

Prognosis of chronic clinical pain conditions : the example of complex regional pain syndrome 1 and low back pain

Citation for published version (APA):

Wertli, M. M. (2014). *Prognosis of chronic clinical pain conditions : the example of complex regional pain syndrome 1 and low back pain*. [Doctoral Thesis, Maastricht University]. Datawyse / Universitaire Pers Maastricht. <https://doi.org/10.26481/dis.20140926mw>

Document status and date:

Published: 01/01/2014

DOI:

[10.26481/dis.20140926mw](https://doi.org/10.26481/dis.20140926mw)

Document Version:

Publisher's PDF, also known as Version of record

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy

If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

Prognosis of chronic clinical pain conditions

The example of complex regional pain syndrome 1 and low back pain

Maria M. Wertli



© Maria Monika Wertli, Maastricht 2014

Production: Datawyse | Universitaire Pers Maastricht

ISBN 978 94 6159 354 2

Prognosis of chronic clinical pain conditions:

The example of complex regional pain syndrome 1 and low back pain

DISSERTATION

to obtain the degree of doctor at Maastricht University,
on the authority of the Rector Magnificus, Prof. Dr. L.L.G. Soete
in accordance with the decision of the Board of Deans,
to be defended in public on
Friday 26th of September 2014 at 12.00 hours

by

Maria Monika Wertli



Promotor

Prof. Dr. Maarten van Kleef

Co-promotores

PD. Dr. Florian Brunner (University of Zurich)

Dr. Roberto Perez (VUMC Amsterdam)

Beoordelingscommissie

Prof. Dr. Bert Joosten, voorzitter

Prof. Dr. Lodewijk van Rhijn

Prof. Dr. Rob Smeets

Prof. Dr. Harry Steinbusch

Prof. Dr. Wouter Zuurmond (VUMC Amsterdam)

Table of contents

Chapter 1	General introduction	7
Chapter 2	Incomplete reporting of baseline characteristics in clinical trials: An analysis of randomized controlled trials and systematic reviews involving patients with chronic low back pain	13
Chapter 3	The role of fear avoidance beliefs as a prognostic factor for outcome in patients with non-specific low back pain: A systematic review	31
Chapter 4	The influence of fear avoidance beliefs on treatment outcome in patients with low back pain: A systematic review	77
Chapter 5	Concordance of qualitative bone scintigraphy results with presence of clinical complex regional pain syndrome 1: Meta-analysis of test accuracy studies	131
Chapter 6	Prognostic factors in complex regional pain syndrome 1: A systematic review	151
Chapter 7	Rational pain management in complex regional pain syndrome 1 (CRPS 1): A network meta-analysis	169
Chapter 8	General discussion	193
	Summary	203
	Samenvatting	207
	Valorization addendum	211
	Dankwoord	217
	Curriculum vitae	219
	List of publications	221

CHAPTER 1

General introduction

Every fifth European suffers from chronic pain that seriously affects the quality of life [1, 2]. Whereas for specific pain conditions (e.g. radiculopathy) an underlying cause can be identified, a large proportion of chronic pain conditions are non-specific with no underlying causal explanation. While in specific pain conditions a targeted therapy of the underlying somatic anomaly may be initiated, in non-specific pain conditions, no causal derived therapies are possible and the therapy remains symptom-based [3]. In order to avoid chronicity, it is important to identify early patients at risk for delayed recovery. However, when no clear underlying cause can be found, the identification of patients at risk represents a major challenge in clinical practice.

Non-specific low back pain (LBP) and complex regional pain syndrome 1 (CRPS 1) are two important examples of non-specific pain conditions. While non-specific LBP is common [4] the Complex Regional Pain Syndrome type 1 (CRPS 1) is less known but economically very important [5, 6]. Both illnesses share a high patient burden and lead to substantial healthcare expenditures for pain management [1, 7].

Non-specific LBP is defined as lumbar pain without a specific morphological correlate and accounts for over 90 percent of all LBP cases [8, 9]. In the absence of a structural abnormality explaining the patient complaint, the diagnosis is solely based on clinical findings. Further examinations—in particular imaging—do not give a causal explanation [8]. CRPS 1 is defined as a pain state following injury, which exceeds in magnitude and duration the expected clinical course when other underlying diseases are excluded [10]. Clinical manifestations include a broad spectrum of sensory, autonomic, motor, and trophic changes usually associated with significant impairment of motor function [10]. For diagnosis, the use of the “Budapest criteria” is recommended [11]. These criteria consider signs and symptoms of the clinical manifestation. In the past, several attempts have been made to improve the diagnostic work-up by using imaging, in particular qualitative bone scintigraphy [12]. To date, the diagnostic value of qualitative bone scintigraphy is controversial [13] and has not been assessed systematically.

From the clinical point of view, individualized and targeted treatment interventions are needed to address individual patients’ needs. At present, the prognosis of acute LBP has been considered to be favorable in the majority of patients without specific therapy. However, a recent systematic review has challenged this assumption because every second patient reported persistent LBP after one year [14]. Therefore, it is important to identify patients at risk for delayed recovery at an early stage. Once identified, these patients require targeted interventions to prevent chronicity (e.g. cognitive behaviorally-oriented physical therapy [15]). Currently, several prognostic factors for chronicity are available [16, 17] and it may be hypothesized that these factors potentially influence treatment efficacy and outcome.

For example, the fear of pain and re-injury is considered as a potential factor for delayed recovery in non-specific LBP. The fear avoidance model introduced in the 1990s [18] is commonly used to explain the link between psychological factors, pain experience, and ultimately the development of chronic pain and disability [19]. According to

the model, negative beliefs about pain may lead to a catastrophizing response in which the worst possible outcome is imagined. Consequently, the fear of activity and avoidance results in disuse, which then reinforces the original negative appraisal in a deleterious cycle [19]. However, current research questions the impact of fear avoidance beliefs on the prognosis and treatment efficacy in non-specific LBP [20, 21].

To date, in patients with CRPS 1, the prognostic aspects have been incompletely understood and their influence on the course of the condition remain controversial. In the absence of valid and confirmed prognostic factors, Brunner et al. recently conducted a Delphi consensus study [22]. The findings of this expert survey indicated that a poor prognosis for CRPS 1 is primarily dependent on clinical manifestations, more specifically, on the presence of sensory abnormalities. Since these results represent a consensus between expert opinions, potential prognostic factors available from the literature need to be systematically assessed.

In evidence-based medicine, systematic reviews and meta-analyses are considered to be the gold standard to draw firm conclusions from randomized controlled trials (RCTs). In studies of non-specific LBP, many systematic reviews and meta-analyses have been reported in the literature. However, it is unknown how they account for prognostic factors. Furthermore, it is unclear how RCTs reported prognostic factors. A new technique—a network-analytic approach—allows for comparison of treatment efficacy between different interventions in the absence of a head-to-head comparison [23]. In studies of patients with CRPS 1, meta-analysis is often impeded because of the small number of RCTs and the small patient sample. Therefore, such a network-analytical approach may allow for a rank-ordering of therapeutic regimens.

The challenges outlined above have stimulated my academic work over the last several years. As a result of my scientific inquiry, the research presented in this thesis comprises a series of papers addressing the following research questions:

- How are potentially important prognostic factors for LBP reported in RCTs, and how are those factors addressed in meta-analyses (Chapter 2)?
- Do fear avoidance beliefs influence the prognosis in patients with acute, sub-acute, and chronic non-specific LBP (Chapter 3)?
- Do fear avoidance beliefs influence the efficacy of treatments in patients with non-specific LBP (Chapter 4)?
- What is the diagnostic accuracy of qualitative bone scintigraphy for the diagnosis of CRPS 1 (Chapter 5)?
- Which prognostic factors indicate a good or bad prognosis for patients with CRPS 1 (Chapter 6)?
- What represents a rational treatment approach for pain management of CRPS 1, as determined by using a network meta-analytic approach (Chapter 7)?

The studies were performed at the Horten Centre, University of Zurich, Zurich, Switzerland, and at the Occupational and Industrial Orthopaedic Center (OIOC), New York

University Hospital for Joint Diseases (NYU-HJD), New York, U.S.A, in collaboration with the University Hospital Balgrist, Zurich, Switzerland.

References

1. Breivik, H., et al., Survey of chronic pain in Europe: prevalence, impact on daily life, and treatment. *European Journal of Pain*, 2006. 10(4): p. 287-333.
2. van Hecke, O., N. Torrance, and B.H. Smith, Chronic pain epidemiology and its clinical relevance. *Br J Anaesth*, 2013. 111(1): p. 13-8.
3. Steurer, J., L. Bachmann, and O. Miettinen, Etiology in a taxonomy of illnesses. *European Journal of Epidemiology*, 2006. 21(2): p. 85-89.
4. Vos, T., et al., Years lived with disability (YLDs) for 1160 sequelae of 289 diseases and injuries 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. *The Lancet*, 2012. 380(9859): p. 2163-2196.
5. Allen, G., B.S. Galer, and L. Schwartz, Epidemiology of complex regional pain syndrome: a retrospective chart review of 134 patients. *Pain*, 1999. 80(3): p. 539-44.
6. Duman, I., et al., Reflex sympathetic dystrophy: a retrospective epidemiological study of 168 patients. *Clin Rheumatol*, 2007. 26(9): p. 1433-7.
7. Jänig, W., R. Schaumann, and W. Vogt, CRPS Complex regional pain syndrome, SUVA, Editor 2013, SUVA: Luzern.
8. Koes, B., et al., An updated overview of clinical guidelines for the management of non-specific low back pain in primary care. *European Spine Journal*, 2010. 19(12): p. 2075-2094.
9. Balagu, F., et al., Non-specific low back pain. *Lancet*, 2012. 379(9814): p. 482-491.
10. Harden, R.N., et al., Proposed New Diagnostic Criteria for Complex Regional Pain Syndrome. *Pain Medicine*, 2007. 8(4): p. 326-331.
11. Harden, R.N., et al., Validation of proposed diagnostic criteria (the "Budapest Criteria") for Complex Regional Pain Syndrome. *PAIN*, 2010. 150(2): p. 268-274.
12. O'Donoghue, J.P., et al., Three-phase bone scintigraphy. Asymmetric patterns in the upper extremities of asymptomatic normals and reflex sympathetic dystrophy patients. *Clin Nucl Med*, 1993. 18(10): p. 829-36.
13. Lee, G.W. and P.M. Weeks, The role of bone scintigraphy in diagnosing reflex sympathetic dystrophy. *J Hand Surg Am*, 1995. 20(3): p. 458-63.
14. Itz, C.J., et al., Clinical course of non-specific low back pain: a systematic review of prospective cohort studies set in primary care. *European Journal of Pain*, 2013. 17(1): p. 5-15.
15. Hill, J.C., et al., A primary care back pain screening tool: identifying patient subgroups for initial treatment. *Arthritis and Rheumatism*, 2008. 59(5): p. 632-41.
16. Hayden, J.A., et al., What is the prognosis of back pain? *Best Practice & Research Clinical Rheumatology*, 2010. 24(2): p. 167-179.
17. Hayden, J.A., et al., Systematic reviews of low back pain prognosis had variable methods and results: guidance for future prognosis reviews. *J Clin Epidemiol*, 2009. 62(8): p. 781-796.e1.
18. Waddell, G., 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(2): p. 157-68.
19. Linton, S.J. and W.S. Shaw, Impact of Psychological Factors in the Experience of Pain. *Physical Therapy*, 2011. 91(5): p. 700-711.
20. Pincus, T., et al., A systematic review of psychological factors as predictors of chronicity/disability in prospective cohorts of low back pain. *Spine (Phila Pa 1976)*, 2002. 27(5): p. E109-20.
21. Pincus, T., et al., Fear avoidance and prognosis in back pain: a systematic review and synthesis of current evidence. *Arthritis and Rheumatism*, 2006. 54(12): p. 3999-4010.
22. Brunner, F., M. Nauer, and L.M. Bachmann, Poor prognostic factors in complex regional pain syndrome 1: A Delphi survey. *Journal of Rehabilitation Medicine*, 2011. 43(9): p. 783-6.
23. Kessels, A.G.H., et al., A simple regression model for network meta-analysis. *OA Epidemiology*, 2013. 1(1): p. 7.

CHAPTER 2

Incomplete reporting of baseline characteristics in clinical trials: An analysis of randomized controlled trials and systematic reviews involving patients with chronic low back pain

Maria M. Wertli, Manuela Schöb, Florian Brunner, Johann Steurer
PLOS ONE, www.plosone.org, 1 March 2013, Volume 8, Issue 3, e58512

Abstract

Objective: The aim of this study was to evaluate the reporting of relevant prognostic information in a sample of randomized controlled trials (RCTs) that investigated treatments for patients with chronic low back pain (LBP). We also analysed how researchers conducting the meta-analyses and systematic reviews addressed the reporting of relevant prognostic information in RCTs.

Methods: We searched the Cochrane Database to identify systematic reviews that investigated non-surgical treatments for patients with chronic LBP. The reported prognostic information was then extracted from the RCTs included in the reviews. We used a purpose-defined score to assess the quantity of information reported in the RCTs. We also determined how the authors of systematic reviews addressed the question of comparability of patient populations between RCTs.

Results: Six systematic reviews met our inclusion criteria, and we analysed 84 RCTs. Based on the scores, the reporting of important prognostic variables was incomplete in almost half of the 84 RCTs. Information regarding patients' general health, social support, and work-related conditions was rarely reported. Almost half of the studies included in one of the meta-analyses provided insufficient information that did not allow us to determine whether patients in the primary trials were comparable.

Conclusions: Missing prognostic information potentially threatens the external validity (i.e. the generalizability or applicability) not only of primary studies but also of systematic reviews that investigate treatments for LBP. A detailed description of baseline patient characteristics that includes prognostic information is needed in all RCTs to ensure that clinicians can determine the applicability of the study or review results to their patients.

Background

Assessing the external validity of randomized controlled trials (RCTs) is a key step in the critical appraisal of clinical studies. Many clinicians trust authors and journal editors to verify the high internal validity of the published studies (e.g., concealment of randomization list, information about drop-outs, intention to treat analysis), but physicians must decide for themselves whether the results apply to an individual patient. The information that is needed for this determination is reported in the Methods and Results sections of journal articles. The Methods section reports the eligibility criteria information, which states the patient qualifications for inclusion in the study. Patient characteristics are reported in the Results section; quite often, the article's Table 1 shows the distribution of characteristics of patients included in the study. Guidelines for reporting, e.g., the CONSORT Statement for randomized controlled trials [1], recommend not only a comprehensive description of eligibility criteria but also a list of baseline characteristics for important prognostic factors.

A complete description of relevant prognostic factors is particularly important in otherwise ill-defined diseases, such as chronic low back pain (CLBP). Several prognostic factors have been identified that can affect treatment effects in patients with CLBP, including age, duration of symptoms, first or recurrent episode, employment status, and comorbidities such as depression [2, 3]. For example, a treatment is effective in patients without depression but be less effective or even ineffective in depressed patients [4].

Knowing the patients' baseline characteristics is important for interpreting study results, both for clinicians and for the researchers who conduct systematic reviews and meta-analyses. Pooling the results of primary studies with unknown or different distributions of relevant prognostic factors in the included population may lead to a biased result [5]. It is unclear whether authors report important prognostic information in sufficient detail in primary studies so as to be helpful in rational pooling of data in meta-analyses and systematic reviews.

The aim of the current study was to evaluate the reporting of relevant prognostic information in a selection of randomized controlled trials (RCTs) investigating treatment outcomes in patients with CLBP. We also determined whether the authors of systematic reviews addressed the question of comparability of patient populations between RCTs.

Methods

Study design

Here we analysed primary studies included in CLBP-related systematic reviews in the Cochrane library. For the purpose of the current study, CLBP represents an ill-defined disease with high health care expenditure [6] for which important prognostic information is known to influence the course of the disease [3, 7]. We aimed to include a complete set of trials for each treatment intervention; therefore, we analyzed primary studies that were included in systematic reviews published in the Cochrane library. The Cochrane Collaboration Guideline [8-10] has published guidelines for the standardized assessment of baseline characteristics to facilitate comparison of systematic reviews. While this study is not a systematic review reporting will be based, if applicable, on the recommendations of the PRISMA statement [11].

Eligibility criteria and selection of systematic reviews

All systematic reviews that were published in the Cochrane library from its inception (1996) to December 2010 that investigated non-surgical treatments for CLBP were eligible for inclusion in our analysis. We searched the Cochrane library for the terms “chronic” and “non-specific low back pain” in the title, abstract, or keywords. Of the returned reviews, only RCTs published in English and German were eligible for further analysis due to the authors’ lack of proficiency in other languages. Non-randomized trials and observational studies were excluded.

Two reviewers (MW and MS) independently screened the titles and abstracts of the identified systematic reviews to determine which ones met the pre-defined inclusion criteria. The full text of each RCT included in the systematic reviews were then independently reviewed (MW and MS). Discrepancies between the two reviewers were discussed and resolved by consensus or by a third party (FB).

Data extraction and synthesis

One reviewer (MS) extracted data from the RCTs, including bibliographic data (authors, year of publication), eligibility criteria, and prognostic information. Prognostic information for LBP was defined a priori in collaboration with experienced clinicians (one internist, one rheumatologist, one general practitioner) and one methodologist in the field and by consulting the relevant literature [2, 3].

We used the prognostic domains proposed by Hayden et al. [7] to categorize the information reported in the RCTs. These domains, which are considered to represent clinically meaningful groups, [2] have been used in previous research and are based on expert consensus [12]. The following six main domains were used: general patient

characteristics, baseline health status, work-related factors, current low back pain (LBP), clinical examination findings, and interactions with work/society. Each main domain is divided into subdomains (e.g., current LBP is further divided according to the patient's clinical history, disability related to the complaint, and changes in the complaint over time). There were a total of 16 prognostic subdomains (Table 1). The six main domains represent a spectrum of important information that helps clinicians decide whether the study results are applicable to their patients.

Table 1. Important prognostic risk factor domains and subdomains in patients with low back pain (modified from Hayden et al. [7])

Domain	Subdomain	SQR Score*
General patient characteristics	Socio-demographic status	Minimal requirement: ≥ 1 subdomain reported
	Social support	
Baseline health status	Overall health	Minimal requirement: ≥ 1 subdomain reported
	Overall psychological health	
	Previous LBP	
Work-related factors	Work: psychosocial demands	Minimal requirement: ≥ 1 subdomain reported
	Work: physical demands	
	Work history	
	Work place attributes	
Current LBP	Clinical history	Minimal requirement: ≥ 1 subdomain reported
	Disability related to the complaint	
	Changes related to complaint over time	
Clinical examination findings	Physical examination findings	Minimal requirement: ≥ 1 subdomain reported
	Definition of NSLBP diagnosis†	
	Changes found during the physical exam	
Interactions with work/society	Compensation issues related to LBP	Minimal requirement: ≥ 1 subdomain reported

LBP: low back pain; NSLBP: nonspecific low back pain; † To fulfil this subdomain, at least one more attribute (in addition to pain duration) had to be reported (e.g. disability, severity, pain referral)[30]; *SQR: Score for the quantity of reporting: Scoring SQR high: information reported in one or more subdomains for all six main domains; SQR moderate: information reported in one or more subdomains for five main domains; SQR low: information reported in one or more subdomains for four or fewer main domains.

One reviewer (MW) confirmed all of the extracted information and assigned the data to the appropriate subdomains. To quantify the amount of reported prognostic information for each RCT, we defined a Score for the Quantity of Reporting (SQR) for each one as follows: High SQR, information was reported for one or more subdomain in all six main domains; moderate SQR, information was reported for one or more subdomains in five of the six main domains; and low SQR, information was reported for one or more subdomains in four or fewer main domains (Table 1).

The SQR for each study was then compared to how the baseline characteristics were assessed in the systematic reviews. Assessment of the comparability of baseline characteristics in studies is defined in the Method guidelines for systematic reviews in

the Cochrane Collaboration Back Review Group for Spinal Disorders [8] (first published in 1997). The relevant question is: "Are the baseline characteristics similar with regards to the most important prognostic factors?" The possible answers are "Yes/No/Don't know," and studies were divided into "Yes", "No", or "Can't tell" categories depending on the answer to that question. The updated Method Guidelines in 2003 [9] further stated, "In order to qualify for a "Yes," groups have to be similar at baseline regarding demographic factors, duration and severity of complaints, percentage of patients with neurologic symptoms." When not enough information is reported, the study must be classified as "Can't tell". We would expect that for primary studies with low SQRs, the answer to the above question would be "Can't tell." We also investigated whether studies with low SQR or that were classified as "Can't tell" were included in the meta-analysis.

Statistical analysis

Descriptive statistics were used to summarize findings across the entire set of RCTs. We wished to evaluate changes in the quantity of reporting over time, particularly after the publication of the CONSORT statement in 1996 [1], which aimed to improve the quality of reporting in RCTs. Toward this end, the mean number of reported subdomains before and after 1998 (to allow one year for implementation of CONSORT suggestions) was compared using the t-test. Analyses were conducted with SPSS for Windows version 19 (IBM SPSS; Chicago, IL USA) and R statistical software for Windows (<http://www.R-project.org/>).

Ethics statement

For this study no ethical approval was required. No protocol was published or registered. All methods were determined a priori.

Results

Study selection

Seven systematic reviews met the eligibility criteria (Figure 1). The reviews were published between 2005 and 2010 and included 100 primary studies. A total of 84 primary studies (RCTs) were included in the analysis. The main reason for exclusion was publication in a language other than English (n=16). Figure 1 shows a flow diagram of the study selection process.

Study characteristics

Table 2 summarizes the objectives, the number of included RCTs, and the conclusions of each systematic review. Most RCTs aimed to investigate treatments only for chronic low back pain; few studies included patients with subacute and acute low back pain. The number of RCTs included in each systematic review ranged from four [13] to thirty-two [14] trials. The RCTs were published between 1971 and 2009. More than half of the studies assessed the effects of acupuncture (n=18, 21.4%) or cognitive behavioural therapy (n=27, 32.1%). Most patients in the control groups received placebo (n=26, 30%), sham procedures (n=12, 14%), or usual care (n=12, 14%), or the patients were placed on a waiting list (n=13, 15%). In most studies, the follow-up time was about 6 months (median 6 months, range 1 hour to 5 years). Details are shown in Table 3.

