, … and is usually relieved by rest." Especially the differentiation of fatigue from tiredness/sleepiness and the specific role of fatigue as a consequence of disease and/or treatment should be made clearer.

We added the following text in the introduction:

"In healthy subjects, tiredness or sleepiness are a natural occurrence after physical or mental efforts, and are usually relieved by rest. [1 2] While fatigue is defined as extreme and persistent tiredness, weakness or exhaustion, of mental and/or physical origin [3] that is not relieved by rest. Fatigue is defined in duration as recent (<1 month) prolonged (1 to 6 months) and chronic (>6 months) [4]. When unexplained, chronic fatigue can be considered either as a syndrome (characterized by severe, disabling fatigue and other symptoms, including musculoskeletal pain, sleep disturbance, impaired concentration, and headaches) [5] or as idiopathic (absence of other symptoms).

Fatigue is one of the most common complaints reported in primary care [6] and is associated with a decreased quality of life and increased morbidity and mortality in the general population. [7]"

We focused on recent fatigue as defined in the methods "Fatigue severity during the last week."

2. The "hypothesis" put forward is not really a hypothesis, i.e. such a statement would need more precision than "fatigue would be relatively prevalent and associated with several clinical, biological and socio-demographical characteristcs". At least "the terms "relevant" and "several" would require specification as well as a theoretical/empirical rationale. Else, using "research questions" might be a better choice altogether, also given the cross-sectional research design.

We deleted the sentence on the hypothesis. Rephrasing it to "research questions" would make the sentence look very similar to the previous one.

Population and methods

3. It does not become clear how the study population was recruited. The recruitment period named is quite long, i.e. May 2014 to April 2017 and needs explanation. Also, no information has been provided about whether this was an online survey, a mail survey etc., how respondents were selected/approached. It is stated that "detailed descriptions of the study design have been reported elsewhere", but the very essentials should be repeated here, at least briefly.

We added the following information in the methods:

"At both baseline and subsequent follow-ups, participants were invited to attend a clinical examination at the Lausanne university hospital. Participants received at home a paper questionnaire, which they filled prior to come to the clinical examination. During the clinical examination, a second questionnaire regarding personal and family history of cardiovascular disease and cardiovascular risk factors was applied. For more details, please consult www.colaus-psycolaus.ch."

Due to financial and logistic limitations, it was not possible to recruit quicker. Further, as participants had to be in the fasting state for blood drawing, examinations could only be conducted in the morning. Each examination took an average of one hour.

The time seems adequate compared to other cohorts: the Rotterdam Study I, sixth examination, took one year to examine 1000 participants. The Paris Prospective Study III took 4 years to examine 10,000 participants. Hence, we consider that three years to recruit 4881 participants with a research team of seven (including the PI, the data manager and the statistician) is reasonable.

Measurement

4. The way fatigue was measured is crucial and should be described in a bit more detail than has been provided here. The FSS is one of the oldest fatigue scales but not every reader can be expected to be familiar with it. In particular, what exactly does the questionnaire measure in terms of item content? Example items should be provided. Further, in the background section the authors refer to a definition of fatigue as a multidimensional phenomenon but this is not reflected on the operational level. What has been the rationale for choosing a unidimensional instrument?

   We now provide in annex 1 the French version of the questionnaire.

5. It is mentioned that the questionnaire has been validated in a Swiss setting. This might be misleading, since later, in the limitations section, the authors write that only a German version has been validated (in Switzerland?), not the French one, which presumably has been used in the present study given that the study was conducted in Lausanne. In any case, even if no validation study for the French version was available, to which extent are the populations involved here and in the validation study comparable (beyond the language difference)? Further, information is lacking about which translation procedures were used and in which ways equivalence was ensured.

We added the following statement in the methods: “To our knowledge, the French version of the FSS has not yet been validated in Switzerland. Still, the Cronbach’s alpha for the questionnaire was 0.918, suggesting an excellent internal consistency

6. Information should be added on whether the cut-off score of 4 for the fatigue scale is based on a clinical criterion.

We added the following statement in the methods: “This cutoff was initially proposed because <5% of healthy controls rate their fatigue at that level, whereas 60-90% of patients with medical disorders experience fatigue at or above this level [2]”

7. The text mentions that “lifestyle variables were collected” but there is only information on measurement of smoking status, even though the results/tables suggest that other variables such as caffeine intake were assessed.

We added the following statement in the Covariates:

“Caffeinated drinks consumption was assessed by the question “How many cups or cans of drinks containing caffeine (coffee, tea, coke or similar) do you drink per day?”, with possible answers “None”, “1-3”, “4-6” and “7 or more”.

Participants were asked to report all medications (prescribed or bought over the counter) they took during the last 6 months. Medications were coded using the Anatomical, Therapeutic Chemical (ATC) classification of the world health organization (www.whocc.no/atc_ddd_index/). Antihistamics were defined as any ATC code beginning with “R06”; antidepressants were defined as an ATC code beginning with “N05BD" or N06AA" or “N06AB" or “N06AF" or “N06AG" or “N06AX" or “N06CA"; hypnotics were defined as any ATC code beginning with “N05C”. Antihypertensive drugs were defined by asking the participants if they were taking drugs for hypertension.

Diabetes was defined by a fasting plasma glucose ≥7 mmol/L and/or the presence of an antidiabetic drug treatment (oral or insulin). Personal history of CVD was assessed by asking the participant if he/she had sustained a coronary event (myocardial infarction or angina pectoris) or a stroke.”

