BMJ Open

# Prognostic factors for disability and sick leave in patients with subacute non-malignant pain: a systematic review of cohort studies

Gitte H Valentin,[1] Marc S Pilegaard,[2] Henrik B Vaegter,[3] Marianne Rosendal,[4] Lisbeth Ørtenblad,[1] Ulla Væggemose,[1] Robin Christensen[5]

**To cite:** Valentin GH, Pilegaard MS, Vaegter HB, *et al*. Prognostic factors for disability and sick leave in patients with subacute non-malignant pain: a systematic review of cohort studies. *BMJ Open* 2016;**6**: e007616. doi:10.1136/bmjopen-2015-007616

▸ Prepublication history and additional material is available. To view please visit the journal (http://dx.doi.org/10.1136/bmjopen-2015-007616).

Received 8 January 2015
Revised 20 November 2015
Accepted 26 November 2015

For numbered affiliations see end of article.

**Correspondence to**
Dr Robin Christensen; Robin.Christensen@regionh.dk

## ABSTRACT

**Objective:** This systematic review aims to identify generic prognostic factors for disability and sick leave in subacute pain patients.

**Setting:** General practice and other primary care facilities.

**Participants:** Adults (>18 years) with a subacute (≤3-month) non-malignant pain condition. Eligibility criteria were cohort studies investigating the prediction of disability or long-term sick leave in adults with a subacute pain condition in a primary care setting. 19 studies were included, referring to a total of 6266 patients suffering from pain in the head, neck, back and shoulders.

**Primary and secondary outcome measures:** The primary outcome was long-term disability (>3 months) due to a pain condition. The secondary outcome was sick leave, defined as 'absence from work' or 'return-to-work'.

**Results:** PubMed, EMBASE, CINAHL and PEDro databases were searched from 16 January 2003 to 16 January 2014. The quality of evidence was presented according to the GRADE WG recommendations. Several factors were found to be associated with disability at follow-up for at least two different pain symptoms. However, owing to insufficient studies, no generic risk factors for sick leave were identified.

**Conclusions:** Multiple site pain, high pain severity, older age, baseline disability and longer pain duration were identified as potential prognostic factors for disability across pain sites. There was limited evidence that anxiety and depression were associated with disability in patients with subacute pain, indicating that these factors may not play as large a role as expected in developing disability due to a pain condition. Quality of evidence was moderate, low or very low, implying that confidence in the results is limited. Large prospective prognostic factor studies are needed with sufficient study populations and transparent reporting of all factors examined.

**Trial registration number:** CRD42014008914.

## Strengths and limitations of this study

- This systematic review provides new knowledge on risk factors across pain sites, which may help physicians and researchers when initial referral decisions are made.
- The review also provides a solid foundation for planning future high-quality studies on risk factors for poor outcomes in pain patients.
- The protocol for the systematic review was registered beforehand in PROSPERO and reported according to the PRISMA statement, with the quality of the evidence judged as recommended by the GRADE Working Group.
- Quality of evidence was moderate, low or very low, implying that confidence in the results is limited.

## INTRODUCTION

Pain is the most common reason patients consult general practice,[1] and long-term disability and sick leave due to a pain condition are associated with huge negative consequences for the individual and for society.[2] It would be both costly and unnecessary, however, to offer specialised treatment to all patients presenting in primary care with a pain condition; despite its frequency, pain is in most cases a temporary phenomenon.[3] Still, a small group of patients will develop chronic or recurrent pain causing long-term disability and sick leave. It is estimated that approximately 3–10% of patients with acute pain develop a chronic pain condition.[3 4] Chronic pain conditions are associated with social and family problems, loss of work, and loss of self-esteem and integrity.[5–7] Moreover, chronic pain is often associated with other symptoms or comorbidities such as fatigue, concentration and memory problems, sleep disorders, depression and anxiety.[5] Once pain has become chronic, treatment is complex and difficult.[2] Thus, early identification of pain patients at high risk of developing long-term problems would offer a great opportunity for reducing cost and suffering associated with long-term disability and sick

leave because optimal care could be initiated at an early stage.

Most pain research focuses on one specific pain site (eg, low back or shoulder pain[8 9]). As a result, prognostic factor research is normally conducted on each site separately.[8] For example, substantial prognostic factor studies on back pain have been carried out, with several systematic reviews reporting prognostic factors for back pain.[10–14] However, this single-site approach limits clinical applicability for the general practitioner (GP) because most pain patients have pain at more than one anatomical location.[9 15] Factors that have predictive value across different pain sites (ie, generic factors) may exist,[8] but few attempts have been made to explore prognostic factors across pain sites.[8 16]

This systematic review was conducted as part of a national Danish 'Health Technology Assessment' (HTA) aimed at identifying possibilities for early identification and timely treatment of pain patients across pain sites with relevance to a broad range of stakeholders in Denmark.[17] The specific aim of the evidence synthesis was to identify potential factors for the development of long-term disability or sick leave in patients with subacute, non-malignant pain in primary care.

## METHODS

The review was conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement[18] on the basis of a predefined protocol available from the International Prospective Register of Systematic Reviews (PROSPERO: CRD42014008914).

### Data sources and searches

Studies were identified via a systematic literature search in the following databases: PubMed, EMBASE, CINAHL and PEDro. Additional studies were identified through experts and through a review of the included studies' reference lists. The following search terms were used: 'Pain', 'Prognosis', 'Predictor', 'Prognostic factor', 'Primary Health Care', 'General Practice' and 'Family Practice'. The search string tailored for the PubMed database is presented below: Search(((('Pain'[Mesh]) OR 'Chronic Pain' [Mesh] OR 'persistent pain')) AND ((('Prognosis' [Mesh]) OR 'Outcome Assessment (Health Care)' [Mesh] OR predict* OR prognost* AND ((English[lang] OR Danish[lang] OR Norwegian[lang] OR Swedish[lang] AND adult[Mesh])))) AND ((('Primary Health Care' [Mesh]) OR 'General Practice' [Mesh]) OR 'Family Practice' [Mesh] OR GP OR 'primary care'). Filters: Published in the past 10 years; English; Danish; Norwegian; Swedish; Adult: 19 +years (full search is available on request). As part of the search and selection strategy, according to the HTA protocol, the major outcome was long-term disability (>3 months) due to a pain condition. A secondary outcome was sick leave, defined as 'absence from work' or 'return-to-work'. The search was restricted to identify studies published in English, Danish, Norwegian or Swedish between 16 January 2003 and 16 January 2014.

### Study selection

Studies were eligible for inclusion if they met the following criteria: prospective cohort study (including randomised controlled trials), with at least 3 months of follow-up investigating the prediction of long-term disability and/or sick leave in adults (>18 years) with a subacute (≤3-month) non-malignant pain condition, visiting GPs or other primary care facilities. 'Non-malignant pain condition' was defined as pain conditions of non-cancer origin. If two or more published studies originated from the same patient population, the study with the longest follow-up period was included. Two reviewers (GHV and MSP) independently assessed abstracts and full-text articles for eligibility, and disagreement was solved by a third reviewer (LØ).

### Data extraction and quality assessment

Two review authors (GHV and MSP) independently performed data extraction using a customised data extraction form. To summarise the evidence following the systematic review in the HTA, we applied the 'Grading of Recommendations Assessment, Development, and Evaluation' (GRADE) approach for rating quality of evidence (ie, our confidence in the estimates).[19] Because we anticipated that the evidence base would come from cohort studies, the GRADE approach for prognostic factor research[20] was applied. The risk of bias in the individual studies was assessed by two reviewers (GHV and MSP) using the Quality in Prognosis Studies tool (QUIPS).[21] The overall risk of bias for each of the studies was judged as: (1) low if there were a low risk of bias in all key domains, (2) unclear risk of bias if there were an unclear risk of bias for one or more key domains and (3) high risk of bias if there were a high risk of bias for one or more key domains.[22]

Disagreement was resolved by consensus. Publication bias was explored by funnel plots.

### Data synthesis and analysis

If a baseline factor was associated with outcomes at follow-up in one or more studies of different pain sites, it was considered a 'possible prognostic' factor and the results were presented as part of the evidence profile. When data were available in different formats, data from the 'fully adjusted' analyses were given preference and included in the analysis. For each outcome, we prepared an evidence profile using GRADEpro software.[23] According to the protocol, we also aimed to combine individual study results with a meta-analysis section. However, given the substantial clinical (as well as statistical) heterogeneity in the individual studies, we decided to downplay the importance of the results from the meta-analysis and focus on the narrative interpretation of the results. A description of the statistical methods

and the corresponding results are presented in online supplementary material S1. The narrative synthesis of the results is presented as proposed by Huguet *et al*.[20]

## RESULTS

### Results of the literature search

The search in the selected databases returned a total of 3533 references. A total of 32 references were identified through the additional search. After removing duplicates, 1841 references remained. The 1841 references were screened for eligibility, and 1641 records were excluded. The remaining 200 articles were read in full text; of these, 181 articles were excluded because they did not satisfy the inclusion criteria. A full list of excluded studies and the reason for exclusion are available from the authors on request. A total of 19 studies satisfied the inclusion criteria and were included in the systematic review. However, only 11 were eligible for inclusion in the evidence profile; the other 8 studies were excluded from the evidence profile due to: (1) inadequate statistical analyses,[3 24] (2) the factors studied were assessed in only one study[25–27] or (3) the factors studied were assessed for only one pain site (eg, only studies on back pain).[28–30] See figure 1 for a flow diagram of the included studies.

### Included studies

The 19 included studies consisted of 17 cohort studies and 2 randomised controlled trials. Fourteen of the studies referred to patients with back pain[4 24–26 28 29 31–38] and one referred to patients with pain in the neck or back.[3] Two studies referred to patients with neck pain,[27 39] one referred to patients with headache[40] and one referred to patients with shoulder or back pain.[30] From this last-mentioned study, only the cohort with back pain was included in the synthesis because the cohort with shoulder pain comprised both patients with subacute pain and patients with chronic pain.[30] Outcome measures were disability in 16 studies,[3 4 24 28–40] sick leave in 3 studies[3 27 37] and return-to-work in 2 studies.[25 26]

Characteristics of the 19 included studies are presented in table 1. The total number of patients included in the 19 studies was 6266. The median number of patients per study was 184 (range 56–2662). In 11 of the studies, more women than men participated. Age was reported in 18 studies, with a median average of 42 years (range 34–52 years). Pain duration at baseline was reported in eight studies with a median average of 12.6 days (range 1–27 days). The follow-up period ranged from 3 months to 22 years (with a median of 9 months). Most of the studies recruited patients from general practice.[3 4 24 26–28 30–33 35–38 40] The remaining studies recruited patients from physiotherapy or chiropractor clinics[29 34 39] and the workers compensation board.[25]

### Risk of bias within studies

Risk of bias in the included studies was assessed using QUIPS (figure 2). Overall, the agreement between the two assessors (GHV and MSP) on the different aspects of risk-of-bias assessment was 85.5% (weighted κ 0.49), which corresponds to a moderate degree of agreement. In all cases, any disagreement between the assessors was settled by consensus discussion. The domain 'Study Confounding' carried the highest risk of bias. In this domain, 3 studies were judged as having a high risk of bias, and 11 studies were assessed as having a moderate risk of bias. The high number of studies judged as having a high or moderate risk of bias in this domain was mainly due to the insufficient description of the factors that were included in the multivariable analysis. On the basis of the judgement of the 6 domains, 11 studies were judged to have a low risk of bias,[4 26 28–30 33 35–39] 3 studies had a moderate risk of bias[27 31 32] and 5 studies had a high risk of bias.[3 24 25 34 40]

Publication bias was assessed by funnel plots for all eight prognostic factors. No obvious asymmetry was found (see online supplementary material S2).