Figure 1. Study flow

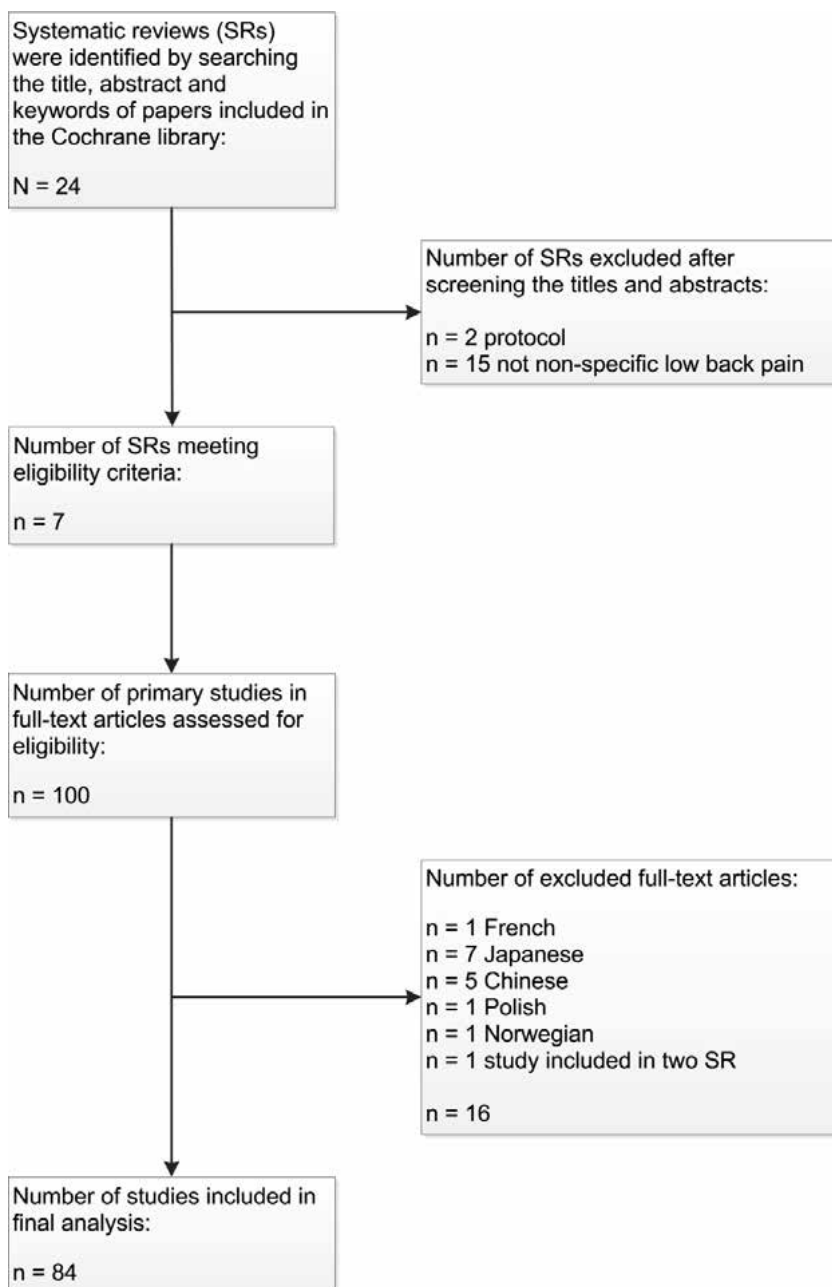


Table 2. Summaries of the systematic reviews in our analysis

Author	Year	Objective	Number of studies analysed	Conclusion
Furlan et al.[32]	2005	To assess the effects of AC for the treatment of NSLBP and the effects of dry-needling for myofascial pain syndrome in the low-back region.	20	Acute LBP: no firm conclusions about the effectiveness of AC. Chronic LBP: AC more effective for pain relief and functional improvement than no treatment or sham treatment and in the short-term only. AC is not more effective than other conventional treatments.
Urquhart et al.[33]	2008	To determine whether antidepressants are more effective than placebo for the treatment of NSLBP	9	No clear evidence that antidepressants are more effective than placebo in the management of patients with CLBP.
Henschke et al.[14]	2010	To determine the effects of behavioural therapy for CLBP and the most effective behavioural approach	32	Short-term: moderate quality evidence that operant therapy is more effective than being placed on a waiting list and that behavioural therapy is more effective than usual care for pain relief. No specific type of behavioural therapy is more effective than another. Intermediate- to long-term: Little or no difference between behavioural therapy and group exercises for pain or depressive symptoms.
Staal et al.[34]	2008	To determine if injection therapy is more effective than placebo or other treatments for patients with subacute or chronic LBP.	10	Insufficient evidence to support the use of injection therapy in subacute and chronic LBP. Insufficient data to answer whether specific subgroups of patients respond to a specific type of injection therapy.
Deshpande et al.[35]	2007	To determine the efficacy of opioids in adults with CLBP.	4	Quality remark: Although high internal validity scores, the study showed a lack of generalizability, inadequate description of study populations, a poor intention-to-treat analysis, and limited interpretation of functional improvement. The benefits of opioids in clinical practice for the long-term management of CLBP remain questionable.
Dagenais et al.[36]	2007	To determine the efficacy of prolotherapy in adults with CLBP.	5	When used alone, prolotherapy is not an effective treatment for CLBP. When combined with spinal manipulation, exercise, and other co-interventions, prolotherapy may improve CLBP and disability. Quality remark: Conclusions are confounded by clinical heterogeneity amongst studies and by the presence of co-interventions.
Khadilkar et al.[37]	2008	To determine whether TENS is more effective than placebo for the management of CLBP.	4	The current evidence from a small number of placebo-controlled trials does not support the use of TENS in the routine management of CLBP.

CLBP/NSLBP: chronic low back pain/nonspecific low back pain; LBP: low back pain; AC: acupuncture; UC: usual care; TENS: transcutaneous electrical nerve stimulation

Table 3. Characteristics of primary studies included in the systematic reviews

Systematic review	Year	Patient population (number of studies)	Recruitment and setting (number of studies)	Duration of LBP (years): mean* (range)	Follow-up (months) mean* (range)	Subjects (n) mean* (range)	Age (years) mean* (range)	Outcomes	Data-pooling
Furlan et al.	2005	Chronic SLBP + NSLBP (1), spinal pain syndrome (1), disabling LBP (1), NSLBP (17)	Specialized Unit (13), GP based (2), GP referral (1), Advertisement (3)	8.95 (0.5–12), 12 n.r.	4.15 (13 days–12 months)	74.5 (17–262)	49.67 (38–73)	VAS (9), McGill (2), ODI (2), others (7)	Yes
Urquhart et al.	2008	NSLBP (6), SPLBP + NSLBP (2)	Specialized Unit (3), Advertisement + Specialized Unit (5), GP based (1)	13.94 (3 months–20.3 years)	2.19 (6 weeks–8 months)	65.56 (16–121)	43.9 (26–53)	McGill (1), VAS (2), RMQ (1), others (7)	Yes
Henschke et al.	2010	SP: non-working disabling (1), chronic NSLBP (17), SLBP + NSLBP (1), Quebec taskforce (1), Topographic criteria of Kuorinka et al. (2), hospitalized (2); LBP: employees on sick leave (1), mildly dysfunctional (3), high paraspinal EMG level (1), FAB (1)	Advertisement (4), GP referral (7), GP referral + Advertisement (5), PT (1), Insurance (2), Specialized Unit (11)	7.46 (1.5–12)	12.49 (3 weeks–5 years)	112.19 (20–409)	43.39 (38–50)	NRS (3), McGill (5), RMQ (4), VAS (2), SF-36 (1), others (12)	Yes
Staal et al.	2008	Chronic SLBP + NSLBP (3), Radiologic severity of Osteoarthritis (Kellgren) (1), NSLBP (6)	Specialized Unit (7), GP setting + Specialized Unit (1), 2 no information (2)	3.29 (1–8.5)	2.32 (after treatment–6 months)	84.5 (16–206)	47 (40–65)	VAS (4), McGill (1), RMQ (1), others (4)	No
Deshpande et al.	2007	Chronic NSLBP (4)	3 GP setting + Specialized Unit, 1 Specialized Unit	Only in one reported (6.6)	3.73 (6 weeks–6 months)	239 (48–336)	50.3 (42–57)	VAS (2), McGill (2), others (2)	Yes
Dagenais et al.	2007	Chronic NSLBP (3), SPLBP + NSLBP (1), SLBP (1)	GP referral (2), Specialized Unit (1), no information (2)	9.55 (0 days–14 years)	9 (624)	171.4 (74–513)	44.4 (39–50)	VAS (4), McGill (1), RMQ (3), NRS (1)	No

Systematic review	Year	Patient population (number of studies)	Recruitment and setting (number of studies)	Duration of LBP (years): mean* (range)	Follow-up (months): mean* (range)	Subjects (n) mean* (range)	Age (years) mean* (range)	Outcomes	Data-pooling
Khadilkar et al.	2008	CLBP: nonspecific (3), acute + chronic LBP (1)	GP-based (1), Advertisement (1), Specialized Unit (1)	3.84 (1.4–6)	1.13 (after treatment–3 months)	133.5 (30–324)	42.88 (31–51)	VAS (3), McGill (1), ODI (1), others (2)	No

LBP: low back pain; SPLBP: specific low back pain; NSLBP: non-specific low back pain; SP: spine pain; EMG: electromyography; GP: general practitioner; includes all primary care-based physician care; Specialized unit: specialist, specialized unit at a hospital; VAS: visual analogue scale; RMQ: Roland Morris Disability Questionnaire; SF-36: Multi-purpose short-form health survey; McGill: McGill Pain Questionnaire; ODI: Oswestry Disability Index; FAB: fear avoidance beliefs.

*Values are calculated as the mean of the mean values reported in the primary trials; the range of the mean values are shown

Reporting of important prognostic factors in primary studies

The information reported for the domains and subdomains is summarized in Table 4. The data reported most often were data about socio-demographic status and the history of the current LBP. Information about the patient's general health status, social support, and work-related information was rarely reported.

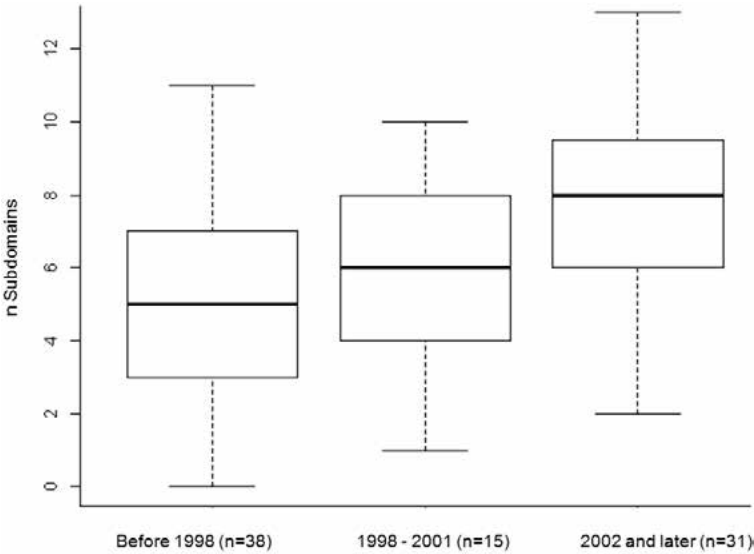
Table 4. Quantity of information in the prognostic subdomains in the 84 RCTs

Domain	Subdomain	Total	%
General patient characteristics	Sociodemographic information	80	95
	Social support	8	10
Baseline health status	Overall health	22	26
	Overall psychological health	47	56
	Previous LBP	33	39
Work-related factors	Work: psychosocial demands	1	1
	Work: physical demands	6	7
	Work history	39	46
	Work place attributes	3	4
Current LBP	Clinical history	67	80
	Disability related to the complaint	48	57
	Changes related to complaint over time	38	46
Clinical examination findings	Physical examination findings	25	30
	Definition of NSLBP diagnosis	56	67
	Change found during the physical exam	21	25
Interactions with work/society	Compensation issues related to LBP	36	43

LBP: low back pain; NSLBP: non-specific low back pain

Although statistically significant (p -value=0.01), the mean number of subdomains with reported information increased after 1998 by fewer than two subdomains (from a mean of 5.4 subdomains to 7.0 subdomains). In studies published after 2001, the median number of subdomains with reported information increased to 8 (of a possible total of 16 subdomains), reflecting a trend towards improved reporting of prognostic important information in recent years (Figure 2).

Figure 2. Reporting of prognostic important information



Median and inter quartile range (IQR) subdomains reported.

In 17 of the 84 studies (20%), information was reported for all six of the main domains (high SQR). Information was reported for five of the six main domains (moderate SQR) in 30 studies (36%) and for four or fewer domains (low SQR) in 37 studies (44%). The 27 studies investigating cognitive behavioural or educational therapy (termed CBT) provided information for more domains on average (high or moderate SQR for 82%) than studies of other interventions. There was poor reporting in the main domains in studies investigating acupuncture, injection therapy, antidepressants, and opioids (SQR poor in 72–100% of the RCTs) (Table 5).

Table 5. Summary of the Score for Quantity of reporting (SQR) types for the RCTs

	SQR high	SQR moderate	SQR low
All studies (n=84, 100%)	17 (20%)	30 (36%)	37 (44%)
Acupuncture (n=18, 21%)	0 (0%)	5 (28%)	13 (72%)
Antidepressants (n=9, 11%)	1 (5.5%)	1 (5.5%)	8 (89%)
Opioids (n=4, 5%)	0 (0%)	0 (0%)	4 (100%)
CBT (n=27, 32%)	11 (41%)	11 (41%)	5 (18%)
TENS: (n=5, 6%)	1 (20%)	1 (20%)	3 (60%)
EMG (n=4, 5%)	0 (0%)	2 (50%)	2 (50%)
Reflexology (n=1, 1%)	0 (0%)	1 (100%)	0 (0%)
Injection Therapy (n=11, 13%)	2 (18%)	0 (0%)	9 (82%)
Prolotherapy (n=5, 6%)	2 (40%)	0 (0%)	3 (60%)

CBT: cognitive behavioural therapy or educational therapy; TENS: transcutaneous electrical nerve stimulation; EMG: electromyography; Prolotherapy: Repeated injections of irritant solutions to strengthen lumbosacral ligaments; SQR: Score for quantity of reporting, scoring SQR high: information reported in one or more subdomains for all six main domains; SQR moderate: information reported in one or more subdomains for five main domains; SQR low: information reported in one or more subdomains for four or fewer main domains

Comparison of the classification systems for reporting prognostic factors using SQR and the Cochrane Collaboration Guidelines for baseline characteristics (CCG-baseline)

In the systematic reviews, the reporting of baseline characteristics was classified as “Can’t tell” in 17 of the 84 studies (20%). The CCG-baseline rating was “Similar” for 59 studies and “Not similar” for 8 studies, indicating that sufficient information for classification was available in most of the studies. The baseline characteristics were classified by the reviewers either as “Similar” or “Not similar” in almost two thirds of the studies with low SQRs (34 studies, 40%) (Table 6). There was thus moderate agreement between the two rating systems, i.e. SQR and CCG-baseline.

Of the 44 studies pooled for meta-analysis, the SQR was low in 22 studies (50%), and 8 (18%) of the studies were classified as “Can’t tell” according to the CGC-baseline system (Table 7). Five (11%) of the 44 pooled studies were classified as low SQR and “Can’t tell” according to the CGC-baseline system.

Table 6. Comparison of the Score for Quantity of Reporting (SQR) categories (high/moderate/low) for the 84 RCTs and the Cochrane Collaboration Guideline-baseline characteristics (CCG-baseline) categories (Similar/Not similar/Can’t tell)

All RCTs (n=84, 100%)	CCG-baseline: Similar or Not similar (n=67, 80%)	CCG-baseline: “Can’t tell” (n=17, 20%)
SQR high (n=17, 20%)	17 (20%)	0 (0%)
SQR moderate (n=20, 24%)	16 (19%)	4 (5%)
SQR low (n=47, 56%)	34 (40%)	13 (15%)

Table 7. Comparison of the Score for Quantity of Reporting (SQR) categories (high/moderate/low) for the 44 RCTs included in meta-analyses and the Cochrane Collaboration Guideline-baseline characteristics (CCG-baseline) categories (Similar/Not similar/Can't tell)

All RCTs (n=44, 100%)	CCG-baseline: Similar or Not similar (n=36, 81%)	CCG-baseline "Can't tell" (n=8, 19%)
SQR high (n=11, 25%)	9 (20%)	2 (5%)
SQR moderate (n=11, 25%)	10 (23%)	1 (2%)
SQR low (n=22, 50%)	17 (39%)	5 (11%)

References

- Schulz, K.F., Altman, D.G., Moher, D., et al., CONSORT 2010 statement: updated guidelines for reporting parallel group randomized trials. *Ann Intern Med*, 2010. 152(11): p. 726-32.
- Hayden, J.A., Chou, R., Hogg Johnson, S., et al., Systematic reviews of low back pain prognosis had variable methods and results: guidance for future prognosis reviews. *J Clin Epidemiol*, 2009. 62(8): p. 781-796.e1.
- Pincus, T., Santos, R., Breen, A., et al., A review and proposal for a core set of factors for prospective cohorts in low back pain: a consensus statement. *Arthritis Rheum*, 2008. 59(1): p. 14-24.
- Nour, K., Laforest, S., Gauvin, L., et al., Behavior change following a self-management intervention for housebound older adults with arthritis: an experimental study. *International Journal of Behavioral Nutrition and Physical Activity*, 2006. 3(1): p. 12.
- Laupacis, A., Wells, G., and Richardson, W., Users' Guides to the Medical Literature: V. how to use an article about prognosis. *JAMA: The Journal of the American Medical Association*, 1994. 272(3): p. 234-237.
- Balagu, F., Mannion, A., Pellis, F., et al., Non-specific low back pain. *Lancet*, 2012. 379(9814): p. 482-491.
- Hayden, J.A., Dunn, K.M., van der Windt, D.A., et al., What is the prognosis of back pain? *Best Practice & Research Clinical Rheumatology*, 2010. 24(2): p. 167-179.
- van Tulder, M.W., Assendelft, W.J., Koes, B.W., et al., Method guidelines for systematic reviews in the Cochrane Collaboration Back Review Group for Spinal Disorders. *Spine (Philadelphia, Pa. 1976)*, 1997. 22(20): p. 2323-2330.
- van Tulder, M., Furlan, A., Bombardier, C., et al., Updated method guidelines for systematic reviews in the cochrane collaboration back review group. *Spine (Phila Pa 1976)*, 2003. 28(12): p. 1290-9.
- Furlan, A.D., Pennick, V., Bombardier, C., et al., 2009 updated method guidelines for systematic reviews in the Cochrane Back Review Group. *Spine (Phila Pa 1976)*, 2009. 34(18): p. 1929-41.
- Liberati, A., Altman, D.G., Tetzlaff, J., et al., The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. *BMJ*, 2009. 339.
- Manchikanti, L., Singh, V., Cash, K.A., et al., A comparative effectiveness evaluation of percutaneous adhesiolysis and epidural steroid injections in managing lumbar post surgery syndrome: a randomized, equivalence controlled trial. *Pain Physician*, 2009. 12(6): p. E355-68.
- Trigkilidas, D., Acupuncture therapy for chronic lower back pain: a systematic review. *Ann R Coll Surg Engl*, 2010. 92(7): p. 595-8.
- Henschke, N., Ostelo, R.W., van Tulder, M.W., et al., Behavioural treatment for chronic low-back pain. *Cochrane Database Syst Rev*, 2010(7): p. CD002014.
- Hayden, J.A., Côté, P., and Bombardier, C., Evaluation of the Quality of Prognosis Studies in Systematic Reviews. *Annals of Internal Medicine*, 2006. 144(6): p. 427-437.
- Deeks, J.J., Dinnes, J., D'Amico, R., et al., Evaluating non-randomised intervention studies. *Health Technol Assess*, 2003. 7(27): p. iii-x, 1-173.
- Research, C.o.S.f.S.R.o.C.E. and Medicine, I.o., Finding What Works in Health Care: Standards for Systematic Reviews, ed. J. Eden, Levit, L., Berg, A., et al. 2011: The National Academies Press.
- Higgins, J.P.T. and Green, S., *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0* The Cochrane Collaboration, 2011.
- Dissemination, C.f.R.a., CRD's guidance for undertaking reviews in health care. 2009. <http://www.york.ac.uk/inst/crd/SysRev/!SSL/!WebHelp/SysRev3.htm>.
- Egger, M., Ebrahim, S., and Smith, G.D., Where now for meta-analysis? *Int J Epidemiol*, 2002. 31(1): p. 1-5.
- Chou, R., Atlas, S.J., Loeser, J.D., et al., Guideline Warfare Over Interventional Therapies for Low Back Pain: Can We Raise the Level of Discourse? *The Journal of Pain*, 2011. 12(8): p. 833-839.
- Jane wit, D., Horwitz, R., and Concato, J., Variation in results from randomized, controlled trials: stochastic or systematic? *J Clin Epidemiol*, 2010. 63(1): p. 56-63.