8. The description about the Insomnia Severity Index is missing information about the type of response scale, i.e. is it assessing frequency, intensity?

We changed the text to

“Insomnia was assessed using the Insomnia Severity Index (ISI).[20] The questionnaire has 16 items evaluating the nature, severity, and impact of insomnia over the last month; namely difficulties falling asleep, sleep maintenance problems, and early morning awakening, sleep dissatisfaction, interference of sleep disturbances with daytime functioning, noticeability of sleep problems by others, and distress caused by the sleep difficulties. Responses range from 0 (not at all) to 4 “Extremely”. Items were scaled 0-4 and then summed to obtain the global ISI score (range: 0-28). The questionnaire is provided in annex 2. Clinically significant…”

We also provide the French version of the ISI in annex 2.

9. The frequency response scale of the CESD is unipolar, thus not a Likert scale.

We deleted the text “using a Likert scale”

10. Under exclusion criteria it is stated that those persons who lacked either one of the relevant variables, i.e. for instance the fatigue questionnaire score or clinical covariates etc. were excluded. (Supplementary figure 1 adds helpful information about this). At first glance, these missing participants might seem more like drop-outs but in any case, information is needed about the reasons for these quite large numbers of “excluded” participants.

As the CoLaus questionnaire included a large number of questions (>350), a likely explanation is that participants got tired and did not respond to all items. Still, we did not collect the reasons why participants did not fill up the questionnaires and any explanation would be speculation. Hence, we decided not to develop this topic.

Results

11. The presentation of the tables is a bit confusing because the annotations (text under the tables) have been set in normal text font so that they appear as part of the “normal” text.

This is the requirement of BMJ Open and we have to comply with the journal’s guidelines.

12. In general, the description of the statistical analyses should be under the respective paragraph in the methods section and not be repeated (or appear first) in the results section.

We changed the statistical methods to

“Bivariate analyses were performed using chi-square or Fisher’s exact test for categorical variables and student’s t-test or Kruskal-Wallis test for continuous variables. All categorical variables significantly ($p<0.05$) associated with fatigue in the bivariate analysis were included in the multivariable analysis. Multivariable analysis was performed using analysis of variance or logistic regression with fatigue (dichotomized into yes/no) as dependent variable; results were expressed as multivariable-adjusted mean±standard error for continuous variables or as Odds ratio (OR) and 95% CI for categorical variables.”

13. Table 1 may be expendable, since the relevant variables have also been included (in categorized form) in table 2, which contains the findings of the logistic regression model.

As the other reviewers and the managing editor did not ask for the removal of the table, we decided to keep it.

14. In table 1, why have age and BMI not been carried forward to the multivariable analysis?

As indicated in the table footnote, “Multivariable analysis conducted using analysis of variance adjusting for gender, age group, BMI categories, insomnia categories, educational level, diabetes, presence of antihistaminic, antidepressive or hypnotic drugs, quality of life and depression.” Hence, age and BMI have been carried forward. We decided not to put the values as age and BMI were adjusted for themselves, but the results were the following

	Multivariable		
	No fatigue	Fatigue	p-value
Age (years)	61.4 ± 0.1	61.2 ± 0.1	0.167
BMI (kg/m^2)	26.4 ± 0.1	26.6 ± 0.1	0.049

15. Among the covariates tested, quality of life is listed, but does not appear in the results table and neither has any information on measurement been provided.

Quality of life (as indicated in the table footnote) stands for “Self-rated health”. We replace the occurrences of “quality of life” by “self-rated health”.

Discussion

16. The comparison with other prevalence findings is difficult, as the authors themselves note, since different assessment instruments have been used. It would be very interesting if some information could be added on what these differences are and how they might have affected findings.

We changed the first paragraph of chapter “Prevalence of fatigue” to

“Using the cut-off of ≥4, fatigue was present in one out of five participants (22.1%), a finding in agreement with the cross-sectional study by Loge et al. [1], which reported a prevalence of 22% among 2323 participants using the Chalder fatigue scale. Conversely, while the cross-sectional study by Lerdal et al. [10], which used the FSS in a sample of 1893 participants, reported a prevalence of fatigue of 46.7% and 23.1% using a cut-off of ≥4 and ≥5, respectively (to compare with 22.1% and 10.9% in our study). A study conducted in general practice attendees reported a prevalence of fatigue of 38% using the Chalder fatigue scale,[11] and a study conducted in the Dutch working population reported a prevalence of fatigue of 22% using other fatigue measures. [12] Overall, our results suggest that the prevalence of fatigue in the Lausanne population is similar or lower than reported previously, although the use of different scales to assess fatigue complicates comparison between studies.”

Reviewer #2 also commented on this.

17. The initially found gender differences disappeared in the multivariable analysis but then only an explanation for the bivariate result is offered. Either only the multivariable results should be interpreted. Or - if the relevance of gender as an influence factor is to be debated on a larger scale - explanations about which of the covariates (for instance depression?) might have made the difference should have been provided.

In the discussion, we only consider the variables that are significant in the multivariable analysis as clinically meaningful.

18. The authors explain their finding that older people reported less fatigue than younger ones, which stands in contrast to prior studies from 1990/2000, with increased quality of life in the elderly. This is certainly interesting, but is there evidence for a significant change in Quol in the elderly within the last 15 years? Could there also be other explanations? To which extent could, for instance, the restricted age range in the present study (45 upwards) have contributed to the finding?