### Prognostic factors for disability

Prognostic factors for disability were assessed in 16 studies.[3 4 24 28–40] A total of 81 factors were assessed in the unadjusted analysis (see online supplementary material S3). Of the 81 factors assessed, 53 were included in the multivariable analysis of the primary studies. Of these factors, the following eight factors were assessed in two or more studies and for at least two different pain sites: multiple site pain, higher baseline pain severity, previous pain episodes, older age, longer baseline pain duration, baseline disability, anxiety and depression. A total of 11 studies were included in the evidence profile (table 2).[4 31–40] For these potential prognostic factors, the results are synthesised and summarised below.

#### Multiple pain sites

The association between multiple pain sites and disability at follow-up was assessed in three studies, including patients with headache,[40] low back pain[36] and neck pain.[39] Multiple pain sites were significantly associated with disability in all three studies when adjusted for: age,[36 40] sex,[36 40] baseline disability[36] and recruitment.[36] In the study by Leaver *et al*,[39] factors adjusted for were not described.

#### Higher pain severity

Six studies including patients with headache[40] and low back pain[31 32 34 36 37] investigated the association between higher pain severity at baseline and disability. Higher pain severity was consistently associated with disability at follow-up when adjusting for: age,[36 40] sex,[36 40] baseline disability[36] and recruitment[36] (in three studies, factors adjusted for were not specified[31 34 37]). However, a substantial heterogeneity between study results was detected. Consequently, the quality of evidence was downgraded due to inconsistency.

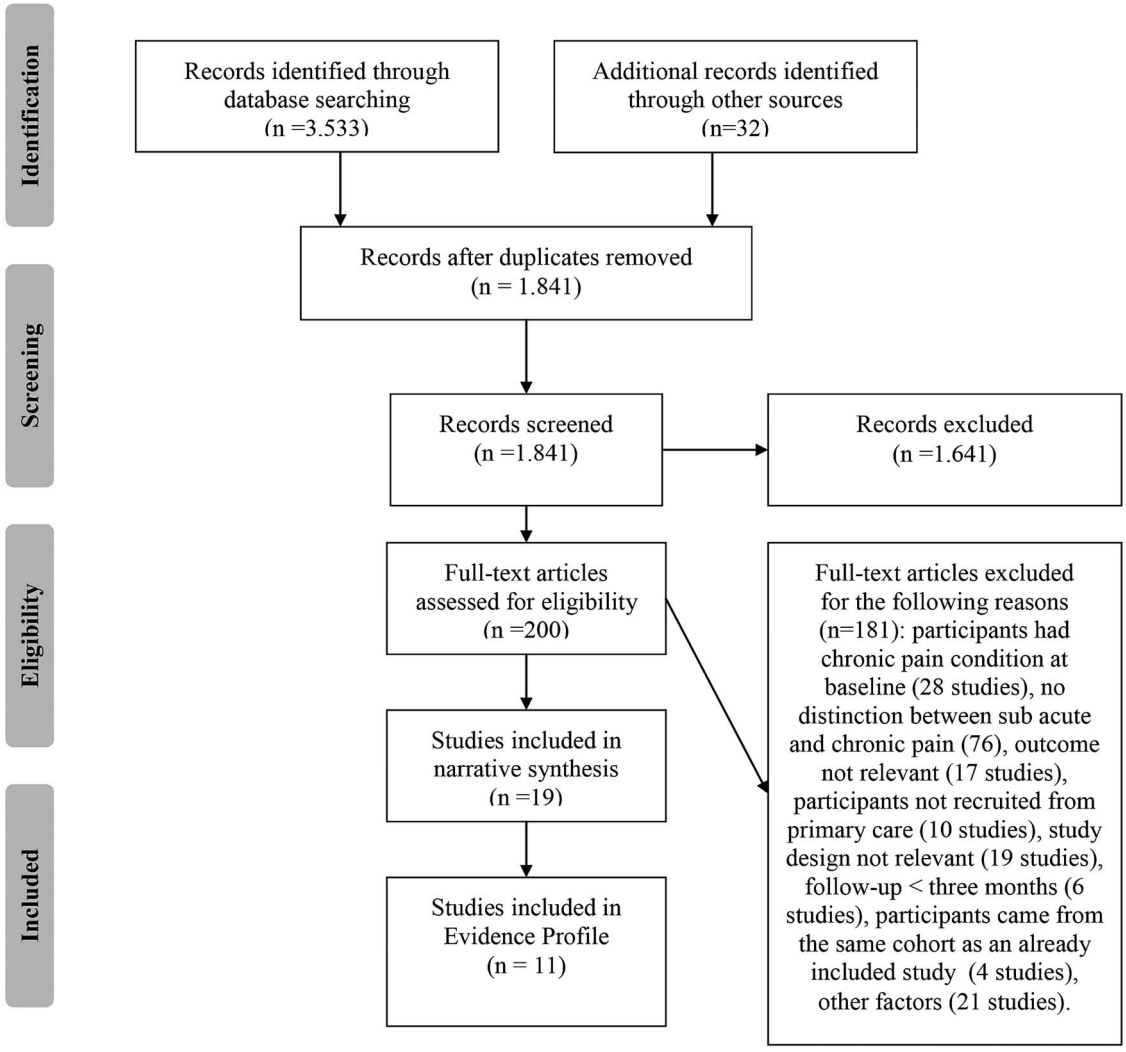



**Figure 1** Flow diagram of the literature search.

### Baseline disability

The association between baseline disability and disability at follow-up was examined in seven studies relating to patients with neck pain[39] and low back pain.[32 34–38] The following factors were adjusted for in the multivariable analysis of the primary studies: age,[36 38] sex,[36 38] body mass index (BMI),[38] duration (hours) between pain debut and inclusion,[35] job status,[35] previous spine surgery,[35] compensation status[35] and self-rated health status.[35] In four studies, the factors adjusted for were not specified.[32 34 37 39] Baseline disability was associated with a higher risk of disability at follow-up in five studies.[35–39] Two studies did not demonstrate a statistically significant association between baseline disability and disability at follow-up.[32 34] The lack of statistical significance in the study by Heneweer *et al* may reflect the lower power of the analysis (small study sample).

### Higher age

The association between higher age and the risk of disability was assessed in seven studies concerning patients with neck pain[39] and low back pain.[4 31–33 36 37]

Higher age was significantly associated with disability at follow-up in four studies.[32 33 37 39] Three studies did not demonstrate significant associations between higher age and poor outcome.[4 31 36] However, although statistically insignificant, visual inspection of the remaining three studies' results indicated a similar trend. A dose–response effect was observed; an increase by 10 years of age was associated with a higher risk of disability than was a 1 year increase.[37] The following factors were adjusted for in the multivariable analysis of the primary studies: sex,[4 33] job,[33] BMI,[33] baseline pain severity,[33] recruitment,[36] depression,[33] anxiety,[33] fear avoidance,[33] activity level prior to current pain episode[33] and baseline disability.[32 36] Three studies did not specify the factors adjusted for.[31 37 39]

### Previous episodes

Four studies including patients with headache[40] and low back pain[4 31 32] investigated the association between previous episodes and disability. The association between previous episodes and disability was inconsistent. Two studies detected a significant association between previous episodes and disability.[31 40] One reported a non-

**Table 1** Characteristics of the included studies

| Author (publication year) | Country of origin | Participant eligibility criteria | Number of participants at baseline | Age at baseline Mean (SD) | Pain site | Recruitment | Outcome measure | Follow-up (months) |
|---|---|---|---|---|---|---|---|---|
| Boardman (2006) | UK | Adults >18 years | 730* | 52 (18–90)† | Head | GP‡ | Disability (Migraine Disability Assessment) | 12 |
| Boersma (2005) | Sweden | No information | 363 | 47 (10.2) | Back or neck | GP‡ | Disability (Örebro Musculoskeletal Pain Screening Questionnaire) and sick leave (>15 days) | 12 |
| Childs (2004) | USA | Patients 18–60 years; with a primary symptom of LBP, with or without referral into the lower extremity; and an Oswestry Disability Questionnaire (ODQ) score of at least 30% | 131 | 33.9 (10.9) | LPB§ | Physiotherapy | Disability (Modified Oswestry Disability Index) | 6 |
| Coste (2004) | France | Patients >18 years, self-referring to GP (n: 40) or rheumatologists (n: 7) for a primary symptom of LBP with pain duration <72 h and without radiation below the gluteal fold | 113 | 44.3 (13.7) | LBP§ | GP‡ | Disability (VAS and Roland Morris Disability Questionnaire) | 3 |
| Grotle (2007) | Norway | Patients 18–60 years; acute LBP of <3 weeks' duration, with or without radiating pain to the limb; and had not been treated for LBP earlier | 123 | 37.9 (10.1) | LBP§ | GP‡ | Disability (Roland Morris Disability Questionnaire) | 12 |
| Grotle (2010) | Norway | Patients consulting GP with non-specific LBP of varying duration and localisation | 258 | 46 (9) | LBP§ | GP‡ | Disability (Roland Morris Disability Questionnaire) | 12 |
| Hancock (2008) | Australia | Primary symptom of pain in the area between the 12th rib and buttock crease causing moderate pain and moderate disability (measured by adaptations of items 7 and 8 of the SF-36) | 240 | 40.7 (15.6) | LPB§ | GP‡ | Disability (Roland Morris Disability Questionnaire) | 3 |
| Hendrick (2013) | New Zealand | Patients aged 18–65 years with an episode of LBP of ≤6 weeks, preceded by a minimum period of 3 months during which participants had | 101 | 38.8 (14.6) | LBP§ | GP‡, Physiotherapy clinics and newspaper advertisement | Disability (Roland Morris Disability Questionnaire) | 3 |