23. Fourney, D., Andersson, G., Arnold, P., et al., Chronic low back pain: a heterogeneous condition with challenges for an evidence-based approach. *Spine (Philadelphia, Pa. 1976)*, 2011. 36(21 Suppl): p. S1-S9.
24. Ostelo, R., Croft, P., van der Weijden, T., et al., Challenges in using evidence to inform your clinical practice in low back pain. *Best Practice & Research Clinical Rheumatology*, 2010. 24(2): p. 281-289.
25. Freeman, M.D., Clinical practice guidelines versus systematic reviews; which serve as the best basis for evidence-based spine medicine? *The Spine Journal*, 2010. 10(6): p. 512-513.
26. Willis, B. and Quigley, M., The assessment of the quality of reporting of meta-analyses in diagnostic research: a systematic review. *BMC Medical Research Methodology*, 2011. 11(1): p. 163.
27. Pambrun, E., Bouteloup, V., Thibaut, R., et al., On the validity of meta-analyses: exhaustivity must be warranted, exclusion of duplicate patients too. *J Clin Epidemiol*, 2010. 63(3): p. 342-343.
28. Kirkham, J.J., Dwan, K.M., Altman, D.G., et al., The impact of outcome reporting bias in randomised controlled trials on a cohort of systematic reviews. *BMJ*, 2010. 340.
29. Bischoff-Ferrari, H.A., Willett, W.C., Orav, E.J., et al., A pooled analysis of vitamin D dose requirements for fracture prevention. *N Engl J Med*, 2012. 367(1): p. 40-9.
30. Dionne, C., Dunn, K., Croft, P., et al., A consensus approach toward the standardization of back pain definitions for use in prevalence studies. *Spine (Philadelphia, Pa. 1976)*, 2008. 33(1): p. 95-103.
31. Tricco, A., Tetzlaff, J., Pham, B., et al., Non-Cochrane vs. Cochrane reviews were twice as likely to have positive conclusion statements: cross-sectional study. *J Clin Epidemiol*, 2009. 62(4): p. 380-386.e1.
32. Furlan, A.D., van Tulder, M.W., Cherkwin, D.C., et al., Acupuncture and dry-needling for low back pain. *Cochrane Database Syst Rev*, 2005(1): p. CD001351.
33. Urquhart, D.M., Hoving, J.L., Assendelft, W.W., et al., Antidepressants for non-specific low back pain. *Cochrane Database Syst Rev*, 2008(1): p. CD001703.
34. Staal, J.B., de Bie, R., de Vet, H.C., et al., Injection therapy for subacute and chronic low-back pain. *Cochrane Database Syst Rev*, 2008(3): p. CD001824.
35. Deshpande, A., Furlan, A., Mailis-Gagnon, A., et al., Opioids for chronic low-back pain. *Cochrane Database Syst Rev*, 2007(3): p. CD004959.
36. Dagenais, S., Yelland, M.J., Del Mar, C., et al., Prolotherapy injections for chronic low-back pain. *Cochrane Database Syst Rev*, 2007(2): p. CD004059.
37. Khadilkar, A., Odebiyi, D.O., Brosseau, L., et al., Transcutaneous electrical nerve stimulation (TENS) versus placebo for chronic low-back pain. *Cochrane Database Syst Rev*, 2008(4): p. CD003008.

CHAPTER 3

The role of fear avoidance beliefs as a prognostic factor for outcome in patients with non-specific low back pain:

A systematic review

Maria M. Wertli, M.D, Eva Rasmussen-Barr, RPT, Ph.D., Sherri Weiser, Ph.D., Lucas M. Bachmann M.D., Ph.D., Florian Brunner M.D., Ph.D.

The Spine Journal, accepted for publication, available online October 18., 2013

Abstract

Background context: Psychological factors including fear avoidance beliefs are believed to influence the development of chronic low back pain (LBP).

Purpose: The purpose of this study was to determine the prognostic importance of fear avoidance beliefs as assessed by the Fear Avoidance Beliefs Questionnaire (FABQ) and the Tampa Scale of Kinesiophobia for clinically relevant outcomes in patients with nonspecific LBP.

Design/setting: The design of this study was a systematic review of observational studies.

Methods: In October 2011, the following databases were searched: BIOSIS, CINAHL, Cochrane Library, Embase, OTSeeker, PeDRO, PsycInfo, PubMed/Medline, Scopus, and Web of Science. To ensure the completeness of the search, a hand search and a search of bibliographies was conducted and all relevant references included. A total of 2,031 references were retrieved, leaving 566 references after the removal of duplicates. For 53 references, the full-text was assessed and, finally, 21 studies were included in the analysis.

Results: The most convincing evidence was found supporting fear avoidance beliefs to be a prognostic factor for work-related outcomes in patients with subacute LBP (i.e., 4 weeks–3 months of LBP). Four cohort studies, conducted by disability insurance companies in the United States, Canada, and Belgium, included 258 to 1,068 patients mostly with nonspecific LBP. These researchers found an increased risk for work-related outcomes (not returning to work, sick days) with elevated FABQ scores. The odds ratio (OR) ranged from 1.05 (95% confidence interval [CI] 1.02–1.09) to 4.64 (95% CI, 1.57–13.71). The highest OR was found when applying a high cutoff for FABQ Work subscale scores. This may indicate that the use of cutoff values increases the likelihood of positive findings. This issue requires further study. Fear avoidance beliefs in very acute LBP (<2 weeks) and chronic LBP (>3 months) was mostly not predictive.

Conclusions: Evidence suggests that fear avoidance beliefs are prognostic for poor outcome in subacute LBP, and thus early treatment, including interventions to reduce fear avoidance beliefs, may avoid delayed recovery and chronicity.

Introduction

Low back pain (LBP) is a pain not attributed to a recognizable pathology. The lifetime prevalence for LBP acute low back pain exceeds 80% [1]. While the overall prognosis is benign, 10-15% of these patients develop chronic low back pain (LBP). This small percentage accounts for three-quarters of the total direct and indirect costs of medical care and lost productivity associated with LBP [2]. There is a consensus in the literature to use a “wait-and-see” strategy in acute low back pain to avoid overtreatment [3]. However, pain that persists beyond the acute stage is an indication for the development of chronic low back pain, a condition where complete recovery and return to 100% function are often difficult to achieve [4]. Current research aims at identifying sub-groups at risk for delayed recovery in patients with sub-acute LBP (>4 weeks) in order to optimize treatment. Timely initiated and targeted multifaceted treatments in patients at risk for delayed recovery have shown health care cost reductions and tend to facilitate recovery[5].

Psychological factors are believed to influence the development of chronic low back pain. The Fear Avoidance Model is widely used to explain how psychological factors affect the experience of pain and the development of chronic pain and disability [6]. It is theorized that for some individuals with LBP, negative beliefs about pain and/or negative illness information leads to a catastrophizing response in which the worst possible outcome is imaged. This leads to fear of activity and avoidance which in turn causes disuse and resultant distress, reinforcing the original negative appraisal in a deleterious cycle [6]. The Fear Avoidance Model suggests that patients without fear avoidance beliefs are more likely to confront pain problems and are more active in the coping process. This type of “good” coping has been used to develop interventions for those with high fear avoidance beliefs.

While the Fear Avoidance Model is generally accepted, it is a matter of debate as to how and when to best assess fear avoidance beliefs in clinical practice. Current treatment guidelines for low back pain recommend the timely identification and initiation of multidisciplinary treatment for other psychological factors (e.g. depression, distress, job dissatisfaction) associated with increased risk for delayed recovery [3, 7, 8]. Whether and how fear avoidance beliefs specifically should be assessed remains unclear. Many questionnaires have been developed to identify fear avoidance beliefs. The two most commonly used are the Fear Avoidance Questionnaire (FABQ [9]) and the Tampa Scale of Kinesiophobia (TSK [10, 11]). Their usefulness for detecting fear avoidance beliefs that influence the transition from acute to chronic LBP is not clearly determined and has been debated in the literature.

To date, the role of fear avoidance beliefs identified by FABQ or TSK as prognostic factors for LBP has not been reviewed systematically. The aim of this systematic review is twofold. First, we review the existing literature on the role of fear avoidance beliefs

as a prognostic factor in acute, sub-acute, and chronic LBP. Second, we analyze the available data in terms of an optimal cut-off value for the FABQ and the TSK.

Methods

These systematic reviews followed the recommendation of the MOOSE statement (Figure 1) on conducting systematic reviews of observational studies [12].

Literature search

We identified all observational studies meeting our eligibility criteria (defined in detail below) published between 1990 and October 2011. The following databases were searched in October 2011: BIOSIS, CINAHL, Cochrane Library, Embase, OTSeeker, PeD-RO, PsycInfo, PubMed/Medline, Scopus, and Web of Science. The search was conducted with the help of an experienced librarian. Search terms included fear and avoidance as Medical Subject Headings (MeSH) as well as other subject headings and different combinations. Three detailed search strategies are depicted in Appendix 1. To ensure the completeness of the literature search, one reviewer (MW) conducted an electronic hand search of the six most often retrieved journals and added all potentially eligible references not retrieved by the systematic search. In addition, bibliographies of included studies relevant to the research question were searched and potentially eligible references were included then in the full text review (inclusion and exclusion criteria applied).

Eligibility criteria

All cohort studies were considered eligible for inclusion in this investigation that met the following criteria: they reported research concerning patients seeking care for LBP; they demonstrated at least moderate study quality; they investigated the prognostic value of the two most often used questionnaires - the Fear Avoidance Questionnaire (FABQ) and the Tampa Scale for Kinesiophobia (TSK); and they were published between 1990 and October 2011. We focused on cohort studies that included at least 300 subjects with a minimal follow-up of three months because of a concern about sample size. Assuming a baseline risk of 20% for chronicity following a bout of acute LBP [1]; a sample size of 316 patients in a two level exposure study (fear avoidance beliefs high vs. low) would generate a relative risk (RR) of 1.75 for the outcome recovery at three months ($\alpha = 0.05$, $\beta = 0.2$) [13].

However, the inclusion of cohorts of more than 300 patients would have included almost only cohorts based on studies by insurance companies, and therefore we believe the results would be less generalizable. Because of this concern, we also consid-

ered studies with between 100 and 300 subjects. In order to allow for a comprehensive overview of the current knowledge in the field, we included prospective and retrospective cohorts as well as secondary analyses of RCTs. Excluded were studies with fewer than 100 patients, with follow-up of less than three months, or reports from conference proceedings. No limits for the study setting or language of the publication were applied.

Study selection, data extraction, and synthesis

The bibliographic details of all retrieved articles were stored in an Endnote file. Two reviewers (MW and ERB) independently screened all 569 references by title and abstract. Full-text was reviewed by both reviewers independently (MW and ERB) in all studies meeting the pre-defined eligibility criteria ($n = 53$). Disagreements were discussed and resolved by consensus or by third party arbitration (SW). Alternative researchers with specific language proficiencies were used for non-English language references. In the case of several publications for the same cohort without change in outcome or follow-up duration, the most recent publication was chosen and missing information from the previous publication added.

Quality assessment

The quality of each study was assessed using the Scottish Intercollegiate Guidelines Network (SIGN) Methodology checklist for cohort studies [14]. All information needed to describe the study population and allowing for operationalization of outcome measures was extracted. To assess the baseline characteristics of the study population, important prognostic factors for the course of LBP were extracted and assigned into the 16 domains proposed by Hayden et al. [15]. The following prognostic information was extracted: general characteristics, social environment, overall health status, overall psychological health, previous LBP, work-psychosocial demands, work-physical demands, work-history and attributes, disability related to LBP, time change of LBP, physical examination findings, change in physical examination, diagnosis of LBP, and compensation related to LBP. The number of domains reported and statistical adjustment done for those domains is reported for each study. Additional measured psychological risk factors were extracted and studies were sub-categorized into the following: 1, investigating fear avoidance beliefs only; 2, measuring fear avoidance beliefs in addition to one or two additional psychological risk factors (e.g. depression, catastrophizing); 3, or investigating more than three psychological constructs.

Based on this information, the risk of bias and quality of each study was rated overall according to the SIGN recommendation into high, moderate, or low quality. The ratings were:

- High quality (++): most of the criteria fulfilled. If not fulfilled, the conclusions of the study are very unlikely to alter.
- Moderate quality (+): some criteria fulfilled. Criteria not adequately described are unlikely to alter the conclusions.
- Low quality (-): few or no criteria fulfilled. The conclusions are likely to alter.

As recommended, studies rated by both reviewers as low quality were excluded from further analysis.

Outcome definition and operationalization of fear avoidance beliefs as a prognostic factor

All investigated outcomes were extracted and categorized into work-related (e.g. sick days, return to work) and non-work-related outcomes (e.g. pain, perceived disability). Each method of outcomes was appraised and operationalized (e.g. perceived disability measured by Oswestry Disability Index (ODI) or by Roland Morris Disability Questionnaire (RMQ)).

The term “prognostic factor” is used to describe factors that influence or predict a course or outcome of LBP. The prognostic value of fear avoidance beliefs is based on the statistical method used in primary studies and reported accordingly. “Non-prognostic” refers to the fact that fear avoidance beliefs did not influence the outcome in the univariate analysis and was not included in the final model.

Psychometric properties and description of the investigated questionnaires

The FABQ [9] is a 16 item questionnaire, each item scored zero to six. High levels indicate increased levels of fear avoidance beliefs. Two subscales exist: a seven-item work subscale (FABQ-W, range 0 to 42) and a four-item physical activity subscale (FABQ-P, range 0 to 24). FABQ and the two sub-scales have been shown to be reliable and valid for the measurement of fear avoidance beliefs. The Cronbach’s alpha for the FABQ-P was 0.75, test-retest reliability $r=0.64$. For the FABQ-W, the Cronbach’s alpha was 0.82, test-retest $r=0.80$ [16].

The Tampa Scale for Kinesiophobia (TSK) [10, 17] is a valid and reliable 17-item questionnaire [16]. Each item is measured on a four-point Likert scale (1= strongly disagree, to 4= strongly agree). The scale ranges from 17 (no fear) to 68 (strong fear of re-injury). Two subscales are used consisting of a “harm factor” (items 3,5,6,9,11,15 with a range of 6-24) and an “activity-avoidance” factor (items 1,2,7,10,13,14,17 with a range of 7-28) [16]. The “harm factor” reflects the beliefs that something is seriously wrong with the body, while the “activity-avoidance” factor indicates that avoiding exercise or activity might prevent an increase in pain.

While the two scales measure different concepts (i.e. FABQ measures fear of pain caused by physical activity; TSK measures fear of movement and re-injury), a moderate

overlap has been shown [16, 18, 19]. The correlation between TSK and FABQ ranged from 0.53 (FABQ-P) and 0.76 (FABQ-W) [19] to 0.39 (FABQ-P) and 0.33 (FABQ-W) [16].

Statistical analysis

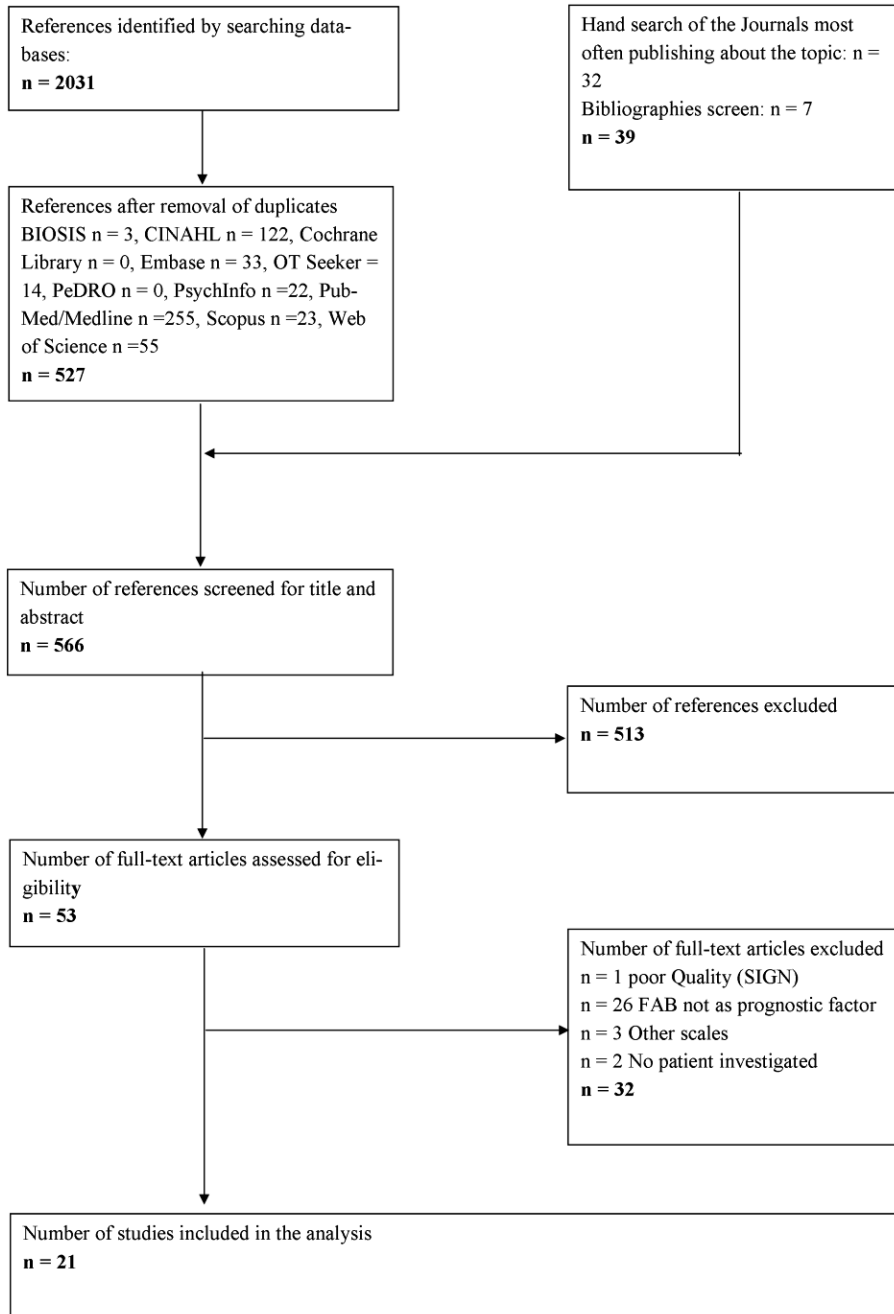
Due to heterogeneous study populations, measurements and scales used as well as outcomes investigated, only descriptive statistics (ranges) were used to summarize findings across all cohort studies for baseline fear avoidance mean values. We calculated mean number of reported domains within prognostic and non-prognostic categories in acute, acute-sub-acute, sub-acute, chronic, and acute – chronic LBP cohorts. Forest plots were generated using the R statistical software for Windows (<http://www.R-project.org/>) [20]. Forest plots were based on the values (odds ratio, 95% confidence interval) given in the study reports.

Results

Study selection

The search and inclusion process is summarized in Figure 1. Out of 2,032 records, 53 were reviewed in full text. The full text assessment utilizing the inclusion and exclusion criteria resulted in the exclusion of 32 studies. The main reasons for exclusion were not investigating fear avoidance beliefs as a prognostic factor, use of questionnaires other than FABQ and TSK, follow-up less than three months, and poor study quality. Two studies were excluded because they investigated the prevalence of pain in the general population. In total 21, studies were included in the analysis.

Figure 1.



Study characteristics

Study designs included prospective cohort studies (n=15, 71%), retrospective cohort studies (n=1, 5%), secondary analyses of randomized controlled trials (n=4, 19%), and cohort studies following an RCT (n=1, 5%). Follow-up duration was mostly one year (three months to two years). More than half of the studies included fewer than 300 patients (58%). These studies were mostly conducted in the primary care setting. Studies including more than 300 patients were mainly based on insurance company data (n=5, 24%) and only one study was conducted in a primary care setting. SIGN quality was good in the three studies (14%) conducted in a primary care setting. A detailed summary of the assessment of the SIGN quality is reported in Appendix 2 and an overall rating is given for each study.

The results will be presented according to the disease duration provided in the studies: acute (one day to six weeks, five studies), acute to sub-acute (one day to three months, four studies), sub-acute (six weeks to three months, two studies), chronic (more than three months, five studies), and acute to chronic (no sub-grouping for disease duration, eight studies).

Acute low back pain

Baseline characteristics are summarized in Table 1. Five prospective cohorts and one secondary analysis of an RCT included between 123 and 1,885 patients with specific and non-specific LBP (n=4 [21-24]) or LBP only (n=2 [25, 26]). The study setting was mainly primary care [21, 23-26]. One study included patients receiving workers' compensation payments [22]. The duration of the LBP episode varied between five and 28 days. The follow-up was mostly one year. Perceived disability was the most common outcome measured (Table 2). One study investigated return to work as outcome.

Table 1. Baseline characteristics of studies investigating acute to sub-acute low back pain

Acute low back pain										
Study	Year	Design	Setting	Population	Diagnostic criteria	N (female)	Age (years), mean (SD)	Disease duration: Days: mean (SD)	First episode (%)	Follow-up duration: mean days (SD)
Sieben[25]	2005	Prospective Cohort Study	Primary Care	All Patient, Netherlands	NSLBP	222 (97)	Range: <24 (5.9%); > 55 (12.6%); 45-54 (38.7%)	5, IQR 2-10	25.2	684, range 396-1116
Grotle[21]	2007	Prospective Cohort Study	Primary Care	All Patient, Norway, same population as 2006	Acute SLBP / NSLBP	123 (68)	37.9 (10.1)	8.1 (6.6)	38	360 (N.R.)
Grotle[23]	2006	Prospective Cohort Study	Primary Care	All Patient, Norway, comparison acute vs. chronic	Acute SLBP / NSLBP	123 (68)	37.9 (10.1)	8.1 (6.6)	N.R.	360 (N.R.)
Turner [22]	2008	Prospective Cohort Study	Disability Insurance	Workers on sick leave and receiving replacement benefits, Washington Workers' Compensation Disability Risk Identification Study Cohort, U.S.A	Acute SLBP / NSLBP	1885 (622)	Range <24>55	N.R. inclusion criteria 83% <14 days between injury and claim receipt	N.R.	360 (N.R.)
Swinkels[26]	2006	Prospective Cohort Study	Primary Care	Netherlands	NSLBP	555 (233)	42.4 (11.3)	<14: 352 (63.4%), <28: 203 (36.6)	22.5	180 (N.R.)
Hancock [24]	2009	Secondary Analysis of RCT	Primary Care	N.R., Sydney, Australia	Acute SLBP / NSLBP	240 (105)	40.7 (15.6)	9.13 (9.3)	93.6	84 (N.R.)

Acute and sub-acute low back pain

Study	Year	Design	Setting	Population	Diagnostic criteria	N (female)	Age (years), mean (SD)	Disease duration: Days: mean (SD)	First episode (%)	Follow-up duration: mean days (SD)
van der Windt[27]	2007	Prospective Cohort Study	Primary Care	All Patient, Netherlands	SLBP / NSLBP	171 (81)	42 (12)	0-14: 105 (61.4%), 21-42: 49 (28.7%), 42-91: 17 (9.9%)	33	90 (N.R.)
Truchon[28]	2010	Prospective Cohort Study	Disability Insurance	Workers on sick leave and receiving replacement benefits, Quebec, Canada	NSLBP	439 (178)	38 (10)	63, range 28-83	N.R.	360 (N.R.)
Turner [29]	2006	Prospective Cohort Study	Disability Insurance	Workers on sick leave and receiving replacement benefits, Washington Workers' Compensation Disability Risk Identification Study Cohort, U.S.A	SLBP / NSLBP	1068 (328)	39.2 (11.1)	After claim submission: 21, range 10-58	N.R.	180 (N.R.)
George [30]	2008	Secondary Analysis of RCT	Outpatient PT	All Patient included in 2 RCT for Manual Therapy and PT, U.S.A	SLBP / NSLBP	160 (71)	34.6 (10-5)	26.5, IQR 11-55	N.R.	180 (N.R.)