We failed to find objective regarding changes in quality of life among elderly subjects in Switzerland. A study conducted in the canton of Valais among 3635 participants (VLV study) indicated that quality of life among Swiss elderly had increased in the last 30 years, but provided no objective data. We added the following statement in the text

“Although there is little information regarding trends in quality of life among Swiss elderly. The VLV study [21] concluded that quality of life among Swiss elderly increased in the last 30 years [22].”

We would be very glad if the reviewer could provide us more references on this topic.

19. The association with depression – even though found before and seemingly “obvious” – deserves more discussion particularly regarding the “symptom overlap”.

We added the following text to the discussion : « Furthermore, fatigue is a common side effect of antidepressant therapy and a symptom of depression, making the identification of the cause of fatigue difficult with a possibility of reverse causality (fatigue leading to depression and vice versa). We used a one-dimensional tool to evaluate fatigue (FSS); hence, we cannot distinguish between physical and mental fatigue. There is considerable overlap in phenomenology of fatigue and depression or anxiety but there are some important differences. People with fatigue without psychiatric symptoms tend to attribute their symptoms to external causes. Conversely, most depressed people experience self-blame or lowered self-esteem [23]. Fatigue and depression appear commonly together. A study conducted in 2009 by Harvey et al. [24], showed that 7% of fatigued persons have no psychiatric symptoms but remained at increased risk of later psychiatric disorder independently of the severity of fatigue."

20. As for the positive association between depression medication and fatigue, it is probably not only – as suggested by the authors – that medication does not always remove fatigue symptoms but in fact a common side effect of some types of antidepressant medication is tiredness/sleepiness.

We added in the text the following statement: "Furthermore, fatigue is a common side-effect of antidepressant therapy."

21. The suggestions for clinical practice are interesting but the rationale for the suggested sequence should be made explicit. If this is based on the ORs found, why for instance prioritize depression over health status (particularly if the latter one can be assessed with a single item)?

We completely rewrote the chapter "implications for clinical practice", which is now "implications of the study". The chapter now reads:

"Based on our study findings, we propose to focus on specific clinical and biological factors amenable to treatment at the individual level. Regarding clinical factors, sleep disturbances such as insomnia and sleep apnea (namely in presence of a patient with obesity) and the presence of depression should be assessed. Hence, lifestyle measures to improve sleep quality and quantity should be preferred to medication.[13] In case of depression, it will be important to warn patients that antidepressor medication might not necessarily lead into rapid relief of fatigue. Regarding biological factors, anemia should be ruled out, while screening for hypothyroidism is not recommended as a first step.

At the population level, preventive measures such as stress management and health promotion like relaxation, time management and cognitive reframing (for example within the work environment) could improve sleep quality, increase self-rated health {Hasson, 2005 #615} and so reduce fatigue."

Reviewer #2 also commented on this topic.

Limitations

22. In general, the limitations of the study and their implications deserve a more throrough discussion. One issue not mentioned is the fact that the study was conducted in Lausanne only. But how representative is the population of one city in the French speaking part of Switzerland for Switzerland as a whole?

We modified the text in the limitations paragraph as follows:

“This study has also several limitations. Firstly, its cross sectional setting precludes the identification of the causes of fatigue, as reverse causality is possible (i.e. fatigue leading to depression and vice-versa).[14] All participants of the CoLaus study are currently being re-contacted and re-examined, so that a prospective analysis of the causes of fatigue will be feasible within two years. Secondly, there is no gold standard for the evaluation of fatigue and no official definition of fatigue. Hence, results might vary according to the scale applied or how participants interpret the term “fatigue”. In this study, we chose to use a scale that was previously applied by other authors to facilitate comparisons. Thirdly, only the German version of the FSS has been validated in Switzerland; the French version used in this study has not yet been validated. Hence, it is possible that the true prevalence levels of fatigue might be under- or over-estimated, or that some items of the questionnaire might not be informative. Still, the Cronbach’s alpha for the questionnaire was 0.918, suggesting an excellent internal consistency. Further, our results provide a first estimation of the prevalence of fatigue in the Swiss French-speaking general population, which could serve as a reference for further studies. Fourthly, a sizable fraction of the sample was excluded, both between the baseline and the second follow-up, and within the current study, which might limit the generalizability of the findings. For instance, excluded participants were more frequently women; as women reported more frequently fatigue, this might lead to an underestimation of prevalence rates or a decrease in the strength of the associations. Still, an analysis using a propensity score weighting for the probability of being excluded led to similar findings. Conversely, it was not possible to assess the reasons why participants did not complete the questionnaire. Fifthly, no information was available regarding shift work or the presence of very young children. Still, as a sizable fraction (almost 70%) of the sample was aged over 55 and over 36% of the sample was aged over 64, it is likely that the number of participants either on shift work or with very young children would be small. Sixthly, the FSS explored fatigue during the previous week while the ISI score explored the sleep during the previous month. Hence, it is possible that the time association between the two variables might not be optimal. Still, as the FSS lies within the period encompassed by the ISI, we believe that the associations obtained are clinically relevant. Seventhly, the study is limited to the population of Lausanne aged 45 to 86, and its generalizability remains to be assessed. For instance, no information was collected regarding other confounders among younger subjects, where prevalence of fatigue might be higher due to parental and professional duties.[15]. Finally, possible biases related to the self-reporting of fatigue could not be avoided, such as over- or under-estimation of symptoms or misunderstanding of what the term “fatigue” meant; still, this dilution bias would lead to a decrease in the strength of the associations, and it would be too restrictive in our opinion to provide a definition of the term “fatigue” to the participants, as different interpretations of the definition itself could also occur.”

