PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Development and internal validation of a multivariable prognostic model to predict chronic pain after a new episode of non-specific idiopathic, non-traumatic neck pain in physiotherapy primary care practice.
AUTHORS	Verwoerd, Martine; Wittink, Harriët; Maissan, Francois; Teunis, Marc; van Kuijk, Sander; Smeets, Rob

VERSION 1 - REVIEW

REVIEWER NAME	Kowalski, Katie
REVIEWER AFFILIATION	Western University, School of Physical Therapy
REVIEWER CONFLICT OF INTEREST	None.
DATE REVIEW RETURNED	01-May-2024

GENERAL COMMENTS	Thank you for the opportunity to review this research study that aimed to identify which modifiable factors are independent prognostic factors of the development of chronic neck pain in patients with acute- or subacute neck pain, and to develop and internally validate a model to predict development of chronic pain. Owing to the high and increasing prevalence and burden of chronic neck pain, these research aims are clinically important for physiotherapy. This study was conducted in accordance with a published protocol and gold standard frameworks for studies seeking to develop and internally validate prognostic models. Results support the need for future studies to further validate the derived model with potential clinical implications in assessment to inform treatment. Overall, the manuscript is well-written, though I have numerous minor points throughout to enhance clarity with some points to enhance conciseness of the discussion section. Whole manuscipt • Throughout the whole of the manuscript including the title, please re-consider use of the term ‘chronification’. This terminology is unclear as uncommon. The manuscript’s published protocol describing a model to ‘predict chronic pain’ is much clearer and suggest considering use of this phrase. • Throughout the whole of the manuscript, please use the same terminology to describe the population as at present it is variable (e.g. Line 28: “Non-specific, non-traumatic neck pain”; Line 104: “nonspecific idiopathic, non-traumatic neck pain) Title • Please add the setting to physiotherapy practice to enhance clarity (i.e., primary care) Abstract

	• Line 24: Please clarify the contrast between Line 24 indicating emphasis on psychosocial factors and lines 51 and 78 indicating biomedical and psychosocial factors • Line 27: Clarify the setting is 30 primary care physiotherapy practices • Line 30: Please align categories of candidate prognostic variables with that described in methods within the manuscript (e.g., not clear what 'prior conditions' is referring to) Intro The intro generally sets up the need for the research study nicely. • The authors may want to considering referencing the newly published Lancet article on the global burden of neck pain which also has projections to 2050, as this is likely to provide the most recent estimates of neck pain prevalence and burden, along with data to support the need for this research study (e.g., projected increase in prevalence of neck pain) • In para 2 on Line 61, the link between moderate effect of physiotherapy on neck pain and need for identifying prognostic factors can be strengthened to better set up the need for how this study will have clinical impact. • The use of "(sub)acute" in the introduction is not immediately clear what this is referring to; please consider use of "acute or subacute neck pain" or at least defining in the introduction Methods Mostly clear and concise. Includes all relevant methodological points. • Line 111: Please define body regions • Line 144: Please edit the section on psychological and behavioral factors to enhance readability. For example, use of one sentence structure to describe the measure and construct being assessed would be helpful. • Line 181: clarify 'missing's remaining available for further analyses' Results Addresses research objectives and are mostly clear with a couple points to consider in addressing potential risk for bias. • Were there any differences in baseline characteristics between the participants who completed follow up surveys and those who did not? May be a potential bias in the observed relationships between predictors and outcome • Please standardize decimal points in table 1 • Line 238: Is "increased scores on catastrophizing" intended to mean higher catastrophizing scores or truly an increase? If an increase, over what time frame? • Figure 2 is useful to demonstrate raw data. Please clarify the difference between the 2 figures within Fig 2. If possible, please also increase the font size for the OR and CI as very difficult to read owing to small size. • Employment / work status is indicated as a prognostic factor in the univariable and multivariable models, however it is not included in the list of prognostic factors in either the protocol or the methods of the manuscript – please clarify. • Lines 263-264: Please clarify as I'm interpreting this sentence to suggest work status is a phrase equivalent to ability to change posture at work, but in results not working (i.e. work status) is distinct from ability to change posture. • Line 289: Incomplete sentence?