Continued

**Table 1** Continued

| Author (publication year) | Country of origin | Participant eligibility criteria | Number of participants at baseline | Age at baseline Mean (SD) | Pain site | Recruitment | Outcome measure | Follow-up (months) |
|---|---|---|---|---|---|---|---|---|
| | | not sought treatment for LBP, and no other pre-existing conditions that limited their mobility | | | | | | |
| Heneweer (2007) | Holland | Patients aged 21–60 years with sufficient knowledge of the Dutch language to complete the questionnaires | 56 | 42 (9.2) | LBP§ | Physiotherapy clinics | Disability (recovery yes/no and sick leave yes/no) | 3 |
| Karjalainen (2003) | Finland | Patients aged 25–60 years having disabling LBP for the preceding 4–12 weeks | 164 | 44 (8.8) | LBP§ | GP‡ | Disability (Oswestry Disability Index) and sick leave (1: 0 days, 2: 1–30 days, 3: >30 days) | 12 |
| Leaver (2013) | Australia | Patients aged 18–70 years with a new episode of non-specific neck pain | 181 | 38.8 (10.7) | Neck | Physiotherapy and chiropractor clinics | Disability (Neck Disability Index) | 3 |
| Lonnberg (2010) | Denmark | Patients seeking care for the first time regarding an episode of LBP | 78 | 57¶ | LBP§ | GP‡ | Disability (Limitations —no further information) | 264 |
| Melloh (2013) | New Zealand | Patients 18–65 years | 315 | 34.9 (12.6) | LBP§ | GP‡ | Disability (Oswestry Disability Index) | 6 |
| Schultz (2004) | Canada | Participants aged 18–60 years remaining off work 4–6 weeks post-injury (subacute group) or remaining off work 6–12 months after injury (chronic) | 253 | 40.3 (11.4) | LBP§ | Workers' Compensation Board | Return-to-work status | 3 |
| Sieben (2005) | Holland | Patients aged 18–60 years with a new episode of non-specific LBP | 222 | No information | LBP§ | GP‡ | Disability (Graded Chronic Pain Scale) | 12 |
| Storheim (2005) | Norway | Patients sick listed from a permanent job and receiving between 50% and 100% compensation for non-specific LBP for 8–12 weeks, but with no sick leave due to LBP during a period of 12 weeks before the current sick-listing period; aged between 20 and 60 years | 93 | RTW: 40.5 (9.8) NRTW: 42.3 (11.7) | LBP§ | GP‡ and National Insurance Offices | Return-to-work status | 12 |

Continued

Table 1 Continued

| Author (publication year) | Country of origin | Participant eligibility criteria | Number of participants at baseline | Age at baseline Mean (SD) | Pain site | Recruitment | Outcome measure | Follow-up (months) |
|---|---|---|---|---|---|---|---|---|
| Swinkels-Meewisse (2006) | Holland | Patients aged 18–65 years having an episode of non-specific LBP independent of radiation | 374** | 42.4 (11.3) | LBP§ | GP‡ and Physiotherapy clinics | Disability (Roland Morris Disability Questionnaire) | 6 |
| Van der Windt (2007) | Holland | Patients 18–65 years with a duration of LBP <12 weeks at presentation, or exacerbation of mild symptoms of back pain | 171 (Back group) | 42.0 (12.0) (back group) | LBP§ | GP‡ | Disability (Roland Morris Disability Questionnaire) | 3 |
| Vos (2008) | Holland | Patients >18 years with neck pain <6 weeks | 187 | 40.7 (14.1) | Neck | GP‡ | Sick leave (>7 days) | 12 |

*In total 2662 patients were included in the study but only 730 respondents are included in the relevant analyses.
†Median (range).
‡General practitioner.
§Low back pain.
¶Median (range).
**555 Participants are included in the trial, but data regarding disability were available from only 374 of the participants.
GP, general practitioner; LBP, low back pain; SF-36, Short Form 36; VAS, visual analogue scale.

significant positive association,[32] and one study reported a non-significant negative association (previous episodes were not associated with disability at follow-up).[4] The following factors were adjusted for in the multivariable analysis of the primary studies: age,[4 40] sex[4 40] and baseline disability.[32] One study did not specify the factors adjusted for.[31]

### Longer pain duration

The association between longer pain duration at baseline and disability was assessed in three studies including patients with headache[40] and low back pain.[32 38] Longer pain duration was defined as <4 vs >24 h in one study.[40] The two other studies did not provide information on their definition of pain duration at baseline.[32 38] All three studies found that longer pain duration was associated with disability at follow-up when adjusted for: age,[38 40] sex,[38 40] BMI[38] and baseline disability.[32]

### Anxiety

The association between anxiety and disability was assessed in two studies concerning patients with headache[40] and low back pain.[36] Both studies reported a significant association between anxiety when adjusted for age,[36 40] sex,[36 40] recruitment[36] and baseline disability.[36] Although 'anxiety' was significantly associated with outcome in the study by Grotle *et al*,[36] the size of the association was fairly low, accounting for <2% of the explained variation, implying that the prognostic value in clinical practice may be limited.

### Depression

Three studies assessed the association between depression and disability. Depression was found to be associated with disability at follow-up in two studies when adjusted for age,[36 40] sex,[36 40] recruitment[36] and baseline disability.[36] However, although significant, Grotle *et al*[36] stated that the size of the association for the factor 'depression' was low, accounting for <2% of the explained variance in the study. No association was found in the remaining study when adjusted for age, sex and BMI.[38]

## Quality of evidence for the risk of developing disability

The quality of evidence for the potential prognostic factors for the risk of developing disability is presented in table 2. All the included studies in the evidence profile were phase 1 studies, which are characterised as predictive modelling studies or explanatory studies conducted to generate a hypothesis.[20] Thus, the quality of evidence was moderate as a starting point.[20] Reasons for upgrading or downgrading the quality of evidence for the given prognostic factor are described below in table 2. The quality of evidence was graded as moderate, low or very low.

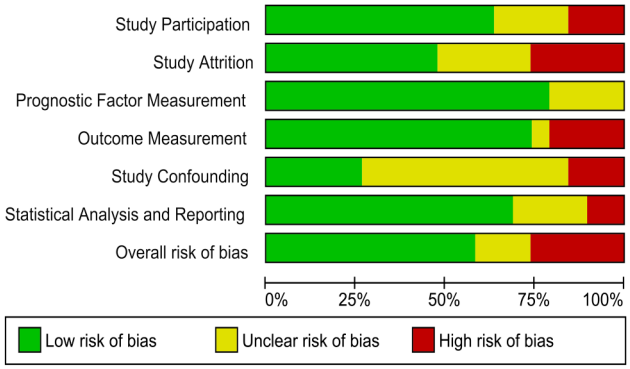



**Figure 2** Risk of bias of the six domains in the Quality in Prognosis Studies tool (QUIPS).

### Prognostic factors for long-term sick leave or return-to-work

Three of the included studies had long-term sick leave as an outcome.[3 27 37] Two of the studies, referring to patients with pain in the neck[27] and back[37] had performed multivariable analysis. The follow-up period was 3 months in both studies. Long-term sick leave was defined as sick leave for more than 30 days in one study[37] and more than 7 days in the other.[27] Baseline disability was the only potential prognostic factor that was assessed in both studies. In the study by Karjalainen *et al*,[37] however, the results were described only as being 'non-significant'. Thus, it was not possible to synthesise the results. In the study by Karjalainen *et al*, factors such as blue-collar work and long-term sick leave at baseline were associated with an increased risk of long-term sick leave at follow-up in patients with subacute back pain. Vos *et al*[27] found that factors such as previous pain episodes, a follow-up appointment scheduled with a GP, and the GPs referring the patient to treatment and baseline disability were all associated with an increased risk of long-term sick leave among patients with acute neck pain. Two studies described 'return-to-work' as an outcome.[25 26] Follow-up was 3[25] and 12 months.[26] Both studies related to patients with back pain, so potential generic factors could not be extracted.

## DISCUSSION

Five potential generic prognostic factors for developing disability following a subacute pain condition were identified. Risk factors across different pain sites included multiple site pain, higher pain severity, higher age, baseline disability and pain duration at baseline. There was inconsistent evidence regarding the association between previous pain episodes and disability. Although a few studies found anxiety, and depression to be associated with disability at follow-up the prognostic value of these factors may be low. Owing to the limited number of studies, it was not possible to identify potential generic risk factors for long-term sick leave or return-to-work. Quality of evidence was low or very low, implying that confidence in the estimate is low.

### Comparison with other studies or reviews

Despite the sparse literature in this field, there is some evidence to support our findings. In concurrence with our findings, a previous review reported factors such as multiple site pain, higher pain severity, higher age, baseline disability and longer pain duration at baseline as being potential prognostic factors for a poor outcome in patients with musculoskeletal pain.[8] A strong association between the number of pain sites and disability was also demonstrated in a previous cross-sectional study.[9] Similarly, a prospective cohort study from 2008 found that the number of pain sites were a strong predictor of work disability 14 years later, regardless of the diagnosis.[42] Furthermore, a recent systematic review found some evidence suggesting that the number of somatic symptoms and baseline severity of the condition influenced the future course in patients with medically unexplained symptoms.[43] Thus, despite the low quality of evidence of the results, we find it reasonable to believe that the factors identified in our systematic review may act as central prognostic factors for the development of disability across pain sites. Therefore, future research should focus on confirming the role of these factors.

Interestingly, our review found limited evidence that psychosocial factors were associated with disability at follow-up. These findings are surprising because psychosocial factors, also known as 'yellow flags', are widely accepted as being key factors in the transition from acute to chronic pain conditions.[44] 'Yellow flags' include depression, anxiety, catastrophic thoughts and pain-related fear of movement/fear avoidance among others.[45 46] Several national and international guidelines recommend that clinicians screen for the presence of these factors in the early phase.[47–49] In addition, several well-established screening tools for the risk of chronicity are based on the presence of these factors (eg, the Orebro Musculoskeletal Pain Questionnaire,[50] the Fear-Avoidance Health Beliefs Questionnaire[51] and the STarT Back Screening Tool[52]). Most studies, however, have not included pain duration at baseline when the importance of 'yellow flags' was assessed.[16] A plausible explanation for the apparent discrepancy between our results and the widely accepted 'yellow flags', therefore, could be the inclusion criteria in our review regarding short pain duration at baseline. It is likely that psychosocial factors are of greater importance once pain has become chronic. Another explanation for the limited evidence could be that our review focuses on risk factors for future disability and sick leave and not on risk factors for developing a chronic pain condition. Nonetheless, future research should address and clarify the role 'yellow-flag' factors play in the various phases of pain.