23. Also, despite the supplementary analysis for selection bias with propensity scores, implications (or eventually non-relevance) of the considerable exclusion/drop-out from the initial study population should be explained/discussed.

As the CoLaus questionnaire included a large number of questions (>350), a likely explanation is that participants got tired and did not respond to all items. Still, we did not collect the reasons why participants did not fill up the questionnaires and any explanation would be speculation. Hence, we decided not to develop this topic. Please see also our response to query #22.

24. Further, there seems to have been a prior approximate 25% dropout from baseline to the 2nd follow-up, which provided the data used for the present study. This drop-out may also be suspected to have been correlated with fatigue levels, and/or depression and/or health status and should therefore be discussed.

We added a comment in the limitations (see response to query #22). The drop-outs between the baseline and the second follow-up were due to several reasons (moving, death, unavailability due to illness, etc.). We decided not to elaborate on this issue as the editor and the other two reviewers had no query regarding this issue, and also in order to keep the manuscript size within reasonable limits.

25. Given that the study took place in Lausanne, I suppose it was the French version of the FSS which was used, which, however, was not validated yet. The possible implications may warrant a more in- depth discussion than just one sentence that this might have led to over- or underestimation.

    Please see our response to query #22.

26. Recruitment period was 3 years. Why was the period so long and might this have affected the findings?

Due to financial and logistic limitations, it was not possible to recruit quicker. Further, as participants had to be in the fasting state for blood drawing, examinations could only be conducted in the morning. Each examination took an average of one hour.

The time seems adequate compared to other cohorts: the Rotterdam Study I, sixth examination, took one year to examine 1000 participants. The Paris Prospective Study III took 4 years to examine 10,000 participants. Hence, we consider that three years to recruit 4881 participants with a research team of seven (including the PI, the data manager and the statistician) is reasonable.

Conclusion

27. This is a mere summary/repeat of the findings and not a conclusion.

All other reviewers and the handling editor were happy with the conclusion and reviewer 1 even asked for complementing it. Hence, we decided to keep it.

General issues:

28. The title of the paper “A Swiss population study” might be somewhat misleading given that the survey was conducted only in Lausanne and did not include the entire age range of the adult population either but just the middle-aged to elderly.

We changed the title to “Prevalence and factors associated with fatigue in the Swiss Lausanne middle-aged population: A population based cross-sectional survey“.

29. In general, the manuscript needs editing regarding use of English language, quite many missing words, incomplete sentences, e.g. on page 6, 2nd para: “Depression was assessed… (by using?) the CES-D ….(which?) is a 20 item self-report instrument developed for research in the general population … (and which?) is used to assess the severity of depressive symptoms over the past week on a 4-point scale.

The manuscrit has been read by an english-specking collegue

30. Given the cross-sectional character of the study and the conceptual overlap of some of the covariates with fatigue, the wording should (in parts) be more conservative, i.e. it would for instance be more appropriate to speak of associations and potential determinants rather than of "determinants".

We changed "determinants of" to "factors associated with" throughout the text.

Minor issues:

31. The headlines of the tables should describe the contents, i.e. the associations investigated rather than refer to the statistical analysis.

We modified the headlines according to the journal's and the reviewer's recommendations

32. Table 2: Under multivariable "OR" is missing in the headline.

We added "OR (95% CI)" in the column title

VERSION 2 – REVIEW

REVIEWER	Solveig Dohrmann University of Southern Denmark
REVIEW RETURNED	24-Apr-2019

GENERAL COMMENTS	Population and Methods Please consider/address: - Study population: o Page 5, line 50-54: the sentence 'As fatigue…., participants were aged 45-86'. I know it is elaborated elsewhere, but it would be very helpful for the reader if a very short explanation why participants were aged 45-86 at the second follow up as compared to the inclusion criteria (age 35-75) were added. - Fatigue scale: o Thank you for providing FSS (and ISI) in the annex, that is very helpful for the reader, especially if one is not familiar with the instrument(s). I suggest an English version as not all readers are familiar with French. As for the French version of FSS (and ISI), the translation process does not become clear. Did you/your colleagues translate FSS (and ISI) in preparation for the study? If yes, how? If no, please provide a reference. - Statistical analysis: o Page 9, line 56: the sentence 'Sensitivity analyses were conducted using a FSS threshold of 5'. Why a threshold of 5 (I suppose it is because 4 and/or 5 were used in most other studies)? As a follow-up on my initial comments on 'Population and methods' - and my apologies for being unclear. I commented that I

