BMJ Open

Predictive models for short-term and long-term improvement in women under physiotherapy for chronic disabling neck pain: a longitudinal cohort study

Tony Bohman,[1] Matteo Bottai,[2] Martin Björklund[3,4]

[1]Department of Neurobiology, Care Sciences and Society, Karolinska Institutet, Stockholm, Sweden
[2]Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden
[3]Department of Community Medicine and Rehabilitation, Umeå University, Umeå, Sweden
[4]Department of Occupational Health Sciences and Psychology, University of Gävle, Gävle, Sweden

Correspondence to
Tony Bohman;
tony.bohman@ki.se

ABSTRACT

Objectives To develop predictive models for short-term and long-term clinically important improvement in women with non-specific chronic disabling neck pain during the clinical course of physiotherapy.

Design Longitudinal cohort study based on data from a randomised controlled trial evaluating short-term and long-term effects on sensorimotor function over 11 weeks of physiotherapy.

Participants and settings Eighty-nine women aged 31–65 years with non-specific chronic disabling neck pain from Gävle, Sweden.

Measures The outcome, clinically important improvement, was measured with the Patient Global Impression of Change Scale (PGICS) and the Neck Disability Index (NDI), assessed by self-administered questionnaires at 3, 9 and 15 months from the start of the interventions (baseline). Twelve baseline prognostic factors were considered in the analyses. The predictive models were built using random-effects logistic regression. The predictive ability of the models was measured by the area under the receiver operating characteristic curve (AUC). Internal validity was assessed with cross-validation using the bootstrap resampling technique.

Results Factors included in the final PGICS model were neck disability and age, and in the NDI model, neck disability, depression and catastrophising. In both models, the odds for short-term and long-term improvement increased with higher baseline neck disability, while the odds decreased with increasing age (PGICS model), and with increasing level of depression (NDI model). In the NDI model, higher baseline levels of catastrophising indicated increased odds for short-term improvement and decreased odds for long-term improvement. Both models showed acceptable predictive validity with an AUC of 0.64 (95% CI 0.55 to 0.73) and 0.67 (95% CI 0.59 to 0.75), respectively.

Conclusion Age, neck disability and psychological factors seem to be important predictors of improvement, and may inform clinical decisions about physiotherapy in women with chronic neck pain. Before using the developed predictive models in clinical practice, however, they should be validated in other populations and tested in clinical settings.

Strengths and limitations of this study

► Strengths of this study are a well-defined sample and thorough development of predictive models.
► The longitudinal design with a short-term and long-term follow-up time, and the inclusion of biopsychosocial prognostic factors in the analyses further strengthens the study.
► A limitation of this study is the relatively small sample.

INTRODUCTION

Neck pain is a common health problem and a cause of substantial disability which has a considerable social and economic impact throughout the world.[1] Hogg-Johnson *et al* estimated the 12-month prevalence of neck pain to range between 30% and 50%, with a higher prevalence among women.[1] The prevalence of neck pain has increased during the last decades, and in 2015 neck and low back pain were the leading causes of disability.[2,3] Neck pain often is unrelated to any other specific pathology.[4] Although single episodes of neck pain generally dissolve over time, they are likely recurrent and may become chronic.[5] A review from Carroll *et al* concluded that 50%–85% of persons with neck pain do not experience complete resolution of their pain.[5]

Prognostic research can help understand the course and future outcome in individuals with neck pain.[6] Guzman *et al* summarised the work of the 2000–2010 Neck Pain Task Force concluding that younger age, no previous pain, good physiological and psychological health, good coping, good social support, exercise and sports, and no prior sick leave may increase the chance of recovery from neck pain.[7] Walton *et al* surveyed prognostic factors for prolonged recovery from neck

pain in 13 systematic reviews.[8] They found evidence only for history of musculoskeletal disorders and older age to prolong recovery in non-whiplash related neck pain. For the remaining factors evaluated, there was insufficient evidence for the influence on neck pain, and more research was recommended.

Neck-pain patients commonly seek physical therapy, but predicting treatment outcome and prognosis for these patients is challenging.[9] Predictive models based on multiple prognostic factors could guide healthcare providers, such as physiotherapists, in determining which patients are more likely to improve and to help all patients develop informed expectations.[10 11] It is generally recommended to build predictive models that are applicable to a well-defined patient population and the healthcare system.[10] Women have a higher risk of chronic neck pain than men, and it has been suggested that women should be assessed separately from men in prognostic research.[12 13] Using the search strategy for predictive models in physiotherapy and musculoskeletal complaints suggested by van Oort *et al*, we found only one study assessing chronic non-specific neck pain, and none with a model specific for women.[14 15]

The individual perception of pain and disability vary during the course of an episode, and the effect of prognostic factors is therefore expected to change over time following physiotherapy.[15 16] The analysis of longitudinal data is effective in evaluating these potential changes over time, as it can separate the change between individuals from that within any given patient. Longitudinal analyses have been recommended in prospective cohort studies on musculoskeletal problems.[17 18] The aim of this study was to develop predictive models for short-term and long-term clinically important improvement in women with chronic disabling neck pain in the clinical course of physiotherapy.

