

PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (<http://bmjopen.bmj.com/site/about/resources/checklist.pdf>) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Predictors for future activity limitation in women with chronic low back pain consulting primary care: a two-year prospective longitudinal cohort study
AUTHORS	Nordeman, Lena; Thorselius, Lena; Gunnarsson, Ronny; Mannerkorpi, Kaisa

VERSION 1 - REVIEW

REVIEWER	Paul Campbell Keele University, UK
REVIEW RETURNED	25-Oct-2016

GENERAL COMMENTS	<p>This study assesses prognostic factors for the prediction of activity limitation in women with chronic low back pain. The study recruited participants from primary care and they were assessed at baseline and followed up at two years. Results show a measure of physical function, clinical stress and baseline disability predicted outcome at 2 years. Whilst the study does have some strengths (prospective design, good response to follow up), I do have a number of concerns that reduce my confidence in the findings at present.</p> <p>1. I struggle to find the novelty in this project; there are numerous prognostic studies and models and numerous reviews that report prognostic factors similar to the ones used in this study. The authors argue that the evidence is insufficient overall, and that the role of lifestyle factors, stress factors and physical performance is unknown, however I feel the evidence is quite robust in this area, here is just a few reviews (Mallen, C.D., Peat, G., Thomas, E., Dunn, K.M. and Croft, P.R., 2007. Prognostic factors for musculoskeletal pain in primary care: a systematic review. <i>Br J Gen Pract</i>, 57(541), pp.655-661. Hockings, R.L., McAuley, J.H. and Maher, C.G., 2008. A systematic review of the predictive ability of the Orebro Musculoskeletal Pain Questionnaire. <i>Spine</i>, 33(15), pp.E494-E500. Hayden, J.A., Chou, R., Hogg-Johnson, S. and Bombardier, C., 2009. Systematic reviews of low back pain prognosis had variable methods and results—guidance for future prognosis reviews. <i>Journal of clinical epidemiology</i>, 62(8), pp.781-796. Pinheiro, M.B., Ferreira, M.L., Refshauge, K., Maher, C.G., Ordoñana, J.R., Andrade, T.B., Tsathas, A. and Ferreira, P.H., 2016. Symptoms of depression as a prognostic factor for low back pain: a systematic review. <i>The Spine Journal</i>, 16(1), pp.105-116. Manek, N.J. and MacGregor, A.J., 2005. Epidemiology of back disorders: prevalence, risk factors, and prognosis. <i>Current opinion in rheumatology</i>, 17(2), pp.134-140. Ramond, A., Bouton, C., Richard, I., Roquelaure, Y., Baufreton, C., Legrand, E. and Huez, J.F., 2010. Psychosocial risk factors for chronic low back pain in primary care—a systematic review. <i>Family practice</i>, p.cmq072.)</p>
-------------------------	---

	<p>2. Can the authors give a little more information about recruitment, for example they state that patients were identified via a search of medical charts (using coding - M545 ICD-10), how reliable is this coding, is there a chance of misclassification? It would also be helpful to know the total number of patients identified, approached, and contacted to give a sense of baseline response.</p> <p>3. I have some statistical questions. Why did the authors perform correlations as the first step and not just univariate regression analysis, and the authors state that this was a “sorting mechanism” can they explain what this means? Also the use of a stepwise procedure requires justification, especially when including variables that would have potential relationship to each other (i.e. within domains), would it not have been better to conduct separate models for each domain and the forward the significant variables from this into a final multivariable model?</p> <p>4. I feel reporting the stepwise significant variables' independent R square contribution to the variance is important as I suspect that the RMDQ at baseline will represent a significant amount of variance explained, and reporting each variables will place a perspective on the actual effect of the other two variables. I would also produce the table for the percentage change analysis, perhaps as an appendix.</p> <p>5. Given the findings on 6MWT, which may indicate frailty, I would advise adding age as a potential confound in the final stepwise model, even though it was not correlation directly with outcome it may well influence these results.</p> <p>6. The authors claim within the discussion that this study has provided new knowledge, this needs to be expanded with greater consideration to previous work.</p>
--	--