	was a little in doubt how your study/your sample related to the CoLaus study, and suggested a flow diagram. What I had in mind was an elaboration of the flow diagram you already provided; adding information on how many were invited to the overall study - how many was enrolled in the cohort/took part at baseline – how many took part in each follow up - and from there the information you already provided. I looked at the study-webpage, but I'm not familiar with French. I also looked at Fireman et al and I learned about the recruitment process and the study cohort (baseline). So, I lacked an overview of the whole process (ie. follow ups, incl. drop outs), and from where in the process your sample stemmed from. Reading the manuscript again, including your elaborations on the paragraph 'Study population' and 'Exclusion criteria' has helped me gain a better understanding of your sample. Thank you for the elaboration. Further, under statistical analysis I had a comment about that stat. significance was assessed for a two-sided test with $p<0.05$. I agree, it is rather common, another option is adding the info in a bracket. Discussion - Prevalence of fatigue o Page 15: the paragraph 'Prevalence of fatigue', Lerdal et al - who also used FSS to explore the relationship between fatigue and psychosocial variables in the general population - found a somewhat high prevalence score as compared with your findings (and the findings of Loge and Bultman (population and work; mostly healthy individuals) whereas the prevalence rate reported by Pawlikowska tended towards the higher end (which might be explained by the fact that participant were recruited through general practice clinics, though participants were not per se active patients at the time of the study?)). I find it interesting and relevant with respect to your study, the aim of your study. Therefore, I encourage you to elaborate, including to elaborate why results from studies conducted in other settings are relevant to you/your findings. - Implication of the study o Page 18, line 42-44: the reference Hasson, 2005 appears as {Hasson, 2005#615}. - Recommendations for future studies o Page 18, line 53-55: please elaborate on your second recommendation for future studies. I agree that use of non-standardized scales make comparisons difficult. However, you argue in the 'Introduction', ''Prevalence of fatigue' (all the scales mentioned in 'Prevalence of fatigue' are standardized once) and 'Strength and limitations' that use of different scales makes comparison across studies difficult - not that use of non-standardized once are a specific obstacle in this endeavor? Further, use of standardized scales might not eliminate the problem - only if the same standardized scale is used across studies (I suppose you relate to population studies, as that is your point of departure?)? Or if findings from different scales are comparable – if scales/the findings are somehow validated against each other? Further, a selection of valid and reliable fatigue-related scales already exists (see for instance Shahid et al, STOP, THAT and one hundred other sleep scales, 2012), and the scale of choice needs to suit the purpose of the study, the study sample. From that and from the complexity and the difficulties with defining fatigue one could argue that measures should be chosen or even tailormade to the situation and/or that different measures of different types (fatigue scales, cognitive test, hormone levels etc.) should be combined (for instance Aaronson et al, Defining and measuring fatigue, 1999). So, using one /few standardized scales

	may come at the expense of precision/bring about other obstacles, which – I believe – needs to be discussed. o Please consider presenting the paragraph 'Recommendation for future studies' after 'Strength and limitations. That will for instance help the reader better understand why you recommend to 'validate the questionnaires in the population of interest' for future studies. You might consider also moving the paragraph 'Implications of the study'. As a follow-up on my initial comments on 'Strength and limitations' – and my apologies for being unclear. I commented/asked how the definition you have chosen relates to/match the instrument you have chosen (FSS) and how both relates to/match your sample. Your definition indicates fatigue as a multidimensional concept (physical, cognitive and emotional). Further, it indicates fatigue of a more chronic nature (symptoms described as a tiredness not relieved by common strategies). Many of your participants are healthy individuals, while some have physical and mental conditions, including chronic conditions (diabetes, depression, insomnia etc.). One could argue that healthy individuals mostly suffers from fatigue of a more acute nature (mostly physical symptoms which can be relieved by common strategies, such as rest), while patients, especially those with chronical conditions, are more likely to suffer from fatigue of a more chronic nature (physical, cognitive and emotional symptoms which are not easily relieved by common strategies, such as rest). How does your definition match your sample? How might your choice have influenced results? As for FSS. From your definition the instrument of choice must be able to measure fatigue as a multidimensional concept. FSS is a unidimensional instrument. How does that match? How might your choice have influenced results? Further, from your definition and your sample the instrument of choice must be able to measure fatigue of a more acute nature (healthy individuals) and fatigue of a more chronic nature (definition and patients). How does that match? How might your choice have influenced results? The rationales behind your choices have become clearer in the revised version of the manuscript, for instance in 'Fatigue scale' is has been added that FSS has been validated in a healthy population setting in German-speaking Switzerland as well as in Portugal and Norway. As measuring fatigue is vital I encourage you to address these issues in the discussion (I know all this needs to be balanced with the requirements of the journal, which is not always and easy task). An option could be to add in a few sentences that your research was guided by a comprehensible definition of fatigue informing choice about instruments, confounders etc and interpretation of results (as a further strength). After that you might want to highlight strength of FSS (other versions than the French have been validated in healthy individuals and different patient groups, which suits your sample and allows for comparison with findings from other studies also using FSS). From there address, discuss short comings (if any) of the definition (fatigue of a more chronic nature vs your sample and FSS (unidimensional vs. the definition/your sample, validation of the French version) and that might have influenced results. After that maybe shortly discuss the general trouble of defining fatigue/lack of an official definition, and thereby the general trouble with measuring fatigue, lack of a golden standard, difficulties in comparison across studies – and from there present matters to be discussed (which is already presented in your discussion). All this brings me to my initial comment about

	‘other standardized fatigue scales’. Unfortunately, I don’t know of any scales validated in French speaking Switzerland (you might find a suggestion in/via Shahid et al, STOP, THAT and one hundred other sleep scales, 2012). My comment/question was meant to encourage a discussion of weather FSS suited the aim of your study. Or if other fatigue scales/measures would have been more favorable? As a concluding remark I encourage you to revise the conclusion you present in the manuscript. In appears undetailed, incomprehensive considering the revision and compared to the conclusion presented in the abstract.