Sub-acute low back pain

Study	Year	Design	Setting	Population	Diagnostic criteria	N (female)	Age (years), mean (SD)	Disease duration: Days: mean (SD)	First episode (%)	Follow-up duration: mean days (SD)
Du Bois [31]	2009	Prospective Cohort Study	Disability Insurance	Workers on sick leave and receiving replacement benefits, Belgian compulsory health insurance system	NSLBP	346 (163)	41, range 16-64	After start sick-leave: 28-42. > 84 days: 98 (28%)	39	90 (N.R.)
Soucy[32]	2006	Prospective Cohort Study	Disability Insurance	Workers on sick leave and receiving replacement benefits, Quebec, Canada	NSLBP	258 (112)	39 (9.6)	62 (10)	N.R.	180 (N.R.)

NSLBP, non-specific low back pain; SLBP, specific low back pain; RCT, randomized controlled trial; N.R., not reported; IQR, interquartile range; PT, physical therapy

Three studies found FABQ [21] or TSK [23, 25] non-prognostic for non-work-related perceived measurements, two for work-related outcomes (RTW [22] and sick days [24]) independent of whether cut-off values were applied or not (Table 2). The average disease duration in these studies was 5 to 9 days. Only one study found fear avoidance beliefs to be prognostic for perceived disability after six months [26]. In this study, patients had a longer disease duration (in 36% disease duration of two and more weeks). The strength of this study was a re-evaluation of fear avoidance beliefs after 6 weeks. TSK and both sub-scales were prognostic at baseline and at 6 weeks. The limitations of the study were that perceived disability as an outcome was not measured in 40% of the patients, no other psychological risk factors were investigated, and the follow-up was 6 months shorter than the other cohort studies.

Acute and sub-acute LBP

Three prospective cohort studies (n=1 primary care [27], n=2 disability insurance information [28, 29]) and one secondary analysis in an RCT (PT setting [30]) included between 160 and 1,068 patients with acute and sub-acute LBP (Table 1). Disease duration ranged from 1 to 83 days. Patients were followed for 3 months to 1 year. Outcome measures included non-work-related perceived measurements [27, 30] and work-related outcomes [28, 29].

In the two cohort studies conducted with information provided by disability insurance companies (Canada and U.S.A.), FABQ-W was prognostic for RTW (Table 2) [28, 29]. Turner et al. found in a sub-acute LBP population (21 days after claim submission) that FABQ-W was prognostic for RTW. This is in contrast to patients with acute LBP in the same cohort [22] (83% with claim receipt in less than 14 days after injury). In a primary care setting, FABQ-P failed to be prognostic for non-work-related perceived measurements [27]. In this study, a high percentage of acute patients was included (61.4% patients < two weeks of LBP vs. mean disease duration 63 days [28] and 21 days [29] respectively) and the follow-up was 3 months shorter compared to others, which might explain the non-prognostic results. A secondary analysis of an RCT in a PT setting [30] was used to calculate various cut-off values for the FABQ-W. In this study, FABQ-P was non-prognostic for perceived disability at follow-up.

All studies investigating acute and sub-acute patients assessed half or fewer of the 16 important prognostic domains and adjusted on average for four or five of them as potential confounders. In comparison, in acute LBP studies all but one study assessed ten to fourteen important prognostic domains and adjusted on average for more potentially confounding variables. This indicates a less stringent methodology in studies investigating patients without sub-categorization of disease duration and could indicate a potential bias.

Sub-acute low back pain

Two prospective cohort studies investigated 258 and 346 sub-acute LBP patients with a disease duration of 41 [31] to 39 days [32] respectively (Table 1). Both investigated RTW as an outcome. They reported more than half of the 16 important prognostic domains (13 and 9 domains respectively).

Both studies found the FABQ total score, the TSK single item question, and the FABQ-W (Table 2) to be prognostic for work-related outcomes after three and six months. None of the studies investigated cut-off values.

Table 2. Results in patients with acute to sub-acute low back pain

Acute low back pain											
Study	Design	SIGN	PI (adjust)	QA	Cut-off used	MA?	Measure reported	Value (95% CI)	Outcome	Predictive	Authors Conclusion
Sieben 2005	P	++	10 (5)	TSK Avoidance Sub-scale Harm Sub-scale	No	Yes	N.R.	N.R.	Poor Outcome: CPG Grade IV	Non-prognostic	We aimed to test the longitudinal validity of the FAM using rigorous methodological, psychological and statistical methods. The results did not support the FAM in explaining the transition from acute LBP to long-term outcome.
Grotle 2007	P	+	12 (11)	FABQ-P	Low: 0-16 / high: 17-24	U.L.	OR U.L.	1.45 (0.44, 4.81)	Non-recovery (RMQ↓ ≤ 4 points)	Non-prognostic	Our data showed that FAB as measured by the FABQ, was not a significant prognostic factor (not at 3 months (Grotle et al., 2005) or 12 months).
				FABQ-W	Low: 0-20 / high: 21-42			1.66 (0.49, 5.68)		Non-prognostic	
Grotle 2006	P	+	12 (11)	FABQ-P	N.R.	Yes	Beta	0.1 (N.R.)	Perceived Disability (ODI)	Non-prognostic	No association in any of the analysis for FABQ-P.
				FABQ-W	N.R.	Yes		0.15 (N.R.)		Non-prognostic	FAB weaker predictor than emotional distress of future pain and disability. FAB and distress explained only 3% to 7% of the variation of 27%.
Turner 2008	P	+	14 (13)	FABQ-W Average of 2	Very high: 6	Yes	OR bivariate	1.71 (0.88, 3.3)	Not RTW	Non-prognostic	Bivariate Mental health, FAB and catastrophizing were

Acute low back pain

Study	Design	SIGN	PI (adjust)	QA	Cut-off used	MA?	Measure reported	Value (95% CI)	Outcome	Predictive	Authors Conclusion
Swinkels 2006 [26]	P	+	11 (3)	Items (scale 0-6)	High: 3-4,9	Yes	Log.Reg	1.67 (0.89, 3.13)		Non- prognostic	strong predictors of work disability. Multivariate they were only statistically significant when RMQ was excluded from the analysis.
				TSK	No	Yes	Beta	0.23 (N.R.)	RMQ	Prognostic	Fear of movement/(re)injury was the most powerful predictor of future disability, even more powerful than baseline pain intensity.
				TSK-activity avoidance	No	Yes		0.16 (N.R.)		Prognostic	Secondary analyses using the 2 subscales of the TSK as an independent variable yielded comparable results
Hancock 2009	Sec	++	7 (5)	TSK-fear of harm	No	Yes		0.23 (N.R.)		Prognostic	
				FABQ-P	Low ≤14	Yes	N.R.	N.R.	Days to Recovery	Non- prognostic	Univariate: significant associations with recovery rate for catastrophizing and FAB (FABQ-work subscale). In the multivariate model no predictor.
				FABQ-W	Low ≤14	Yes	N.R.	N.R.		Non- prognostic	

Acute and sub-acute low back pain

Study	Design	SIGN	PI (adjust)	QA	Cut-off used	MA?	Measure reported	Value (95% CI)	Outcome	Predictive	Main finding
van der Windt 2007	P	+	8 (5)	FABQ-P	50-75 (<50: low)	Yes	OR	0.77 (0.25, 2.34)	RMQ Decrease < 30%	Non-prognostic	Although often suggested as a potentially important predictor, associations between FAB and persistent pain were not strong.
					>75	Yes	OR	1.73 (0.6, 4.99)		Non-prognostic	
					50-75 (<50: low)	Yes	OR	1.45 (0.67, 3.14)	Persistent Symptoms	Non-prognostic	
					>75	Yes	OR	1.58 (0.7, 3.53)		Non-prognostic	
Truchon 2010	P	+	7 (5)	FABQ-W	No	Yes	OR	3.13 (N.R.)	Not RTW	Prognostic	Cognitive appraisal of LBP (OR 2.57) and fear of work (OR 3.13) were significantly predictive of absence from work.
Turner 2006	P	+	7 (4)	FABQ-W Average of 2 Items (scale 0-6)	High: 5-6	Yes	OR	4.64 (1.57, 13.71)	Not RTW	Prognostic	After controlling for demographics, pain, disability, and other psychosocial variables, workers with high FABQ-W had 5 times higher odds for work disability at 6 months.
					Low-moderate: 3-4.9	Yes	OR	2.96 (0.98, 8.9)		Prognostic	
George 2008	Sec	+	6 (5)	FABQ-P	>14	Yes	N.R.	N.R.	Perceived Disability (ODI)	Non-prognostic	FABQ-P not associated with ODI score and change of ODI at 6-month, as likelihood of experiencing meaningful clinical change (MCID) at

Acute and sub-acute low back pain

Study	Design	SIGN	PI (adjust)	QA	Cut-off used	MA?	Measure reported	Value (95% CI)	Outcome	Predictive	Main finding
				FABQ-W	>20	Yes	+ LR	2.35 (2.2, 2.5)	Perceived Disability (ODI)	Prognostic	6-month The FABQ-W was able to predict 6-month disability outcomes; however, it was not a definitive predictive tool and it should be used with other examination findings in clinical settings.
					> 25	Yes	- LR	0.3 (0.15, 0.45)			
							+ LR	2.67 (2.52, 2.82)			
							- LR	0.65 (0.5, 0.8)			
					> 30	Yes	+ LR	2.5 (2.35, 2.65)			
							- LR	0.82 (0.67,0.97)			

Sub-acute low back pain

Study	Design	SIGN	PI (adjust)	QA	Cut-off used	MA?	Measure reported	Value (95% CI)	Outcome	Predictive	Authors Conclusion
Du Bois 2008	P	+	13 (5)	FABQ	No	Cox Regr.	OR	1.05 (1.02, 1.09)	Not RTW	Prognostic	RF for sickness absence > 3 months were ODI (OR 1.04; CI: 1.02–1.06), FABQ score, blue collar worker (OR 2.18; CI: 1.21–3.92), LBP < 12 weeks (OR: 0.32; CI: 0.17–0.64), pain behaviour (OR: 1.72; CI: 1.25–2.39).
Soucy 2006	P	+	9 (4)	TSK single Item	N.R.	Cox Regr.	OR	1.39 (1.1, 1.76)		Prognostic	TSK single item question is predictive.
				FABQ-W	No	Yes	OR	0.38 (0.25, 0.58)	RTW	Prognostic	Being a man (OR = 0.43), stressed-out at work (OR = 0.44), and worried about work (OR = 0.38) increased the risk of chronic disability. FABQ-W best predictors of chronic disability.

P, prospective cohort study; R, retrospective cohort study; Sec, secondary analysis of RCTs; P RCT, prospective cohort after RCTs; QA, questionnaire used; FAB, fear avoidance beliefs; FABQ, fear avoidance questionnaire; FABQ-P: physical activity sub-scale of the FABQ; FABQ-W, work sub-scale of the FABQ; TSK, tampa scale of kinesiophobia; LBP, low back pain; ODI, Oswestry disability index; RMQ, Roland Morris questionnaire; RTW, return to work; CPG, chronic pain grade (von Korf) questionnaire; PI (adj.): reporting of prognostic indicators of the 16domains (statistical adjustment for PI at baseline); SIGN, SIGN quality rating: ++ high quality, + moderate quality; OR, Odds Ratio; MA, multiple regression analysis; IQR, interquartile range; Cox. Regr., cox regression analysis; Log. Regr., logistic regression; LR: Likelihood ratio + / -; U.L., univariate logistic regression

Chronic low back pain

In total three prospective cohort studies [23, 33, 34], one retrospective cohort study [35], and one secondary analysis of an RCT [36] included between 184 to 681 chronic LBP patients (Table 3). A sub-set of chronic patients from a larger study [23] included 50 patients only. Disease duration ranged from 90 to 580 days. Two studies investigated work-related outcomes [35, 36] and three non-work-related perceived measures [23, 33, 34]. Patients were followed for six to twelve months.

In three prospective cohort studies [23, 33, 34], fear avoidance beliefs was non-prognostic for non-work-related perceived measurements (Table 4). In a retrospective cohort [35], FABQ was found to be prognostic for RTW. This study has several limitations and a potentially biased finding cannot be excluded. Fewer patients were included, the follow up duration was shorter when compared to the non-predictive studies, and fewer prognostic factors were investigated. A secondary analysis of an RCT including 628 patients [36] and found fear avoidance beliefs not to be prognostic for time to RTW.

Table 3. Baseline characteristics in patients with chronic and not sub-categorized low back pain

Chronic low back pain										
Study	Year	Design	Setting	Population	Diagnostic criteria	N (female)	Age (years), mean (SD)	Disease duration: Days: mean (SD)	First episode (%)	Follow-up duration: mean days (SD)
Grotle [23]	2006	Prospective Cohort Study	Primary Care	All Patient, Norway, comparison acute vs. chronic	NSLBP	50 (31)	40.4 (9.5)	579.9 (785)	N.R.	360 (N.R.)
Costa [33]	2011	Prospective Cohort Study	Primary Care	All Patient, Sydney metropolitan area, Australia	NSLBP	184 (88)	43.9 (14)	90 (N.R.)	22	360 (N.R.)
Heinrich [34]	2011	Prospective Cohort Study	Outpatient Multidisciplinary Rehab	Workers on sick leave and receiving replacement benefits, Germany	NSLBP	681 (335)	45.5 (9.5)	336 (813)	N.R.	360 (N.R.)
Heymans [35]	2007	Retrospective Cohort Study	PT Restoration Program	On sick leave, pending compensation claim, Victoria, Australia	SLBP / NSLBP	194 (129)	41.8 (10)	642 (1206)	N.R.	180 (N.R.)
Heymans [36]	2009	Secondary Analysis of RCT	Occupational Health Service	Workers on sick leave and receiving replacement benefits, Netherlands	NSLBP	628 (182)	40.6 (9.5)	174 (389)	N.R.	360 (N.R.)

Acute and chronic low back pain

Study	Year	Design	Setting	Population	Diagnostic criteria	N (female)	Age (years), mean (SD)	Disease duration: Days: mean (SD)	First episode (%)	Follow-up duration: mean days (SD)
Dunn [37]	2011	Prospective Cohort Study	Primary Care	All Patient, North Staffordshire, U.K.	SLBP / NSLBP	389 (211)	46.7 (N.R.)	< 3 months to > 10 years	23	360 (N.R.)
Foster[38]	2010	Prospective Cohort Study	Primary Care	All Patient, North Staffordshire + Central Cheshire, U.K.	SLBP / NSLBP	1591 (930)	43.9 (10.3)	< 28: 601 (37.8%) to > 180: 373 (23.5%)	N.R.	180 (N.R.)
Kovacs[39]	2007	Prospective Cohort Study	Primary Care	Working population only, Spain	SLBP / NSLBP	165 (89)	45.8, IQR 39-54	1-14: 59 (35.8%), 14-90: 37 (22%), >90: 69 (41.8%)	N.R.	360 (N.R.)
Helmhout[40]	2010	Secondary Analysis of RCT	Primary Care	Professional Military (no combat-related activities), Netherlands	NSLBP	273 (4)	39 (10.5)	28-42: 12 (4%), 42-180: 24 (8%), 180-360: 34 (12%), >360: 208(76%)	N.R.	180 (N.R.)
Jensen [41]	2010	Prospective Cohort Study after RCT	Primary Care	All Patient, Denmark	SLBP / NSLBP	196 (106)	42.2 (10.4)	<3 months to > 1 year	N.R.	360 (N.R.)

NSLBP, non-specific low back pain; SLBP, specific low back pain; RCT, randomized controlled trial; N.R., not reported; IQR, interquartile range; PT, physical therapy

Acute to chronic LBP without sub-grouping

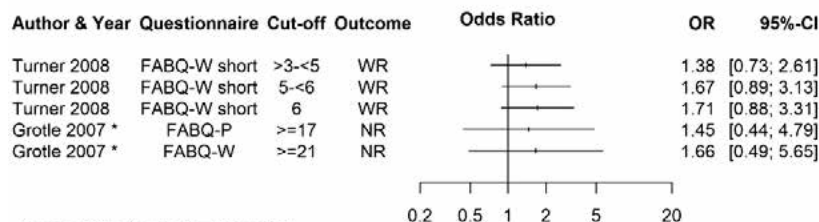
Four cohort studies and one analysis of an RCT conducted in the primary care setting included patients with disease duration ranging from one day to more than ten years (Table 3). Three cohorts [37-39] were of high quality (SIGN quality assessment). One study investigated work-related outcome [39], the other studies investigated non-work-related perceived measurements. Patients were followed between six months and one year. One secondary analysis of an RCT investigated professional military [40].

One high quality prospective cohort study [39] found FABQ and FABQ-W scores to be prognostic for sick days over one year (Table 4). While this study was conducted in a primary care setting, the outcome was collected from a Social Security database and therefore robust work-related outcome measures were available. Further, more than half of the patients suffered less than three months of LBP with almost one-third in the sub-acute phase. In contrast, TSK was found to be non-prognostic in two high quality studies that investigated non-work-related perceived measurements [37, 38]. Dunn et al. [37] included patients with very long disease duration (up to 10 years); therefore, their non-prognostic findings might be due to a high proportion of very chronic subjects with LBP. Foster et al. [38] investigated patients (n=1591) in primary care. This cohort included mainly acute to sub-acute (76%) LBP patients. They found other psychological constructs – the self-regulatory model and the self-efficacy model – to be more prognostic than fear avoidance beliefs. In both models it is theorized that low personal control over the pain or poor self-efficacy and the inevitability of a future with pain could lead to passive coping and avoidance. The interaction of the different tools used to measure these models is not yet completely understood, but might explain the non-prognostic findings in the screening for fear avoidance beliefs in this cohort. Two studies [40, 41] applying cut-off values found high values to be prognostic for perceived disability at six months and one year respectively.

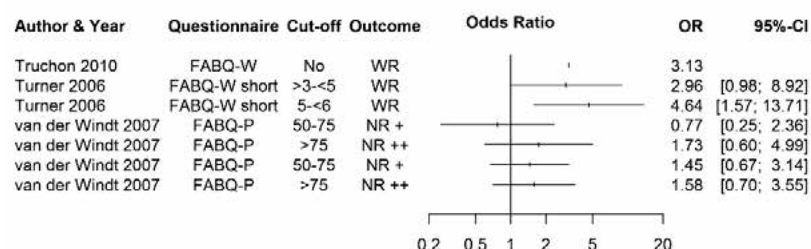
In summary, most studies that included patients with LBP for 2 weeks to 3 months found FABQ or TSK to be prognostic for work-related outcomes. An overview of the studies reporting the results of the multivariate model for binary outcomes (odds ratio) is depicted in Figure 2. Differences in settings, design, population, and outcome measured between prognostic and non-prognostic studies are summarized in Appendix 3.

Figure 2. Overview of fear avoidance beliefs as prognostic factor for binary outcomes

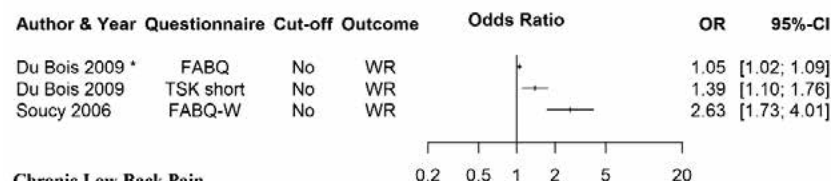
Acute Low Back Pain



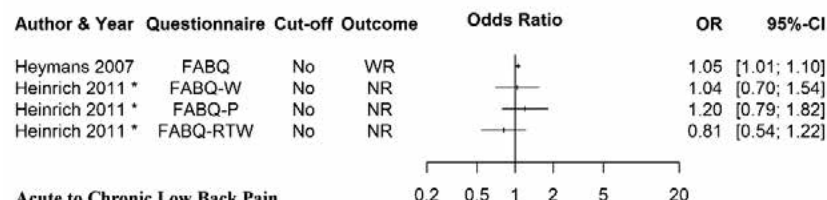
Acute to Sub-Acute Low Back Pain



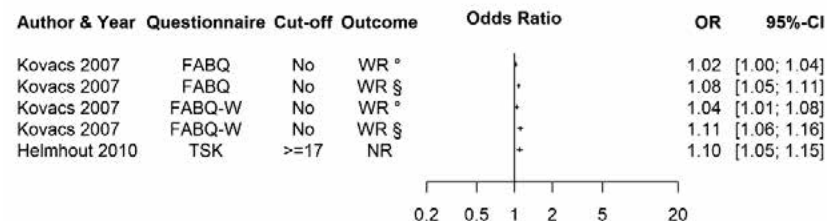
Sub-Acute Low Back Pain



Chronic Low Back Pain



Acute to Chronic Low Back Pain



NR: Non-work-related outcomes; WR: Work-related outcomes

* Linear regression

+ Perceived Disability

++ Persistent Symptoms

° Sick days 1-60

§ Sick days 61-365

Table 4. Results chronic and not sub-categorized low back pain patients

Chronic low back pain											
Study	Design	SIGN	PI (adjust)	QA	Cut-off used	MA?	Measure reported	Value (95% CI)	Outcome	Predictive	Authors Conclusion
Grotle 2006	P	+	12 (11)	FABQ-P	N.R.	Yes	Beta	0.25 (N.R.)	Perceived Disability (ODI)	Non-prognostic	FAB was a weaker predictor of future pain and disability than emotional distress. Contribution of FAB and distress was 22% in the chronic group
Costa 2011	P	+	7 (1)	TSK	N.R.	S.T.	N.R.	N.R.	RMQ	Non-prognostic	At baseline pain self-efficacy and FAB partially mediated the effects of pain on disability. At 12 months only improvements in pain self-efficacy beliefs mediated between changes in pain and disability.
Heinrich 2011	P	+	10 (6)	FABQ-W German	N.R.	Yes	OR	1.04 (0.7, 1.54)	CPG	Non-prognostic	FAB only a moderate ability to predict or explain the outcomes pain intensity, disability and functional capacity. However, depression and body mass index (BMI) were significantly related to pain-related therapeutic outcome.
				FABQ-P German	N.R.	Yes	OR	1.20 (0.79, 1.82)		Non-prognostic	
				FABQ-RTW Prediction German	N.R.	Yes	OR	0.81 (0.54, 1.22)		Non-prognostic	
Heymans 2007	R	+	7 (3)	FABQ	No Cut-off	Yes	OR	0.95 (0.91, 0.99)	RTW	Prognostic	We found in our study that patients with higher FAB were less likely to return to work at 6 months
Heymans 2009	Sec	+	8 (5)	FABQ-W	No Cut-off	Cox Regr.	HR	0.98 (0.97, 1.0)	Time to RTW (HR >1 = shorter time until RTW)	Non-prognostic	Psychological beliefs have been identified as important factors in relation to the course of LBP. The effect of FAB on prolonged sick-leave in our study was small.