METHODS
Design and study population
This longitudinal cohort study sought to develop predictive models, including multiple prognostic factors jointly, for chronic neck pain in women. Predictive models are one of the four themes of the Prognosis Research Strategy Framework for Prognostic Studies.[6 11] Carroll *et al* classify predictive models as 'Phase II' studies in a 3-level hierarchy of evidence from longitudinal studies, and suggests them suitable for predicting recovery from neck pain.[5]

We used data from a randomised controlled trial (RCT) evaluating the short-term and long-term effects, following 11 weeks of physiotherapy interventions (coordination exercise, strength training and massage), on sensorimotor function in 108 women with non-specific chronic disabling neck pain.[19] The trial was carried out in Gävle, Sweden, in 2008.

The RCT included Swedish-speaking women aged 25–65 years with non-specific disabling neck pain lasting for 3 months or more (chronic). The women were recruited via the social insurance agency and with advertisement in local papers and invitations at municipality and county council work sites, primary and occupational healthcare units. Non-specific chronic disabling neck pain was defined as neck pain reported by the patient as the 'most painful area' along with disability, measured as >9 normalised points of the first 19 items in the Disability Arm Shoulder Hand questionnaire.[20 21] These 19 items refer to disability in activities of daily living regarding neck, shoulders and arms. Excluded were individuals with trauma to the head or neck, a diagnosis of rheumatic, neurological, connective tissue, inflammatory or endocrine disease or psychiatric diagnosis affecting their everyday life, fibromyalgia, cancer, stroke, cardiac infarction, diabetes type 1, cervical radiculopathy, vestibular disorders, surgery or fracture to the back, neck or shoulder in the last 3 years or shoulder luxation in the last year. Finally, strenuous exercise more than three times per week during the previous 6 months also led to exclusion.[19]

Patient and public involvement
As this study was based on secondary data from a RCT, patients were not directly involved in the design and completion of the study.

Data collection and variables
Before the start of the intervention (baseline), participants filled in a self-administered questionnaire, containing instruments and questions to measure the potential prognostic factors. Outcome information was collected by follow-up questionnaires at 3, 9 and 15 months from baseline. Participants who did not provide information on the outcome at any of the three follow-ups were excluded, resulting in a study population of 89 participants (figure 1).

Outcome
The Patient Global Impression of Change Scale (PGICS) provided information for the outcome 'global perceived change of general health' by comparing general health at follow-up with general health at baseline. PGICS is a 7-point ordinal Likert scale (very much improved, much improved, minimally improved, not changed, minimally worse, much worse and very much worse).[22] The scale categories were dichotomised into 'improved' (very much improved, much improved) and 'not improved' (all the other categories). According to the Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials (IMMPACT) recommendations for clinical important outcomes in chronic pain, the 7-point scale is recommended when assessing general health. The categories 'very much improved' and 'much improved' reflects what patients consider to be a clinical important improvement in general health.[23] The Neck Disability Index (NDI) provided information for the second outcome.[24] The NDI has 10 items (pain intensity, personal care, lifting, reading, headaches, concentration, work, driving, sleeping and recreational activities) with six possible
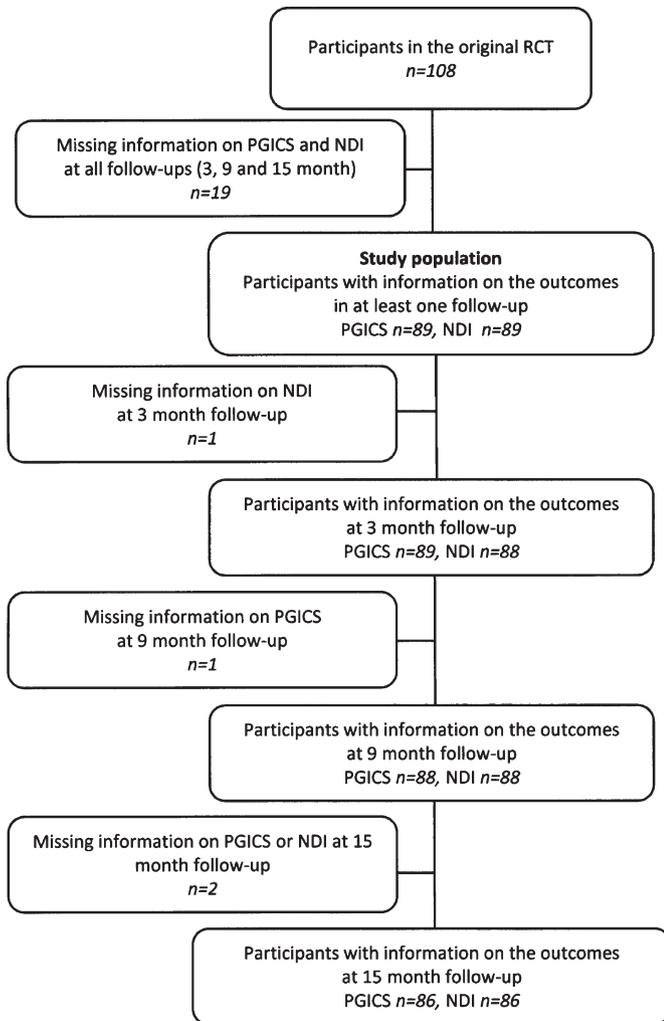
Figure 1 Included participants and progress of participants along the follow-up period concerning the outcome measures. NDI, The Neck Disability Index; PGICS, The Patient Global Impression of Change Scale; RCT, randomised controlled trial.

answers in each item, scored 0 (no limitation) to 5 (major limitations). Total score of the NDI ranged from 0 to 50. The NDI has high reliability, strong internal consistency and strong validity, when compared with other instruments for evaluating patients with neck pain.[24] In the present study, we used the normalised NDI (NDI%, 0–100). The minimal important change (MIC) for the NDI% was set equal to 6.3.[25] Using MIC as a cut-off, a dichotomised NDI outcome was constructed for each follow-up, where a clinical important improvement was defined as a NDI% decrease of more than 6.3 between the baseline and the follow-up, and 'no improvement' as an increase, or a decrease less than or equal to 6.3 NDI%. Participants with a NDI% of 0 at follow-up were also defined as having a clinical important improvement regardless of the baseline score.