	• The published protocol indicates the physiotherapy treatments received by participants will be reported and the possible impact of treatments received will be discussed. However, I do not see these results or a discussion of their potential impact. Please either include or justify why not included, as treatments could be an important confounder. Discussion • Suggest restructuring first para on Line 353 to enhance clarity and alignment with study objectives. For example, at present it appears to discuss the multivariable model, then individual predictors and then back to the multivariable model, which contributes to lack of clarity of which results are being discussed. • Line 361-262: Incomplete sentence? • Line 263: please clarify existing headache as unclear if referring to headache before neck pain onset or after neck pain onset • The first objective of identifying individual prognostic factors is minimally discussed in the discussion section. Suggest adding a short para discussing individual factors and their context within the wider literature. • Employment status was by far the strongest prognostic variable within the model, but not discussed at all. I can appreciate it is because it was categorized as a non-modifiable factor from a PT perspective, but still really useful knowledge to have. Please consider if this can be touched on within the discussion section – perhaps para 2? • Line 368 Para 2: discussion on self-efficacy could be more explicit in that this is likely the opposite relationship one would expect - otherwise left up to reader to interpret in this way. This para can also be strengthened by highlighting the common prognostic factors identified in this study and the wider literature, even if wider MSK populations, as this seems to have a nice message for clinical practice • Line 374 Para 3: The OR for the various belief factors within the model are quite small, and on par with self-efficacy. In the prior paragraph, the low OR for self-efficacy is discussed to aid interpretation of results and suggest the same is done for this paragraph. • Line 378-380: please clarify as my interpretation of this is that the first sentence is describing the value of illness perceptions within prognostic models, but then the second sentence is indicating models often don't include illness beliefs • Line 388: 'did not retain' should likely be something like "were not retained" • Line 398-401: Please clarify as I'm not sure the point trying to be made in this sentence, as this study focused on pain outcomes and not changes in psychological factors • Line 402 Para 5: This short para feels tangential to the discussion of identifying factors predicting development of chronic pain. Please consider whether it is critical to this discussion, or if it can be more concise and integrated with the prior para. • Line 414: Please substitute "participants" for "them" • Line 406 Para 6: I agree with a justification for the definition of chronic pain as there is important implications for incidence and limitations linked to sample size, but this can be substantially more concise to reduce length. Could also consider adding key points to limitations section and deleting para. • Line 418-420: There may be value in adding something briefly along these lines to the introduction, to 'set the scene' and support the definition of outcome – e.g., need for models to better

	differentiate subgroups of chronic neck pain, such as persistent and recurrent • Para on Line 421: This para appears to be justifying a methodological decision, so suggest enhancing the conciseness and shifting key points to the methods section. Alternatively, if it is intended to be a discussion of the limitation of using the 3/10 cut off, please move key points to the limitations section. • Para on line 430: Please consider whether this short paragraph is critical to the discussion of model development to predict chronic pain. While I agree with the point being made, I'm not sure it is critical to the discussion of prognostic factors, when the discussion section is already quite long • While identified as a limitation in the published protocol, please add the limitation of use of non-validated outcome measures as prognostic factors, owing to some non-validated measures being retained in the final model • Line 456: Define NSNP

REVIEWER NAME	Wingbermühle, Roel
REVIEWER AFFILIATION	SOMT
REVIEWER CONFLICT OF INTEREST	None.
DATE REVIEW RETURNED	27-May-2024

GENERAL COMMENTS	Authors developed and internally validated a multivariable prognostic model for chronic pain in patients with a new episode of (sub)acute non-specific non-traumatic neck pain presenting to primary care physiotherapy, with an emphasis on modifiable psychosocial factors. The authors concluded that the developed model, that included mostly potentially modifiable factors for physiotherapy practice, has the potential to obtain a valid prognosis for chronification of (sub)acute non-specific neck pain and they recommend model's external validation. Prediction models are increasingly abundant in health literature and can play an important role in personalised clinical care. I read the study with great interest and wish to compliment the authors for their well conducted study, and efforts made to enhance the knowledge on multivariable prognostic models for neck pain patients. Overall, the writing is clear and follows the TRIPOD statement. Appropriate methods to develop and internally validate the model were used. There are still a few items I would like to see clarified and addressed by the authors. 1. Abstract, line 27: I suggest using 30 primary care physiotherapy practices instead of 30 primary care physiotherapy. 2. Line 130-131. Reference (11) is not referring to the study protocol (this is reference 14). 3. Good that the calculation as recommended by Riley et al. 2019 was used to ensure an adequate sample size. Riley uses a minimum number of events per predictor parameter. The study considered 26 candidate prognostic factors (line 165-166). Was it recognised that predictors with more than 2 categories requires more than 1 parameter for this calculation? If not, how does this affect the samples size calculation and results?

	4. Multicollinearity was assessed before modelling (line 186). Where also the linear relationships between continuous predictor variables and the outcome evaluated, and if necessary, transformations considered? 5. The number of participants at 6 months is inconsistent. The flow-chart (line 214) shows n=391 participants at 6 months and n=603 participants at baseline. The number of loss to follow-up at the 6-months follow-up increased to 213 (line 227). 603-231results in 372 participants and not the shown 391 participants. Furthermore, it is not clear for me how the outcome of n=603 (62 chronic neck pain+541 no chronic neck pain) participants was assessed, since 78 participants did not complete any follow-up measurement (line 224-225). 6. The calculation of the calibration curve (figure 5, line 294) is not completely clear to me, I would expect to see (dots for) groupings of individuals. 7. Please correct small typo's in the intermezzo: Cur – rently (line 307); under- standing (line 311). 8. In line 399, a word after musculoskeletal seems missing. In line 449-450, the word "in" between insight and an seems missing: a physiotherapist can utilize this model to gain insight an individual patient's probability of experiencing chronic neck pain. 9. In line 456 the abbreviation NSNP is used for the first (and last) time, better write in full. 10. The model seems to overestimate the higher risks. It is stated that this overestimation is unlikely to remain visible in an external validation with enough participants at high risk (line 438). Can you substantiate this? 11. The objective of this study was to develop and internally validate a model, with an emphasis on modifiable psychosocial factors (line 23-24) and the model can be clinically beneficial to assess and intervene on the model's modifiable factors (line 450-451). Also, it is mentioned that a physiotherapist can utilize this model to gain insight in an individual patient's probability of experiencing chronic neck pain (499-450). Should readers consider that the emphasis on modifiable psychosocial factors now gives us a model that provides a different individual prognosis compared to a model that would have been developed without this emphasis?