### Strength and limitations

It is considered a strength of our systematic review that we followed a rigorous protocol (registered in PROSPERO) prespecifying all the outcomes and

**Table 2** GRADE evidence profile of the potential prognostic factors for long-term disability in patients with a subacute pain condition

| | | Quality assessment | | | | | | Summary of findings | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | Number of participants included in the analyses | Multivariable analysis* | | | |
| Prognostic factors (number of studies) | Phase | Limitations | Inconsistency | Indirectness | Imprecision | Publication bias | Other considerations | | + | 0 | – | Overall quality |
| Multiple site pain[36 40 41] (3) | 1 | No serious limitations | No serious inconsistency | No serious indirectness | No serious imprecision | Undetected† | | 1164 | 3 | 0 | 0 | Moderate +++ |
| High baseline pain severity[31 32 34 36 37 40] (6) | 1 | Serious limitation (–1) | Serious inconsistency‡ (–1) | No serious indirectness | No serious imprecision | Detected§ (–1) | Dose–response effect detected (+1) | 1711 | 6 | 0 | 0 | Very low + |
| Baseline disability[32 34–39] (7) | 1 | No serious limitations | No serious inconsistency¶ | No serious indirectness | No serious imprecision | Undetected† | | 1263 | 5 | 2 | 0 | Moderate +++ |
| Older age[4 31–33 36 37 39] (7) | 1 | No serious limitations | No serious inconsistency** | No serious indirectness | No serious imprecision | Detected (–1)†† | Dose-response effect detected (+1) | 1296 | 4 | 3 | 0 | Moderate +++ |
| Longer pain duration[32 38 40] (3) | 1 | Serious limitations (–1) | No serious inconsistency | No serious indirectness | No serious imprecision | Detected (–1)‡‡ | | 1236 | 3 | 0 | 0 | Very low + |
| Previous episodes[4 31 32 40] (4) | 1 | Serious limitation (–1) | No serious inconsistency§§ | No serious indirectness | Serious imprecision (–1) | Detected¶¶ (–1) | | 1353 | 2 | 2 | 0 | Very low + |
| Anxiety[36 40] (2) | 1 | Serious limitations (–1) | No serious inconsistency | No serious indirectness | No serious imprecision | Undetected† | | 988 | 2 | 0 | 0 | Low ++ |
| Depression[36 38 40] (3) | 1 | No serious limitations | No serious inconsistency | No serious indirectness | No serious imprecision | Undetected† | | 1157 | 2 | 1 | 0 | Moderate +++ |

*For multivariable analyses: +, number of significant effects with a positive value; 0, number of non-significant effects; –, number of significant effects with a negative value.
†The association between the prognostic factor and disability is assessed only in the trials included in the evidence profile.
‡Substantial heterogeneity between study results was detected. Consequently, the quality of evidence was downgraded due to inconsistency.
§Nine studies assessed the association between high baseline pain and disability at follow-up, but only six studies reported the results in the adjusted analyses.
¶In the study by Karjalainen *et al*, the association between baseline disability and disability at follow-up was reported by an increase by 20% in maximum score at baseline, whereas the other studies reported an increase by 1 point. This difference could be a plausible reason for the inconsistency between the results.
**Karjalainen *et al* reported the association between an increase in age by 10 years and disability, whereas the other studies reported the association between age and disability by an increase of 1 year.
††Nine studies reported the association between age and disability, but only seven studies included the results in the adjusted analyses.
‡‡The association between baseline pain duration and disability was reported in the unadjusted analyses in five studies, but only three studies included the results in the adjusted analyses.
§§Inconsistency in the results between the study by Boardman *et al* and the study by Swinkels-Mewisse *et al* can be explained by differences in the reporting of previous episodes. In the study by Swinkels *et al*, the participants could have experienced pain once 10 years ago, whereas Broadman *et al* look at pain episodes one or more times per week.
¶¶Nine studies reported previous pain episodes in the unadjusted analyses, but only seven studies included the results in the adjusted analyses.

analyses; our adherence to the protocol most likely strengthens the credibility of the evidence synthesis. We reported our findings as recommended by the PRISMA statement[18] and judged the quality of the evidence based on the recommendations from the GRADE Working group. We believe that the GRADE framework applied to prognostic factor research was valuable for assessing and transparently reporting the quality of the evidence of the possible prognostic factors. To the best of our knowledge, this is the first time GRADE has been used in the evaluation of prognostic studies.

Limitations regarding the interpretation of the results from this study should be taken into consideration. A total of 14 of the 19 included studies in our review referred to patients with back pain. The high number of studies concerning patients with back pain may affect the external validity of the results to patients with pain at other sites. However, the vast majority of pain patients visiting general practice suffer from back pain, and the large number of studies on back pain included in the present review therefore reflects the distribution of patients seen in general practice.[53] Future studies assessing prognostic factors for non-spinal pain are needed. Publication bias is one of the most common biases in systematic reviews. Therefore, we conducted funnel plots to explore whether publication bias was present in our analysis. No obvious asymmetry was found. In accordance with current knowledge, the use and appropriate interpretation of funnel plots are, however, controversial because of questions about statistical validity, disputes over appropriate interpretation and low power of the tests.[54] For instance, a funnel plot can be symmetrical even in the presence of publication bias.[54] Hence publication bias might be present in our analyses although undetected. Another common limitation in systematic reviews is the risk of selective reporting of primary study results. Our review was based primarily on observational cohort studies on prognostic factors (phase 1 studies). Such studies harbour a high risk that non-significant findings are not reported or only included in the first (unadjusted) part of the analysis. Any non-reporting of non-significant results invites a risk that the findings in the synthesis were overestimated. We attempted to account for such bias due to selective outcome reporting by listing all the studies that examined a specific prognostic factor in the unadjusted analysis. If a factor was investigated in eight studies, for example, but included only in the adjusted analysis of five, the quality of the evidence was downgraded.

## Implications for clinical practice

No high-quality evidence was provided for any of the potential prognostic factors; therefore, no definite clinical conclusion can be made about how to identify patients at high risk of long-term disability or sick leave at an early stage in general practice. However, the empirical evidence illustrates what kind of prognostic factor research would be relevant to pursue in order to increase value and reduce waste in prognostic factor research on long-term disability among patients with subacute non-malignant pain.[55] It appears that multiple site pain, high-baseline pain severity, older age, baseline disability and pain for a longer duration are associated with future disability across pain sites in subacute pain patients. Therefore, it may be helpful for GPs to have these factors in mind in clinical decision-making.

## Implications for future research on prognostic factors

Correctable weaknesses in biomedical and public health research studies can produce misleading results and waste valuable resources.[56] During the preparation of this review, it became clear that the current literature in this field falls short on a number of counts. As suggested by Ioannidis *et al*,[56] this area of research also has weaknesses, such as selective reporting of results; lack of prespecified defined prognostic factors to be assessed; inadequate description of methods; inadequate or poor quality of statistical analysis; failure to distinguish between prognostic factors among patients with acute, subacute and chronic pain; and lack of published studies on patients suffering from non-spinal pain. We suspect most of these limitations can be related to the absence of detailed written protocols and poor documentation of research in general.[56–59]

Although good research ideas often yield unanticipated but valuable results, much research fails to effect worthwhile achievements. As long as the way in which research projects are prioritised for research is transparent and warranted, the disappointments should not be deemed wasteful; they are simply an inevitable feature of the way science works.[55] In order to gain further knowledge on which factors are central prognostic factors (subacute phase), future studies should take into account baseline pain duration at the time patients are enrolled. Future studies on prognostic factors in chronic pain should be conducted as large, prospective, registered and protocol-based prognostic factor studies with sufficient study populations and transparent reporting of all factors studied. Once sufficient knowledge on risk factors has been obtained, documentation of effective treatment for high-risk pain patients is needed. Further, the effect of offering stratified care to pain patients based on their risk profiles should be tested in randomised controlled trials.[60]

**Author affiliations**
[1]Department of Central Denmark Region, Health Technology Assessment and Health Services Research, CFK-Public Health and Quality Improvement, Aarhus, Denmark
[2]Department of Public Health, Research Initiative for Activity Studies and Occupational Therapy, General Practice, University of Southern Denmark, Odense, Denmark
[3]Pain Research Group, Pain Centre South, Department of Anaesthesiology and Intensive Care Medicine, University Hospital Odense, Odense, Denmark
[4]Research Unit for General Practice, Aarhus University, Denmark
[5]Musculoskeletal Statistics Unit, Department of Rheumatology, The Parker Institute, Copenhagen University Hospital, Copenhagen, Denmark

**Acknowledgements** The authors wish to thank Danish Regions for their support. They also thank Hanne Caspersen from the State and University library for conducting the literature search and Professor Troels S Jensen from the Danish Research Center for commenting on the manuscript. Finally, RC acknowledged The Oak Foundation and Copenhagen University Hospital at Frederiksberg for their support.

**Contributors** All the authors were responsible for the design and search strategy. GHV and MSP were responsible for conducting the search. GHV, MSP and RC conducted the data analysis and produced the tables and graphs. RC provided input into the data analysis and interpretation. The initial draft of the manuscript was prepared by GHV, then circulated among all authors for critical revision. All authors helped to evolve analysis plans, interpret data and critically revise successive drafts of the manuscript.

**Funding** This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

**Competing interests** None declared.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data sharing statement** No additional data are available.