Potential prognostic factors

The selection of potential prognostic factors was based on systematic reviews and prospective studies on the prognosis of non-specific neck pain, clinical considerations and availability in the data (table 1). Eight potential prognostic factors had support from literature; age, neck disability, average pain intensity, depression, fear of movement, catastrophising, social support and leisure physical activity.[5 8 26 27] Chronic widespread pain, recovery expectation, physical work load and dual-working were considered potential prognostic factors based on clinical considerations.

Average pain intensity was not limited to neck pain, but all participants reported neck pain as the 'most painful area', as this was a criterion for inclusion in the RCT.[19] Dual-working was a combination of three items: (1) Working/not working, (2) What does your household look like? (living alone, living alone with children, living with another adult, living with another adult and children) and (3) Who is primarily performing the housework in your family? (myself, someone else and equally shared). The factor dual-working was then dichotomised according to the answers into; yes (working, living with children and/or another adult and performing the housework) and no (any other combination of the three items). Dual-working has, to the best of our knowledge, not been evaluated in previous studies, but in our opinion, it is possibly associated with the prognosis of neck pain in women.

Statistical methods

Descriptive statistics are presented as means and SD or median and IQR when appropriate. The level of the CIs was set at 95% and that of the tests at 5%. All p values were two-sided. Stata V.14.2 was used in the analyses.

Random-effects logistic regression models were used to estimate the probability of a clinically important improvement of the outcomes from baseline to 3 and 15 months follow-up.[17] Data from all three follow-ups, 3, 9 and 15 months, were used in the analyses. Random-effects logistic regression can model longitudinal data efficiently, while taking into account the potential dependence of the repeated measures taken on each participant.[17] The regression models were estimated for each of the two outcomes, PGICS and NDI, separately. Both outcomes showed strong associations between the follow-ups. The linearity of the relationship between the logit of the probability of the outcome variables and numeric independent variables (potential prognostic factors) was verified by means of restricted cubic splines.[17 28] None of the numeric independent variables showed evidence against linearity.

Random-effects logistic regressions, including data from all three follow-ups, were used for all potential prognostic factors one at a time to estimate the time-specific ORs with 95% CI for each of the two outcome measures, separately. If the interaction term (factor-by-time) was not statistically significant, the analyses were repeated without the interaction term. This resulted in a single OR for the factor from 3 to 15 months. When the interaction was significant, the OR at 3 and 15 months were reported

Table 1 Potential prognostic factors and corresponding bibliographical references to definition and psychometric properties of the factors

Potential prognostic factor	Measurement	Categorisation in the analyses
Age	Age at baseline.	Continuous.
Neck disability[24]	Normalised Neck Disability Index (NDI%).	Continuous, 1–100 0 = 'no disability' and 100 = 'complete disability'
Average pain intensity[40]	Average pain intensity during the previous week measured on an 11-point Numeric Rating Scale (NRS).	Continuous, 0–10 0 = 'no pain' and 10 = 'pain as bad as it could be'
Depression[41]	Self-rating version of the Montgomery Åsberg Depression Rating Scale (MADRS-S).	Continuous, 0–54 0 = 'no depressive symptoms' and 54 = 'severe depression'
Fear of movement[42 43]	The Tampa scale of kinesiophobia.	Continuous, 17–68 A sum score of 17–68, where higher score indicate higher fear of movement.
Catastrophising[44]	The six-item catastrophising subscale from the Coping Strategies Questionnaire (CSQ).	Continuous, 0–36 Each item of the subscale had a score of 0–6, where a high score indicated a high degree of catastrophising.
Social support[45]	Item numbers 5 and 14 of the Swedish Multidimensional Pain Inventory (MPI-S).	Continuous, 1–12 Each item with a score of 0–6, where 0 = 'no social support' and 12 = 'high social support'.
Recovery expectation[31]	The answer to the question; 'Do you think that any intervention or exercise will lead to recovery?'	Continuous, 1–5 5-point ordinal Likert scale from 1 = 'No, definitely not' to 5 = 'Yes, completely recovered'.
Chronic widespread pain[46]	Derived from a pain drawing.	Yes; pain in all four quadrants of the body together with spinal pain.
Leisure physical activity[47]	'How much have you, in general, moved or exerted yourself physically during leisure time in the past year'.	▶ No/low intensity; very little/an occasional walk or similar/everyday physical activity like gardening, cleaning/low intensity physical activity eg, walking or similar at least once a week. ▶ Moderate/high intensity; more strenuous physical activity, eg, jogging, swimming, etc, at least once a week/regular high intensity physical activity, eg, running, ball sports, etc.
Physical workload[48]	'How physically strenuous has your work or daily activity been the past 12 months?'	▶ Mostly sedentary. ▶ Low/heavy work load; low physical work load but mobile/quite physically strenuous work/physically strenuous work.
Dual-working	A combination of three items: 1. Working/not working. 2. What does your household look like? 3. Who is primarily responsible for and performing the housework in your family?	▶ Yes; (1) working, (2) living together with children and/or another adult person and (3) having the responsibility for the housework. ▶ No; any other combination of the three items (see the Methods section).