REVIEWER	Annette Becker Department of General Practice / Family Medicine University of Marburg
REVIEW RETURNED	30-Oct-2016

GENERAL COMMENTS	<p>Thank you very much for this interesting study on predictors for future activity limitation in women with low back pain. The authors present a cohort study with 2 years follow-up including 130 women with chronic low back pain. The manuscript is very well written leaving only few points for discussion or – if necessary- for revision:</p> <p>Women are recruited via an ICD-10 driven search in medical charts at eight primary care clinics in South-Western Sweden. Patients suffering from low back pain at time of recruitment and at least for the previous 12 weeks were included (apart from other inclusion criteria). Choosing this way of recruitment the authors primarily addressed women without acute need for treatment and rather low grade of disability – for a long time / almost 10 years of back pain. This can be seen in Table 2, showing rather low RMDQ scores at baseline and low scores of HADS. From this perspective a decrease in RMDQ score of 1,9 points in 2 years is low but understandable. This should be mentioned.</p> <p>As for the recruitment method could you please comment on any ongoing non pharmacological treatments and health care utilization</p>
-------------------------	--

	<p>during follow up? What is meant by „severe co-morbidity“, please give examples.</p> <p>To conclude that interventions aiming to decrease stress are likely to be important to improve women’s prognosis is a bold statement based on the cohort study. This should be attenuated. However similar comments are made in the discussion section. This seems sensible, but there are studies and reviews on stress reduction trainings. They showed no reduction in disability or pain, but in pain acceptance which might mediate the association of stress, physical activity and disability. This should be discussed.</p> <p>The exclusive inclusion of women in this study should not be listed as limitation (page 4), especially when its advantage is emphasized within the manuscript.</p> <p>Methods: The independent variables were chosen from binary regression analyses following statistical criteria only. To me this seems rather precarious given the fact that binary associations may be confounded. I think it might be sensible to include a second block of known influential factors (such as depression a. o.) in the model. If not, please justify your approach.</p>
--	---

REVIEWER	Steven George Duke University, USA
REVIEW RETURNED	07-Nov-2016

GENERAL COMMENTS	<p>Thank you for the opportunity to review this paper.</p> <p>Its strengths include 1) the prospective design with two year follow up after a consultation for chronic (defined as greater than 12 week’s duration of symptoms) low back pain; 2) use of the ICF model for providing structure to predictor variables; and 3) use of a well-accepted measure for the primary outcome (RMDQ). The finding that a measure of stress, the baseline score of the RMDQ, and a physical performance test were predictive of two year outcome was not surprising given the current state of literature for low back pain.</p> <p>Its primary limitations include 1) a narrow sample (Swedish women seeking primary care) that is likely to have limited generalizability to other populations of interest to the field; 2) Even though the authors tried to mitigate the relatively small sample size with the analysis plan, there is still limited number of participants for the candidate predictor variables considered.</p> <p>Comment and suggestions for improvement on this paper include 1) the authors should be more transparent on how much variance the other two predictor variables account for after the baseline RMDQ scores, beta weights for final models are presented but a clear description of how much variance beyond the baseline variable is added would help when comparing to other studies in this area; 2) there is too strong of a recommendation in the Discussion that the variables identified to increase risk should be treatment targets, I am not sure treatment recommendations are warranted based on the results of this study.</p>
-------------------------	---

VERSION 1 – AUTHOR RESPONSE

Reviewer #1.

Reviewer Name: Paul Campbell

Institution and Country: Keele University, UK

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

“This study assesses prognostic factors for the prediction of activity limitation in women with chronic low back pain. The study recruited participants from primary care and they were assessed at baseline and followed up at two years. Results show a measure of physical function, clinical stress and baseline disability predicted outcome at 2 years. Whilst the study does have some strengths (prospective design, good response to follow up), I do have a number of concerns that reduce my confidence in the findings at present.”

We are grateful for your thoughtful comments and questions.