Acute and chronic low back pain

Study	Design	SIGN	PI (adjust)	QA	Cut-off used	MA? Regr.	Measure reported	Value (95% CI)	Outcome	Predictive	Authors Conclusion
Dunn 2011	P	++	10 (5)	TSK single item	N.R.	Cox	Risk Ratio	0.46 (N.R.)	Poor Outcome: Grade IV CPG	Non-prognostic	FAB and catastrophizing were not important. The brief measurement method used could have impacted on the findings.
Foster 2010	P	++	7 (5)	TSK	N.R.	Yes	Beta	-0.05 (-0.11, 0.02)	RMQ	Non-prognostic	Univariate: FAB, catastrophizing and depression explained considerable proportions of the variance in disability. Multivariate: four scales remained significantly associated with outcome, FAB was not predictive.
Kovacs 2007	P	++	9 (8)	FABQ	No	Yes	OR	1.024 (1.01, 1.04)	Sick Days 1-60/year	Prognostic	These results show that, in Spanish LBP patients, baseline FABQ values influence the magnitude of LBP related work loss during the following year.
							OR	1.077 (1.05, 1.11)	Sick Days 61-365/year	Prognostic	
				FABQ-W		Yes	OR	1.042 (1.01, 1.08)	Sick Days 1-60/year	Prognostic	
							OR	1.11 (1.06, 1.16)	Sick Days 61-365/year	Prognostic	
				FABQ-P		Yes	N.R.	N.R.	Both outcomes	Non-prognostic	The baseline score of FABQ-P did not influence the duration of sick leave throughout the 1-year period
Helmhout 2010	Sec	+	7 (4)	TSK	17-68	Yes	OR	0.91 (0.87, 0.96)	Decrease RMQ ≥30%	Prognostic	One dominant prognostic factor for improvement: functional disability (RMQ scores 4 - 7 (OR 2.29) / >11 (OR 2.53)). Baseline TSK score (OR 0.97) associated with improvement.

Acute and chronic low back pain

Study	Design	SIGN	PI (adjust)	QA	Cut-off used	MA?	Measure reported	Value (95% CI)	Outcome	Predictive	Authors Conclusion
Jensen 2010, B	P RCT	+	1.1 (8)	FABQ-P 3 Items	Low: <21	Yes	Beta	0.84 (0.25, 1.44)	Disability and Pain	Prognostic	Four additional variables contributed to the multivariate models: disability in daily life activities, compensation claim, FABQ-P, duration of pain

P, prospective cohort study; R, retrospective cohort study; Sec, secondary analysis of RCTs; P RCT, prospective cohort after RCTs; QA, questionnaire used; FAB, fear avoidance beliefs; FABQ, fear avoidance questionnaire; FABQ-P: physical activity sub-scale of the FABQ; FABQ-W, work sub-scale of the FABQ; TSK, Tampa scale of kinesiophobia; LBP, low back pain; ODI, Oswestry disability index; RMDQ, Roland Morris questionnaire; RTW, return to work; CPG, chronic pain grade (von Korff) questionnaire; PI (adj.): reporting of prognostic indicators of the 16 domains (statistical adjustment for PI at baseline); SIGN, SIGN quality rating: ++ high quality, + moderate quality; OR, Odds Ratio; MA, multiple regression analysis; IQR, interquartile range; Cox. Regr., cox regression analysis; S.T., Schobel's Test; P.R.: Poisson regression Model

Summary of scales used

To assess the fear avoidance beliefs, two-thirds of the studies used the FABQ and its two subscales FABQ-P and FABQ-W (14 studies, 67%) and one-third used the TSK (7 studies, 33%) with varying versions of both scales. FABQ-W and its short versions (Table 5) were most often investigated for work-related outcomes (8 analyses) compared to non-work-related perceived measurements (4 analyses). FABQ-total score was only used in work-related outcomes. FABQ-P and the short versions were mainly for non-work-related perceived measurements (11 analyses) and rarely for work-related outcomes (3 analyses). TSK was also mainly used for non-work-related perceived measurements (5 analyses) and only once for work-related outcome. It is important to note that in sub-acute patients, TSK and FABQ-P were only used once and twice respectively. Therefore, it remains unanswered whether or not they are prognostic in this patient population.

In two studies, low baseline fear avoidance beliefs values (low-risk population) explained non-prognostic findings [21, 23]. In contrast, a study [39] with very high baseline FABQ-scores (mean 60.5, scale 0 – 96) was prognostic while other studies with the same disease duration (acute to chronic) were non-prognostic. Only 23 (67%) of 34 analyses reported baseline values. Therefore, the influence of baseline fear avoidance beliefs on prognostic and non-prognostic findings could not be addressed completely.

Table 5. Overview of scales and cut-off used and baseline fear avoidance beliefs detected

Questionnaire used	All: n (%)	Acute LBP	Acute- to Sub-acute LBP	Sub-acute LBP	Chronic LBP	Acute- to Chronic LBP
Prognostic Results: Yes = +		+	+	+	+	+
FABQ: n	3 (8)	-	-	1	1	1
Baseline Mean Values: range [§]		-	-	N.R.	46.9	60.5
FABQ-W: n	10 (30)	3	2	1	3	1
Baseline Mean Values: range [§]		13.1 - 14.5	16.2 - 21.5	4.2	19.5 - 24.6	26.5
Cut-off used: high		21-42 / >14	Diff. cut-offs	-	-	-
FABQ-W: Average of 2 Items (Scale 0-6)	2 (6)	1	1	-	-	-
Baseline Mean Value: range [§]		N.R.	N.R.	-	-	-
Cut-off used: high		5-6	5-6	-	-	-
Cut-off used: medium		3.0-4.9	3.0-4.9	-	-	-
FABQ-P: n	8 (23)	3	2	-	2	1
Baseline Mean Value: range [§]		12.3 - 17	15.5 - 16.6	-	12.9	20.1
Cut-off used: high		17-24 / >14	>75%	-	-	-
FABQ-P 3 Item: n	1 (3)	-	-	-	-	1
Baseline Mean Value: range [§]		-	-	-	-	N.R.
Cut-off used: high		-	-	-	-	>20

Questionnaire used	All: n (%)	Acute LBP		Acute- to Sub-acute LBP		Sub-acute LBP		Chronic LBP		Acute- to Chronic LBP	
Prognostic Results: Yes = +		+	-	+	-	+	-	+	-	+	-
TSK: n	4 (12)	1	1	-	-	-	-	-	1	1	1
Baseline Mean Value: range [§]		36.4	N.R.	-	-	-	-	-	40.5	38	39.7
Cut-off high		-	-	-	-	-	-	-	-	17-68	-
TSK single item	2 (6)	-	-	-	-	1	-	-	-	-	1
Baseline Mean Value: range [§]		-	-	-	-	43	-	-	-	-	N.R.
TSK Avoidance sub-scale	2 (6)	1	1	-	-	-	-	-	-	-	-
Baseline Mean Value: range [§]		16.5	16.7	-	-	-	-	-	-	-	-
TSK Harm sub-scale	2 (6)	1	1	-	-	-	-	-	-	-	-
Baseline Mean Value: range [§]		8.5	10.4	-	-	-	-	-	-	-	-
Total	34 (100)	-	-	-	-	-	-	-	-	-	-
No Cut-off: n (%)	18 (60)	-	3	1	-	3	-	1	5	2	3
Cut-off used: n (%)	12 (40)	-	6	2	2	-	-	-	-	2	-

FABQ, fear avoidance questionnaire; FABQ-W, FABQ work subscale; FABQ-P, FABQ physical subscale; TSK, Tampa Scale of Kinesiophobia.

[§] The range baseline mean values reported for each study.

Potential influence of important prognostic factors

We found substantial variation in the reporting of most prognostic domains known to potentially influence the course of LBP (Table 6). Almost half of all studies (n=10, 45%) assessed or reported half or fewer of the 16 domains. Statistical adjustment for prognostic factors varied from adjustment for only one to 13. Most variation in reported domains was found between prognostic and the non-prognostic studies in the chronic LBP group. The prognostic retrospective study assessed five domains; the non-prognostic studies reported on average more than nine domains. Measuring additional psychological prognostic information might influence the prognostic value of screening for fear avoidance beliefs (Table 7). One study of a chronic LBP population [35] and one that included acute LBP patients [26] found fear avoidance beliefs to be prognostic, but investigated no additional psychological factors; all non-prognostic studies investigated various additional psychological domains.

Table 6. Important prognostic risk factor domains in patients with low back pain (modified from Hayden et al. [15])

LBP Duration	Acute		Acute - Sub-acute		Sub-acute		Chronic		Acute- to Chronic		% Reporting
Predictive: Yes = +	+	-	+	-	+	-	+	-	+	-	
Characteristics	3	9	3	2	3	-	1	5	4	3	
Social Environment	3	3	1	-	1	-	-	-	3	1	
Overall Health Status	-	7	-	-	2	-	-	4	3	3	
Overall Psychological	-	9	2	1	2	-	-	3	3	3	
Previous LBP	3	8	1	1	2	-	-	3	-	1	
Work Psychosocial Demands	-	5	1	-	3	-	-	4	2	1	
Work Physical Demands	-	1	-	1	2	-	-	4	2	-	
Work History	3	7	-	1	1	-	-	4	2	2	
Work Place Attributes	-	1	-	-	3	-	-	-	-	-	
Disability current LBP	3	9	3	2	3	-	1	5	4	3	
Time Change LBP	3	5	-	1	2	-	-	2	2	2	
Physical Exam Findings	7	-	-	-	2	-	1	2	1	-	
Change in Physical Findings	-	-	-	-	-	-	-	-	-	-	
Definition of LBP	3	8	1	1	2	-	1	4	4	3	
Health Care Received	-	4	1	-	-	-	-	2	-	-	
Compensation Issues	3	5	1	-	3	-	1	4	1	-	
Average domain per study	10.3	9.0	4.7	5.0	10.3		5.0	9.2	7.8	7.3	

Average of reported prognostic domains of studies in each category (acute, acute – sub-acute, sub-acute, chronic, acute – chronic); LBP, low back pain

Table 7. Summary of psychological domains investigated

Psychological Domains	Total: n (%)	Acute LBP		Acute / Sub-acute		Sub-acuteLBP		ChronicLBP		Acute- toChronic	
Prognostic: yes = +	33 (100)	+	-	+	-	+	-	+	-	+	-
FAB only	9 (27)	3	-	1	1	1	-	1	1	1	-
FAB + Depression	7 (21)	-	-	-	-	2	-	-	-	2	3
FAB + Catastrophizing	2 (6)	-	2	-	-	-	-	-	-	-	-
FAB + Pain Coping Inventory	1 (3)	-	-	-	-	-	-	-	1	-	-
FAB + Anxiety + Depression	1 (3)	-	-	-	-	-	-	-	1	-	-
> 3 Concepts	13 (40)	-	7	2	1	-	-	-	2	1	-

FAB, fear avoidance beliefs

Discussion

Main findings

In this systematic review of 21 observational studies, we found convincing evidence that high fear avoidance beliefs, and particularly work beliefs, in patients suffering from sub-acute low back pain (LBP) are prognostic for poor work-related outcomes. High fear avoidance beliefs in patients with LBP for more than three months or less than two weeks were less associated with poor outcome. Neither the FABQ nor the TSK scale can be recommended over the other. Both scales used in the sub-acute population were mainly prognostic, while they were mainly non-prognostic when used in the chronic population. In sub-acute LBP patients, the TSK and FABQ-P were investigated only once and twice respectively. In studies that used cut-off values, four were prognostic and four non-prognostic for poor outcome. Less than half of all studies used cut-off values and when they did, the values varied widely among studies. Therefore, we were unable to identify specific cut-off values.

It has been suggested that scales developed to measure fear avoidance are assessing different constructs and that this makes it difficult to compare findings or draw strong conclusions [42]. The TSK measures fear of reinjury or further damage, the FABQ is largely a measure of fear of pain. While these are distinct constructs, these differences did not seem to matter in empirical studies. Indeed, it has been shown that both are good measures of fear avoidance and there is considerable overlap between the two [16].

One factor that may account for the variable prognostic utility of fear avoidance beliefs may be the low baseline values in some of the studies. A lack of variability in the predictor variable results in low correlations with outcomes. Another reason for negative findings in this population may be that other psychological factors such as outcome expectations, depression and catastrophizing might overshadow the importance of fear

avoidance beliefs in acute LBP patients. This is consistent with the idea that most individuals express a variety of distress-related emotions at the start of an episode. It may be that these beliefs early on are not important for outcome because of a generally high level of distress at this time among all patients. Similarly, a chronic population has already solidified their beliefs and there may be less variation in beliefs about the effect of activity on pain in this group. It therefore makes sense that these beliefs are more important in guiding the course of LBP in the sub-acute phase when individuals are at risk of becoming chronic and beliefs may be more disparate. Since there was an absence of assessment of additional psychological constructs with potential overlap in the studies reviewed here, this is a topic for future study. Recent research aims at identifying predictors for outcome in sub-acute LBP patients that can be addressed in a timely manner to influence outcome. Our findings emphasize the importance of fear avoidance beliefs as a potentially modifiable belief in sub-acute LBP.

Results in light of existing literature

To our knowledge this is the first systematic review summarizing current evidence on the role of fear avoidance beliefs identified by the FABQ or the TSK as a prognostic factor in acute, sub-acute, and chronic LBP. A systematic review published in 2006 found little evidence to link fear of pain with poor prognosis [43]. They included a few studies using the TSK and the FABQ at baseline assessment. Most studies included in the current analysis were published after January 2006; therefore, more evidence was available for this analysis. Our findings are in line with a systematic review finding moderate evidence for fear avoidance beliefs predicting work-related outcomes. While Iles et al. (2008) included studies with smaller samples and a shorter follow-up, in the current review several large cohort studies with a follow-up of at least six months to one year were available and therefore further strengthen the evidence. Our findings contradict an analysis of systematic reviews suggesting that fear avoidance should no longer be used as psychological risk factor in LBP [44]. The authors based their conclusion on a summary of various systematic reviews published between 2001 and 2007 when limited data was available and the Fear Avoidance Model was still mainly untested.

In recent years other models, in addition to the Fear Avoidance Model, have been discussed and investigated in regard to the experience of pain and disability [6, 45]. Pain catastrophizing conceptually falls under the model of behaviors that constitute the Fear Avoidance Model with a considerable overlap. It is believed that catastrophizing is a precursor of pain-related fear [46]. If so, catastrophizing might predict poor outcomes in acute LBP, while fear avoidance beliefs might be more likely to be predictive in sub-acute LBP. However, it is possible to have fear avoidance beliefs without catastrophizing [42]. In fact, in the acute stage of injury the avoidance of certain physical activities may be warranted. Therefore, a distinction should be made between those with fear avoidance beliefs who do not catastrophize and those who do. This is a variation on the

original model that needs testing [42]. In the current systematic review only two studies addressed the interaction of catastrophizing and fear avoidance beliefs. Therefore, the importance of catastrophizing for the prognosis in low back pain patients is beyond the scope of the current analysis and should be addressed in further research.

The self-efficacy model, focusing on more positive coping responses to pain, has increasingly received attention. It has been shown that the fear avoidance beliefs and the self-efficacy model are two distinct models that only partially overlap [6, 47]. It has therefore been postulated that the assessment of both constructs is more likely to identify patients at risk. Screening instruments like the STarT Back tool, incorporating questions covering more than one model, have been shown to identify patients at risk and improved risk-factor based treatment [5].

Strength and limitations

The strength of this systematic review is the assessment of fear avoidance beliefs in light of disease duration and a comprehensive evaluation of currently available studies. The search was inclusive, no language limitations were applied, and a thorough bibliographic search was conducted in order to include all relevant studies. The extraction process was done in accordance with current guidelines.

The main limitation is a possibility of bias in most of the included studies. Many studies were underpowered to detect a difference in a two-level exposure study (fear avoidance beliefs high vs. low). While we aimed at including cohort studies with at least 300 patients, this would have led to the exclusion of many studies in the primary care setting. The results would then be based largely upon insurance company databases with limited generalizability. While including smaller samples in this review led to a more comprehensive overview, we therefore cannot exclude the possibility of biased findings due to chance or selection. We are aware of some limitations with regards to generalizability by limiting our search to studies that used FABQ or TSK only to assess fear avoidance beliefs. Other questionnaires also include questions that address fear avoidance beliefs and were not included into this analysis. The strength of this approach is that FABQ and TSK are used widely and only assess fear avoidance beliefs. This allowed us to make comparisons among studies. Further some of the recent questionnaires that assess different coping strategies including the STarT Back Tool or the Orebro questionnaire included questions that were derived from the FABQ or TSK [48, 49]. Therefore it is likely that our findings would apply to studies that use those questionnaires with regards to the importance of fear avoidance beliefs. In addition, both tools were developed based on Vlayen's theoretical model described here [16].

Another limitation is that prognostic factors shown to be associated with poor outcome in low back pain are often not reported. This information is important to generalize study results and to decide whether or not study populations are comparable. The heterogeneous studies and the limited information reported impeded us from

conducting a meta-analysis. We have tried to balance these limitations by providing a comprehensive comparative description of all the studies included and prognostic factors reported. In recent years, experts have suggested amending or replacing the categories of acute, sub-acute, and chronic with “recurrent” back pain. It has been shown that recurrence is an important prognostic factor[50]. In the currently reviewed literature this concept has not yet been adopted and therefore will not be discussed. Our analysis suggests that high fear avoidance beliefs in pain persisting longer than four weeks are associated (independent of recurrence) with a worse outcome.

Implications for research

This systematic review highlights several important areas for future study. While the evidence for screening for fear avoidance beliefs as a prognostic factor in sub-acute LBP for work-related outcomes is supported, insufficient evidence is available for non-work-related outcomes. In sub-acute LBP patients, FABQ-P and TSK were hardly investigated, and studies in the primary care setting were scarce. Further research should aim at establishing the prognostic value of FABQ-P and TSK for non-work-related outcomes in sub-acute LBP. Furthermore, cut-off values would allow clinicians to identify patients at risk. The findings of this review expand the knowledge of currently proposed cut-off values mainly identified for short term outcomes. Future research should aim at investigating the value of cut-off values, e.g. the proposed cut-off values in the clinical implication section. The current review highlights the importance of a systematic assessment of many prognostic factors. The Multinational Musculoskeletal Inception Cohort Study Collaboration (MMICS) evaluated available instruments and proposed a core set to assess not only prognostic factors but also outcomes in future research [50]. We support the recommendation of a systematic assessment of prognostic factors with comparable measures in future studies. This will allow researchers to conduct subgroup analyses for specific prognostic risk factors and ultimately to pool study populations in future systematic reviews.

Current approaches aim at including variables from complimentary models in screening instruments [5, 51]. Research should also focus on enhancing our understanding about the time-related importance of various prognostic factors for the development of chronic LBP. One method could include using tools incorporating models such as fear avoidance beliefs, catastrophizing, self-efficacy, and satisfaction and then following patients over time. Another approach proposed by researchers is to further categorize patients using the fear avoidance beliefs model (e.g. misinformed avoiders, learned pain avoiders, affective avoiders). Those subjects who fall into these subgroups could then be assigned different treatment strategies which could be assessed in future research.

Implications for clinical practice

Treatment recommendations for LBP vary widely and often reflect the personal beliefs of physicians and other health care providers [52, 53] without taking patient preference or need into consideration. In addition the clinical judgment of therapists in the studies reviewed here generally failed to correlate with patients self-reported fear avoidance beliefs [54]. In clinical practice, a screening instrument can allow clinicians to accurately identify risk factors like fear avoidance beliefs that have been shown to be modifiable [55, 56]. Therefore, standardized measurement methods are needed to improve outcomes. This systematic review supports a proactive standardized assessment of psychological risk factors when LBP persists longer than four weeks. High fear avoidance beliefs can be addressed and a decrease in fear avoidance beliefs leads to better outcome [57-59]. Therefore targeted interventions should be initiated in patients expressing fear avoidance beliefs when pain persists. Since the duration of LBP differed substantially in many studies, and the acute, sub-acute, and chronic patients were sometimes studied together, it is important to practice flexibility with durations and cut-offs. Someone whose pain began 3 weeks ago (acute) may be assessed accurately at that time, as may someone who exceeds the sub-acute phase by a week or two. Clinical judgment along with standardized assessment is the best way to identify and treat a patient at risk.

To identify clinically relevant fear avoidance beliefs in patients with back pain several cut-off values have been proposed. It is reasonable to believe that there is a dose dependent relationship. Cut off values for low, moderate and high fear avoidance beliefs could be an approach to identify relevant fear avoidance beliefs for clinical practice. Based on the study of George et al a cut-off of 20 or less in the FABQ-W scale indicates a low risk, and FABQ-W of 25 is associated with a high risk for long term outcomes in patients with sub-acute LBP. Higher cut-off values for the FABQ-W scores (e.g. >28, 32) were associated with poor short term outcome [60] but not with poor long term outcome [30]. For the FABQ-P a study with a small patient sample (therefor not included in the current review) found a cut-off of 14 and more to be associated with poor long term prognosis [61]. A potential ceiling effect has been observed [62]. Cut-off values for the FABQ-P were oftentimes based on median split. We propose a pragmatic approach by using a cut-off for low risk of less than 14 and for high risk 16 and more. For TSK the cut-off value of 37 points and higher was proposed and was based on a median split in the validation population [11]. In the current analysis a cut-off value of less than 17 points was associated with low risk [40]. We therefore propose for future validation studies the use of low risk (<17 points), moderate risk (17- <37 points) and high risk (> 37 points). While the proposed cut-off values need further validation, they are supported by currently available literature.

Recent research has shown that questionnaires that incorporate several psychological factors (e.g. STarT Back Tool, Orebro questionnaire) which include questions derived from the FABQ or TSK [48, 49] allow for identifying patients at risk for delayed recovery and allow risk based treatment [5].

Conclusion

Evidence suggests that fear avoidance beliefs are prognostic for poor outcome in patients with sub-acute LBP and should be addressed in this population to avoid delayed recovery.