separately. Potential prognostic factors with a p value ≤0.2 for the estimated OR were considered candidate factors for the predictive models.[28] Based on an a priori decision of clinical relevance, statistical effect measure modifications were tested between pain intensity and depression or recovery expectations. If the effect measure modification was statistically significant, this was included in the subsequent analyses.[28 29]

Developing the models

A sequential backward manual selection procedure with the random-effects logistic regressions was used to build the predictive models.[28] All candidate factors were included in the initial model. The factor with the highest p value (Wald-test) were excluded one by one until all factors in the model had a p value ≤0.1.[28] If a candidate factor showed effect-modification with time, then the factor was included together with the effect measure modification in the model. During the development process, all the analyses were adjusted for the assigned RCT interventions. 'Intervention' was then removed from the final predictive models before evaluating the models and presenting the results. The association between

the prognostic factor and the outcome was reported as beta-coefficient (β) with SE, OR with 95% CI and associated p value.

Evaluation of the models

The bias corrected area under the receiver operating characteristic curve (AUC) with 95% CI was obtained by cross-validation based on 100 design-matrix bootstrap replicates and used to determine the internally validated predictive ability of the models.[28 29] The AUC represents a summary of the sensitivity and specificity of the model in distinguishing participants who improve during the follow-up period from those who do not. The AUC ranges from 0.5 (no predictive ability) to 1.0 (perfect predictive ability). Overfitting was assessed by calculating the heuristic shrinkage factor.[30] A shrinkage factor of 1.0 indicates no overfitting of the model. The shrinkage corrected beta-coefficients (Sβ) were calculated by the formula; shrinkage factor x β.

Sensitivity analyses

To compare the study population with that of the non-responders, we used the χ^2 test for categorical variables, and the t-test or the Wilcoxon's rank-sum test for continuous variables.

RESULTS

Study population

Baseline characteristics of the study population are listed in table 2. The women's age ranged between 31 and 65 years, and all women reported neck pain duration of more than 8 months, with a median of 120 months (IQR 60–216). Eighty women (91%) were working and 66 (74%) reported no sick leave because of neck pain during the previous 6 months. Using PGICS as outcome measure, 47 (53%) women were categorised as improved at 3-month follow-up, 31 (35%) at 9-month follow-up and 26 (30%) at 15-month follow-up, while using NDI as outcome measure, 39 (44%) women had improved at 3-month follow-up, 37 (42%) at 9-month follow-up and 31 (36%) at 15-month follow-up. Only one participant reported 'no neck disability' (NDI%=0) at the 3-month and 15-month follow-ups, respectively, while none reported 'no neck disability' at the 9-month follow-up.

Development and evaluation of the predictive model

The analyses on each potential prognostic factor are presented in table 3. The follow-up data observations used for the analyses differed slightly between the two outcome measures because of missing responses (figure 1). No effect measure modification was observed between pain intensity and depression or recovery expectations. In the analyses with PGICS as outcome measure age, neck disability and average pain intensity met the criteria for inclusion into the predictive model analysis. No potential prognostic factor appeared to be modified by time. In the corresponding analyses with NDI as outcome measure, five potential prognostic factors met the criteria for

Table 2 Baseline characteristics of the study population, n=89

Characteristics		Missing (n)*
Age, median (IQR)	52 (47–59)	0
Neck disability, median (IQR)	28 (20–33)	0
Average pain intensity, mean (SD)	5.2 (1.6)	0
Depression, median (IQR)	8 (6–12)	1
Fear of movement, median (IQR)	29 (26–33)	0
Catastrophising, median (IQR)	6 (3–11)	0
Social support, mean (SD)	6.2 (3.4)	1
Recovery expectation, median (IQR)	4 (3–4)	4
Chronic widespread pain, freq (%)		0
Yes	40 (45)	
Leisure physical activity, freq (%)		0
No/low intensity	60 (67)	
Moderate/high intensity	29 (33)	
Physical workload, freq (%)		1
Mostly sedentary	36 (41)	
Low/heavy work load	52 (59)	
Dual-working, freq (%)		4
Yes	38 (45)	

*Missing answer.
freq, frequency.

inclusion: age, neck disability, depression, catastrophising and leisure physical activity. The effect of catastrophising changed over time, therefore an effect measure modification term (catastrophising x time) was added to the predictive model analysis.

The resulting predictive models from the sequential selection procedure are presented in table 4. Internal validation showed acceptable predictive ability for the models of 0.64 and 0.67, respectively (table 4). The calculated shrinkage factor was 0.73 for the PGICS model and 0.65 for the NDI model.

Sensitivity analyses

Non-responders (n=19) showed no statistically significant differences in the characteristics from the study population except for average pain intensity, where non-responders had a higher median of 6.5 compared with a median of 5 (p=0.03) in the study population.