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1

Thank you for the opportunity to review this research study that aimed to identify which modifiable factors are independent prognostic factors of the development of chronic neck pain in patients with acute- or subacute neck pain, and to develop and internally validate a model to predict development of chronic pain. Owing to the high and increasing prevalence and burden of chronic neck pain, these research aims are clinically important for physiotherapy. This study was conducted in accordance with a published protocol and gold standard frameworks for studies seeking to develop and internally validate prognostic models. Results support the need for future studies to further validate the derived

model with potential clinical implications in assessment to inform treatment. Overall, the manuscript is well-written, though I have numerous minor points throughout to enhance clarity with some points to enhance conciseness of the discussion section.

Whole manuscript

Comment 1

Throughout the whole of the manuscript including the title, please re-consider use of the term 'chronification'. This terminology is unclear as uncommon. The manuscript's published protocol describing a model to 'predict chronic pain' is much clearer and suggest considering use of this phrase.

Our protocol also uses the term "chronification" of neck pain, which we use logically in this paper as well. However, we understand that this is not common in the literature. Therefore, we changed "chronification" into "developing chronic pain" and "predicting chronic pain" in the manuscript.

Comment 2

Throughout the whole of the manuscript, please use the same terminology to describe the population as at present it is variable (e.g. Line 28: "Non-specific, non-traumatic neck pain"; Line 104: "nonspecific idiopathic, non-traumatic neck pain)

Thank you for this comment. We have changed the terminology to "acute—or subacute nonspecific idiopathic, non-traumatic neck pain" throughout the manuscript, which aligns with the protocol's terminology.

Title

Comment 3

Please add the setting to physiotherapy practice to enhance clarity (i.e., primary care)

The title has been changed to "Development and internal validation of a multivariable prognostic model to predict chronic pain after a new episode of non-specific idiopathic, non-traumatic neck pain in physiotherapy primary care practice." This title provides clearer information about the patient population and study setting. These changes were applied throughout the manuscript.

Abstract

Comment 4

Line 24: Please clarify the contrast between Line 24 indicating emphasis on psychosocial factors and lines 51 and 78 indicating biomedical and psychosocial factors.

The contrast was not intended to be presented. Where there are more psychological factors than social factors, we include biomedical, psychological, and social factors. We have adjusted every sentence to include biomedical, psychological, and social factors. (see p. 2, line 23, p. 4 , line, 52, p. 6, line 82)

Comment 5

Line 27: Clarify the setting is 30 primary care physiotherapy practices

See comment 3.

Comment 6

Line 30: Please align categories of candidate prognostic variables with that described in methods within the manuscript (e.g., not clear what 'prior conditions' is referring to)

We revised our baseline measures to align with our study protocol and the description of candidate prognostic factors in our paper:

See p. 2, line 30: "Candidate prognostic variables collected from participants included their age and sex, neck pain symptoms, work-related factors, general factors, psychological and behavioural factors, and the remaining factors: therapeutic relation and healthcare provider attitude."

Introduction

The intro generally sets up the need for the research study nicely.

Comment 7

The authors may want to considering referencing the newly published Lancet article on the global burden of neck pain which also has projections to 2050, as this is likely to provide the most recent estimates of neck pain prevalence and burden, along with data to support the need for this research study (e.g., projected increase in prevalence of neck pain)

Thank you for directing us to the latest Lancet article. We have added the reference (see p. 5, line 61-63): The estimated global number of neck pain cases is projected to be 269 million (219–322) by 2050, an increase of 32.5% (23.9–42.3) from 2020 to 2050.(7)

Comment 8

In para 2 on Line 61, the link between moderate effect of physiotherapy on neck pain and need for identifying prognostic factors can be strengthened to better set up the need for how this study will have clinical impact.

Thank you for this comment. I have linked the different sentences in this paragraph more clearly to illustrate the relationship between the steps and the potential clinical impact.

See p. 5, line 64-69:

"Physiotherapy is a common first-line treatment; however, its effectiveness in patients with chronic pain is often only moderate.(8)(9)(10) Consequently, identifying prognostic factors to predict chronic pain is a top priority for neck pain research and for clinical care.(11) By identifying these prognostic factors, especially modifiable factors, physiotherapists can make more informed decisions, potentially target modifiable factors, and prevent the development of chronic idiopathic neck pain."

Comment 9

The use of "(sub)acute" in the introduction is not immediately clear what this is referring to; please consider use of "acute or subacute neck pain" or at least defining in the introduction

Throughout the manuscript, we changed all mentions of '(sub)acute' to 'acute or subacute' and 'acute and subacute' to more accurately describe our researched population.

Methods

Mostly clear and concise. Includes all relevant methodological points.

Comment 10

Line 111: Please define body regions

We provided examples of the different body regions described by the ICD-11 to clarify “body regions”. See p. 8, line 121-123:

“…widespread primary pain (ICD-11) (diffuse musculoskeletal pain in at least 4 of 5 body regions (e.g. shoulder or upper arm, wrist or hand, pelvis, or ankle or food) and in at least three or more body quadrants (as defined by upper-lower / left-right side of the body) and axial skeleton (neck, back, chest and abdomen.”