## REFERENCES

1. Kibsgård K, Abell A, Andersen BH, *et al*. Behandling af kroniske smerter i Region Midtjylland. 2009. https://www.rm.dk/ (accessed May 2014).
2. Guide to Assessing Psychosocial Yellow Flags in Acute Low Back Pain: Risk Factors for Long-Term Disability and Work Loss. http://www.cebp.nl/media/m24.pdf (accessed May 2014).
3. Boersma K, Linton SJ. Screening to identify patients at risk: profiles of psychological risk factors for early intervention. *Clin J Pain* 2005;21:38–43.
4. Grotle M, Brox JI, Glomsrod B, *et al*. Prognostic factors in first-time care seekers due to acute low back pain. *Eur J Pain* 2007;11:290–8.
5. Breivik H, Collett B, Ventafridda V, *et al*. Survey of chronic pain in Europe: prevalence, impact on daily life, and treatment. *Eur J Pain* 2006;10:287–333.
6. Eriksen J, Jensen MK, Sjøgren P, *et al*. Epidemiology of chronic non-malignant pain in Denmark. *Pain* 2003;106:221–8.
7. Harker J, Reid KJ, Bekkering GE, *et al*. Epidemiology of chronic pain in Denmark and Sweden. *Pain Res Treat* 2012;2012:371248.
8. Mallen CD, Peat G, Thomas E, *et al*. Prognostic factors for musculoskeletal pain in primary care: a systematic review. *Br J Gen Pract* 2007;57:655–61.
9. Kamaleri Y, Natvig B, Ihlebaek CM, *et al*. Localized or widespread musculoskeletal pain: does it matter? *Pain* 2008;138:41–6.
10. Chou R, Shekelle P. Will this patient develop persistent disabling low back pain? *JAMA* 2010;303:1295–302.
11. Campbell P, Wynne-Jones G, Muller S, *et al*. The influence of employment social support for risk and prognosis in nonspecific back pain: a systematic review and critical synthesis. *Int Arch Occup Environ Health* 2013;86:119–37.
12. Kuijer W, Groothoff JW, Brouwer S, *et al*. Prediction of sickness absence in patients with chronic low back pain: a systematic review. *J Occup Rehabil* 2006;16:439–67.
13. Nicholas MK, Linton SJ, Watson PJ, *et al*. "Decade of the Flags" Working Group. Early identification and management of psychological risk factors ("yellow flags)" in patients with low back pain: a reappraisal. *Phys Ther* 2011;91:737–53.
14. Pincus T, Vogel S, Burton AK, *et al*. Fear avoidance and prognosis in back pain: a systematic review and synthesis of current evidence. *Arthritis Rheum* 2006;54:3999–4010.
15. Mallen CD, Thomas E, Belcher J, *et al*. Point-of-care prognosis for common musculoskeletal pain in older adults. *JAMA Intern Med* 2013;173:1119–25.
16. Laisne F, Lecomte C, Corbiere M. Biopsychosocial predictors of prognosis in musculoskeletal disorders: a systematic review of the literature (corrected and republished)*. *Disabil Rehabil* 2012;34:1912–41.
17. Ørtenblad L, Valentin GH, Carstensen K, *et al*. SMERTER-En MTV om muligheder for tidlig målrettet indsats til smertepatienter. Health Tecnology Assessment and Health Services, CFK ▪ Public Health and Quality Improvement ▪ Central Denmark Region. 2014.
18. Liberati A, Altman DG, Tetzlaff J, *et al*. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *J Clin Epidemiol* 2009;62:e1–34.
19. Guyatt GH, Oxman AD, Schunemann HJ, *et al*. GRADE guidelines: a new series of articles in the Journal of Clinical Epidemiology. *J Clin Epidemiol* 2011;64:380–2.
20. Huguet A, Hayden JA, Stinson J, *et al*. Judging the quality of evidence in reviews of prognostic factor research: adapting the GRADE framework. *Syst Rev* 2013;2:71.
21. Hayden JA, van der Windt DA, Cartwright JL, *et al*. Assessing bias in studies of prognostic factors. *Ann Intern Med* 2013;158:280–6.
22. Cochrane handbook for systematic reviews of interventions. http://handbook.cochrane.org/ (accessed 28 Oct 2014).
23. GRADEpro. http://www.guidelinedevelopment.org/ (accessed 28 Nov 2014).
24. Lonnberg F, Pedersen PA, Siersma V. Early predictors of the long-term outcome of low back pain—results of a 22-year prospective cohort study. *Fam Pract* 2010;27:609–14.
25. Schultz IZ, Crook J, Meloche GR, *et al*. Psychosocial factors predictive of occupational low back disability: towards development of a return-to-work model. *Pain* 2004;107:77–85.
26. Storheim K, Brox JI, Holm I, *et al*. Predictors of return to work in patients sick listed for sub-acute low back pain: a 12-month follow-up study. *J Rehabil Med* 2005;37:365–71.
27. Vos CJ, Verhagen AP, Passchier J, *et al*. Clinical course and prognostic factors in acute neck pain: an inception cohort study in general practice. *Pain Med* 2008;9:572–80.
28. Hancock MJ, Maher CG, Latimer J, *et al*. Independent evaluation of a clinical prediction rule for spinal manipulative therapy: a randomised controlled trial. *Eur Spine J* 2008;17:936–43.
29. Childs JD, Fritz JM, Flynn TW, *et al*. A clinical prediction rule to identify patients with low back pain most likely to benefit from spinal manipulation: a validation study. *Ann Intern Med* 2004;141:920–8.
30. van der Windt DA, Kuijpers T, Jellema P, *et al*. Do psychological factors predict outcome in both low-back pain and shoulder pain? *Ann Rheum Dis* 2007;66:313–19.
31. Sieben JM, Vlaeyen JW, Portegijs PJ, *et al*. A longitudinal study on the predictive validity of the fear-avoidance model in low back pain. *Pain* 2005;117:162–70.
32. Swinkels-Meewisse IE, Roelofs J, Schouten EG, *et al*. Fear of movement/(re)injury predicting chronic disabling low back pain: a prospective inception cohort study. *Spine (Phila Pa 1976)* 2006;31:658–64.
33. Hendrick P, Milosavljevic S, Hale L, *et al*. Does a patient's physical activity predict recovery from an episode of acute low back pain? A prospective cohort study. *BMC Musculoskelet Disord* 2013;14:126.
34. Heneweer H, Aufdemkampe G, van Tulder MW, *et al*. Psychosocial variables in patients with (sub)acute low back pain: an inception cohort in primary care physical therapy in The Netherlands. *Spine (Phila Pa 1976)* 2007;32:586–92.
35. Coste J, Lefrancois G, Guillemin F, *et al*. French Study Group for Quality of Life in Rheumatology. Prognosis and quality of life in patients with acute low back pain: insights from a comprehensive inception cohort study. *Arthritis Rheum* 2004;51:168–76.
36. Grotle M, Foster NE, Dunn KM, *et al*. Are prognostic indicators for poor outcome different for acute and chronic low back pain consulters in primary care? *Pain* 2010;151:790–7.
37. Karjalainen K, Malmivaara A, Mutanen P, *et al*. Outcome determinants of subacute low back pain. *Spine* 2003;28(23):2634–40.
38. Melloh M, Elfering A, Chapple CM, *et al*. Prognostic occupational factors for persistent low back pain in primary care. *Int Arch Occup Environ Health* 2013;86:261–9.
39. Leaver AM, Maher CG, McAuley JH, *et al*. People seeking treatment for a new episode of neck pain typically have rapid improvement in symptoms: an observational study. *J Physiother* 2013;59:31–7.
40. Boardman HF, Thomas E, Millson DS, *et al*. The natural history of headache: predictors of onset and recovery. *Cephalalgia* 2006;26:1080–8.
41. Leaver AM, Maher CG, Herbert RD, *et al*. A randomized controlled trial comparing manipulation with mobilization for recent onset neck pain. *Arch Phys Med Rehabil* 2010;91:1313–18.
42. Kamaleri Y, Natvig B, Ihlebaek CM, *et al*. Does the number of musculoskeletal pain sites predict work disability? A 14-year prospective study. *Eur J Pain* 2009;13:426–30.

43. olde Hartman TC, Borghuis MS, Lucassen PL, *et al*. Medically unexplained symptoms, somatisation disorder and hypochondriasis: course and prognosis. A systematic review. *J Psychosom Res* 2009;66:363–77.
44. Kent P, Kjaer P. The efficacy of targeted interventions for modifiable psychosocial risk factors of persistent nonspecific low back pain—a systematic review. *Man Ther* 2012;17:385–401.
45. Fritz JM, George SZ, Delitto A. The role of fear-avoidance beliefs in acute low back pain: relationships with current and future disability and work status. *Pain* 2001;94:7–15.
46. Pincus T, Vlaeyen JW, Kendall NA, *et al*. Cognitive-behavioral therapy and psychosocial factors in low back pain: directions for the future. *Spine (Phila Pa 1976)* 2002;27:E133–8.
47. Accident Compensation Corporation. New Zealand acute low back pain guide, incorporation the guide to assessing psychosocial yellow flags in acute low back pain. http://www.nzgg.org.nz/guidelines/0072/acc1038_col.pdf (accessed Oct 2007).
48. Management of soft tissue injuries: treatment providers. http://www.workcover.nsw.gov.au/NR/rdonlyres/A3708AE4-E902-4DA8-B692-5E1AE3710728/0/management_soft_tissue_injuries_treatment_providers_4811.pdf (accessed Oct 2007).
49. Klinisk retningslinje for udredning og klassificering af personer med nakkebesvær. http://fysio.dk/Upload/Graphics/PDF-filer/Kliniske_retningslinjer/Nakkebesv%C3%A6r/Nakke2.pdf (accessed 06 Jun 2014).
50. Hockings RL, McAuley JH, Maher CG. A systematic review of the predictive ability of the Orebro Musculoskeletal Pain Questionnaire. *Spine (Phila Pa 1976)* 2008;33:E494–500.
51. Waddell G, Newton M, Henderson I, *et al*. A Fear-Avoidance Beliefs Questionnaire (FABQ) and the role of fear-avoidance beliefs in chronic low back pain and disability. *Pain* 1993;52:157–68.
52. Hill JC, Dunn KM, Lewis M, *et al*. A primary care back pain screening tool: identifying patient subgroups for initial treatment. *Arthritis Rheum* 2008;59:632–41.
53. Brooks PM. The burden of musculoskeletal disease—a global perspective. *Clin Rheumatol* 2006;25:778–81.
54. Sterne JA, Sutton AJ, Ioannidis JP, *et al*. Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *BMJ* 2011;343:d4002.
55. Chalmers I, Bracken MB, Djulbegovic B, *et al*. How to increase value and reduce waste when research priorities are set. *Lancet* 2014;383:156–65.
56. Ioannidis JP, Greenland S, Hlatky MA, *et al*. Increasing value and reducing waste in research design, conduct, and analysis. *Lancet* 2014;383:166–75.
57. Al-Shahi Salman R, Beller E, Kagan J, *et al*. Increasing value and reducing waste in biomedical research regulation and management. *Lancet* 2014;383:176–85.
58. Chan AW, Song F, Vickers A, *et al*. Increasing value and reducing waste: addressing inaccessible research. *Lancet* 2014;383:257–66.
59. Glasziou P, Altman DG, Bossuyt P, *et al*. Reducing waste from incomplete or unusable reports of biomedical research. *Lancet* 2014;383:267–76.
60. Hingorani AD, Windt DA, Riley RD, *et al*. Prognosis research strategy (PROGRESS) 4: stratified medicine research. *BMJ* 2013;346:e5793.

# Supplementary material

**Supplementary Material S1:** Description of meta-analysis including forest plots of the association between potential prognostic factors and long-term disability in patients with a sub-acute pain condition.