Comment 1. “I struggle to find the novelty in this project; there are numerous prognostic studies and models and numerous reviews that report prognostic factors similar to the ones used in this study. The authors argue that the evidence is insufficient overall, and that the role of lifestyle factors, stress factors and physical performance is unknown, however I feel the evidence is quite robust in this area, here is just a few reviews (Mallen, C.D., Peat, G., Thomas, E., Dunn, K.M. and Croft, P.R., 2007. Prognostic factors for musculoskeletal pain in primary care: a systematic review. *Br J Gen Pract*, 57(541), pp.655-661. Hockings, R.L., McAuley, J.H. and Maher, C.G., 2008. A systematic review of the predictive ability of the Orebro Musculoskeletal Pain Questionnaire. *Spine*, 33(15), pp.E494-E500. Hayden, J.A., Chou, R., Hogg-Johnson, S. and Bombardier, C., 2009. Systematic reviews of low back pain prognosis had variable methods and results—guidance for future prognosis reviews. *Journal of clinical epidemiology*, 62(8), pp.781-796. Pinheiro, M.B., Ferreira, M.L., Refshauge, K., Maher, C.G., Ordoñana, J.R., Andrade, T.B., Tsathas, A. and Ferreira, P.H., 2016. Symptoms of depression as a prognostic factor for low back pain: a systematic review. *The Spine Journal*, 16(1), pp.105-116. Manek, N.J. and MacGregor, A.J., 2005. Epidemiology of back disorders: prevalence, risk factors, and prognosis. *Current opinion in rheumatology*, 17(2), pp.134-140. Ramond, A., Bouton, C., Richard, I., Roquelaure, Y., Baufreton, C., Legrand, E. and Huez, J.F., 2010. Psychosocial risk factors for chronic low back pain in primary care—a systematic review. *Family practice*, p.cmq072.)”

Response: Thank you for all the valuable references! We have read them with interest but unfortunately the duration of low back pain are too short in some of the references. We agree that there are numerous prognostic studies presenting prognostic factors for the transition from non-specific acute or subacute (<12 weeks duration) low back pain to chronic low back pain. In the present study we wanted to evaluate prognostic factors for patients with chronic low back pain. In this study only patients with chronic low back pain with duration >12 weeks were included at baseline. We wanted to evaluate factors associated with recovery or recovery delay for patients with chronic low back pain (> 12 weeks duration at baseline) in the long-term.

This study is an extended analysis of the material from the 2-year longitudinal study of prognostic factors for work ability in women with chronic low back pain (Nordeman et al 2014). In the present study, we aimed to focus on the prognostic value of lifestyle behavioural factors, stress symptoms and physical performance for future activity limitation using the same material and measurements. We have rewritten the last paragraph the section Introduction (Page 6, Line 6-10).

New references included

Hayden JA, Chou R, Hogg-Johnson S, et al. Systematic reviews of low back pain prognosis had variable methods and results: guidance for future prognosis reviews. *J Clin Epidemiol*. 2009;62(8):781-96.e1.

Mallen CD, Peat G, Thomas E, Dunn KM, Croft PR. Prognostic factors for musculoskeletal pain in primary care: a systematic review. *Br J Gen Pract.* 2007 Aug;57(541):655-61.

Comment 2. "Can the authors give a little more information about recruitment, for example they state that patients were identified via a search of medical charts (using coding - M545 ICD-10), how reliable is this coding, is there a chance of misclassification? It would also be helpful to know the total number of patients identified, approached, and contacted to give a sense of baseline response."

Response: A systematic search in medical charts for low back pain diagnoses M545 according to ICD-10 was performed using a report tool for primary care at the eight primary health care centres. All patients with diagnosis M545 identified in this report during the period 2004-2005 were contacted (n=476). All patients fulfilling the inclusion criteria and accepting participation were invited to enroll in the study (n=182). Fifty-two participants declined to participate, leaving 130 participants to assess at baseline. All participants (n=130) were contacted after two years. Ninety-five percent (n=123/130) could be followed up after two years.

The total number of patient identified, approached, and contacted is presented in Flow Chart, Figure 1.

Comment 3. "I have some statistical questions. Why did the authors perform correlations as the first step and not just univariate regression analysis, and the authors state that this was a "sorting mechanism" can they explain what this means? Also the use of a stepwise procedure requires justification, especially when including variables that would have potential relationship to each other (i.e. within domains), would it not have been better to conduct separate models for each domain and then forward the significant variables from this into a final multivariable model?"