Comment 11

Line 144: Please edit the section on psychological and behavioral factors to enhance readability. For example, use of one sentence structure to describe the measure and construct being assessed would be helpful.

We understand the difficulty of reading this long list of constructs and measurement methods. We removed the constructs mentioned earlier in the paragraph to reduce the text length. Additionally, I have given each sentence the exact same structure, with slight variations in the use of "measured," "assessed," and "evaluated" to maintain readability instead of listing them all at once. See p. 10, line 155-164:

Psychological and behavioral factors: Illness perceptions were assessed using the Dutch version of the Brief Illness Perception Questionnaire (IPQ-DLV).(27) Catastrophizing was measured with the short version of the Pain Catastrophizing Scale (PCS).(28) Depression and distress were assessed with the 21-item version of the Depression Anxiety Stress Scale (DASS-21).(29) Kinesiophobia was measured using the 11-item version of the Tampa Scale for Kinesiophobia (TSK).(30) Coping strategies were evaluated with the Pain Coping Inventory (PCI).(31)(32) Hypervigilance was assessed using the Pain Vigilance and Awareness Questionnaire (PVAQ)(33), and self-efficacy in managing pain was measured with the 2-item version of the Pain Self-Efficacy Questionnaire.(34)

Comment 12

Line 181: clarify ‘missing’s remaining available for further analyses’

We understand that this sentence was not clear. We changed it to: “As a result, we had to proceed with the available data during the subsequent analysis, even though a significant portion was missing.” (see p. 11, line 190-191)

Results

Addresses research objectives and are mostly clear with a couple points to consider in addressing potential risk for bias.

Comment 13

Were there any differences in baseline characteristics between the participants who completed follow up surveys and those who did not? May be a potential bias in the observed relationships between predictors and outcome

We did not compute and test differences between complete and incomplete records, as all participants were used for the analysis, regardless of missing data. Our methodology used multiple imputation (MI) with a fully conditional specification to impute incomplete records, so that all participants were completed before analyses. Assuming that the data are at least missing at random (MAR).Before the MI, we conducted Little's MCAR test, which resulted in a p-value greater than 0.05, indicating that the probability of missing data depends only on the observed data, not the unobserved data. These findings justify the application of MI to fill in the missing follow-up data, minimizing potential bias. MI is a robust method that reduces bias from missing data.

We initially did not report the results of Little's MCAR test but have now included this in our results section (see p. 14, line 232-233)

The Little's MCAR test yielded a p-value greater than 0.05, supporting the appropriateness of multiple imputations.

We also add the following literature to support this:

Pedersen A, Mikkelsen E, Cronin-Fenton D, Kristensen N, Pham TM, Pedersen L, et al. Missing data and multiple imputation in clinical epidemiological research. Clin Epidemiol. 2017 Mar;Volume 9:157–66.

Comment 14

Please standardize decimal points in table 1

We standardized decimal points to one decimal and adjusted the number accordingly. However, it is not necessary to add decimal points to the median on the 0-10 point scale.

Comment 15

Line 238: Is "increased scores on catastrophizing" intended to mean higher catastrophizing scores or truly an increase? If an increase, over what time frame?

Thank you for this comment. It meant "higher scores" instead of "increase". We have changed "increased" to "higher" in the manuscript.

Comment 16

Figure 2 is useful to demonstrate raw data. Please clarify the difference between the 2 figures within Fig 2. If possible, please also increase the font size for the OR and CI as very difficult to read owing to small size.

Due to the limit on the number of figures allowed, we had to merge the two figures. They are now included in the article with full clarity, improving the visibility of the OR and CI. The first figure displays the continuous variables, while the second illustrates the categorical and dichotomous outcomes. These two variables yield distinctly different odds ratios, so combining them into a single overview would not represent the data clearly. I have noted this clarification beneath the figure in the article:

"The first figure displays the continuous variables, while the second illustrates the categorical and dichotomous variables"

Comment 17

Employment / work status is indicated as a prognostic factor in the univariable and multivariable models, however it is not included in the list of prognostic factors in either the protocol or the methods of the manuscript – please clarify.

During our analysis, we encountered an issue where 92 of our 603 participants were not employed. This posed a challenge because our list of potential prognostic factors included three work-related factors: "Happiness at Work," "Job Satisfaction," and "Potential to Self-Modify Posture." Consequently, this resulted in unavoidable missing data within our dataset because of a category we did not intend to measure. Imputation was neither possible nor appropriate.

Excluding these participants would have meant constructing a prognostic model solely applicable to those employed, which was not our intention. To include these individuals, we did assign an additional category, 'Not Working', to the three work-related variables. This allowed seamless inclusion in the multivariable analysis.

We did not describe this choice in our paper; we now address it on p. 14, line 226-228:

Of our 603 participants, 92 (15.3 %) did not work. We included these participants as not working in all the work-related factors in our multivariable analyses.

However, this approach naturally led to multicollinearity. Therefore, we decided to focus solely on the factor "Ability to Change Posture at Work," as detailed in our manuscript see p. 18, line: 261-268.

