Author’s response to reviews

Title: Stress levels predict substantial improvement in pain intensity after 10 to 12 years in women with fibromyalgia and chronic widespread pain. A cohort study.

Authors:

Anna Bergenheim (anna.c.bergenheim@vgregion.se)
Sofia Johannesson (sofia.e.johannesson@vgregion.se)
Lena Nordeman (lena.nordeman@vgregion.se)
Monica Joelsson (Monica.joelsson@vgregion.se)
Kaisa Mannerkorpi (Kaisa.Mannerkorpi@neuro.gu.se)

Version: 1 Date: 28 Dec 2018

Author’s response to reviews:

To Reviewers

Thank you for your valuable comments which we found to have improved the manuscript. Below are our answers in bold, after each of your questions/comments. In the enclosed manuscript, the changes are written in red. The answers below are also enclosed as a Word-file.

Reviewer reports:

Marcus Beasley (Reviewer 1): Thank you for this well written and interesting paper that aimed to look at the course of symptoms for people with fibromyalgia or chronic widespread pain, and identify some predictors of improvement. There are however two main problems with the methods and analyses, one to do with each of the aims, that limit the conclusions that can be made.

Firstly, when looking out how symptoms progress after 10 to 12 years for people with fibromyalgia or chronic widespread pain, you have not considered regression toward the mean. This occurs if you recruit people into your study according to whether they meet certain eligibility criteria related to the severity of their symptoms. The severity of symptoms can be treated like a random variable, they fluctuate over time. A group of people recruited when their symptoms are at their worst will tend to improve anyway.
Reply: Thank you for pointing this out, you are right, and we have added a sentence about this in limitations in the discussion section, p. 15.

Regarding the analyses of predictors we have chosen a large improvement (at least 50 %) as dependent variable, which most likely is not a result of regression towards the mean.

Secondly, a related problem in looking at predictors of recovery or improvement in symptoms has not been considered. That is, if a measure has random fluctuation over time (in this case, pain intensity) and if recruitment to the cohort is related to this measure (having CWP or FM diagnosis), and the characteristics considered as predictors are related to the measure at the time of recruitment (current stress is related to current pain status), then the characteristics can seem to predictive even without a change in outcome. A better explanation of this is given in this recent paper: Sorjonen, K., Lundberg, M., & Melin, B. (2018). Using logistic regression to predict onset and recovery with tau equivalency. Frontiers in Psychology, 9, 1849.

Reply: Thank you for sharing this interesting article, which I have read and I have also consulted an experienced statistician. However, we don´t find the conclusion of the article applicable in the analyses of predictors in the present study, since the dependent variable in our analysis is a cut off for change: at least 50 % improvement in pain intensity, and not a cut-off value of post-test pain intensity. Please let us know if you think differently.

I also have a few minor comments and suggestions:

- The word 'contradicting' in the abstract and page 3, line 25 of the Background section might need to be replaced with 'contradictory'.

Reply: Thank you, we have changed it to Contradictory.

- Would 'communicated' be a better word than 'mediated' in the abstract and page 14, line 19 in the Conclusions?

Reply: Thank you for this suggestion, we have changed it to Communicated in the abstract and in the conclusions.

- Rather than use p-values in the abstract and in Table 2 of the results, it might be better to give changes along with confidence intervals.

Reply: We have now added mean and Bootstrapped 95 % C.I in the abstract and also in the column of change in Table 2.
- As you have given the number of people who got had moderate or substantial improvement, you could also show the number of people who had moderate (30-50%) or substantial (<50%) worsening of pain intensity, and the number remaining unchanged (<30% better or worse).

Reply: You are right. We have now added this information in the result section, p 10, line 12 to 18.

- I'm not sure that you have used the right measure of improvement (i.e. relative improvement from baseline). The IMMPACT recommendations you reference are for measuring treatment outcomes in clinical trials, and this study isn't looking at treatment outcomes in a clinical trial.