VERSION 2 – AUTHOR RESPONSE

Answers to Population and Methods

Please consider/address:

- Study population: o Page 5, line 50-54: the sentence ‘As fatigue…., participants were aged 45-86’. I know it is elaborated elsewhere, but it would be very helpful for the reader if a very short explanation why participants were aged 45-86 at the second follow up as compared to the inclusion criteria (age 35-75) were added.

The second follow up was 10 years after the initial study. We added the following statement in the methods, study population, first paragraph: “second follow-up between May 2014 and April 2017, 10.9 years on average after the baseline study”

- Fatigue scale: o Thank you for providing FSS (and ISI) in the annex, that is very helpful for the reader, especially if one is not familiar with the instrument(s). I suggest an English version as not all readers are familiar with French. As for the French version of FSS (and ISI), the translation process does not become clear. Did you/your colleagues translate FSS (and ISI) in preparation for the study? If yes, how? If no, please provide a reference.

We now provide an English version of the FSS in Annex 2.

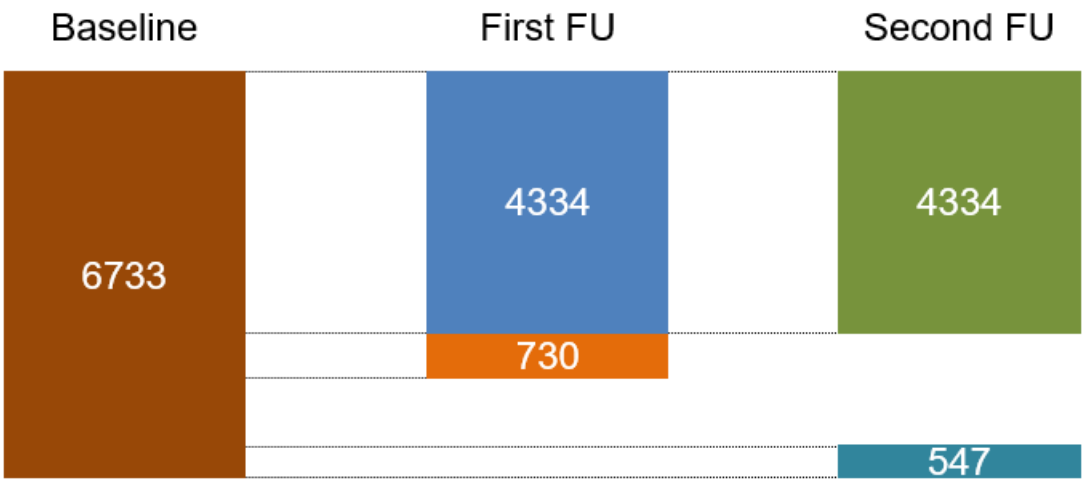
The French version of the FSS was used in a French study (Henry, Tourbah et al. 2019) and in a Canadian study (Laberge, Gagnon et al. 2005) , but no validation was made.

- Statistical analysis: o Page 9, line 56: the sentence ‘Sensitivity analyses were conducted using a FSS threshold of 5’. Why a threshold of 5 (I suppose it is because 4 and/or 5 were used in most other studies)?

The threshold of 5 has been used in other studies (i.e. reference 19 in the paper) and was also requested by reviewer 1 James N. Baraniuk, MD (query #5)

As a follow-up on my initial comments on ‘Population and methods’ - and my apologies for being unclear. I commented that I was a little in doubt how your study/your sample related to the CoLaus study, and suggested a flow diagram. What I had in mind was an elaboration of the flow diagram you already provided; adding information on how many were invited to the overall study - how many was enrolled in the cohort/took part at baseline – how many took part in each follow up - and from there the information you already provided. I looked at the study-webpage, but I’m not familiar with French. I also looked at Fireman et al and I learned about the recruitment process and the study cohort (baseline). So, I lacked an overview of the whole process (ie. follow ups, incl. drop outs), and from where in the process your sample stemmed from. Reading the manuscript again, including your elaborations on the paragraph ‘Study population’ and ‘Exclusion criteria’ has helped me gain a better understanding of your sample. Thank you for the elaboration. Further, under statistical analysis I had a comment about that stat. significance was assessed for a two-sided test with $p<0.05$. I agree, it is rather common, another option is adding the info in a bracket.

Both reviewers 1 and 3 found the flow diagram adequate. We also believe that making a flow diagram from the start of the study will be too detailed and incomprehensive. For information, we provide in the next page the flowchart from baseline to the second follow-up (the figures relate to the number of participants).



Discussion

- Prevalence of fatigue o Page 15: the paragraph ‘Prevalence of fatigue’, Lerdal et al - who also used FSS to explore the relationship between fatigue and psychosocial variables in the general population - found a somewhat high prevalence score as compared with your findings (and the findings of Loge and Bultman (population and work; mostly healthy individuals) whereas the prevalence rate reported by Pawlikowska tended towards the higher end (which might be explained by the fact that participant were recruited through general practice clinics, though participants were not per se active patients at the time of the study?)). I find it interesting and relevant with respect to your study, the aim of your study. Therefore, I encourage you to elaborate, including to elaborate why results from studies conducted in other settings are relevant to you/your findings.