DISCUSSION

We developed and internally validated predictive models for clinically important short-term and long-term improvement from chronic disabling neck pain among women in the clinical course of physiotherapy. The models were developed for the 'global perceived change

Table 3 Random-effects regression analyses* for potential short-term and long-term prognostic factors with PGICS and NDI as outcome, n=89

Potential prognostic factor	n/obs†	OR 3–15 m‡	95% CI	P value	P-eff.m.§
PGICS					
Age	*89/263*	0.92	0.84 to 1.01	0.08	0.75
Neck disability	*89/263*	1.08	1.01 to 1.16	0.03	0.26
Average pain intensity	*89/263*	1.60	1.00 to 2.55	0.05	1.00
Depression	*88/260*	1.04	0.92 to 1.19	0.50	0.07
Fear of movement	*89/263*	0.98	0.87 to 1.09	0.67	0.48
Catastrophising	*89/263*	1.06	0.94 to 1.20	0.34	0.20
Social support	*88/260*	1.08	0.88 to 1.33	0.47	0.57
Recovery expectations	*85/253*	1.74	0.71 to 4.25	0.22	0.47
Chronic widespread pain	*89/263*				0.45
No		1.0			
Yes		1.99	0.46 to 8.66	0.36	
Leisure physical activity	*89/263*				0.36
No/low intensity		1.0			
Moderate/high intensity		0.77	0.16 to 3.65	0.75	
Physical workload	*88/260*				0.64
Mostly sedentary		1.0			
Low to heavy workload		1.35	0.31 to 5.86	0.69	
Dual-working	*85/251*				0.41
No		1.0			
Yes		0.77	0.17 to 3.39	0.73	
NDI					
Age	89/262	0.96	0.90 to 1.02	0.20	0.68
Neck disability	89/262	1.05	1.01 to 1.11	0.02	0.31
Average pain intensity	89/262	1.17	0.84 to 1.63	0.35	0.67
Depression	88/259	0.94	0.86 to 1.03	0.17	0.75
Fear of movement	89/262	0.99	0.92 to 1.07	0.87	0.88
Catastrophising	89/262				0.03
3 m¶		1.10	0.98 to 1.24	0.12	
15 m¶		0.93	0.83 to 1.05	0.23	
Social support	88/259	1.08	0.93 to 1.26	0.32	0.72
Recovery expectations	85/253	0.73	0.39 to 1.37	0.33	0.37
Chronic widespread pain	89/262†				0.10
No		1.0			
Yes		0.72	0.26 to 1.99	0.53	
Leisure physical activity	89/262				0.36
No/low intensity		1.0			
Moderate/high intensity		0.48	0.16 to 1.43	0.19	
Physical workload	88/259				0.06
Mostly sedentary		1.0			
Low to heavy workload		1.16	0.41 to 3.25	0.78	
Dual-working	85/250				0.16
No		1.0			
Yes		0.55	0.19 to 1.58	0.27	

An OR>1 reflects a higher odds of clinical important improvement and an OR<1 reflects a lower odds of clinical important improvement.
Potential prognostic factors included in the development of the predictive models are presented in italics.
*Data from all three follow-ups (3, 9 and 15 months) were used in the analyses.
†Number of participants in the analyses/number of outcome observations in the analyses.
‡ OR for clinically important improvement between 3 and 15 months.
§P value for the effect modification over time between 3 and 15 months from baseline.
¶Results (OR, 95% CI and p values) for clinically important improvement at 3 and 15 months, respectively.
NDI, normalised Neck Disability Index; PGICS, Patient Global Impression of Change Scale.

Table 4 The final predictive models* of clinically important improvement in chronic disabling neck pain with PGICS and NDI as outcome

Prognostic factor	β	SE	OR	95% CI	P value	Sβ
PGICS (89/263)†						
Neck disability	0.08	0.04	1.08	1.01 to 1.16	0.03	0.06
Age	−0.08	0.05	0.92	0.84 to 1.01	0.08	−0.06
AUC (95% CI)‡	0.64 (0.55 to 0.73)					
NDI (88/259)†						
Neck disability	0.08	0.03	1.08	1.02 to 1.14	0.01	0.05
Depression	−0.11	0.05	0.89	0.81 to 0.98	0.02	−0.07
Catastrophising						
3 months	0.07	0.06	1.08	0.96 to 1.21	0.21	0.05
15 months	−0.09	0.06	0.92	0.82 to 1.03	0.14	−0.06
x time§	−0.01	0.01	0.99	0.97 to 1.00	0.03	−0.01
AUC (95% CI)‡	0.67 (0.59 to 0.75)					

An OR>1 reflects a higher odds of clinical important improvement and an OR<1 reflects a lower odds of clinical important improvement.

The intercept (β_0) for the PGICS model was 1.27, and for the NDI model −1.89.

*Data from all three follow-ups (3, 9 and 15 months) were used in the models.

†Number of participants in the analysis/number of outcome observations in the analysis.

‡Internally validated receiver operating characteristics curve (AUC) with 95% CI.

§Effect measure modification term in the model (catastrophising x time).

β, beta-coefficient; NDI, normalised Neck Disability Index; PGICS, Patient Global Impression of Change Scale; Sβ, shrinkage corrected beta-coefficient (shrinkage factor x β).

of general health' (PGICS) and the NDI, separately. Both models had acceptable predictive ability.[29] The results showed that, when assessing potential biopsychosocial prognostic factors, perceived disabilities related to neck pain, depression, catastrophising and age are of importance for predicting short-term and long-term improvements in these women. Further, we found that prognostic factors could predict different results depending on the follow-up time, as for catastrophising in the NDI model. This is important in the management of patients in clinical settings. Similarly to previous studies, we found that the prognostic factors in the models could differ depending on the outcome.[16 31]

Our results indicate that the odds for clinically important improvement, with PGICS as outcome measure, decrease with age and increase with higher baseline levels of disability related to neck pain. With NDI as an outcome measure, the odds for improvement increased with higher baseline levels of neck pain-related disability and decrease with higher levels of depression. These results were valid over the total follow-up time of 15 months. However, the NDI model indicated increased odds for short-term improvement, but decreased odds for long-term improvement with higher levels of catastrophising at baseline. This interesting finding was made possible by the longitudinal nature of our study design. To the best of our knowledge, this is the first study using longitudinal analyses in predictive models for neck pain. Interestingly, depression was found to be a predictor even though women with a psychiatric diagnosis affecting their everyday life were excluded from the study.