## Data synthesis and analysis

If a baseline factor was associated with outcomes at follow-up in two or more studies of different pain sites, it was considered a "possible prognostic" factor and the results were combined and subsequently presented as part of the evidence profile. When data were available in different formats, data from the "fully adjusted" analyses were given preference and included in the meta-analysis. The studies in the meta-analyses were heterogenic regarding statistical methods. Hence, it was not possible directly to compare the results from the original studies. In order to assed the statistical power between a prognostic factor and outcome Fishers z-transformation was used. From each cohort study, the "statistical signal" (The "Wald-test"—the ratio between signal and noise) was derived from the effect size and the standard error of the estimate (SE). These were subsequently handled using Fisher's z-transformation.[23] This z-transformation was used to communicate the statistical power for any given association (i.e., correlation) between a given prognostic factor and an outcome. In one study the outcome was "absence of disability" as opposed to "disability". Thus the effect estimate (HR: 0.97) was reversed (HR: 1.03) before transformed into logscale and entered in RevMan.[24]

Summary estimates of associations across studies were derived from random effects meta-analysis, anticipating clinical heterogeneity, with modelling allowing for differences in the association from study to study.[25]

Heterogeneity across studies was statistically assessed using the Q-test and quantified by the inconsistency ($I^2$) index.[26] $I^2$ represents the percentage of total variation across studies attributable to heterogeneity rather than (statistical) chance.[27] In cases with substantial heterogeneity across studies ($I^2 > 50\%$), the robustness of the results was checked using the "fixed effects" model. A result was

considered robust if the point estimate for "fixed effects" was within the confidence interval of "random effects"; as a consequence, the risk of "small-study" bias was considered to be high if the point estimate was outside the confidence estimate. If this was the case, the evidence for the given prognostic factor was downgraded due to inconsistency. In order to explore the robustness of the pooled estimates sensitivity analyses were conducted in cases were three or more studies in the meta-analysis had similar effect value. Meta-analyses were performed using Review Manager (RevMan) provided by the Cochrane Collaboration.[28] A two-sided P-value of $\leq 0.05$ (and 95% confidence interval excluding the null) was considered to be statistically significant in all analyses. For each outcome, we prepared an evidence profile based on the GRADE profiler software.[29]

**Prognostic factors for disability**

Prognostic factors for disability were assessed in 16 studies.[3, 4, 24, 30, 34-45] A total of 81 factors were assessed in the unadjusted analysis (Supplementary Material S1). Of the 81 factors assessed, 53 were included in the multivariable analysis of the primary studies. Of these factors, the following eight were assessed in two or more studies and for at least two different pain sites: multiple-site pain, higher baseline pain severity, previous pain episodes, older age, longer baseline pain duration, baseline disability, anxiety, and depression. A total of eleven studies were included in the evidence profile (figure ?).[4, 24, 37-45] The association between multiple-pain sites and disability at follow-up was assessed in three studies, including patients with headache,[45] low back pain,[41] and neck pain[44] (figure ?). All three studies found a statistically significant association between multiple-pain sites at baseline and risk of disability at follow-up. The combined estimate showed a statistically significant association between multiple-pain sites at baseline and risk of disability at follow-up ($p = 0.010$). The test for heterogeneity was $I^2 = 57\%$. However, "fixed effects" did not change the result significantly.

The association between higher pain severity at baseline and disability was assessed in six studies including patients with headache[45] and low back pain[37, 38, 40-42] (figure 3b). The combined estimate showed a statistically significant association between higher baseline pain severity and disability ($p < 0.001$). $I^2$ was 93% and the result was not robust when using the "fixed effects" model, so small-study bias is therefore likely. Consequently, the quality of evidence was downgraded due to inconsistency. The association between baseline disability and disability at follow-up was examined in seven studies relating to patients with neck pain[44] and low back pain[24, 38, 40-43] (figure 3c). Baseline disability was associated with a higher risk of disability at follow-up ($p = 0.007$). $I^2$ was 96%. The "fixed effects" model did not change the result significantly.The association between higher age and the risk of disability was assessed in seven studies concerning patients with neck pain[44] and low back pain[4, 37-39, 41, 42] (figure 3d). A significant association was seen between higher age and risk of disability ($p = 0.04$). $I^2$ was 89%. The "fixed effects model" did not change the result significantly. A dose-response effect was observed; an increase by 10 years of age was associated with a higher risk of disability than was a one-year increase. No statistical significant associations between disability and previous pain episodes ($p = 0.08$; figure 4a), longer baseline pain duration ($p= 0.12$; figure 4b), anxiety ($p = 0.25$; figure 4c), or depression ($p = 0.14$; figure 4d) were observed. The "fixed effects" model did not change the results notably for the four potential prognostic factors in figure 4.

Sensitivity analyses were conducted for the results of the meta-analysis presented in figure 3C and 3D. In figure 3C the study by Karjalainen et al. was omitted and in figure 3D the studies by Grotle et al, 2007 and Karjalainen et al. were omitted. The statistical signal did not change substantially after selective omitting studies with different effect value (Supplementary material S3

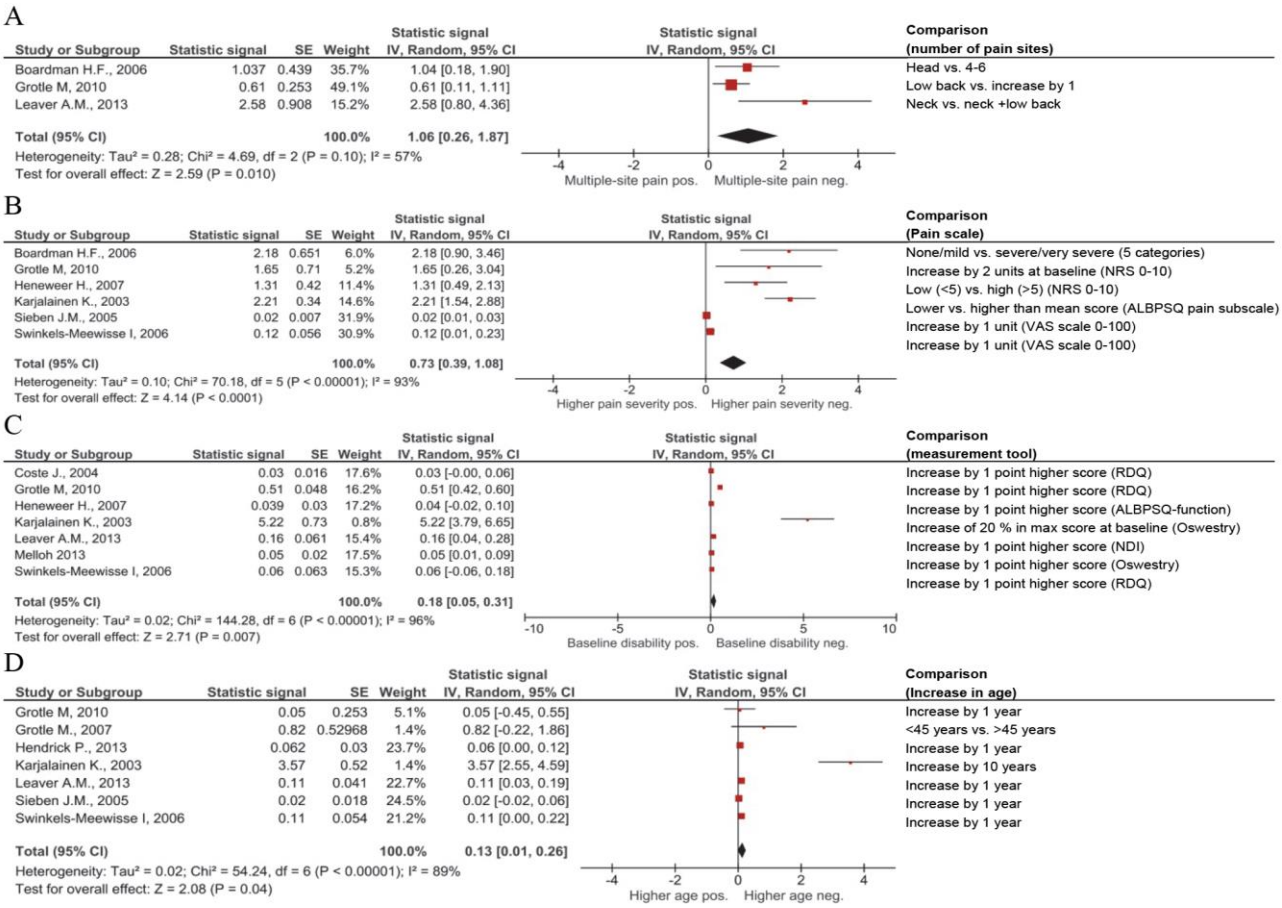



**Figure 3:** Forest plots of the association between potential prognostic factors and long-term disability in patients with a sub-acute pain condition. 3A: Adjusted for: Age [41, 45], sex [41, 45], baseline disability [41], recruitment [41], factors adjusted for not described [44]. 3B: Adjusted for: Age [41, 45], sex [41, 45], baseline disability [41], recruitment [41], factors adjusted for not described [37, 40, 42].3C: Adjusted for: Age [41, 43], Sex [41, 43], BMI [43], duration (hours) between pain debut and inclusion [24], job status [24], previous spine surgery [24], compensation status [24], self rated health status [24], factors adjusted for not described [38, 40, 42, 44].3D: Adjusted for sex [4, 39], job [39], BMI [39], Baseline pain severity [39], recruitment [41], depression [39], anxiety [39], fear avoidance [39], activity level prior current pain episode [39], baseline disability [38, 41], factors adjusted for not described [37, 42, 44].