Response: Thank you for this comment! Spearman Rank correlation between RMDQ at two years and each of the independent variables at baseline was performed to evaluate independent variables and reduce the number of independent variables of interest. This analysis was also performed between RMDQ percentage change and each of the independent variables at baseline. In this revision, independent variables with $p < 0.20$ were included in next multivariate regression step. We have rewritten a paragraph in section Method, Statistical analysis (Page 11, Line 10-18).

We have thoroughly considered different alternatives for analysis. Stepwise regression could be considered as an exploratory method for model building. The present study aimed to assess several potential predictors of interest for included group of women with chronic low back pain, thus stepwise regression was considered appropriate to use. Aware of the assumptions for multivariate regression, the correlation matrix, values of Tolerance (0.52-0.74) and VIF (1.3-1.9) were checked for multicollinearity.

Comment 4. "I feel reporting the stepwise significant variables' independent R square contribution to the variance is important as I suspect that the RMDQ at baseline will represent a significant amount of variance explained, and reporting each variables will place a perspective on the actual effect of the other two variables. I would also produce the table for the percentage change analysis, perhaps as an appendix."

Response: We have taken this comment into consideration and have added information about each variables R square contribution to the variance. Section Results, Predictors for activity limitation (RMDQ) at the two-year follow-up (Page 16, Line 8-11).

We have also taken the other comment into consideration and have added a new table (Table 4, Page 18), showing Spearman's correlation coefficient and p-values between each of the independent variables and percentage change in RMDQ. As previously presented, no statistically significant association was found when entered in the stepwise multivariate regression (Table 4). We have added information in the Results section, Predictors for activity limitation at the two-year follow-up, referring to this new table (Table 4).

Comment 5. "Given the findings on 6MWT, which may indicate frailty, I would advise adding age as a potential confound in the final stepwise model, even though it was not correlation directly with outcome it may well influence these results."

Response: We have taken this comment into consideration and we included age as an additional independent variable in the final models. This did not influence the results. We have added information about this in section Methods (Page 11, Line 23-24), in Table 3 and 4, and in section Discussion (Page 19, Line 20-22, Page 20, Line 1-2).

Comment 6. "The authors claim within the discussion that this study has provided new knowledge, this needs to be expanded with greater consideration to previous work."

Response: We have taken this comment into consideration and have added a paragraph in section Discussion (Page 22, Line 15-20).

Reviewer #2.

Reviewer Name: Annette Becker

Institution and Country: Department of General Practice / Family Medicine, University of Marburg

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

"Thank you very much for this interesting study on predictors for future activity limitation in women with low back pain. The authors present a cohort study with 2 years follow-up including 130 women with chronic low back pain. The manuscript is very well written leaving only few points for discussion or – if necessary- for revision:"

We are grateful for your thoughtful comments and questions.

Comment 1. "Women are recruited via an ICD-10 driven search in medical charts at eight primary care clinics in South-Western Sweden. Patients suffering from low back pain at time of recruitment and at least for the previous 12 weeks were included (apart from other inclusion criteria). Choosing this way of recruitment the authors primarily addressed women without acute need for treatment and rather low grade of disability – for a long time / almost 10 years of back pain. This can be seen in Table 2, showing rather low RMDQ scores at baseline and low scores of HADS. From this perspective a decrease in RMDQ score of 1,9 points in 2 years is low but understandable. This should be mentioned."

Response: We have taken this comment into consideration and added information about this in section Discussion, Strengths and limitations (Page 20, Line 20-21).

Comment 2. "As for the recruitment method could you please comment on any ongoing non pharmacological treatments and health care utilization during follow up? What is meant by „severe co-morbidity“, please give examples."

Response: We agree that additional information about ongoing non pharmacological treatments and health care utilization should have strengthened the study. Unfortunately, this was not collected in this study. However, pharmacological treatment at the first assessment were collected and are presented in Table 1. We have taken the latter comment into consideration and added information with examples of severe co-morbidity in section, Methods, Selection of patients (Page 7, Line 20-21).