Additionally, we conducted multivariable analyses with each of the three work-related variables separately, which minimally affected the quality of the model.

Comment 18

Lines 263-264: Please clarify as I'm interpreting this sentence to suggest work status is a phrase equivalent to ability to change posture at work, but in results not working (i.e. work status) is distinct from ability to change posture.

We hope our response to comment 17 and the additional text in our paper clarifies how we handled the factor 'work' and contrasted it with the variable 'Ability to change posture at work'.

Comment 19

Line 289: Incomplete sentence?

This was indeed an incomplete sentence; the reference to the intermezzo section was already given earlier in the paragraph.

Comment 20

The published protocol indicates the physiotherapy treatments received by participants will be reported and the possible impact of treatments received will be discussed. However, I do not see these results or a discussion of their potential impact. Please either include or justify why not included, as treatments could be an important confounder.

You are right, and thank you for including the protocol in the review of our results paper. We have added a detailed overview of the interventions applied across the study population in a Table and a Figure (see Appendix 3 on page 40).

Additionally, in the results section, we described the most frequently applied therapies (p. 14, line 232-238): "The interventions most frequently applied were (1) joint mobilization, manipulation, traction, and nerve mobilization techniques, with an application rate of 85.4%, and (2) information and

advice, with an application rate of 86.7%. Exercise and massage were applied to 58.1% and 54.7% of the study population, respectively. For a detailed overview of the interventions applied across the study population, see Appendix 3."

In our discussion section (p. 25, line 438-446):

"Furthermore, in our study's protocol discussion, we noted that our study did not influence the therapies participants received; however, these therapies could potentially affect both the outcomes and the accuracy and generalizability of the developed model. Participants were treated according to the Dutch Physiotherapy Guideline for neck pain, which might modify our candidate prognostic factors and potentially reduce chronicity risks. Given the diversity of factors, the variety of modalities used by physiotherapists, and the therapists' varied backgrounds, we considered the impact of these therapies on our study results minimal. Ideally, these therapies would not be applied or analyzed within the multivariable prognostic model to assess their impact; however, this was not feasible due to sample size constraints."

Discussion

Comment 21

Suggest restructuring first para on Line 353 to enhance clarity and alignment with study objectives. For example, at present it appears to discuss the multivariable model, then individual predictors and then back to the multivariable model, which contributes to lack of clarity of which results are being discussed.

The revised description of the study results is now aligned with the dual objectives of the study. Previously, the results addressing the first objective were somewhat overlooked. The first paragraph of the discussion has been updated accordingly (see p. 22, line 350-357)

"In this prospective cohort study, we (1) identified which (modifiable factors) are independent prognostic factors of the development of chronic neck pain, and we (2) developed and internally validated a prognostic model for predicting chronic pain after a new episode of acute- or subacute nonspecific idiopathic, non-traumatic neck pain. We found several significant associations between non- and modifiable factors and chronic pain: being female, higher pain intensity at baseline, longer duration of neck pain, experiencing pain in different body regions, the onset of headache since the neck pain began, higher disability scores, unemployment, higher scores on catastrophizing, illness beliefs about recovery (concerned and duration), depression, distress, and lower treatment beliefs."

Comment 22

Line 361-262: Incomplete sentence?

We add "outcome of chronic pain" to this sentence (see p. 22, line 361-363):

"Nonetheless, the model's corrected R2 of 0.24 suggests that the model provides a meaningful but limited explanation of the probability distribution of the outcome of chronic pain."

Comment 23

Line 263: please clarify existing headache as unclear if referring to headache before neck pain onset or after neck pain onset

We changed the headache description to "since the onset of neck pain" and "headache(s) before the neck pain" to match the description of the candidate prognostic factors and our protocol.

See p. 18, line 270-271:

“These included sex (female), higher pain intensity at baseline, reported pain in different body regions, headache since the onset of neck pain, headache(s) before the neck pain, an inability….”

Comment 24

The first objective of identifying individual prognostic factors is minimally discussed in the discussion section. Suggest adding a short para discussing individual factors and their context within the wider literature.

See also our reply and the changes that we made in comment 21. We also compared these variables in the discussion section. We clarified the text further. See, for example, p. 22, line 369-370:

“When comparing our individual prognostic factors and those included in our prognostic model with existing prognostic studies in musculoskeletal pain, several common factors emerge…”

And p. 23 line: 393-397:

“Furthermore, it is important to note that several psychological factors, such as depression, kinesiophobia, catastrophizing, and poor coping skills, are commonly recognized as associated with and prognostic for chronic pain.(57)(14) These factors were not retained in our final prognostic model. Although these factors showed an association in our univariable analysis, they did not improve the predictive accuracy of our model…”

Comment 25

Employment status was by far the strongest prognostic variable within the model, but not discussed at all. I can appreciate it is because it was categorized as a non-modifiable factor from a PT perspective, but still really useful knowledge to have. Please consider if this can be touched on within the discussion section – perhaps para 2?

Indeed, this represents a significant prognostic value we previously did not address. We now discuss this in more detail in our discussion session (see p. 22-23, line 372-375).