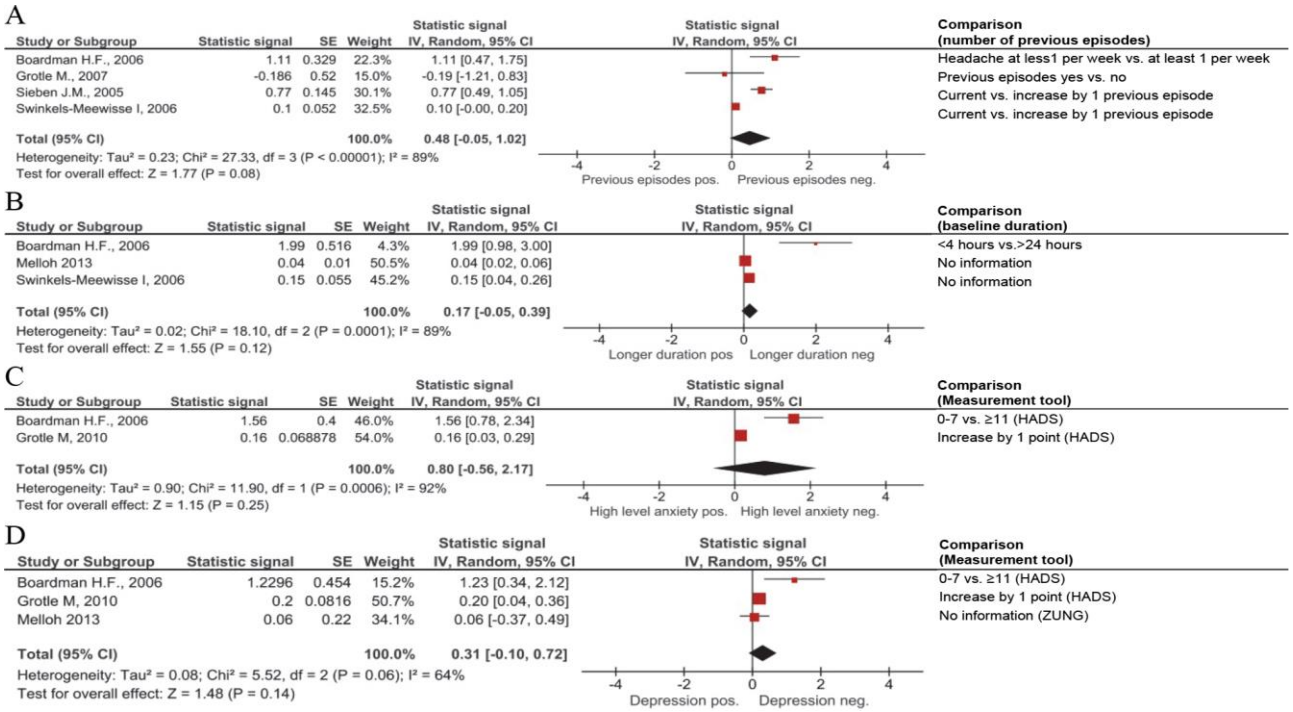
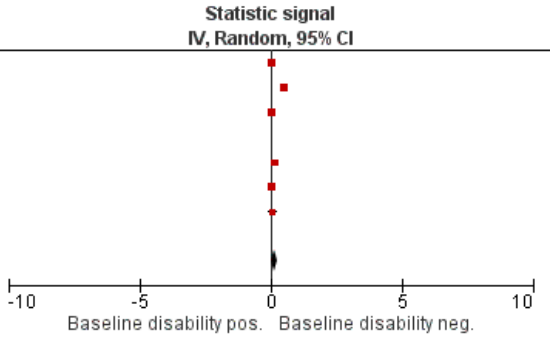



**Figure 4:** Forest plots of the association between potential prognostic factors and long-term disability in patients with a sub-acute pain condition. 4A: Adjusted for: age [4, 45], sex [4, 45], baseline disability [38], factors adjusted for not described [37]. 4B: Adjusted for: Age [43, 45], sex [43, 45], BMI [43], baseline disability [38]. 4C: Adjusted for: Age [41, 45], sex [41, 45], recruitment [41], baseline disability [41]. 4D: Adjusted for: Age [41, 43, 45], sex [41, 43, 45], BMI [43], recruitment [41], baseline disability [41].

**Supplementary Material S1 (continued):** Sensitivity analysis omitting studies with different effect value.

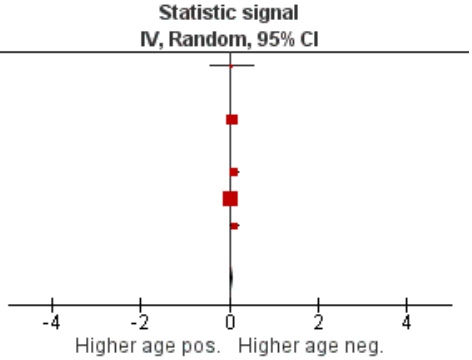
| Study or Subgroup | Statistic signal | SE | Weight | Statistic signal IV, Random, 95% CI |
|---|---|---|---|---|
| Coste J., 2004 | 0.03 | 0.016 | 18.3% | 0.03 [-0.00, 0.06] |
| Grotle M, 2010 | 0.51 | 0.048 | 16.2% | 0.51 [0.42, 0.60] |
| Heneweer H., 2007 | 0.039 | 0.03 | 17.6% | 0.04 [-0.02, 0.10] |
| Karjalainen K., 2003 | 5.22 | 0.73 | 0.0% | 5.22 [3.79, 6.65] |
| Leaver A.M., 2013 | 0.16 | 0.061 | 15.0% | 0.16 [0.04, 0.28] |
| Melloh 2013 | 0.05 | 0.02 | 18.1% | 0.05 [0.01, 0.09] |
| Swinkels-Meewisse I, 2006 | 0.06 | 0.063 | 14.8% | 0.06 [-0.06, 0.18] |
| **Total (95% CI)** | | | **100.0%** | **0.14 [0.03, 0.24]** |

Heterogeneity: Tau² = 0.02; Chi² = 94.45, df = 5 (P < 0.00001); I² = 95%
Test for overall effect: Z = 2.55 (P = 0.01)

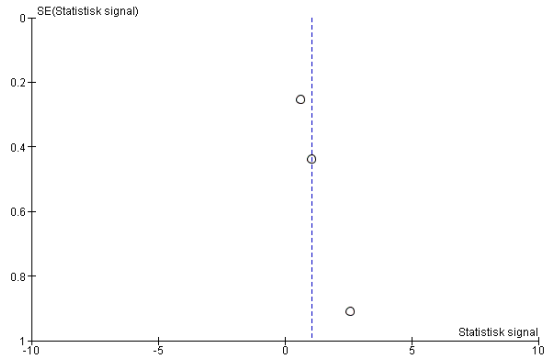



| Study or Subgroup | Statistic signal | SE | Weight | Statistic signal IV, Random, 95% CI |
|---|---|---|---|---|
| Grotle M, 2010 | 0.05 | 0.253 | 0.7% | 0.05 [-0.45, 0.55] |
| Grotle M., 2007 | 0.82 | 0.52968 | 0.0% | 0.82 [-0.22, 1.86] |
| Hendrick P., 2013 | 0.062 | 0.03 | 27.1% | 0.06 [0.00, 0.12] |
| Karjalainen K., 2003 | 3.57 | 0.52 | 0.0% | 3.57 [2.55, 4.59] |
| Leaver A.M., 2013 | 0.11 | 0.041 | 18.4% | 0.11 [0.03, 0.19] |
| Sieben J.M., 2005 | 0.02 | 0.018 | 41.7% | 0.02 [-0.02, 0.06] |
| Swinkels-Meewisse I, 2006 | 0.11 | 0.054 | 12.2% | 0.11 [0.00, 0.22] |
| **Total (95% CI)** | | | **100.0%** | **0.06 [0.02, 0.10]** |

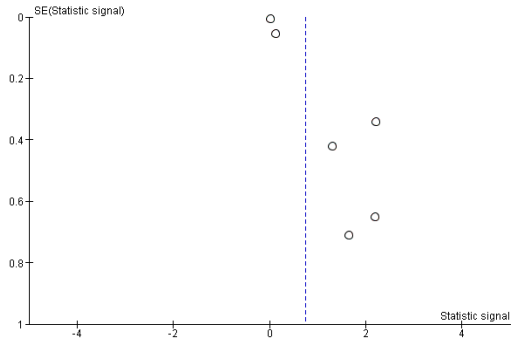
Heterogeneity: Tau² = 0.00; Chi² = 6.21, df = 4 (P = 0.18); I² = 36%
Test for overall effect: Z = 2.80 (P = 0.005)



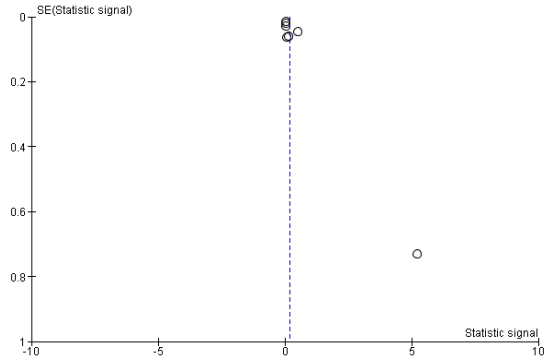
**Supplementary Material S2:** Funnel plots of studies included in the meta-analyses.



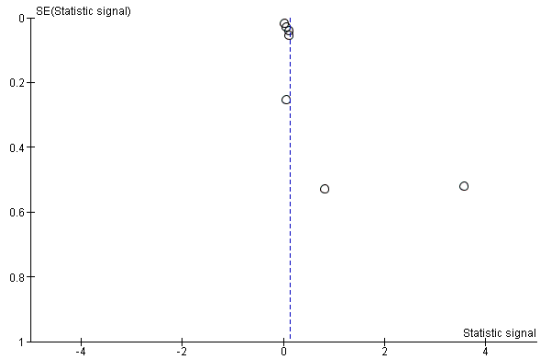

**S2A:** Multiple-site pain



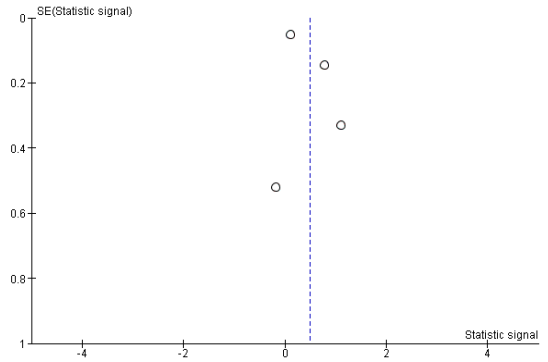

**S2B:** Higher pain severity



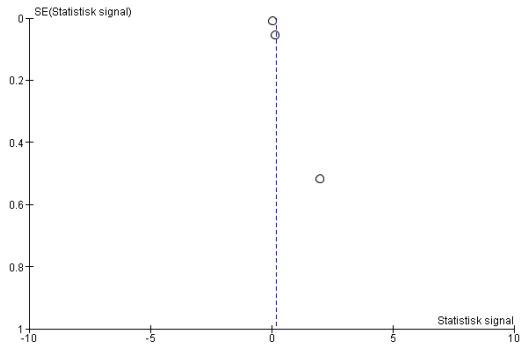

**S2C:** Disability



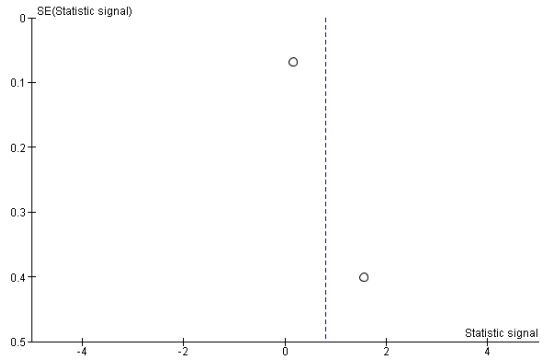

**S2D:** Higher age



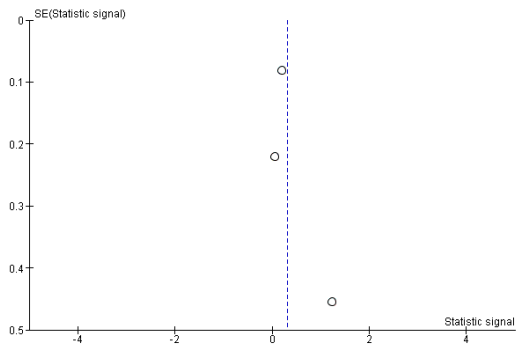

**S2E:** Previous episodes




**S2F:** Longer duration




**S2G:** High level anxiety




**S2H:** Depression

**Supplementary Material S3:** Potential prognostic factors for disability assessed in the **unadjusted analysis** of the included studies. A plus sign indicates a statistically significant association between the given factor and the outcome. A minus sign indicates a non-significant association.