Acknowledgements

This study was performed in collaboration with many experienced researchers and clinicians. We thank the research team of the Occupational and Industrial Orthopaedic Center (OIOC), New York University (NYU) Langone Medical Center in New York for their very valuable discussions and input. Thanks are extended to Reto Kofmehl from Horten Centre of patient oriented research, University of Zurich, Switzerland for compiling the figures.

References

1. Balague, F., Mannion, A.F., Pellise, F., et al., Non-specific low back pain. *Lancet*, 2012. 379(9814): p. 482-91.
2. Fourney, D., Andersson, G., Arnold, P., et al., Chronic low back pain: a heterogeneous condition with challenges for an evidence-based approach. *Spine (Philadelphia, Pa. 1976)*, 2011. 36(21 Suppl): p. S1-S9.
3. van Tulder, M., Becker, A., Bekkering, T., et al., Chapter 3 European guidelines for the management of acute nonspecific low back pain in primary care. *European Spine Journal*, 2006. 15(0): p. s169-s191.
4. Krismer, M. and van Tulder, M., Strategies for prevention and management of musculoskeletal conditions. Low back pain (non-specific). Best practice & research. *Clinical rheumatology*, 2007. 21(1): p. 77-91.
5. Hill, J., Whitehurst, D.G.T., Lewis, M., et al., Comparison of stratified primary care management for low back pain with current best practice (STarT Back): a randomised controlled trial. *Lancet*, 2011. 378(9802): p. 1560-1571.
6. Linton, S.J. and Shaw, W.S., Impact of Psychological Factors in the Experience of Pain. *Physical Therapy*, 2011. 91(5): p. 700-711.
7. Chou, R. and Huffman, L., Nonpharmacologic therapies for acute and chronic low back pain: a review of the evidence for an American Pain Society/American College of Physicians clinical practice guideline. *Annals of Internal Medicine*, 2007. 147(7): p. 492-504.
8. Airaksinen, O., Brox, J., Cedraschi, C., et al., Chapter 4 European guidelines for the management of chronic nonspecific low back pain. *European Spine Journal*, 2006. 15(0): p. s192-s300.
9. 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(2): p. 157-68.
10. Miller, R.P., Kori, S.H., and Todd, D.D., Development of the Tampa Scale of Kinesiophobia (TSK) - unpublished report. 1991.
11. Vlaeyen, J.W., Kole-Snijders, A.M., Boeren, R.G., et al., Fear of movement/(re)injury in chronic low back pain and its relation to behavioral performance. *Pain*, 1995. 62(3): p. 363-72.
12. Stroup, D.F., Berlin, J.A., Morton, S.C., et al., Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. *JAMA*, 2000. 283(15): p. 2008-12.
13. Schlesselman, J.J., Sample size requirements in cohort and case-control studies of disease. *American Journal of Epidemiology*, 1974. 99(6): p. 381-384.
14. Harbour, R., Lowe, G., and Twaddle, S., Scottish Intercollegiate Guidelines Network: the first 15 years (1993-2008). *J R Coll Physicians Edinb*, 2011. 41(2): p. 163-168.
15. Hayden, J.A., Dunn, K.M., van der Windt, D.A., et al., What is the prognosis of back pain? Best practice & Research Clinical Rheumatology, 2010. 24(2): p. 167-179.
16. Swinkels-Meewisse, E.J.C.M., Swinkels, R.A.H.M., Verbeek, A.L.M., et al., Psychometric properties of the Tampa Scale for kinesiophobia and the fear-avoidance beliefs questionnaire in acute low back pain. *Man Ther*, 2003. 8(1): p. 29-36.
17. Vlaeyen, J.W.S., Kole-Snijders, A.M.J., Rotteveel, A.M., et al., The role of fear of movement/(re)injury in pain disability. *Journal of Occupational Rehabilitation*, 1995. 5(4): p. 235-252.
18. Ostelo, R.W.J.G., Swinkels-Meewisse, I.J.C.M., Knol, D., et al., Assessing pain and pain-related fear in acute low back pain: what is the smallest detectable change? *International Journal of Behavioral Medicine*, 2007. 14(4): p. 242-248.
19. Crombez, G., Vlaeyen, J.W., Heuts, P.H., et al., Pain-related fear is more disabling than pain itself: evidence on the role of pain-related fear in chronic back pain disability. *Pain*, 1999. 80(1-2): p. 329-339.
20. R Development Core Team. R: A language and environment for statistical computing. 2011; Available from: <http://www.R-project.org/>.
21. Grotle, M., Brox, J.I., Glomsrød, B., et al., Prognostic factors in first-time care seekers due to acute low back pain. *European Journal of Pain*, 2007. 11(3): p. 290-298.

22. Turner, J., Franklin, G., Fulton Kehoe, D., et al., ISSLS prize winner: early predictors of chronic work disability: a prospective, population-based study of workers with back injuries. *Spine* (Philadelphia, Pa. 1976), 2008. 33(25): p. 2809-2818.
23. Grotle, M., Vollestad, N.K., and Brox, J.I., Clinical course and impact of fear-avoidance beliefs in low back pain: prospective cohort study of acute and chronic low back pain: II. *Spine* (Phila Pa 1976), 2006. 31(9): p. 1038-46.
24. Hancock, M.J., Maher, C.G., Latimer, J., et al., Can rate of recovery be predicted in patients with acute low back pain? Development of a clinical prediction rule. *European Journal of Pain*, 2009. 13(1): p. 51-55.
25. Sieben, J.M., Vlaeyen, J.W., Portegijs, P.J., et al., A longitudinal study on the predictive validity of the fear-avoidance model in low back pain. *Pain*, 2005. 117(1-2): p. 162-70.
26. Swinkels-Meewisse, I.E.J., Roelofs, J., Schouten, E.G.W., et al., Fear of movement/(re)injury predicting chronic disabling low back pain: a prospective inception cohort study. *Spine* (Philadelphia, Pa. 1976), 2006. 31(6): p. 658-664.
27. van der Windt, D.A., Kuijpers, T., Jellema, P., et al., Do psychological factors predict outcome in both low-back pain and shoulder pain? *Annals of the Rheumatic Diseases*, 2007. 66(3): p. 313-9.
28. Truchon, M., Cote, D., Schmouh, M.E., et al., Validation of an adaptation of the stress process model for predicting low back pain related long-term disability outcomes: a cohort study. *Spine* (Phila Pa 1976), 2010. 35(13): p. 1307-15.
29. Turner, J.A., Franklin, G., Fulton-Kehoe, D., et al., Worker recovery expectations and fear-avoidance predict work disability in a population-based workers' compensation back pain sample. *Spine* (Phila Pa 1976), 2006. 31(6): p. 682-9.
30. George, S.Z., Fritz, J.M., and Childs, J.D., Investigation of elevated fear-avoidance beliefs for patients with low back pain: a secondary analysis involving patients enrolled in physical therapy clinical trials. *Journal of Orthopaedic and Sports Physical Therapy*, 2008. 38(2): p. 50-8.
31. Du Bois, M., Szpalski, M., and Donceel, P., Patients at risk for long-term sick leave because of low back pain. *Spine J*, 2009. 9(5): p. 350-9.
32. Soucy, I., Truchon, M., and Cote, D., Work-related factors contributing to chronic disability in low back pain. *Work*, 2006. 26(3): p. 313-326.
33. Costa, L.d.C.M., Maher, C.G., McAuley, J.H., et al., Self-efficacy is more important than fear of movement in mediating the relationship between pain and disability in chronic low back pain. *European Journal of Pain*, 2011. 15(2): p. 213-219.
34. Heinrich, M., Hafenbrack, K., Michel, C., et al., Measures of success in treatment of chronic back pain: pain intensity, disability and functional capacity. Determinants of treatment success in multimodal day clinic setting. *Schmerz*, 2011. 25(3): p. 282-289.
35. Heymans, M.W., Ford, J.J., McMeeken, J.M., et al., Exploring the contribution of patient-reported and clinician based variables for the prediction of low back work status. *J Occup Rehabil*, 2007. 17(3): p. 383-97.
36. Heymans, M.W., Anema, J.R., van Buuren, S., et al., Return to work in a cohort of low back pain patients: development and validation of a clinical prediction rule. *J Occup Rehabil*, 2009. 19(2): p. 155-65.
37. Dunn, K.M., Jordan, K.P., and Croft, P.R., Contributions of prognostic factors for poor outcome in primary care low back pain patients. *Eur J Pain*, 2011. 15(3): p. 313-9.
38. Foster, N.E., Thomas, E., Bishop, A., et al., Distinctiveness of psychological obstacles to recovery in low back pain patients in primary care. *Pain*, 2010. 148(3): p. 398-406.
39. Kovacs, F.M., Muriel, A., Castillo Sanchez, M.D., et al., Fear avoidance beliefs influence duration of sick leave in Spanish low back pain patients. *Spine* (Phila Pa 1976), 2007. 32(16): p. 1761-6.
40. Helmhout, P.H., Staal, J.B., Heymans, M.W., et al., Prognostic factors for perceived recovery or functional improvement in non-specific low back pain: secondary analyses of three randomized clinical trials. *European Spine Journal*, 2010. 19(4): p. 650-9.
41. Jensen, O.K., Nielsen, C.V., and Stengaard-Pedersen, K., One-year prognosis in sick-listed low back pain patients with and without radiculopathy. Prognostic factors influencing pain and disability. *Spine J*, 2010. 10(8): p. 659-75.

42. Pincus, T., Smeets, R.J.E.M., Simmonds, M.J., et al., The Fear Avoidance Model Disentangled: Improving the Clinical Utility of the Fear Avoidance Model. *Clinical Journal of Pain*, 2010. 26(9): p. 739-746.
43. Pincus, T., Vogel, S., Burton, A.K., et al., Fear avoidance and prognosis in back pain: a systematic review and synthesis of current evidence. *Arthritis and Rheumatism*, 2006. 54(12): p. 3999-4010.
44. Lakke, S.E., Soer, R., Takken, T., et al., Risk and prognostic factors for non-specific musculoskeletal pain: a synthesis of evidence from systematic reviews classified into ICF dimensions. *Pain*, 2009. 147(1-3): p. 153-64.
45. Main, C.J., Foster, N., and Buchbinder, R., How important are back pain beliefs and expectations for satisfactory recovery from back pain? *Best Pract Res Clin Rheumatol*, 2010. 24(2): p. 205-17.
46. Leeuw, M., Goossens, M.E., Linton, S.J., et al., The fear-avoidance model of musculoskeletal pain: current state of scientific evidence. *Journal of Behavioral Medicine*, 2007. 30(1): p. 77-94.
47. Woby, S., Urmston, M., and Watson, P., Self-efficacy mediates the relation between pain-related fear and outcome in chronic low back pain patients. *European Journal of Pain*, 2007. 11(7): p. 711-718.
48. Hill, J.C., Dunn, K.M., Lewis, M., et al., A primary care back pain screening tool: identifying patient subgroups for initial treatment. *Arthritis and Rheumatism*, 2008. 59(5): p. 632-41.
49. Linton, S.J. and Halld  n, K., Can we screen for problematic back pain? A screening questionnaire for predicting outcome in acute and subacute back pain. *The Clinical Journal of Pain*, 1998. 14(3): p. 209-215.
50. Pincus, T., Santos, R., Breen, A., et al., A review and proposal for a core set of factors for prospective cohorts in low back pain: a consensus statement. *Arthritis and Rheumatism*, 2008. 59(1): p. 14-24.
51. Hill, J.C. and Fritz, J.M., Psychosocial Influences on Low Back Pain, Disability, and Response to Treatment. *Physical Therapy*, 2011. 91(5): p. 712-721.
52. Rainville, J., Carlson, N., Polatin, P., et al., Exploration of physicians' recommendations for activities in chronic low back pain . . . including commentary by Volinn E. Spine, 2000. 25(17): p. 2210-2220.
53. Werner, E.L., Cote, P., Fullen, B.M., et al., Physicians' determinants for sick-listing LBP patients: a systematic review. *Clinical Journal of Pain*, 2012. 28(4): p. 364-71.
54. Calley, D.Q., Jackson, S., Collins, H., et al., Identifying patient fear-avoidance beliefs by physical therapists managing patients with low back pain. *Journal of Orthopaedic and Sports Physical Therapy*, 2010. 40(12): p. 774-83.
55. George, S.Z., Fritz, J.M., Bialosky, J.E., et al., The effect of a fear-avoidance-based physical therapy intervention for patients with acute low back pain: results of a randomized clinical trial. *Spine (Phila Pa 1976)*, 2003. 28(23): p. 2551-60.
56. Jellema, P., van der Horst, H.E., Vlaeyen, J.W., et al., Predictors of outcome in patients with (sub)acute low back pain differ across treatment groups. *Spine (Phila Pa 1976)*, 2006. 31(15): p. 1699-705.
57. Staal, J.B., Hlobil, H., Koke, A.J., et al., Graded activity for workers with low back pain: who benefits most and how does it work? *Arthritis and Rheumatism*, 2008. 59(5): p. 642-9.
58. George, S.Z., Zeppieri, G., Jr., Cere, A.L., et al., A randomized trial of behavioral physical therapy interventions for acute and sub-acute low back pain (NCT00373867). *Pain*, 2008. 140(1): p. 145-57.
59. Burton, A.K., Waddell, G., Tillotson, K.M., et al., Information and advice to patients with back pain can have a positive effect. A randomized controlled trial of a novel educational booklet in primary care. *Spine (Phila Pa 1976)*, 1999. 24(23): p. 2484-91.
60. Fritz, J.M. and George, S.Z., Identifying psychosocial variables in patients with acute work-related low back pain: the importance of fear-avoidance beliefs. *Physical Therapy*, 2002. 82(10): p. 973-83.
61. Klaber Moffett, J.A., Carr, J., and Howarth, E., High fear-avoiders of physical activity benefit from an exercise program for patients with back pain. *Spine (Phila Pa 1976)*, 2004. 29(11): p. 1167-72; discussion 1173.
62. Kovacs, F., Muriel, A., Medina, J., et al., Psychometric characteristics of the Spanish version of the FAB questionnaire. *Spine (Philadelphia, Pa. 1976)*, 2006. 31(1): p. 104-110.

Appendix 1:**Search history for PubMed, CINAHL, psychINFO October week 2 2011**

PubMed

#	Query	Results
1	("low back pain"[MeSH Terms] OR ("low"[All Fields] AND "back"[All Fields] AND "pain"[All Fields]) OR "low back pain"[All Fields]) OR ("low back pain"[MeSH Terms] OR ("low"[All Fields] AND "back"[All Fields] AND "pain"[All Fields]) OR "low back pain"[All Fields] OR "lumbago"[All Fields]) OR ("low back pain"[MeSH Terms] OR ("low"[All Fields] AND "back"[All Fields] AND "pain"[All Fields]) OR "low back pain"[All Fields] OR ("low"[All Fields] AND "back"[All Fields] AND "ache"[All Fields]) OR "low back ache"[All Fields])	<u>21548</u>
2	("fear"[MeSH Terms] OR "fear"[All Fields]) AND avoidance[All Fields]	<u>3531</u>
3	#3 AND #4	<u>299</u>
4	Search #3 AND #4 Limits: Humans, Publication Date from 1990 to 2011	<u>255</u>

CINAHL

#	Query	Results
1	(MH "Low Back Pain") OR "low back pain"	10957
2	"lumbago"	40
3	"low back ache or low backache"	0
4	S1 or S2	10977
5	"fear avoidance"	400
6	(MH "Avoidant Personality Disorder") OR "avoidance"	5477
7	"avoidance"	5454
8	S5 or S6 or S7	5477
9	S3 or S4	10977
10	S8 and S9	274

#	Searches	Results
1	exp Avoidance/ or avoidance.mp. orexp Approach Avoidance/ or exp Avoidance Conditioning.	34914
2	fear.mp. or exp Fear/ or exp Conditioned Fear	43029
3	1 and 2	3676
4	fear avoidance.mp.	333
5	3 or 4	3676
6	low back pain.mp.	2010
7	pain.mp. or exp Back Pain	3370
8	low.mp. and 7 [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures]	2160
9	6 or 8	2160
10	lumbago.mp.	26
11	(low and (backache or back ache)).mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures]	25
12	or/9-11	2196
13	5 and 12	113
14	limit 13 to (human and yr="1990 -Current")	111

Appendix 2: Summary of the SIGN quality assessment

Study	2.1.: Research Question	2.2 Definition	2.3 Same source?	2.4 Inclusion process	2.7 Comp. Participant drop-out	2.8. Outcome defined?	2.9. Blinding to exposure status	2.10 Recogn pot. Influence of	2.11 Reliable assessment of exposure?	2.12 Valid Measure Exposure	2.13 Reliable Method	2.14 valid Method	2.15 Prognostic factor?	2.16 Confounder?	2.17 CI provided?	3.1 SIGN
Du Bois 2009	Well covered	Adequately addressed	Adequately addressed	Well covered	Adequately addressed	Well covered	Not addressed	Not addressed	Poorly addressed	Poorly addressed	Adequately addressed	Adequately addressed	Poorly addressed	Adequately addressed	Yes	+
Dunn 2011	Well covered	Adequately addressed	Well covered	Well covered	Adequately addressed	Well covered	Not reported	Not reported	Adequately addressed	Adequately addressed	Adequately addressed	Adequately addressed	Poorly addressed	Well covered	No	+
Foster 2010	Adequately addressed	Adequately addressed	Well covered	Well covered	Poorly addressed	Poorly addressed	Not addressed	Not addressed	Adequately addressed	Adequately addressed	Adequately addressed	Adequately addressed	Adequately addressed	Adequately addressed	Yes	+
Grotle 2007	Well covered	Well covered	Adequately addressed	Adequately addressed	Poorly addressed	Well covered	Not addressed	Not addressed	Well covered	Well covered	Well covered	Well covered	Adequately addressed	Adequately addressed	Yes	+
Grotle 2006	Well covered	Well covered	Adequately addressed	Adequately addressed	Poorly addressed	Well covered	Not addressed	Not addressed	Well covered	Well covered	Well covered	Well covered	Adequately addressed	Adequately addressed	Yes	+
Hancock 2009	Adequately addressed	Well covered	Adequately addressed	Well covered	Not applicable	Well covered	Adequately addressed	Adequately addressed	Adequately addressed	Adequately addressed	Adequately addressed	Adequately addressed	Poorly addressed	Adequately addressed	Yes	++
Heinrich 2011	Well covered	Adequately addressed	Adequately addressed	Well covered	Well covered	Well covered	Not addressed	Not addressed	Not addressed	Not addressed	Adequately addressed	Adequately addressed	Adequately addressed	Adequately addressed	Yes	++
Helmhout 2010	Adequately addressed	Adequately addressed	Well covered	Poorly addressed	Well covered	Well covered	Not addressed	Not addressed	Adequately addressed	Adequately addressed	Adequately addressed	Adequately addressed	Poorly addressed	Adequately addressed	Yes	+
Heymans 2009	Adequately addressed	Adequately addressed	Adequately addressed	Not addressed	Not addressed	Adequately addressed	Adequately addressed	Not addressed	Adequately addressed	Adequately addressed	Adequately addressed	Adequately addressed	Poorly addressed	Well covered	Yes	+
Heymans 2007	Adequately addressed	Adequately addressed	Adequately addressed	Not addressed	Not addressed	Adequately addressed	Adequately addressed	Not addressed	Adequately addressed	Adequately addressed	Adequately addressed	Adequately addressed	Poorly addressed	Adequately addressed	Yes	+
Jensen 2010	Adequately addressed	Adequately addressed	Adequately addressed	Poorly addressed	Adequately addressed	Adequately addressed	Adequately addressed	Not addressed	Adequately addressed	Adequately addressed	Adequately addressed	Adequately addressed	Poorly addressed	Well covered	Yes	+

Study	2.1.: Research Question	2.2 Definition	2.3 Same source?	2.4 Inclusion process	2.7 Comp. Participant drop-out	2.8. Outcome defined?	2.9. Blinding exposure status	2.10 Recogn of pot. Influence of	2.11 Reliable assessment of exposure?	2.12 Valid Measure of Exposure	2.13 Reliable Method	2.14 valid Method	2.15 Prognostic factor?	2.16 Confounder?	2.17 CI provided?	3.1 SIGN
Kovacs 2007	Adequately addressed	Adequately addressed	Adequately addressed	Adequately addressed	Not addressed	Adequately addressed	Adequately addressed	Not applicable	Adequately addressed	Adequately addressed	Adequately addressed	Adequately addressed	Adequately addressed	Adequately addressed	Yes	++
Mannion 2001	Adequately addressed	Adequately addressed	Adequately addressed	Well covered	Adequately addressed	Adequately addressed	Not addressed	Not addressed	Not addressed	Not addressed	Adequately addressed	Adequately addressed	Poorly addressed	Adequately addressed	No	-
Sieben 2005	Adequately addressed	Well covered	Adequately addressed	Adequately addressed	Adequately addressed	Adequately addressed	Not addressed	Not addressed	Adequately addressed	Adequately addressed	Adequately addressed	Adequately addressed	Adequately addressed	Adequately addressed	Yes	++
Soucy 2006	Adequately addressed	Poorly addressed	Adequately addressed	Adequately addressed	Not addressed	Adequately addressed	Not addressed	Not addressed	Adequately addressed	Adequately addressed	Adequately addressed	Adequately addressed	Poorly addressed	Adequately addressed	Yes	+
Truchon 2010	Adequately addressed	Adequately addressed	Adequately addressed	Adequately addressed	Adequately addressed	Adequately addressed	Not addressed	Not addressed	Adequately addressed	Adequately addressed	Adequately addressed	Adequately addressed	Poorly addressed	Adequately addressed	Yes	+
Turner 2006	Adequately addressed	Poorly addressed	Adequately addressed	Adequately addressed	Adequately addressed	Adequately addressed	Not addressed	Not addressed	Poorly addressed	Poorly addressed	Adequately addressed	Adequately addressed	Poorly addressed	Adequately addressed	Yes	+
van der Windt 2001	Adequately addressed	Adequately addressed	Adequately addressed	Poorly addressed	Not addressed	Adequately addressed	Not addressed	Not addressed	Adequately addressed	Adequately addressed	Adequately addressed	Adequately addressed	Adequately addressed	Well covered	Yes	+
Costa 2011	Well covered	Well covered	Adequately addressed	Adequately addressed	Adequately addressed	Adequately addressed	Not addressed	Not addressed	Adequately addressed	Adequately addressed	Adequately addressed	Adequately addressed	Poorly addressed	Adequately addressed	to determine	+
Turner 2008	Adequately addressed	Adequately addressed	Adequately addressed	Adequately addressed	Adequately addressed	Adequately addressed	Not addressed	Not addressed	Poorly addressed	Poorly addressed	Poorly addressed	Poorly addressed	Poorly addressed	Well covered	Yes	+
Swinkels 2006	Adequately addressed	Adequately addressed	Adequately addressed	Adequately addressed	Adequately addressed	Adequately addressed	Not addressed	Not addressed	Well covered	Well covered	Well covered	Well covered	Adequately addressed	Well covered	Yes	+
George 2008	Adequately addressed	Adequately addressed	Adequately addressed	Poorly addressed	Adequately addressed	Adequately addressed	Not addressed	Not addressed	Adequately addressed	Adequately addressed	Adequately addressed	Adequately addressed	Poorly addressed	Adequately addressed	Yes	+