Reply: We agree about the problem, and we now have addressed this in the discussion section, p.12.

We have not found any other references regarding measure of this type of “natural” change over time.

I think all your logistic regression models of predictors should be adjusted for baseline pain intensity. If not, it might be helpful to explain why.

Reply: We followed your advice and all regression models are now adjusted for pain intensity. The results in Table 4 are now slightly different, as well as the results of the multivariable regression analysis described in the text at p. 11 and in the abstract.

Reviewer 2 (Reviewer 2):

PEER REVIEWER COMMENTS:

GENERAL COMMENTS: The study quality is moderate. The authors did a great job to perform a 10-12 years follow up. The statistical analyses need to be substantially improved.

REQUESTED REVISIONS:
Design and execution: I strongly suggest that the authors do not present mean change scores, but responder (clinically relevant improvement) of all secondary outcomes. e.g, as suggested by Arnold. Arthritis Rheum. 2012 Mar;64(3):885-94. In addition, I suggest to present data of the FIQ total score. For this score, responder criteria (>=14 % or 20% improvement from baseline are available, too).

Reply: Thank you for this suggestion. We have added the results of FIQ total score in the Result section p. 10, as well as in Table 2 and 3, and we also discussed it in the Discussion section, p.14.

The responder criteria in the article you refer to are based on intervention studies, which is not quite applicable on within-group course of symptoms over time.

We therefore find it sufficient to relate only the results of the primary outcome FIQ pain and the FIQ total score, to MCID in the Discussion section and also address the limitation of the use of the MCID references (p.14).

I strongly suggest that the authors present data how many patients reported a moderate (>=30%) and substantial (>= 50%) deterioration of their pain scores and of the secondary outcomes. The presentation of the data is biased towards reporting the positive outcomes.

Reply: Thank you for pointing this out, we have added this information in the result section, p.10.

Please add a multivariate logistic regression analysis with all predictors included.

Reply: We have made a stepwise multivariable logistic regression analysis with the predictors included that were statistically significant in the univariable analyses. The results of the multivariable analysis is presented in the text in the Result section, not in table.

Please present data how many patients met the criteria of mild, moderate and severe FM according to the FIQ total score at baseline and at follow-up.

Reply: We have now added a table with the results of this calculation in the result section (Table 3), as well as a section addressing this in the discussion, p.14.

Interpretation: 1. Major limitations of the analyses are not discussed. Other variables than the baseline variables tested might have influenced the results at follow-up. These variables might have changed during the follow-up period. For example: No data are available which drug and non-drug therapies the patients have received during the follow up time (btw: the data on drugs in table 1 need to be presented in a more detailed way, e.g. opioids, antidepressants, anticonvulsants), the level of physical activity, the satisfaction of the patients with their work and family, comorbidities which might have developed (e.g. diabetes).
Reply: Thank you, we have now addressed these limitations in the section about Limitations in the Discussion, p.15.

Regarding the drugs, we unfortunately do not have any more specific information of the medications from baseline 10 to 12 years ago. We will have this in mind when planning future studies.

2. A study with an 11-year follow-up are not included in INTRODUCTION and DISCUSSION (Wolfe Eur J Pain 17(2013) 581-586 ). 3. The divergent results of the studies might be due to different study settings (primary or secondary or tertiary care) in which the patients have been recruited.

Reply: Thank you for sharing this interesting reference, we have included it in the introduction and in the discussion.

ADDITIONAL REQUESTS/SUGGESTIONS:

Different definitions of CWP are used in epidemiology studies leading to divergent rates of prevalence. The ACR 1990 criteria used in the study are the least restrictive ones.

Reply: We agree. However, we included both FM and CWP in the study and found it to be best to use the same reference of criteria for FM and CWP. At the time of baseline 2009, the most established definition of FM were the 1990 criteria. Later definitions were developed from the year 2010 and onwards.