The relative strength of the incorporated factors in the model should be interpreted with great caution as their independent effect on the outcome was not thoroughly examined.

Our findings are somewhat different from those of other similar studies, because of the different study populations, potential prognostic factors examined, outcomes and follow-up times. We found only one prediction study on patients with chronic non-specific neck pain in physiotherapy, but with a study population also including males.[15] Poor outcome was predicted by pain medication at discharge and at 1 year follow-up, and by catastrophising at 1 year follow-up, the latter in line with our results. Similar to our study, the outcome was assessed as a minimal clinical important difference, but was based on the Northwick neck pain questionnaire.

We found four reports on non-specific neck pain predictive models with reasonably similar outcomes, methods and follow-up time as in the present study.[31–34] However, all four reports included both male and female patients with different durations of pain (acute, subacute and chronic), and were not performed in the course of physiotherapy only. Similar to us, Hill *et al* and Verhagen *et al* found the psychological factors catastrophising and depression/distress to predict their outcomes.[31 34] Also, baseline neck disability was included in the models by Hill *et al* and by Kjellman *et al*.[31 33] Other factors in their models were pain intensity, treatment expectations, concomitant back pain, manual social class, well-being and somatisation, while we found that treatment expectations and pain intensity were not important predictors in our study.[31 33 34]

Shellingerhout *et al* presented a predictive model with a set of nine predictors for global perceived recovery in an adult primary care population.[32] With the exception of age, none of the predictors found were similar to our predictors, a discrepancy that may be explained by different study settings, population and outcome compared with our study. In their population, only 34% reported chronic neck pain and 60% were women. Further, only 28% of the patients were referred to physiotherapy while the others were referred to 'usual care', spinal manipulation or a behaviour graded activity programme. Even though Shellingerhout *et al* reported a model including nine predictors, their predictive ability (AUC=0.66) was similar to that of our models (AUC of 0.64 and 0.67). One could speculate that as prognostic factors in research related to neck pain most often are weak, reaching higher levels of predictive ability may be hard even with large samples and well conducted analyses.[5 8]

Our study meets most of the important criteria for deriving predictive models in relation to musculoskeletal disorders and physiotherapy suggested by Beneciuk *et al*: a well-defined study population, longer follow-up times than 6 months and psychological and psychosocial assessments incorporated in the model development.[35] Other strengths of our study are a follow-up rate of 82% and a small number of missing data. We used longitudinal analyses, an advantage when assessing individual change over time.[18] The outcome, clinical important improvement, was based on the IMMPACT recommendations (PGICS), and on responsiveness analyses in which the same sample as in the present study was included (NDI).[23 25] The use of conservative 'cut-points' for the Wald test in the selection procedure in order to decrease the risk of type II errors is also supported in the literature.[28 29]

Our study also has limitations. The sample size of 89 participants could be considered small when conducting a prediction study. A small sample size increases the risk of overdispersion during analyses. However, the sample size was large enough to ensure the recommendations of 10–15 participants for each parameter estimate (coefficient) included in the analyses.[28 35] Furthermore, our results are based on secondary analyses of data, possibly limiting information on additional prognostic factors which could have influenced the derivation of the models. Of the 108 women included at baseline, 19 did not respond to any of the three follow-ups and were excluded from the analyses. They showed similar characteristics as the included women except for an average pain intensity with a median of 6.5 compared with a median of 5 in the study population. Therefore, attrition might have influenced the results. Some of the participants were recruited by advertisement and could have different characteristics than ordinary healthcare patients.[36] However, the baseline levels of average pain intensity, neck pain-related disability and fear of movement in the study were similar to those in the study by Shellingerhout *et al*, who investigated prognostic factor for neck pain in a sample of patients referred to primary care.[32]

The clinical implication of our study findings is that baseline neck disability seems to be an important factor to consider, as a higher baseline disability was associated with higher odds for clinical important improvement in both our models. This is in line with the findings of Bot *et al* in exploring predictors for patients with neck or shoulder pain in general practice.[37] They found that being more disabled at baseline predicted a larger reduction in disability at 3 and 12 months follow-up. Furthermore, psychological factors are important to consider in the clinical course of physiotherapy concerning women with chronic neck pain. As psychological factors are modifiable, they may be targeted in order to increase the possibility for short-term and long-term improvement in the management of these women. The information about the prognostic factors and the outcomes included in our models is easily collected at the first consultation by a physiotherapist, and the models could be a valuable tool to help physiotherapists manage these patients.

While the predictive ability of our models was acceptable, it also indicated that there may still be other factors that could help predict the outcome more precisely. It would therefore be valuable to include variables in future investigations that were not included in the present study, like for example lifestyle factors such as tobacco and alcohol use, and psychological factors related to work.[38 39] Our models will make clinicians aware of what factors are important to consider when predicting which patients will have the best chance of a clinical important improvement. Also notable to clinicians is that factors can vary depending on the outcome measured, and that the chance of improvement can change over time for some factors, as for catastrophising in the current study. Furthermore, our results facilitate future prognostic research related to chronicity in neck pain. A predictive model development, including internal validation, is only the first step in the process of deriving a model that could be implemented in clinic. A second step should be external validation, for example, by using data from RCTs including other types of interventions targeting chronic neck pain. Finally, further external validation should be done by investigating the impact of the model in clinical practice before implementation.[11] Because all women in our sample underwent physiotherapy, our predictive models should be restricted to this population and setting. The validity in other populations and other healthcare providers will remain unknown until external validations will be available.