Comment 3. "To conclude that interventions aiming to decrease stress are likely to be important to improve women's prognosis is a bold statement based on the cohort study. This should be attenuated. However similar comments are made in the discussion section. This seems sensible, but there are studies and reviews on stress reduction trainings. They showed no reduction in disability or pain, but in pain acceptance which might mediate the association of stress, physical activity and disability. This should be discussed."

Response: We have taken this comment in consideration and have attenuated this recommendation in the Abstract, Conclusions and in section Discussion (Page 23, Line 5-8). We have added information about pain acceptance in section Discussion (Page 22, Line 12-14).

New reference included

Jeong S, Cho S. Acceptance and patient functioning in chronic pain: the mediating role of physical activity. Qual Life Res.2016; Sept 1.

Comment 4. "The exclusive inclusion of women in this study should not be listed as limitation (page 4), especially when its advantage is emphasized within the manuscript."

Response: Thank you for the comment. We have rewritten the sentence in section, Strengths and limitations of this study (Page 4).

Comment 5. "Methods: The independent variables were chosen from binary regression analyses following statistical criteria only. To me this seems rather precarious given the fact that binary associations may be confounded. I think it might be sensible to include a second block of known influential factors (such as depression a. o.) in the model. If not, please justify your approach."

Response: Thank you for the comment. There are a lot of assumptions about the data before analysing with multivariate regression models. Prior to the multiple regression, the variables were evaluated for assumptions of multivariate analysis including checking for multicollinearity and singularity. Tolerance values were checked (0.52-0.74) indicating low correlation between the independent variables and the questionnaires total scores were used to avoid singularity.

Spearman Rank correlation between RMDQ at two years and each of the independent variables at baseline was performed to evaluate independent variables and reduce the number of independent variables of interest. This analyse was also performed between RMDQ percentages change and each of the independent variables at baseline. In this revision of the analysis, independent variables with $p < 0.20$ were included in next multivariate regression step.

The final models have in this revision been adjusted for age (Table 3, Table 4). Additional information

has been added in Methods, Statistical analysis (Page 11, Line 10-18, Line 23-24)
An additional paragraph to discuss this has been added in the section Discussion, Strengths and limitations (Page 19, Line 20-22, Page 20, Line 1-2)

Reviewer #3.

Reviewer Name: Steven George

Institution and Country: Duke University, USA

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

"Its strengths include 1) the prospective design with two year follow up after a consultation for chronic (defined as greater than 12 week's duration of symptoms) low back pain; 2) use of the ICF model for providing structure to predictor variables; and 3) use of a well-accepted measure for the primary outcome (RMDQ). The finding that a measure of stress, the baseline score of the RMDQ, and a physical performance test were predictive of two year outcome was not surprising given the current state of literature for low back pain."

We are grateful for your thoughtful comments and questions.

Comment 1. "Its primary limitations include:

1a) "a narrow sample (Swedish women seeking primary care) that is likely to have limited generalizability to other populations of interest to the field"

Response: Thank you for your comment! We agree that the characteristics of the patients in the study used to derive the predictive model have to be similar to those in whom the model will be used in clinical practise. Previous studies have shown that chronic pain is more common in women. Moreover, gender is suggested to be a predictor of chronic pain and disability. Various factors can impact disability in patients with chronic low back pain and function and functional demands often differs between women and men. Therefore, when evaluating predictive factors for activity limitation a selection of female patients was considered appropriate. However, we agree with referee that in future studies of prognostic factors for activity limitation in men with chronic low back pain should be considered.

We have rewritten the sentence about the limitation of the study according to sample size and the inclusion of solely female patients which limits the generalizability to men in the Section, Strengths and limitations of this study (Page 4).

1b) "Even though the authors tried to mitigate the relatively small sample size with the analysis plan, there is still limited number of participants for the candidate predictor variables considered."

Response: We agree that the sample size should be commented. In the first assessment 130 patients were included and 95% (123/130) were followed up after two years. Each variable were assessed in a bivariate analysis and built the multivariate model by using independent variables with $p < 0.20$ entered in the model. There are different guidelines concerning the number of cases required for multiple regression analysis varying between 5-15 cases per independent variable. Aware of problems which may occur when there are too few cases relative to the number of independent variables, the number of participants were considered to be sufficient for the model used in the study. We have added information about this in the section Discussion (Methodological aspects), (Page 20, Line 3-6).