“In our study, not working showed a high OR in both univariable and multivariable analyses. A physiotherapist cannot directly modify this factor; however, attention could be given to potentially modifiable factors associated with unemployment, such as physical disability and mental health.(50,51)”

Comment 26

Line 368 Para 2: discussion on self-efficacy could be more explicit in that this is likely the opposite relationship one would expect - otherwise left up to reader to interpret in this way. This para can also be strengthened by highlighting the common prognostic factors identified in this study and the wider literature, even if wider MSK populations, as this seems to have a nice message for clinical practice.

We believe we address the common factors prevalent in the broader musculoskeletal literature in the second, third, and fourth paragraphs. We provide a more detailed analysis of the notable self-efficacy outcome, specifically clarifying the potential sources of bias for this result. See p. 23, line 375-380:

“In addition, in our study, a higher score on the Pain Self-Efficacy Questionnaire 2-item version was associated with higher odds of chronic neck pain. Notably, this association was characterized by a low regression coefficient and OR and was insignificant with a small CI. Moreover, this outcome may be biased using this short questionnaire, where the largest group of our population scored above 10

on a 0-12 point scale for self-efficacy, exhibiting a known ceiling effect. (REF) This notable outcome might, therefore, be questioned."

Comment 27

Line 374 Para 3: The OR for the various belief factors within the model are quite small, and on par with self-efficacy. In the prior paragraph, the low OR for self-efficacy is discussed to aid interpretation of results and suggest the same is done for this paragraph.

Indeed, the various belief factors exhibit relatively small ORs (i.e. close to 1), despite the fact that some of these prognostic factors (such as concerns, duration beliefs, and treatment beliefs) are significant. It is crucial to consider the scale of the outcome measures. It is crucial to consider the scale of the outcome measures. An OR of 1.2 or 1.3 on an 11-point scale indicates a relatively modest association, as the continuous nature of the scale typically results in smaller OR values, as the OR applies to each increase of one unit of the predictor. In contrast, an OR of approximately 3 for a dichotomous outcome (e.g., work status, headache, or pain in different body regions) signifies a much stronger association. The dichotomous nature amplifies the OR. Therefore, when interpreting OR values, the scale of the outcome measure must be considered to accurately assess the strength of the observed associations.

Comment 28

Line 378-380: Please clarify. My interpretation is that the first sentence describes the value of illness perceptions within prognostic models, but the second sentence indicates that models often don't include illness beliefs.

We do understand that our sentences are confusing. Initially, we describe our findings regarding illness perceptions and note similar associations compared to longitudinal studies on low back pain, which do not involve prognostic modelling. Subsequently, we discuss how illness perceptions contribute to the robustness of a series of models. However, it is important to note that in prognostic studies on neck pain, illness perceptions are generally not included.

To clarify this, we have rewritten this part in our discussion (p. 23, line 381-389):

"Our model incorporated four illness perception factors: beliefs about recovery (including concerns and duration), identity, and treatment beliefs. Longitudinal studies on low back pain have yielded similar findings, illustrating individual associations between illness beliefs (e.g., duration and treatment beliefs) and negative clinical outcomes over various time periods.(53–55) However, in prognostic multivariable models, the contribution of illness perceptions to the robustness of a prognostic model varies.(55,56) Notably, illness beliefs are often excluded from the candidate prognostic factors in models developed and externally validated for neck pain models.(12,57–59)"

Comment 29

Line 388: 'did not retain' should likely be something like "were not retained"

We changed this sentence to a grammatically correct sentence (see. p. 23, line 395):

These factors were not retained in our final prognostic model.

Comment 30

Line 398-401: Please clarify as I'm not sure the point trying to be made in this sentence, as this study focused on pain outcomes and not changes in psychological factors

We understand your concerns. We want to emphasize that we observe changes in pain intensity when these variables are treated. However, these measures do not indicate whether there is also a change in the variables themselves when they are treated. To ensure effective treatment of these variables, it is important to determine if changes in these variables are the reason for the reduction in pain. We have tried to clarify this on p. 24, line 405-409.

“In contrast, it is important to note that the most studies involving patients with acute- and subacute musculoskeletal pain have mainly focused on pain and disability as outcomes. However, these studies, which investigate the effectiveness of treating physiological factors, should also examine whether identified changes in these psychological factors contribute to the reduction in pain intensity or disability observed in their study population.”(63)(64)(65)

Comment 31

Line 402 Para 5: This short para feels tangential to the discussion of identifying factors predicting development of chronic pain. Please consider whether it is critical to this discussion, or if it can be more concise and integrated with the prior para.

Although we think it is an important point, it is not the most relevant for the discussion section of this research paper. Therefore, we deleted this small paragraph from the discussion section.

Comment 32

Line 414: Please substitute “participants” for “them”

Thank you for this comment. We changed “them” to “participants”.

Comment 33

Line 406 Para 6: I agree with a justification for the definition of chronic pain as there is important implications for incidence and limitations linked to sample size, but this can be substantially more concise to reduce length. Could also consider adding key points to limitations section and deleting para.