We added the following text in the discussion. “Comparison between studies is hampered by the small number of studies assessing the prevalence of fatigue in non-selected samples, the different fatigue scales used and the somewhat different settings (i.e. general population vs. general practice). Still, they provide a first basis for comparison, and it would be important that future studies use similar

assessment methods to facilitate comparisons. Overall, our results suggest that the prevalence of fatigue in the Lausanne population is comparable or lower than reported previously, although the use of different scales to assess fatigue complicates comparison between studies."

- Implication of the study o Page 18, line 42-44: the reference Hasson, 2005 appears as {Hasson, 2005#615}.

We thank the reviewer for noting this. It has been corrected.

- Recommendations for future studies

o Page 18, line 53-55: please elaborate on your second recommendation for future studies. I agree that use of non-standardized scales make comparisons difficult. However, you argue in the 'Introduction', ''Prevalence of fatigue' (all the scales mentioned in 'Prevalence of fatigue' are standardized once) and 'Strength and limitations' that use of different scales makes comparison across studies difficult - not that use of non-standardized once are a specific obstacle in this endeavor? Further, use of standardized scales might not eliminate the problem - only if the same standardized scale is used across studies (I suppose you relate to population studies, as that is your point of departure?)? Or if findings from different scales are comparable – if scales/the findings are somehow validated against each other? Further, a selection of valid and reliable fatigue-related scales already exists (see for instance Shahid et al, STOP, THAT and one hundred other sleep scales, 2012), and the scale of choice needs to suit the purpose of the study, the study sample. From that and from the complexity and the difficulties with defining fatigue one could argue that measures should be chosen or even tailormade to the situation and/or that different measures of different types (fatigue scales, cognitive test, hormone levels etc.) should be combined (for instance Aaronson et al, Defining and measuring fatigue, 1999). So, using one /few standardized scales may come at the expense of precision/bring about other obstacles, which – I believe – needs to be discussed.

The aim of the study was to examine the prevalence and the factors associated with fatigue in a population-based sample from the city of Lausanne, Switzerland, not to compare the validity of fatigue scales. We used the only scale available for Switzerland, and the other two reviewers had no issues regarding this. Our paper is an epidemiological paper, not a methodological one.

o Please consider presenting the paragraph 'Recommendation for future studies' after 'Strength and limitations. That will for instance help the reader better understand why you recommend to 'validate the questionnaires in the population of interest' for future studies. You might consider also moving the paragraph 'Implications of the study'.

The paragraph "recommendations for future studies" has been moved as requested

As a follow-up on my initial comments on 'Strength and limitations' – and my apologies for being unclear. I commented/asked how the definition you have chosen relates to/match the instrument you have chosen (FSS) and how both relates to/match your sample. Your definition indicates fatigue as a multidimensional concept (physical, cognitive and emotional). Further, it indicates fatigue of a more chronic nature (symptoms described as a tiredness not relieved by common strategies). Many of your participants are healthy individuals, while some have physical and mental conditions, including chronic conditions (diabetes, depression, insomnia etc.). One could argue that healthy individuals

mostly suffers from fatigue of a more acute nature (mostly physical symptoms which can be relieved by common strategies, such as rest), while patients, especially those with chronical conditions, are more likely to suffer from fatigue of a more chronic nature (physical, cognitive and emotional symptoms which are not easily relieved by common strategies, such as rest). How does your definition match your sample? How might your choice have influenced results? As for FSS. From your definition the instrument of choice must be able to measure fatigue as a multidimensional concept. FSS is a unidimensional instrument. How does that match? How might your choice have influenced results? Further, from your definition and your sample the instrument of choice must be able to measure fatigue of a more acute nature (healthy individuals) and fatigue of a more chronic nature (definition and patients). How does that match? How might your choice have influenced results? The rationales behind your choices have become clearer in the revised version of the manuscript, for instance in 'Fatigue scale' is has been added that FSS has been validated in a healthy population setting in German-speaking Switzerland as well as in Portugal and Norway. As measuring fatigue is vital I encourage you to address these issues in the discussion (I know all this needs to be balanced with the requirements of the journal, which is not always and easy task). An option could be to add in a few sentences that your research was guided by a comprehensible definition of fatigue informing choice about instruments, confounders etc and interpretation of results (as a further strength). After that you might want to highlight strength of FSS (other versions than the French have been validated in healthy individuals and different patient groups, which suits your sample and allows for comparison with findings from other studies also using FSS). From there address, discuss short comings (if any) of the definition (fatigue of a more chronic nature vs your sample and FSS (unidimensional vs. the definition/your sample, validation of the French version) and that might have influenced results. After that maybe shortly discuss the general trouble of defining fatigue/lack of an official definition, and thereby the general trouble with measuring fatigue, lack of a golden standard, difficulties in comparison across studies – and from there present matters to be discussed (which is already presented in your discussion). All this brings me to my initial comment about 'other standardized fatigue scales'. Unfortunately, I don't know of any scales validated in French speaking Switzerland (you might find a suggestion in/via Shahid et al, STOP, THAT and one hundred other sleep scales, 2012). My comment/question was meant to encourage a discussion of weather FSS suited the aim of your study. Or if other fatigue scales/measures would have been more favorable?

To our knowledge, there are no other fatigue scales available for the French-speaking Switzerland. Regarding the limitations of the scale, we had included in the limitations paragraph that "Secondly, there is no gold standard for the evaluation of fatigue and no official definition of fatigue. Hence, results might vary according to the scale applied or how participants interpret the term "fatigue". In this study, we chose to use a scale that was previously applied by other authors to facilitate comparisons."