CONCLUSION

The developed predictive models evaluating clinical important improvement from chronic non-specific neck pain among women in the clinical course of physiotherapy showed acceptable predictive ability. Age, neck pain-related disability, depression and catastrophising seem to be factors that can be of guidance for physiotherapists trying to predict short-term and long-term

clinical important improvement in these patients. With the exception of neck pain-related disability, the different outcome measures had different sets of prognostic factors. The effect of some factors may be modified by time. The present study is the first step towards developing predictive models for clinical practice. The next steps will include external validations in different neck pain populations and clinical settings.

REFERENCES

1. Hogg-Johnson S, van der Velde G, Carroll LJ, et al. The burden and determinants of neck pain in the general population: results of the Bone and Joint Decade 2000-2010 Task Force on Neck Pain and Its Associated Disorders. *Spine* 2008;33(4 Suppl):S39–51.
2. Vos T, Allen C, Arora M, et al. Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015. *The Lancet* 2016;388:1545–602.
3. Hoy D, March L, Woolf A, et al. The global burden of neck pain: estimates from the global burden of disease 2010 study. *Ann Rheum Dis* 2014;73:1309–15.
4. WHO Scientific Group on the Burden of Musculoskeletal Conditions at the Start of the New Millennium. The burden of musculoskeletal conditions at the start of the new millennium. *World Health Organ Tech Rep Ser* 2003;919:1–218. back cover.
5. Carroll LJ, Hogg-Johnson S, van der Velde G, et al. Course and prognostic factors for neck pain in the general population: results of the Bone and Joint Decade 2000-2010 Task Force on Neck Pain and Its Associated Disorders. *Spine* 2008;33(4 Suppl):S75–82.
6. Hemingway H, Croft P, Perel P, et al. Prognosis research strategy (PROGRESS) 1: a framework for researching clinical outcomes. *BMJ* 2013;346:e5595.
7. Guzman J, Haldeman S, Carroll LJ, et al. Clinical practice implications of the Bone and Joint Decade 2000-2010 Task Force on Neck Pain and Its Associated Disorders: from concepts and findings to recommendations. *Spine* 2008;33(4 Suppl):S199–213.
8. Walton DM, Carroll LJ, Kasch H, et al. An Overview of Systematic Reviews on Prognostic Factors in Neck Pain: Results from the International Collaboration on Neck Pain (ICON) Project. *Open Orthop J* 2013;7:494–505.
9. Hurwitz EL, Carragee EJ, van der Velde G, et al. Treatment of neck pain: noninvasive interventions: results of the Bone and Joint Decade 2000-2010 Task Force on Neck Pain and Its Associated Disorders. *J Manipulative Physiol Ther* 2009;32(2 Suppl):S141–75.
10. Beattie P, Nelson R. Clinical prediction rules: what are they and what do they tell us? *Aust J Physiother* 2006;52:157–63.
11. Steyerberg EW, Moons KG, van der Windt DA, et al. Prognosis Research Strategy (PROGRESS) 3: prognostic model research. *PLoS Med* 2013;10:e1001381.
12. Bohman T, Alfredsson L, Hallqvist J, et al. The influence of self-reported leisure time physical activity and the body mass index on recovery from persistent back pain among men and women: a population-based cohort study. *BMC Public Health* 2013;13:385.
13. Messing K, Stock SR, Tissot F. Should studies of risk factors for musculoskeletal disorders be stratified by gender? Lessons from the 1998 Québec Health and Social Survey. *Scand J Work Environ Health* 2009;35:400–00.
14. van Oort L, van den Berg T, Koes BW, et al. Preliminary state of development of prediction models for primary care physical therapy: a systematic review. *J Clin Epidemiol* 2012;65:1257–66.
15. Cecchi F, Molino-Lova R, Paperini A, et al. Predictors of short- and long-term outcome in patients with chronic non-specific neck pain undergoing an exercise-based rehabilitation program: a prospective cohort study with 1-year follow-up. *Intern Emerg Med* 2011;6:413–21.
16. Pool JJ, Ostelo RW, Knol D, et al. Are psychological factors prognostic indicators of outcome in patients with sub-acute neck pain? *Man Ther* 2010;15:111–6.
17. Fitzmaurice GM. *Applied longitudinal analysis*. Hoboken, N.J: Wiley-Interscience, 2004:325–71.
18. Hoogendoorn WE, Bongers PM, de Vet HC, et al. Comparison of two different approaches for the analysis of data from a prospective cohort study: an application to work related risk factors for low back pain. *Occup Environ Med* 2002;59:459–65.
19. Rudolfsson T, Djupsjöbacka M, Häger C, et al. Effects of neck coordination exercise on sensorimotor function in chronic neck pain: a randomized controlled trial. *J Rehabil Med* 2014;46:908–14.
20. Margolis RB, Tait RC, Krause SJ. A rating system for use with patient pain drawings. *Pain* 1986;24:57–65.
21. Atroshi I, Gummesson C, Andersson B, et al. The disabilities of the arm, shoulder and hand (DASH) outcome questionnaire: reliability and validity of the Swedish version evaluated in 176 patients. *Acta Orthop Scand* 2000;71:613–8.
22. Dworkin RH, Turk DC, Farrar JT, et al. Core outcome measures for chronic pain clinical trials: IMMPACT recommendations. *Pain* 2005;113:9–19.
23. Dworkin RH, Turk DC, Wyrwich KW, et al. Interpreting the clinical importance of treatment outcomes in chronic pain clinical trials: IMMPACT recommendations. *J Pain* 2008;9:105–21.
24. Vernon H. The Neck Disability Index: state-of-the-art, 1991-2008. *J Manipulative Physiol Ther* 2008;31:491–502.
25. Björklund M, Wiitavaara B, Heiden M. Responsiveness and minimal important change for the ProFitMap-neck questionnaire and the Neck Disability Index in women with neck-shoulder pain. *Qual Life Res* 2017;26:161–70.
26. Carroll LJ, Hogg-Johnson S, Côté P, et al. Course and prognostic factors for neck pain in workers: results of the Bone and Joint Decade 2000-2010 Task Force on Neck Pain and Its Associated Disorders. *Spine* 2008;33(4 Suppl):S93–100.
27. Bruls VE, Bastiaenen CH, de Bie RA. Prognostic factors of complaints of arm, neck, and/or shoulder: a systematic review of prospective cohort studies. *Pain* 2015;156:765–88.
28. Vittinghoff E, Glidden DV, Shiboski SC, et al. *Regression methods in biostatistics, linear, logistic, survival, and repeated measures models*. 2nd ed. New York, USA: Springer New York, 2012:261–308. 395-429.
29. Hosmer DW, Lemeshow S. *Applied logistic regression*. 2nd ed. New York; Chichester: Wiley, 2000:91–142.
30. Steyerberg EW, Borsboom GJ, van Houwelingen HC, et al. Validation and updating of predictive logistic regression models: a study on sample size and shrinkage. *Stat Med* 2004;23:2567–86.
31. Hill JC, Lewis M, Sim J, et al. Predictors of poor outcome in patients with neck pain treated by physical therapy. *Clin J Pain* 2007;23:683–90.
32. Schellingerhout JM, Heymans MW, Verhagen AP, et al. Prognosis of patients with nonspecific neck pain: development and external validation of a prediction rule for persistence of complaints. *Spine* 2010;35:E827–35.
33. Kjellman G, Skargren E, Oberg B. Prognostic factors for perceived pain and function at one-year follow-up in primary care patients with neck pain. *Disabil Rehabil* 2002;24:364–70.
34. Verhagen AP, Karels CH, Schellingerhout JM, et al. Pain severity and catastrophising modify treatment success in neck pain patients in primary care. *Man Ther* 2010;15:267–72.