Appendix 3: Summary characteristics in prognostic and non-prognostic studies

LBP Duration	All		Acute		Acute to Sub-acute		Sub-acute		Chronic		Acute to chronic	
Predictive Results: Yes = +	+	-	+	-	+	-	+	-	+	-	+	-
Studies: n (%)	24 (100)	1 (4)	5 (20)	3 (13)	2 (8)	2 (8)	-	1 (4)	4 (17)	3 (13)	3 (13)	-
Number of Patient included: range	50 - 2677	555 (337)	123 - 1885	160 - 1068	160 - 171	258-346	-	194	50 - 681	165 - 273	165 - 1591	-
Age: range	16 - 65	42.5	<24 to >55	28-45	25-55	16-64	-	30-50	30-55	25-54	33-54	-
Setting												
Primary care: n (%)	13 (56)	1 (4)	4 (17)	-	1 (4)	-	-	-	2 (8)	3 (13)	3 (13)	-
Disability insurance, workers on sick leave: n (%)	7 (30)	-	1 (4)	2 (8)	-	2 (8)	-	-	2 (8)	-	-	-
Outpatient PT: n (%)	3 (13)	-	-	1 (4)	1 (4)	-	-	1 (4)	-	-	-	-
Design other than prospect. Cohort	5 (21)	-	-	Sec (1)	Sec (1)	-	-	R (1)	Sec (1)	P-RCT (1)	-	-
Studies: n (%)												
Follow-up (days): range	84 - 360	180	84 - 1116	180 - 360	90 - 180	258 - 346	-	180	360	180 - 360	180 - 360	-
Outcomes analyzed: total 25 (100%)												
RTW: n (%)	6 (24)	-	1	2	-	2	-	1	1	-	-	-
Sick days: n (%)	4 (16)	-	-	-	-	-	-	-	-	2	2	-
Perceived disability: n (%)	12 (48)	1	3	1	2	-	-	-	2	2	1	-
Pain: n (%)	2 (8)	-	-	-	-	-	-	-	1	-	1	-
Recovery: n (%)	1 (4)	-	1	-	-	-	-	-	-	-	-	-
Diagnosis: total 23 (100%)												
NSLBP: n (%)	9 (40)	1 (4)	1 (4)	1 (4)	-	2 (8)	-	-	4 (17)	1 (4)	-	-

LBP Duration		All	Acute		Acute to Sub-acute		Sub-acute		Chronic		Acute to chronic	
Predictive Results: Yes = +			+	-	+	-	+	-	+	-	+	-
SLBP and NSLBP: n (%)		14 (60)	-	4 (17)	2 (8)	2 (8)	-	-	1 (4)	-	2 (8)	3 (13)
First episode: not reported: n (%)		16 (69)	-	2 (8)	3 (13)	1 (4)	1 (4)	-	1 (4)	3 (13)	3 (13)	2 (8)
Percent first episode reported: n (%)		7 (31)	22.5%	25-93.6%	-	33%	39%	-	-	22%	-	23%
MA / Cox Regr: n (%)		21 (92)	1	4	3	2	2	-	1	3	3	3
Other statistics: n (%)		2 (8)	-	U.L. (1)	-	-	-	-	-	S.T. (1)	-	-
Health Care System: total 23 (100%)												
Europe: n (%)		14 (64)	1	3 (13)	-	1 (4)	1 (4)	-	-	3 (13)	3 (13)	3 (13)
USA: n (%)		4 (16)	-	1 (4)	2 (8)	1 (4)	-	-	-	-	-	-
Canada: n (%)		2 (8)	-	-	1 (4)	.	1 (4)	-	-	-	-	-
Australia: n (%)		3 (12)	-	1 (4)	-	.	-	-	1 (4)	1 (4)	-	-

CHAPTER 4

The influence of fear avoidance beliefs on treatment outcome in patients with low back pain: A systematic review

Maria M. Wertli, M.D., Eva Rasmussen-Barr, RPT; Ph.D., Ulrike Held, Ph.D., Sherri Weiser, Ph.D., Lucas M. Bachmann, M.D.; Ph.D., Florian Brunner, M.D.; Ph.D.
Submitted

Abstract

Background context: Psychological factors are believed to influence the development of chronic low back pain. To date it is not known how fear avoidance beliefs influence the treatment efficacy in low back pain.

Purpose: To summarize the evidence examining the influence of fear avoidance beliefs measured with the Fear Avoidance Beliefs Questionnaire (FABQ) or the Tampa Scale of Kinesiophobia (TSK) on treatment outcomes in patients with low back pain.

Design / Setting: Systematic Review of randomized controlled trials.

Patient Sample: Patients with low back pain.

Outcome Measures: Work related outcomes and self-reported measures including return to work, pain and disability

Methods: In January 2013, the following databases were searched: BIOSIS, CINAHL, Cochrane Library, Embase, OTSeeker, PeDRO, PsycInfo, PubMed/Medline, Scopus and Web of Science. A hand search of the six most often retrieved journals and a bibliography search completed the search.

Study eligibility criteria, participants, and interventions: Research studies that included patients with low back pain who participated in randomized controlled trials (RCTs) investigating non-operative treatment efficacy. Out of 646 records 78 papers were assessed in full-text and 17 RCTs were included. Study quality was high in five studies, moderate in 12 studies. This study was not funded and the authors have no conflict of interest to declare.

Results: In patients with low back pain of up to six months duration, high fear avoidance beliefs were associated with more pain and / or disability (four RCTs) and less return to work (three RCTs) (GRADE high quality evidence, 831 patients vs. 322 in non-predictive studies). A decrease in fear avoidance beliefs values during treatment was associated with less pain and disability at follow-up (GRADE moderate evidence, two RCTs with moderate quality, 242 patients). Interventions that addressed fear avoidance beliefs were more effective than control groups based on biomedical concepts (GRADE moderate evidence, 1051 vs. 227 patients in studies without moderating effects). In chronic patients with LBP, the findings were less consistent. Two studies found baseline fear avoidance beliefs to be associated with more pain and disability and less RTW (339 patients) while three others (832 patients) found none (GRADE low evidence).

Limitations: Heterogeneity of the studies impeded a pooling of the results.

Conclusion: Evidence suggests that fear avoidance beliefs are associated with poor treatment outcome in patients with LBP of less than six months and thus early treatment, including interventions to reduce fear avoidance beliefs, may avoid delayed recovery and chronicity. Patients with high fear avoidance beliefs are more likely to improve when fear avoidance beliefs are addressed in treatments than when these beliefs are ignored, and treatment strategies should be modified if fear avoidance beliefs are present.

Introduction

Low back pain is one of the leading causes globally of years lived with disability [1]. Almost all adults once in their lifetime complain about LBP, but 10-15 percent develop chronic LBP [2]. This small percentage accounts for three-quarters of the total LBP associated costs [3]. There is a consensus in the literature in acute low back pain to avoid unnecessary investigation and overtreatment by treating symptomatically with encouragement to return to normal activity [4]. Persisting pain for several weeks predicts the development of chronic low back pain, a condition where complete recovery are often difficult to achieve [5]. Current research aims to identify risk indicators for delayed recovery in patients with sub-acute LBP in order to optimize treatment and avoid chronification. Targeted and timely interventions in patients at risk for chronic pain facilitate recovery and may reduce health care costs [6].

The Fear Avoidance Model (FAM) is a theoretical model that describes how psychological factors affect the experience of pain and the development of chronic pain and disability [7]. It is theorized that for some individuals with LBP, negative beliefs about pain and/or negative illness information leads to a catastrophizing response in which the worst possible outcome of activity is imagined. This leads to fear of activity and avoidance which in turn causes disuse and resultant distress, reinforcing the original negative appraisal in a deleterious cycle [7]. The FAM suggests that patients without fear avoidance beliefs are more likely to confront pain problems and are more active in the coping process. This type of “good” coping has been used to develop interventions for those with high fear avoidance beliefs.

While the Fear Avoidance Model is generally accepted, it is a matter of debate as to how and when to best assess fear avoidance beliefs in clinical practice. Current treatment guidelines for low back pain recommend the timely identification and initiation of multidisciplinary treatment for other psychological factors (e.g. depression, distress, job dissatisfaction) associated with increased risk for delayed recovery [4, 8, 9]. There are however no recommendations for the assessment of fear avoidance beliefs. In a recent systematic review (Wertli et al. submitted) high fear avoidance beliefs, identified by the most frequently used questionnaires, the Fear Avoidance Questionnaire (FABQ [10]) and the Tampa Scale of Kinesiophobia (TSK [11, 12]), were prognostic in patients with sub-acute low back pain for not returning to work in observational studies. Whether high fear avoidance beliefs identified by these two questionnaires specifically influences treatment efficacy in currently used treatment strategies is unknown [13-17].

How fear avoidance beliefs influences treatment outcomes and moderate treatment efficacy in LBP has not been reviewed systematically. The aim of the current review is therefore to assess the influence of fear avoidance beliefs on the outcome of various treatments in randomized controlled trials (RCTs) for patients with LBP.

Methods

The review was conducted following the PRISMA statement on conducting systematic reviews of randomized controlled trials [18]. To allow for comparison between studies and to clarify the moderating role of fear avoidance beliefs on treatment efficacy in patients with LBP, we restricted this review to studies that assessed fear avoidance beliefs with the most often used questionnaires, the fear avoidance questionnaire (FABQ) and the Tampa Scale of Kinesiophobia (TSK). For both questionnaires a moderate overlap has been shown.

Literature search

We identified all RCTs meeting our eligibility criteria (defined in detail below) published between January 1990 and January 2013. The following databases were searched in January 2013: BIOSIS, CINAHL, Cochrane Library, Embase, OTSeeker, PeDRO, PsycInfo, PubMed/Medline, Scopus, and Web of Science. The search was conducted with the help of an experienced librarian. Search terms included “fear” and “avoidance” as Medical Subject Headings (MeSH) as well as other subject headings and different combinations. Two detailed search strategies are depicted in Appendix 1. To ensure the completeness of the literature search, one reviewer (MW) conducted an electronic hand search of the six most often retrieved journals and added all potentially eligible references not retrieved by the systematic search. In addition, bibliographies of included studies relevant to the research question were searched and potentially eligible references were included in the full text review (inclusion and exclusion criteria applied). Authors of studies that were eligible but did not report sufficient information were contacted and, where available, this information was included in the analysis.

Eligibility criteria

All RCTs were considered eligible for inclusion in this investigation that met the following criteria: they reported research concerning patients seeking care for LBP; they demonstrated at least moderate study quality; they investigated the effect of fear avoidance beliefs assessed by the two most often used questionnaires - the Fear Avoidance Questionnaire (FABQ) and the Tampa Scale for Kinesiophobia (TSK); and they were published between January 1990 and January 2012. We focused on RCTs with at least 30 patients per group because of a concern about sample size. Assuming a 50 percent reduction in perceived disability that was one third greater in the treatment group when compared to the reference group, a sample size of 37 patients per group would be sufficient to detect the difference in allowing a drop-out rate of 15% (alpha 0.80, significance level 0.05). No limits for the study setting or language of the publication were applied. Excluded were conference proceedings.

Study selection, data extraction, and synthesis

The bibliographic details of all retrieved articles were stored in an Endnote file. Two reviewers (MW and ERB) independently screened all 691 references by title and abstract. The full text was reviewed by both reviewers independently (MW and ERB) in all studies meeting the pre-defined eligibility criteria ($n = 78$). Disagreements were discussed and resolved by consensus or by third party arbitration (SW). Researchers with specific language proficiencies reviewed non-English language references. In the case of several publications for the same RCT without change in outcome or follow-up duration, the most recent publication was chosen and missing information from the previous publication added.

Quality assessment

The quality of each study was assessed using the Scottish Intercollegiate Guidelines Network (SIGN) methodology checklist for RCTs [19]. In addition all quality criteria not covered in this checklist but considered to be important according to a consensus study on methodological criteria for the assessment of moderators in systematic reviews were added [20]. Meta-analyses were only considered when all the following criteria were met: moderators were measured prior to randomisation; adequate quality of measurement of baseline factors; explicit test of interaction between moderator and treatment. All information needed to describe the study population was extracted. To assess the baseline characteristics of the study population, important prognostic factors for the course of low back pain were extracted and assigned to one of the 16 domains used in previous research [21] and proposed by Hayden et al. [22] general characteristics, social environment, overall health status, overall psychological health, previous LBP, work-psychosocial demands, work-physical demands, work-history and attributes, disability related to LBP, time change of LBP, physical examination findings, change in physical examination, diagnosis of LBP, and compensation related to LBP. The number of domains reported is given for each study.

This information was used to rate the overall bias risk and study quality according to the SIGN recommendations. The ratings included high quality (++; most criteria fulfilled and if not fulfilled, the study conclusions are very unlikely to be altered), moderate quality (+; some criteria fulfilled and if not fulfilled, the study conclusions are unlikely to be altered), low quality (–; few or no criteria fulfilled, conclusions likely to be altered). Studies rated as low quality by both reviewers were excluded from further analysis.

Outcome definition

All investigated outcomes were extracted and operationalized [(e.g. perceived disability measured by Oswestry Disability Index (ODI) or by Roland Morris Disability Questionnaire (RMQ)].

Operationalization of fear avoidance beliefs as predictor, mediator and moderator

The definitions for predictor, mediator and moderator were adopted from Pincus and colleagues [20]. Predictor: baseline fear avoidance beliefs affect outcome but do not interact with the allocated treatment intervention. Mediator: change in fear avoidance beliefs during treatment impacts outcome, with or without interacting with allocated treatment. Moderator: fear avoidance beliefs at baseline interact with treatment. The quality of the moderator analysis was assessed for each study by two reviewers (EB and MW) and discussed with an experienced statistician (UH). The following factors were considered: 1) when multiple comparisons were conducted, was the significance level adjusted from $p < 0.05$ to $p < 0.001$; 2) was an interaction test between treatment and moderator conducted; 3) were moderators equally distributed between groups at baseline. Predictive” or “moderating” means that fear avoidance beliefs were significantly associated with outcome in an univariate analysis or in a stepwise procedure and therefore were included in the final multiple model.

Treatment definition

All information given about the treatments delivered were extracted. Based on this information each treatment was categorized into: treatment based on a biomedical approach (e.g. PT without CBT approach, the Handy Hints Back Book); treatments aimed to address fear avoidance beliefs (e.g. psychological informed PT, cognitive behavioural therapy (CBT), active advice including information about psychological factors, graded activity). Whether a treatment addresses fear avoidance beliefs or not was rated independently by two reviewers (MW, EB) and disagreements were discussed. If no consensus could be achieved an independent third reviewer (SW) decided whether or not a treatment addressed fear avoidance beliefs.

Psychometric properties and description of the investigated questionnaires

The FABQ [10] is a 16-item questionnaire, each item scored zero to six. High levels indicate increased levels of fear avoidance beliefs. Two subscales exist: a seven-item work subscale (FABQ-W, range 0 to 42) and a four-item physical activity subscale (FABQ-P, range 0 to 24). FABQ and the two sub-scales have been shown to be reliable and valid for the measurement of FAB. The Cronbach’s alpha for the FABQ-P was 0.75, test-retest

reliability $r=0.64$. For the FABQ-W, the Cronbach's alpha was 0.82, test-retest $r=0.80$ [23].

The Tampa Scale for Kinesiophobia (TSK) is a 17-item questionnaire [23]. Each item is measured on a four-point Likert scale (1 = strongly disagree, to 4 = strongly agree). The scale ranges from 17 (no fear) to 68 (strong fear of re-injury). Several studies found support for the construct and predictive validity and reliability of the TSK [11, 24]. Further research found that a two factor model based on 13 items to explain up to 70% of the variation and its validity was confirmed [25]. Two factor model consisted of a "harm factor" (items 3, 5, 6, 9, 11, 15 with a range of 6-24) and an "activity-avoidance" factor (items 1, 2, 7, 10, 13, 14, 17 with a range of 7-28) [23]. The "harm factor" reflects the beliefs that something is seriously wrong with the body, while the "activity-avoidance" factor indicates that avoiding exercise or activity might prevent an increase in pain.

While the two scales measure different concepts (i.e. FABQ measures fear of pain caused by physical activity; TSK measures fear of movement and re-injury), a moderate overlap has been shown [23, 26, 27]. The correlation between TSK and FABQ ranged from 0.39 (FABQ-P) and 0.33 (FABQ-W) [23] to 0.53 (FABQ-P) and 0.76 (FABQ-W) [27].

Statistical analysis

Due to heterogeneous study populations, measurements and scales used as well as outcomes investigated, no meta-analysis was performed. Descriptive statistics (ranges) were used to summarize findings across all RCTs for baseline fear avoidance mean values. Forest plots were generated using the R statistical software [28]. Forest plots were based on the values (odds ratio, hazard ratio, beta with corresponding 95% confidence interval) reported in the study reports.

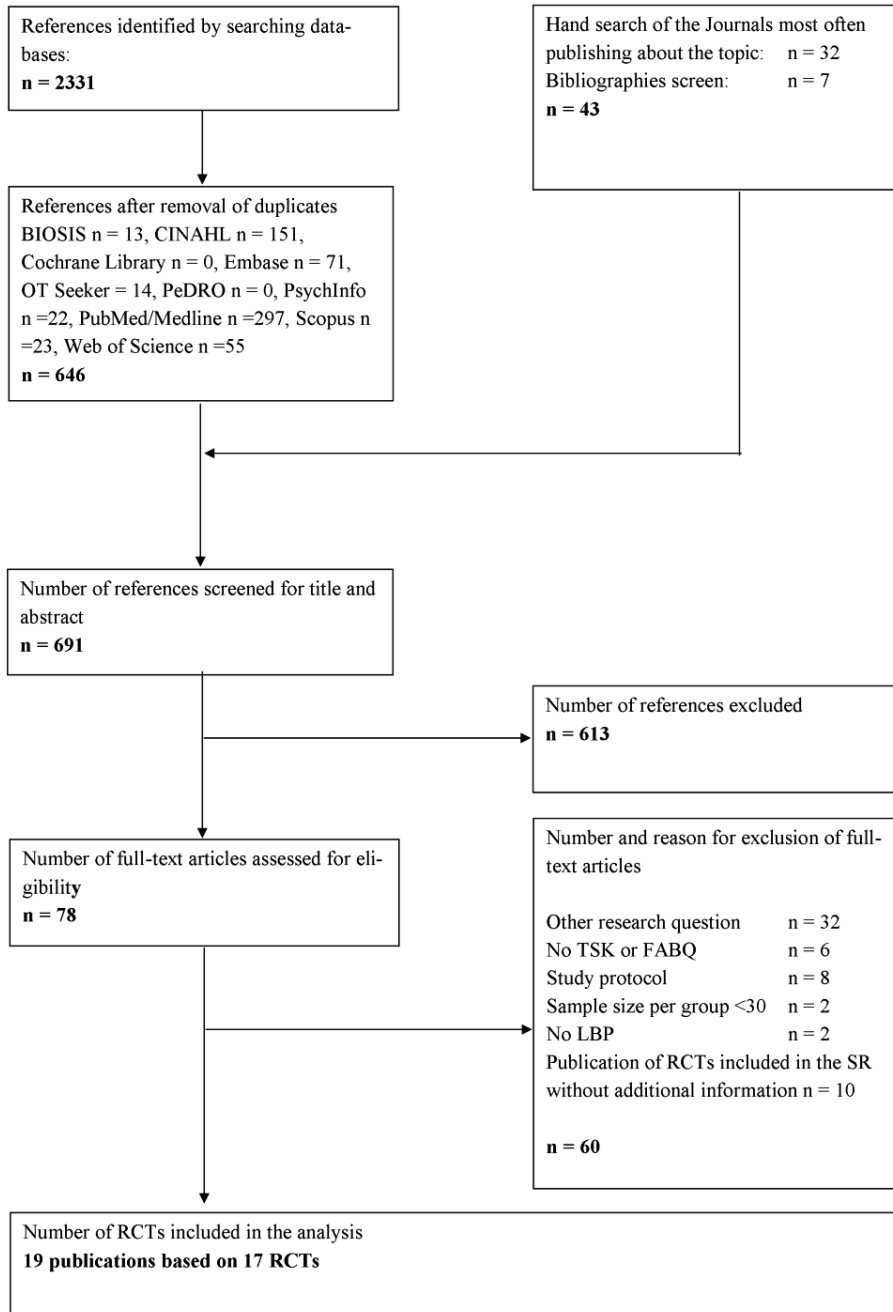
For appraising the evidence of the results the GRADE system [29] was used which was also adopted by the Cochrane Back Review Group [30]: high quality (75% of the RCTs with consistent findings, direct and precise data, no suspected publication biases, further research is very unlikely to change our confidence in the estimate of effect), moderate quality (several RCTs of moderate quality, further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate), low quality (only a few RCTs of moderate quality available, further research is very likely to have an important impact on the estimate of effect and is likely to change the estimate), very low quality (effect is very uncertain).

Results

Study selection

The search and inclusion process is summarized in Figure 1. Out of 2,331 records, 78 were reviewed in full text. The full text assessment utilizing the inclusion and exclusion criteria resulted in the exclusion of 60 studies. The main reasons for exclusion were not investigating fear avoidance beliefs as a predictor or moderator ($n = 32$), use of questionnaires other than FABQ and TSK ($n = 6$), duplicate publications ($n = 10$) and study protocol only ($n = 8$). In total, 18 publications based on 17 RCTs were included in the analysis. One study led to two publications reporting different interventions. Therefore, both publications were included [31, 32]. Study quality was mostly moderate ($n = 12$, 70%) and good ($n = 5$, 30%).