We have significantly condensed this paragraph. The reduced sample size due to our choice of definition is indeed a limitation of the study, which we discuss in our limitation section. However, we aim to highlight a strength of our study through this paragraph: the use of a comprehensive approach combined with the development of more effective models through our chosen methodology. See our revised paragraph on p. 24, line 410-416:

“The incidence of chronic pain in our participants differed from our systematic review findings. Our preliminary sample size calculation assumed a 45% chronicity rate for neck pain, which divided the number of patients by the non-recovery cases.(12) This disparity can be attributed to our definition of chronic pain and the definition of the measurement approach. Unlike most studies that use single time point assessment (e.g. 3, 6, or 12 months) with specific pain score threshold(71), including those in our review(12), our study used a more comprehensive approach. This approach provides a precise representation of chronic pain as a continuous experience.”

Comment 34

Line 418-420: There may be value in adding something briefly along these lines to the introduction, to ‘set the scene’ and support the definition of outcome – e.g., need for models to better differentiate subgroups of chronic neck pain, such as persistent and recurrent

Thank you for this suggestion. We have now added the varying definition of the outcome as a potential explanation for the low performance of the existing models. See p. 5, line 75-77.

“Furthermore, the varying definitions of the outcome, including persistent and/or recurrent pain groups, contribute to the low performance of these models.”

Comment 35

Para on Line 421: This para appears to be justifying a methodological decision, so suggest enhancing the conciseness and shifting key points to the methods section. Alternatively, if it is intended to be a discussion of the limitation of using the 3/10 cut off, please move key points to the limitations section.

We have relocated this point to the limitations section and revised the paragraph to more clearly address our arbitrary choice of a cut-off point of 3 and its influence on our study results. See p. 25, line 436-437:

“….. Therefore, choosing a threshold of 3 is debatable, and selecting a different threshold could yield different study results.”

Comment 36

Para on line 430: Please consider whether this short paragraph is critical to the discussion of model development to predict chronic pain. While I agree with the point being made, I’m not sure it is critical to the discussion of prognostic factors, when the discussion section is already quite long

We understand your point. Thanks to your critical input. We will discuss this topic in our next paper, as we believe it is important to consider standardizing an international threshold for chronic pain. Consequently, we have deleted this paragraph in our discussion section.

Comment 37

While identified as a limitation in the published protocol, please add the limitation of use of non-validated outcome measures as prognostic factors, owing to some non-validated measures being retained in the final model

We did not address this in our results paper. However, in the limitations section, we now reflect on including a factor in our final model that was measured using a subjective, non-validated method. See p. 25-26, line 447-454:

“Our final prognostic model retained the factor 'self-modifying posture during work'. This factor was measured subjectively using a non-validated question, which poses a limitation as it may not distinguish between perceived and actual behavior during work. The limitation of this subjective measurement lies in its inability to clearly distinguish whether individuals perceive that they can change positions during work or are changing their positions. Additionally, this type of questioning prevents us from confirming the accuracy of reports, such as whether a patient who claims they cannot change positions is indeed unable to do so. Establishing the validity and discriminative ability of the different concepts being tested is important to investigate.”

Comment 38

Line 456: Define NSNP

See comment 2 and comment 9 of reviewer 2.

Reviewer: 2

Authors developed and internally validated a multivariable prognostic model for chronic pain in patients with a new episode of (sub)acute non-specific non-traumatic neck pain presenting to primary

care physiotherapy, with an emphasis on modifiable psychosocial factors. The authors concluded that the developed model, that included mostly potentially modifiable factors for physiotherapy practice, has the potential to obtain a valid prognosis for chronification of (sub)acute non-specific neck pain and they recommend model's external validation.

Prediction models are increasingly abundant in health literature and can play an important role in personalised clinical care. I read the study with great interest and wish to compliment the authors for their well conducted study, and efforts made to enhance the knowledge on multivariable prognostic models for neck pain patients. Overall, the writing is clear and follows the TRIPOD statement. Appropriate methods to develop and internally validate the model were used. There are still a few items I would like to see clarified and addressed by the authors.

We would like to thank the reviewer for this positive feedback.

Comment 1

Abstract, line 27: I suggest using 30 primary care physiotherapy practices instead of 30 primary care physiotherapy.

See also comment 3 of reviewer 1. We changed all instances of the settings to "physiotherapy primary care practices" to clearly indicate the setting in which the research was conducted.

Comment 2

Line 130-131. Reference (11) is not referring to the study protocol (this is reference 14).

Thank you for your sharp review of this manuscript. We changed the reference.

Comment 3

Good that the calculation as recommended by Riley et al. 2019 was used to ensure an adequate sample size. Riley uses a minimum number of events per predictor parameter. The study considered 26 candidate prognostic factors (line 165-166). Was it recognised that predictors with more than 2 categories requires more than 1 parameter for this calculation? If not, how does this affect the samples size calculation and results?

We dichotomized predictors with more parameters and included only the predictors in the sample size calculation. We only took more parameters in our analysis for the predictors 'Headache' and work-related factors ('not working') instead of considering one predictor in the sample size calculation. This resulted in 31 predictors in our statistical analysis, compared to 26 in the sample size calculation. Details on the scale range and our approach to variable handling in statistical analysis are outlined in Table 1 of our protocol:

Verwoerd MJ, Wittink H, Maissan F, van Kuijk SMJ, Smeets RJEM. A study protocol for the validation of a prognostic model with an emphasis on modifiable factors to predict chronic pain after a new episode of acute- or subacute nonspecific idiopathic, non-traumatic neck pain presenting in primary

care. PLoS One. 2023 Jan 17;18(1):e0280278. doi: 10.1371/journal.pone.0280278. PMID: 36649242; PMCID: PMC9844852.