Comment 2. "Comment and suggestions for improvement on this paper include:

2a) "The authors should be more transparent on how much variance the other two predictor variables

account for after the baseline RMDQ scores, beta weights for final models are presented but a clear description of how much variance beyond the baseline variable is added would help when comparing to other studies in this area”

Response: We have taken this comment into consideration and have added information about each variables R square contribution to the variance. Section Results, Predictors for activity limitation (RMDQ) at the two-year follow-up (Page 16, Line 8-11) and in section Discussion (Page 22, Line 21-22, Page 23, Line 1-2)

2b) “There is too strong of a recommendation in the Discussion that the variables identified to increase risk should be treatment targets, I am not sure treatment recommendations are warranted based on the results of this study.”

Response: We have taken this comment into consideration and have rewritten the paragraph in the section Discussion (Page 23, Line 5-8).

VERSION 2 – REVIEW

REVIEWER	Paul Campbell Keele University, UK
REVIEW RETURNED	04-Jan-2017

GENERAL COMMENTS	The authors have successfully addressed my initial points and I have no further comments to add.
-------------------------	--

REVIEWER	Annette Becker Department of General Practice / Family Medicine, University of Marburg, Germany
REVIEW RETURNED	22-Jan-2017

GENERAL COMMENTS	<p>Very well written paper with sound methodology. However, despite an extensive revision by the authors (the paper improved a lot) I still cannot ignore some problems:</p> <ul style="list-style-type: none"> - The paper does not add much to the current knowledge. Baseline disability accounts for almost the total amount of variance and there is no association of the independent variables with change in disability. - I still cannot support the selection process of variables, which strictly relies on significance in bivariate analyses (apart from age) instead on clinical or epidemiological importance. It simplifies the highly complex, often bidirectional relationships of influential variables in chronic pain. To my opinion a back and fourth selection process should be performed with all those variables who showed evidence in previous studies. The chosen process is prone to bias. However, sample size might be too low then. <p>I do not think that further revision is necessary. It is now a question of whether the paper fits to BMJ Open or rather a more specialised journal.</p>
-------------------------	---

VERSION 2 – AUTHOR RESPONSE

Reviewer: 1

Reviewer Name: Paul Campbell

Institution and Country: Keele University, UK Please state any competing interests or state 'None declared': None Declared

"The authors have successfully addressed my initial points and I have no further comments to add."

Response: Thank you for this comment!

Reviewer: 2

Reviewer Name: Annette Becker

Institution and Country: Department of General Practice / Family Medicine, University of Marburg, Germany Please state any competing interests or state 'None declared': None

"Very well written paper with sound methodology. However, despite an extensive revision by the authors (the paper improved a lot) I still cannot ignore some problems:

- The paper does not add much to the current knowledge. Baseline disability accounts for almost the total amount of variance and there is no association of the independent variables with change in disability.

- I still cannot support the selection process of variables, which strictly relies on significance in bivariate analyses (apart from age) instead on clinical or epidemiological importance. It simplifies the highly complex, often bidirectional relationships of influential variables in chronic pain. To my opinion a back and fourth selection process should be performed with all those variables who showed evidence in previous studies. The chosen process is prone to bias. However, sample size might be too low then.

I do not think that further revision is necessary. It is now a question of whether the paper fits to *bmj* open or rather a more specialised journal."

Response: We appreciate the objections from you. The initial decision to collect variables was based on previous studies indicating suitable variables of potential interest. However, this decision led to a large number of variables and a further sorting mechanism was needed before the final multivariate regression model. This sorting mechanism could be done either by further using clinical reasoning and prior knowledge or by looking at statistical significance. Each of these approaches will have a different risk for bias. Using clinical reasoning and prior knowledge may make us blind to new knowledge that previous studies missed. Using the approach we finally chose, bivariate correlation in Spearman's rank correlation may cause clinically insignificant findings to be put forward. An additional paragraph to discuss this has been added in the section Discussion, Strengths and limitations (Page 19, Line 20-23, Page 20, Line 1-5).