As a concluding remark I encourage you to revise the conclusion you present in the manuscript. In appears undetailed, incomprehensive considering the revision and compared to the conclusion presented in the abstract.

Both reviewers 1 and 3 found the conclusion adequate. Hence, we decided to keep it as it suited the majority of the reviewers.

Henry, A., et al. (2019). "Anxiety and depression in patients with multiple sclerosis: The mediating effects of perceived social support." Mult Scler Relat Disord 27: 46-51.

BACKGROUND: Social support has been identified as a buffering or intervening variable in stressful life events. Research has demonstrated that greater social support is associated with better mental health in multiple sclerosis (MS), but little is known about its links to specific aspects of mental health. We therefore investigated if and how perceived social support modulates depression, anxiety and fatigue in patients with MS. METHODS: We recruited 112 patients with MS from three French hospitals and administered a demographic and clinic interview, and self-report measures of perceived social support (Multidimensional Scale of Perceived Social Support), depression and anxiety (Hospital Anxiety and Depression Scale), and fatigue (Fatigue Severity Scale). We then analyzed the relationships between these domains using path analysis. RESULTS: The causal path model provided an excellent fit for the data (chi(2)=9.8, p=.778, standardized root mean square residual=0.043, comparative fit index=1.00). Results indicated that the level of social support from friends is a predictor of anxiety symptomatology. Thus, anxiety may have both a direct and an indirect impact on fatigue and depression levels. CONCLUSIONS: This study highlights the important roles played by perceived social support and anxiety in MS. These should be key pharmacological and non-pharmacological targets for optimizing patient care. (NCT 02-880-553).

Laberge, L., et al. (2005). "Fatigue and daytime sleepiness rating scales in myotonic dystrophy: a study of reliability." J Neurol Neurosurg Psychiatry 76(10): 1403-1405.

OBJECTIVES: To assess the reliability of the Epworth Sleepiness Scale (ESS), Daytime Sleepiness Scale (DSS), Chalder Fatigue Scale (CFS), and Krupp's Fatigue Severity Scale (KFSS) in patients with myotonic dystrophy type 1 (DM1). METHODS: In total, 27 patients with DM1 were administered the questionnaires on two occasions, with a 2 week interval. Internal consistency and test retest reliability were measured using intraclass correlation coefficients (ICCs), and Cronbach's alpha, Cohen's kappa, and Goodman-Kruskal's gamma coefficients. RESULTS: Internal consistency of the CFS and KFSS were adequate (alpha > 0.70) but that of the ESS was weak (alpha = 0.24). Both daytime sleepiness and fatigue rating scales showed significant test retest reliability. Test retest reliability for individual items revealed inconsistencies for some ESS and CFS items. CONCLUSIONS: Reliability of the CFS, DSS, and KFSS was high, allowing their use for individual patients with DM1, but that of the ESS was lower, rendering its current usage in DM1 questionable. Fatigue rating scales such as the KFSS, which are based on the behavioural consequences of fatigue, may constitute a more accurate and comprehensive measure of fatigue severity in the DM1 population.

VERSION 3 - REVIEW

REVIEWER	Solveig Dohrmann The Centre of Maritime Health and Society, The University of Southern Denmark
REVIEW RETURNED	03-Jul-2019

GENERAL COMMENTS	Dear all Thank you for presenting a revised version of your manuscript, it has improved a lot. I have only two comment; I believe that the manuscript will benefit from: 1: a few sentences to be added to the discussion about why it might be that Lerdal et al (also using FSS in a population based study) found somewhat higher prevalence rates (46.7% (cut-off=4)

	and 23.1% (cut-off=5)) as compared with your results (22.1% (cut-off=4) and 10.9% (cut-off=5)) (page 15, line 11-12). 2: more details to be added to the conclusion (the one in the abstract is more detailed). I wish you all a very nice summer.

VERSION 3 – AUTHOR RESPONSE

Reviewer(s)' Comments to Author:

Reviewer: 2

Reviewer Name: Solveig Dohrmann

Institution and Country: The Centre of Maritime Health and Society, The University of Southern Denmark Please state any competing interests or state ‘None declared’: None declared

A few sentences to be added to the discussion about why it might be that Lerdal et al (also using FSS in a population based study) found somewhat higher prevalence rates (46.7% (cut-off=4) and 23.1% (cut-off=5)) as compared with your results (22.1% (cut-off=4) and 10.9% (cut-off=5)) (page 15, line 11-12).

We added the following text in the discussion, page 15, after the citation of Lerdal et al.:

“The investigated population was aged 19-81 years, included younger patients (women of childbearing age with menstruation and young parents) compared to our study aged between 45 and 86 years; that could explain this difference in prevalence of fatigue.”

More details to be added to the conclusion (the one in the abstract is more detailed).

We changed the conclusion page 21 to:

“In a population-based sample aged 45 to 86, fatigue was present in one out of five subjects. Regarding clinical factors, sleep disturbances such as insomnia and sleep apnea should be assessed first, followed by depression. Regarding biological factors, anemia should be ruled out, while screening for hypothyroidism is not recommended as a first step. Sleep complaints and fatigue in older subjects are not due to aging and should prompt the identification of the underlying cause.”