35. Beneciuk JM, Bishop MD, George SZ. Clinical prediction rules for physical therapy interventions: a systematic review. *Phys Ther* 2009;89:114–24.

36. Woodhouse A, Pape K, Romundstad PR, *et al*. Health care contact following a new incident neck or low back pain episode in the general population; the HUNT study. *BMC Health Serv Res* 2016;16:81.

37. Bot SD, van der Waal JM, Terwee CB, *et al*. Predictors of outcome in neck and shoulder symptoms: a cohort study in general practice. *Spine* 2005;30:E459–70.

38. Bohman T, Alfredsson L, Jensen I, *et al*. Does a healthy lifestyle behaviour influence the prognosis of low back pain among men and women in a general population? A population-based cohort study. *BMJ Open* 2014;4:e005713.

39. Svedmark Å, Björklund M, Häger CK, *et al*. Impact of Workplace Exposure and Stress on Neck Pain and Disabilities in Women-A Longitudinal Follow-up After a Rehabilitation Intervention. *Ann Work Expo Health* 2018;62:591–603.

40. Turk DC, Melzack R. *Handbook of pain assessment*. 3rd ed. New York; London: ednGuilford, 2011.

41. Svanborg P, Asberg M. A new self-rating scale for depression and anxiety states based on the Comprehensive Psychopathological Rating Scale. *Acta Psychiatr Scand* 1994;89:21–8.

42. Cleland JA, Fritz JM, Childs JD. Psychometric properties of the Fear-Avoidance Beliefs Questionnaire and Tampa Scale of Kinesiophobia in patients with neck pain. *Am J Phys Med Rehabil* 2008;87:109–17.

43. Vlaeyen JW, Kole-Snijders AM, Rotteveel AM, *et al*. The role of fear of movement/(re)injury in pain disability. *J Occup Rehabil* 1995;5:235–52.

44. Jensen IB, Linton SJ. Coping strategies questionnaire (CSQ): Reliability of the swedish version of the CSQ. *Scandinavian Journal of Behaviour Therapy* 1993;22:139–45.

45. Bergström G, Jensen IB, Bodin L, *et al*. Reliability and factor structure of the Multidimensional Pain Inventory--Swedish Language Version (MPI-S). *Pain* 1998;75:101–10.

46. Wolfe F, Smythe HA, Yunus MB, *et al*. The American College of Rheumatology 1990 criteria for the classification of fibromyalgia. *Arthritis & Rheumatism* 1990;33:160–72.

47. Ekblom-Bak E, Hellénius ML, Ekblom O, *et al*. Independent associations of physical activity and cardiovascular fitness with cardiovascular risk in adults. *Eur J Cardiovasc Prev Rehabil* 2010;17:175–80.

48. Leijon O, Wiktorin C, Härenstam A, *et al*. Validity of a self-administered questionnaire for assessing physical work loads in a general population. *J Occup Environ Med* 2002;44:724–35.