| | **Prognostic factor** | **Boardman** | **Boersma** | **Childs** | **Coste** | **Grotle, 2007** | **Grotle, 2010** | **Hancock** | **Hendrick,** | **Heneweer** | **Karjalainen** | **Leaver** | **Lonnberg** | **Melloh** | **Sieben** | **Swinkels-Meewisse** | **Van der Windt** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **Pain site** | **Head** | **Back or neck** | **LBP** | **LBP** | **LBP** | **LBP** | **LBP** | **LBP** | **LBP** | **LBP** | **Neck** | **LBP** | **LBP** | **LBP** | **LBP** | **LBP** |
| Socio demo-graphic | **Age** | | | | ÷ | ÷ | ÷ | | + [a] | | + | + | ÷ | ÷ | ÷ [a] | + [a] | |
| | Gender | | | | ÷ | ÷ | ÷ | | | | ÷ | + | ÷ | ÷ | | ÷ [a] | |
| | BMI | | | | | | | | | | + | + | | + | | | |
| | Smoking | | | | | ÷ | | | | | | | | ÷ | | | |
| | Alcohol | ÷ [a] | | | | | | | | | | | | ÷ | | | |
| | Caffeine | ÷ [a] | | | | | | | | | | | | | | | |
| | Civil status | | | | | | | | | | | | | | | ÷ [a] | |
| | Origin | | | | ÷ | | | | | | | | | | | | |
| | Education | | | | | ÷ | + | | | | ÷ | ÷ | | ÷ | ÷ [a] | ÷ [a] | |
| | Social class | | | | | | ÷ | | | | | | | | | | |
| Pain | **Baseline intensity** | + [a] | | | ÷ | | + | | | + | + | ÷ | | + | + [a] | + [a] | |
| | **Baseline duration** | + [a] | | | | | | | | | | + | ÷ | ÷ | | + [a] | |
| | **Previous episodes** | + [a] | | | ÷ | ÷ | | | | | | ÷ | ÷ | | + [a] | + [a] | |
| | Associated symptoms | + [a] | | | | | | | | | | | | | | | |
| | Duration of previous episodes | | | | | | | | | | | | | + | | | |
| | Permanent pain at nigth | | | | ÷ | | | | | | | | | | | | |
| | Pain increased by coughing | | | | + | | | | | | | | | | | | |
| | Pain increased by back movement | | | | ÷ | | | | | | | | | | | | |
| | Pain worse when standing | | | | ÷ | | | | | | | | | | | | |
| | Pain worse when lying down | | | | + | | | | | | | | | | | | |
| | Pain on passive movements | | | | ÷ | | | | | | | | | | | | |
| | Pain radiation | | | | | ÷ | + | | | | ÷ | | ÷ | | | + [a] | |
| | Neurological signs | | | | | ÷ | | | | | | | ÷ | | | | |
| | Onset (sudden) | | | | | | | | | | | | | | | ÷ [a] | |
| | Sensory pain | | | | | | | | | | | | | + | | | |
| | Affective pain | | | | | | | | | | | | | + | | | |
| | **Multiple-pain sites** | + [a] | | | | | + | | | | | + | | | | | |
| | Bothersomeness | | | | | | + | | | | | | | | | | |

a: Only adjusted estimates presented

**Supplementary Material S3 (continued):** Potential prognostic factors for disability assessed in the **unadjusted analysis** of the included studies. A plus sign indicates a statistically significant association between the given factor and the outcome. A minus sign indicates a non-significant association.

| Prognostic factor | | Boardman | Boersma | Childs | Coste | Grotle, 2007 | Grotle, 2010 | Hancock | Hendrick | Heneweer | Karjalainen | Leaver | Lomberg | Melloh | Sieben | Swinkels-Meewisse | Van der Windt |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Pain site** | | **Head** | **Back or neck** | **LBP** | **LBP** | **LBP** | **LBP** | **LBP** | **LBP** | **LBP** | **LBP** | **Neck** | **LBP** | **LBP** | **LBP** | **LBP** | **LBP** |
| Functional status at baseline | Forward lumbar bending (ROM[b]) | | | | | ÷ | | | | | | | ÷ | | | | |
| | Scoliosis | | | | | | | | | | | | ÷ | | | | |
| | Disability | | | | + | | + | | | ÷ [a] | + | + | | ÷ | | ÷ [a] | |
| | Physical activity | | | | | | | | ÷ | | | | | ÷ | | + [a] | |
| | Participation | | | | | | | | | | | | | | | | |
| | General health | | | | + | | | | ÷ [a] | | ÷ | + | | | | | |
| | Limitations physical | | | | ÷ | | | | | | | | | + | | | |
| | Disability days last month | | | | | + | | | | | | | | | | | |
| Work related | Employed at baseline | | | | ÷ | | + | | | | | + | | | | | |
| | Changed work status | | | | | | | | | | | ÷ | | | | | |
| | Physically demanding work | | | | ÷ | | | | | | ÷ | | ÷ | + | | | |
| | Job unsatisfaction | | | | ÷ | ÷ | ÷ | | | | ÷ | | | + | | | |
| | Resigned attitude towards the job | | | | | | | | | | | | | + | | | |
| | Uncertainty | | | | | | | | | | | | | ÷ | | | |
| | Organization | | | | | | | | | | | | | ÷ | | | |
| | Interruptions | | | | | | | | | | | | | ÷ | | | |
| | Concentration | | | | | | | | | | | | | ÷ | | | |
| | Time pressure | | | | | | | | | | | | | ÷ | | | |
| | Ergonomics | | | | | | | | | | | | | ÷ | | | |
| | Emotions (job related) | | | | | | | | | | | | | ÷ | | | |
| | Resources (job related) | | | | | | | | | | | | | ÷ | | | |
| | Social support at work | | | | | | | | | | | | | + | | | |
| | Compensation status | | | | + | | | | | | | ÷ | | | | | |
| | Job difficulty | | | | ÷ | | | | | | | | | | | | |
| | Pain onset at work | | | | | | | | | | | | + | | | | |
| | Sick leave due to pain | | | | | | + | | | | ÷ | | ÷ | ÷ | | | |
| | Past sick leave for pain | | | | | | | | | | | + | | | | | |

a: Only adjusted estimates presented

b: Range of motion

**Supplementary Material S3 (continued):** Potential prognostic factors for disability assessed in the **unadjusted analysis** of the included studies. A plus sign indicates a statistically significant association between the given factor and the outcome. A minus sign indicates a non-significant association.

| Prognostic factor | | Boardman | Boersma | Childs | Coste | Grotle, 2007 | Grotle, 2010 | Hancock | Hendrick | Heneweer | Karjalainen | Leaver | Lonnberg | Melloh | Sieben | Swinkels-Meewisse | Van der Windt |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Pain site** | | **Head** | **Back or neck** | **LBP** | **LBP** | **LBP** | **LBP** | **LBP** | **LBP** | **LBP** | **LBP** | **Neck** | **LBP** | **LBP** | **LBP** | **LBP** | **LBP** |
| Psychological | Psychiatric disorder | | | | ÷ | | | | | | | | | | | | |
| | **Anxiety** | + [a] | | | | | + | | | | | | | | | | |
| | **Depression** | + [a] | | | | | + | | | | | | | + | | | |
| | Emotional distress | | | | | + | | | | | | | | | | | ÷ [a] |
| | Fear of pain | | | | | | + | | | | | | | | | | |
| | Fear of movement | | | | | | | | | ÷ | | | | | | + [a] | |
| | Fear avoidance | | ? | | | ÷ | + | | ÷ [a] | ÷ [a] | | | | | | | ÷ [a] |
| | Coping | | | | | | ÷ | | | ÷ | | | | | | | |
| | Catastrophising | | | | | | + | | | | | | | + | | | ÷ [a] |
| | Somatisation | | | | | | | | | | | | | + | | | ÷ [a] |
| | Perceived risk of not recovering | | | | | | | | | | + | | | | | | |
| | Expectation regarding the effectiveness of treatment | | | | | | | | | | ÷ | | | | | | |
| | Social functioning | | | | ÷ | | | | | | | | | | | | |
| | Vitality | | | | ÷ | | | | | | | | | | | | |
| | Physical problems | | | | ÷ | | | | | | | | | | | | |
| | Mental health | | | | ÷ | | | | | | | + | | + | | | |
| Other | Prescription of bed rest | | | | ÷ | | | | | | | | | | | | |
| | Prescription of sick leave | | | | ÷ | | | | | | | | | | | | |
| | Recruitment group | | | | | ÷ | ÷ | | | | | | | | | | |
| | Sleeping/relaxation medication | | | | | ÷ | | | | | | | | | | | |
| | Pain medication | | | | | ÷ | | | | | | + | | | | | |
| | Prior low back surgery | | | | ÷ | | | | | | | | | | | | |
| | Other co morbidity | | | | ÷ | | | | | | | | | | | | |
| | "Clinical prediction rule" | | | + | | | | + | | | | | | | | | |
| | ALBPSQ* / CPG[d] | | | | + | | + | | | ÷ [a] | | | | | | | |

a: Only adjusted estimates presented

c: GP, chiropractor or physiotherapist.

d: Acute low back screening questionnaire