Comment 4

Multicollinearity was assessed before modelling (line 186). Where also the linear relationships between continuous predictor variables and the outcome evaluated, and if necessary, transformations considered?

To prevent overfitting, we did not use data-driven methods (e.g. restricted cubic splines) to assess (non-)linearity of the association. Based on hypotheses, some predictors were dichotomized. No other continuous predictor variables were transformed, or modeled using more complex associations, because we did not have prior hypotheses on non-linear associations.

Comment 5

The number of participants at 6 months is inconsistent. The flow-chart (line 214) shows n=391 participants at 6 months and n=603 participants at baseline. The number of loss to follow-up at the 6-months follow-up increased to 213 (line 227). 603-231results in 372 participants and not the shown 391 participants.

Furthermore, it is not clear for me how the outcome of n=603 (62 chronic neck pain+541 no chronic neck pain) participants was assessed, since 78 participants did not complete any follow-up measurement (line 224-225).

Thank you for your alertness. The number of participants who did not submit the required forms at 6 months was 211. We changed the 231 to 211 in the text and updated the figure from 391 to 392. We reviewed all the missing data in our flowchart and text. See our text at p. 14 and line. 231-232.

“This number changed to 224 at the 3-months follow-up, and to 211 at the 6-month mark.

We had a maximum of 3.8% missing data across all baseline variables. We imputed the missing outcome data, allowing us to accurately count the number of chronic and non-chronic patients in our entire dataset. Multiple imputation is a robust and justified method for handling missing data. For further details, please refer to our response to comment 13 of reviewer 1 and the additional literature cited in both our response and our paper.

Comment 6

The calculation of the calibration curve (figure 5, line 294) is not completely clear to me, I would expect to see (dots for) groupings of individuals.

Because grouping of individuals produces a calibration plot that may heavily depend on the arbitrary number of groups that are chosen, we chose to plot a smooth curve.

For clarity, we have explained the tick marks below Figure 4.

“The tick marks at the bottom of the Calibration curve represent the distribution of predicted probabilities. Each tick mark indicated a predicted probability for an individual observation. A dense cluster of tick marks indicated more observations with that specific predicted probability.”

Comment 7

Please correct small typo's in the intermezzo: Cur – rently (line 307); under- standing (line 311).

Thank you for making us aware of these typos. We have corrected them.

Comment 8

In line 399, a word after musculoskeletal seems missing. In line 449-450, the word "in" between insight and an seems missing: a physiotherapist can utilize this model to gain insight an individual patient's probability of experiencing chronic neck pain.

The word "pain" is indeed missing after musculoskeletal; we have corrected this.

We also changed the second sentence to read (see p. 26, line 456-457): "In clinical practice, a physiotherapist can utilize this model to gain insight into a patient's probability of experiencing chronic neck pain."

Comment 9

In line 456 the abbreviation NSNP is used for the first (and last) time, better write in full.

See also comment 2 of reviewer 1. We changed the description of the population to "acute- or subacute nonspecific idiopathic, non-traumatic neck pain", in line with our protocol.

Comment 10

The model seems to overestimate the higher risks. It is stated that this overestimation is unlikely to remain visible in an external validation with enough participants at high risk (line 438). Can you substantiate this?

No, we do not. Our expectation is based on experience. However, we removed this sentence as it was overly assertive without supporting literature.

Comment 11

The objective of this study was to develop and internally validate a model, with an emphasis on modifiable psychosocial factors (line 23-24) and the model can be clinically beneficial to assess and intervene on the model's modifiable factors (line 450-451). Also, it is mentioned that a physiotherapist can utilize this model to gain insight in an individual patient's probability of experiencing chronic neck pain (499-450). Should readers consider that the emphasis on modifiable psychosocial factors now gives us a model that provides a different individual prognosis than a model that would have been developed without this emphasis?

No, the individual prognosis remains the same. However, we now have a model that holds potential, incorporating factors that may be modifiable. Nonetheless, with this model, we do not yet know whether treating these variables would impact the outcome. This may prompt further attention to these variables, potentially leading to therapeutic interventions. However, based on this study, we do not know if changing these variables will alter the prognosis. Further research is necessary to explore this.

We hope to have described this adequately in our clinical application and further research section.

VERSION 2 – REVIEW

REVIEWER NAME	Kowalski, Katie
REVIEWER AFFILIATION	Western University, School of Physical Therapy
REVIEWER CONFLICT OF INTEREST	None.
DATE REVIEW RETURNED	21-Jul-2024

GENERAL COMMENTS	Thank you for the opportunity to review the revised version of this manuscript. The authors have done a nice job responding to comments and making edits to the manuscript as needed, which has improved the clarity and quality of the manuscript. Congratulations authors!

REVIEWER NAME	Wingbermühle, Roel
REVIEWER AFFILIATION	SOMT
REVIEWER CONFLICT OF INTEREST	None.
DATE REVIEW RETURNED	22-Jul-2024

GENERAL COMMENTS	The writing is clear and follows the TRIPOD statement and appropriate methods to develop and internally validate the model were used. The reviewers comments are adequately addressed. I'd like to congratulate the authors with their nice work.