Figure 1. PRISMA study flow



Study characteristics

Baseline characteristics are summarized in Table 1. Most studies included patients with non-specific low back pain (NSLBP, n=15). One study also included patients with specific low back pain (SLBP) [33]. In two studies, LBP was not further specified [34, 35]. Ten studies included patients with LBP of up to 6 months (range of patients included 66 to 314) and followed them for four weeks to 12 months. Patients with LBP of six months and longer were the scope in six studies (range of patients included 71 to 598, disease duration 9 to 19 years) with a follow-up between three months [34] to three years [36]. Most RCTs compared interventions addressing fear avoidance beliefs to usual care or studies based on biomedical approaches (n= 14, 82%). Two RCTs compared treatments based on biomedical approaches [17, 37] and one RCT compared two fear avoidance addressing strategies [38]. Outcome measures were based in 14 studies (78%) on self-perceived measures (e.g. pain, disability), and in four studies (22%) on return to work (RTW) information.

Table 1. Baseline characteristics

Low back pain of less than six months duration									
Study	Setting	Diagnostic criteria	DD: days mean (SD)	Age mean (SD)	Important co-morbid condition	Treatment	n (f)	First episode (%)	FU (days) SIGN PF
Fritz 2002, [17] ID 5	Health care workers, recruited from occupational health care providers, U.S.A	Work-related NSLBP ≥ 3 , requiring modification of employment duties. 18% symptoms distal to the knee.	5.5 (4.6)	37.4 (10.4)	None	Control (n 37): PT (according to AHCPR clinical practice guidelines); Intervention (n = 41), PT according to Classification System by Delitto	78 (30)	50	1 month + 10
George, Brennan, 2006, [39] ID 7	2 RCT in general population and 1 RCT military personnel, recruited from outpatient physical therapy clinics communities, U.S.A	Acute NSLBP (≤ 30 days) + in treatment arms. Leg pain in 32.3%	13.3 (10.3)	36.2 (10.3)	None	3 RCT pooled: Trial 1 (n 75): manipulation + exercise; Trial 2 (n 59) Stabilization exercise, Trial 3 (n 31): direction-specific exercise	165 (83)	23.7	1 month + 7
Jellema, 2006, [45] ID 12	Randomization on level GP practice. Participating GPs selected 10 consecutive patients, Netherlands	NSLBP, 14.6 % radiating below the knee	UC 14 (10.4); MIC 13 (11.9)	UC 42 (12); MIC 43.4 (11.1)	None	UC (n 171) vs. MIC (n 143)	314 (149)	N.R.	6 weeks + 8

Low back pain of less than six months duration

Study	Setting	Diagnostic criteria	DD: days mean (SD)	Age mean (SD)	Important co-morbid condition	Treatment	n (f)	First episode (%)	FU (days)	SIGN	PF
Hancock, James, 2009, [31] ID 9B	Patient presenting in 40 general practitioners working in primary practice across Sydney, Australia	NSLBP: 12. rib to buttock crease with moderate pain and moderate disability (SF-36, item 7/8)	9.13 (9.31)	40.7 (15.6)	None	Diclophenac / Placebo + sham manipulation / spinal manipulative therapy;	239 (105)	previous episodes mean (sd): 3.7 (6.4)	Until to recovery	++	6
Hancock, Davies, 2009, [32] ID 9C	Same study as Hancock, James 2009					Spinal manipulation + Diclophenac / Placebo (n 119) vs. Placebo Spinal manipulation + Diclophenac / Placebo (n 120)			Time to recovery	++	6
Staal, 2008, [40] ID 17	Occupational health care setting Dutch airline workers, Netherlands	NSLBP (≥ 4 weeks duration), full or partial absence from work	Median 8 weeks (IQR 6-12 weeks)	38 (8.5)	None	UC (n 67) vs. Behaviorally-oriented graded activity intervention (n 67)	134 (8)	N.R.	1 year	++	6
George, Zepplier, 2008, [33,44] ID 8	Patients referred for rehabilitation to three participating University of Florida affiliated clinics, U.S.A.	NSLBP + SLBP Quebec Task Force on Spinal Disorders (QTFSD). Leg pain TBC 38%, GA 49%, GX 42%	TBC 47 (34), GA 41 (39), GX 69 (49)	37.5 (14.9)	None	Treatment-based classification (TBC) physical therapy vs. TBC + graded activity (GA) vs. TBC + graded exposure (GX)	108 (74)	TBC 50%, GA 31%, GX 33%	1 and 6 months	+	8

Low back pain of less than six months duration

Study	Setting	Diagnostic criteria	DD: days mean (SD)	Age mean (SD)	Important co-morbid condition	Treatment	n (f)	First episode (%)	FU (days)	SIGN	PF
Storheim, 2005, [41] ID 18A	Recruitment: National Insurance Offices, GPs in 2 counties. Total population of about 150,000 people near Oslo, Norway	NSLBP, sick-listed from a permanent job and receiving between 50% and 100% compensation listing for LBP for 12 weeks	8-12 weeks sick listed due to LBP, prior no sick listing for LBP for 12 weeks	41 (10)	None	UC, PT, CBT as co- variant: no influence on the predictor model	93 (51)	N.R.	1 year	+	12
Georg Donald, 2003, [42] ID 6	Patients referred by physicians to outpatient PT for evaluation and treatment, Rehab Center Florida, U.S.A	NSLBP ± leg pain (non-radicular) < 8 weeks. Leg pain 62%, Onset sudden 53%	27.3 days SD 16.2 days	38.2 (10.1)	None	UC (n 32) vs. Fear Avoidance Treatment (n 34)	66 (38)	32-47%	1 and 6 months	+	7
Burton, 1999, [47] ID 2	Six primary care practices in the northeast section of England, 5 National Health Service GP, 1 pain osteopathic practice	New episode of acute or recurrent NSLBP, with or without referred leg pain	Acute to sub-acute (<3 months)	Control 44.7 (12.2); Experimental 42.6 (10.9)	None	Control Booklet (n 79) vs. Experimental Booklet (n 83)	162 (89)	N.R.	3 months	+	5
Woods, 2008, [46] ID 20B	Recruited via newspaper, e-mail advertisements, posters	NSLBP + TSK score ≥ 38	Acute to sub-acute (<7 weeks)	46.5 (11.9)	None	GivE protocol + GA, GA only and Waiting list	83 (29)	N.R.	2 months	+	6

Low back pain of less than six months duration

Study	Setting	Diagnostic criteria	DD: days mean (SD)	Age mean (SD)	Important co-morbid condition	Treatment	n (f)	First episode (%)	FU (days)	SIGN	PF
Klaber Moffett, 2004, [43] ID 13A	Referral by 87 participating GPs in the York area, UK	NSLBP, 4 weeks to 6 months duration	4 weeks to 6 months	41.9 (8.9)	None	UC GP (n 98) vs. Exercise program (n 89)	187 (106)	N.R.	1 year	+	5

Low back pain of six months or longer

Study	Setting	Diagnostic criteria	DD	Age mean (SD)	Important co-morbid condition	Treatment	n (f)	First episode (%)	FU (days)	SIGN	PF
Ang, 2010, [34] ID 1	Electronic medical record identification, invitation. Enrolment GP clinic Indiana University (IU) + Veterans Administration (VA) MC, U.S.A.	LBP (60%), hip or knee (40%), ≥ 3 months despite analgesic treatment, at least moderate severity (BPI score of ≥/5)	Intervention 8 years; UC 10 years	Intervention 55.2 (12.6); UC 55.8 (11)	Depression, Hip and Knee pain 40%	Intervention (n 123): step 1 optimized antidepressant therapy (12 weeks), step 2 pain self-management program (12 weeks) vs. UC (n 127) vs. Control: not depressed	250 (132)	N.R.	3 months	++	7
Rasmussen-Barr, 2012, [36] ID 3B	Subjects with LBP seeking care at a primary health care setting, a private physiotherapy clinic or primary health care setting were included	NSLBP (costal margin - inferior gluteal folds, no leg pain), recurrent for 8 weeks, still at work. Before episode ≥ 1 year pain-free.	10 years (range 1-38)	38 (11)	None	Reference group (advice) (n 35), vs. Exercise Group (n 36)	71 (36)	27%	6, 12 and 36 months	+	10
Mannion, 1999, [37] ID 16	Hospital based outpatient treatment, recruitment by advertisement, Switzerland	NSLBP± referred pain (non-radicular), continual or recurrent, ≥3 months, causing absence from work or solicitation of medical attention	10.9 years (9.4)	45.1 (10.0)	None	Modern active PT (n 49) vs. muscle conditioning on training devices (n 49) vs. low-impact aerobics (n 50)	148 (84)	N.R.	6 months	+	14

Low back pain of six months or longer

Study	Setting	Diagnostic criteria	DD	Age mean (SD)	Important co-morbid condition	Treatment	n (f)	First episode (%)	FU (days)	SIGN	PF
van der Hulst, 2008, [49] ID 19	Patients admitted to outpatient multidisciplinary Back Rehabilitation Program (MR), Netherlands	NSLBP at least 3 months	Median 60 months, UC 48months, range 559 days, MR 72months, range 380d	39 (10)	None	UC (n 84) vs. Multidisciplinary rehabilitation(MR)(n 79)	163 (64)	N.R.	End of treatment and 4 months	++	7
Underwood, 2011, [38] ID 14C	56 general practices in seven regions across England	NSLBP, at least moderate troublesomeness for > 6 weeks	13 years since first onset (13)	53.8 (15)	None	AM vs. AM+CBT: treatment as co-variant	598 (356)	N.R.	1 year	++	10
Magnussen, 2007, [35] ID 15	Participants on disability pension >1 year, county of Hordaland, Norway; identified through the National Insurance Administration	LBP not further specified	Control group: Disability pension for 11.6 years (5.8), Intervention 9.8 years (4.8)	49 (5.45)	All on disability pension for a mean 11 years	Control group (n 44) vs. Intervention including FAB reducing information (n 45). 29 patients completed the intervention	89 (56)	N.R.	1 year	+	8

LBP: low back pain; NSLBP, non-specific low back pain; SLBP, specific low back pain; PF, prognostic factor reporting of the total of 16 possible domains; SIGN, SIGN quality rating: ++ high quality, + moderate quality; DD, Disease Duration; FU, Follow-up; TBC, treatment-based classification; GA: graded activity; GX, graded exposure; PT, physical therapy; UC, usual care; MIS, minimal intervention strategy; AM, active management; CBT, cognitive behavioral therapy; GiVE, Graded in vivo exposure; QA, Questionnaire used; FAB, Fear Avoidance Beliefs; FABQ, fear avoidance questionnaire; FABQ-P, FABQ physical activity sub-scale; FABQ-W: FABQ work sub-scale; TSK, Tampa Scale of Kinesiophobia; RTW, return to work

Influence of baseline fear avoidance beliefs on treatment outcome in patients with LBP of up to 6 months

Seven out of nine RCTs (77%) found baseline fear avoidance beliefs to influence treatment outcome (Table 2) [17, 33, 39-43]. Patients with high baseline scores had more pain and / or disability (four studies [33, 39, 42, 43], Figure 2) and were less likely to return to work (three studies [17, 40, 41]). Fear avoidance work scores (FABQ-W) was assessed in seven studies and predicted poor work related outcomes in four studies [17, 40, 41] and disability in one study [39]. High FABQ-W scores were not associated with RTW in one study [42] and did not predict pain in two studies [31, 32]. Of the seven studies that used FABQ-P, high FABQ-P scores were associated with less RTW in one study [40] and more disability in two studies [42, 43]. One study found a combination of high fear avoidance beliefs and catastrophizing to be associated with more pain and disability while fear avoidance (FABQ-P) alone was not [33, 44]. Three studies found no influence of FABQ-P on RTW [17, 41] and pain [45]. TSK was only used once. In this study only patients with high TSK scores (>38) were included and in this group there was no association between TSK and pain [46].

Mediating effect of change in fear avoidance beliefs on treatment outcome in patients with LBP of up to 6 months

In all three RCTs that investigated mediating effects found an association of decrease in fear avoidance beliefs during treatment with more RTW [40] (FABQ-P, FABQ-W, TSK), less pain [33, 44] (FABQ-P, not FABQ-W) and disability [47] (FABQ-P) (Table 2, Figure 3)

Table 2. Summary of influence of fear avoidance beliefs on treatment outcomes and efficacy in patients with LBP of less than six months

ID	Study, Year	Scale: mean (SD)	Predictor	Mediator	Moderator
5	Fritz 2002	FABQ-W (mean 27.9, SD 8.6), FABQ-P (mean 18.9, SD 5.8)	+ High FABQ-W score predict non-RTW (return to work without restriction) at 4 weeks (OR 1.2 (CI 1.07-1.34)).	∅	- RTW: No differences between control PT (AHCPR clinical practice guidelines) and intervention (PT Classification System by Delitto)
7	George, Brennan, 2006	FABQ-W (mean 15.1, SD 10.5) FABQ-P (mean 16.6, SD 4.6)	- FABQ-P non-predictive (N.R.) for RTW (return to work without restriction). + High FABQ-W predicted less disability (ODI) reduction in men in all treatment groups (Beta 0.43 (CI 0.03-0.83))	∅	- No moderator for disability (ODI) reduction in all treatments (manipulation + exercise; stabilization exercise; direction-specific exercise)
12	Jellema, 2006	FABQ-P (mean 18.8, SD 1.8)	- FABQ-W women not predictive for ODI (N.R.). FABQ-P not reported ∅	∅	+ High FABQ-P in MIS more likely to have a CMID in disability (RMQ, OR 1.07, CI 0.99-1.16) compared to UC
9B+ C	Hancock, James, 2009	FABQ-P (mean 17.0, SD 5.4) FABQ-W (mean 14.5, SD 10.4)	- FABQ-P non-predictive for days to recovery from pain (NRS) - FABQ-W univariate associated, multivariate not associated with days to recovery from pain (N.R.)	∅	- FABQ-P in UC non-predictive for RMQ (N.R.) + In high FABQ-P NSAID more effective than placebo: HR to have a day without pain (NRS) 1.059 (1.002-1.118) and for 7 days HR 1.042 (0.988-1.10) - FABQ-W non-predictive: HR for 1 day without pain (NRS) 0.996 (CI 0.972-1.02), HR for 7 days without pain 0.992 (CI 0.967 – 1.018))

ID	Study, Year	Scale: mean (SD)	Predictor	Mediator	Moderator
17	Staal, 2008	FABQ-P (mean 16.9, SD 5.1) modified FABQ-W (range 0-36, mean 21.8, SD 9.2) TSK (mean 39.4, SD 6.8) Cut-offs: FABQ-P ≥ 16 , FABQ-W ≥ 26	<p>FABQ-P ≥ 16 less likely to RTW (>27 days full return to regular work) (N.R.)</p> <p>FABQ-W ≥ 26 predictive: less likely to RTW (N.R.). TSK not reported</p>	<p>Reduction of FABQ-P at 3 months increase the chance of RTW (FABQ-P HR 2.0 (CI 1.2-3.5))</p> <p>Reduction in FABQ-W at 3 months increase the chance of RTW FABQ-W 2.2 (CI 1.3-3.7)</p> <p>Reduction in TSK at 3 months increased chance of RTW, TSK 2.3 (CI 1.4-3.9)</p>	<p>Compared to UC behavioural oriented GA increased the chance for RTW (>27 days full return to regular work). The effect was decreased by fear avoidance beliefs: low fear avoidance RTW HR 2.6-3.0 (FABQ-P <16 [HR 3.0 (CI 1.6-5.5)], FABQ-W <26 [HR 1.5-4.6]]) compared to high fear avoidance HR for RTW 1.6 (FABQ-P ≥ 16 [HR 1.6 (1.0-2.7)], FABQ-W ≥ 26 [HR 1.6 (CI 1.0-2.7)]).</p>
8A + B	George, Zeppieri, 2008; Benecuk 2012	FABQ-P (mean 13.3, SD 11.1) FABQ-W (mean 37.3, SD 26.7) Cut-off FABQ-P >14	<p>Patients with high FABQ-P and high catastrophizing experienced more pain (NRS) and disability (ODI) than patients with high FABQ-P alone or patients with low FABQ-P. Further they experienced more decrease in pain and disability at follow-up</p>	<p>Reduction in FABQ-P was associated with less pain (NRS) at 1 month [beta= 0.26 (CI -0.46 to -0.06)], 6 months [beta= 0.25 (-0.47 to -0.03)]</p>	\emptyset
18A	Storheim, 2005	FABQ-W (patients with RTW mean 25.5, SD 9.0; patients not RTW mean 33.6, SD 5.8)	<p>Higher FABQ-W predictive for longer time to RTW (returned to full-time work according to data from the National Insurance Offices) at 1 year [HR 1.082 (CI N.R.)] for all treatments \downarrow</p> <p>FABQ-P non-predictive (N.R.)</p>	<p>Reduction in FABQ-W was not associated with less pain (NRS)</p>	<p>No differences between treatments (UC, PT, CBT)</p>

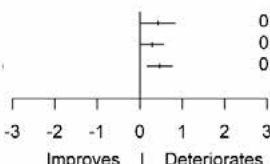
ID	Study, Year	Scale: mean (SD)	Predictor	Mediator	Moderator
6	Georg Donald, 2003	FABQ-W (mean 14.3, SD 12.1) FABQ-P (mean 15.3 (5.5))	High FABQ-P predicted more disability (ODI) at 1 month (beta 0.292, p= 0.042) and 6 months (beta 0.469, p= 0.002). FABQ-W non-predictive for disability (ODI) (N.R.) (Graf 12).	\emptyset	+ FAB-treatment reduces FABQ-P baseline effect on disability (ODI) at 1 month (beta -0.13) and 6 months (beta -0.15) compared to standard PT in the FAB-PT subgroup treatment showed more improvement in disability (ODI) that in the PT subgroup(Figure 4)
2	Burton, 1999	FABQ-P (mean N.R.) cut-off m/s >14	- \emptyset	+ FABQ-P baseline >14 and reduction of FABQ-P of >4 points at 2 weeks and 3 months are more likely to improve in RMQ (≥ 3 points) at 3 months [RR to deteriorate at 0.56 (CI 0.31-1.03) and 0.46 (CI 0.26-0.8) respectively]	+ Experimental booklet more likely to have CMID in FABQ-P (>4 points): at 2 weeks (RR 2.72), 3 months at 2 weeks (RR 1.53), 1 year (RR 1.47) compared to control booklet.
20B	Woods, 2008	TSK cut-off > 38 points = inclusion criteria	- TKS non-predictive for pain related disability (PDI) at 2 months	\emptyset	(+) Mean effect Give for pain related disability (PDI) pre- to 6 month 20.4 -> 12.2, GA 23.3 -> 23.7, Control 24 -> 18.8. Give better than GA
13A	Klaber Moffett, 2004	FABQ-P (mean 1.8, SD 5), FABQ-W (14-24 points 50%) Cut-off high FABQ-P: ≥ 14	+ FABQ-P <14 non-predictive for disability (RMQ) in all treatments [Exercise group (EG) and UC]] at 1 year	\emptyset	+ FABQ-P ≥ 14 : EG more likely to have CMID RMQ (≥ 2 points) compared to UC. The effect increases over time [OR 6 months 1.98 (CI 0.72-5.48), 12 months 3.58 (CI 1.30-9.84)].

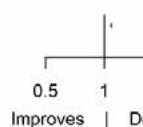
\emptyset , not investigated LBP: N.R., not reported; m/s, median split; low back pain; NSLBP, non-specific low back pain; SLBP, specific low back pain; FU, Follow-up; FAB, fear avoidance beliefs; FABQ, fear avoidance questionnaire; FABQ-P, FABQ physical activity sub-scale; FABQ-W, FABQ work sub-scale; TSK, Tampa Scale of Kinesiophobia; TBC, treatment-based classification; GA: graded activity; GX, graded exposure; PT, physical therapy; UC, usual care; MIS, minimal intervention strategy; AM, active management; CBT, cognitive behavioral therapy; Give, Graded in vivo exposure; OR, odds ratio; RR, relative risk CMID, clinical meaningful important difference; ODI, Oswestry Disability Index: higher score indicates more disability, range 0-100% (Fairbank 1980): CMID = ≥ 12 points reduction; RTW, return to work; PDI, pain disability index, range 0-70 (Pollard, 1984): higher score indicates more disability; RMQ, Roland Morris Questionnaire: higher score indicates more disability: CMID = ≥ 2 -3 point reduction, $\geq 30\%$ change, range 0-24 (Morris 1984)

Figure 2. Influence of fear avoidance beliefs on treatment outcomes

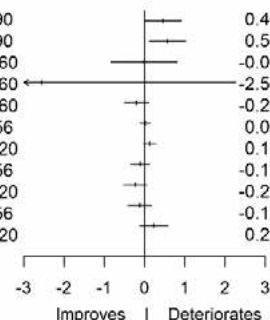
Low Back Pain up to 6 months

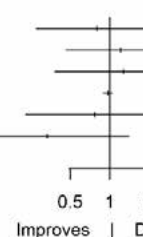
Author & Year	Questionnaire	Cut-off	Outcome	Odds Ratio	OR	95%-CI
Fritz 2002	FABQ-W	No	WR		1.2	[1.07; 1.34]

Author & Year	Questionnaire	Cut-off	Outcome	FU	β	β	95%-CI
George, B 2006	FABQ-W	No	Disability	30		0.43	[0.03; 0.83]
George, D 2003	FABQ-P	No	Disability	28		0.29	[0.01; 0.57]
George, D 2003	FABQ-P	No	Disability	180		0.47	[0.17; 0.77]

Author & Year	Questionnaire	Cut-off	Outcome	FU	Hazard Ratio	HR	95%-CI
Storheim 2005	FABQ-W	No	WR	360		1.08	

Low Back Pain 6 months and longer

Author & Year	Questionnaire	Cut-off	Outcome	FU	β	β	95%-CI
Ang 2010	TSK short	No	Pain	90		0.46	[0.01; 0.91]
Ang 2010	TSK short	No	Disability	90		0.57	[0.12; 1.02]
Underwood 2011	FABQ-P	≥ 14	Disability(#)	360		-0.01	[-0.84; 0.82]
Underwood 2011	FABQ-P	≥ 14	Disability(\$)	360		-2.56	[-10.35; 5.23]
Underwood 2011	FABQ-P	≥ 14	Pain	360		-0.20	[-0.50; 0.10]
v.d. Hulst 2008	TSK	No	Disability	56		0.02	[-0.12; 0.15]
v.d. Hulst 2008	TSK	No	Disability	120		0.13	[-0.03; 0.29]
v.d. Hulst 2008	TSK	No	SF36D	56		-0.11	[-0.35; 0.13]
v.d. Hulst 2008	TSK	No	SF36D	120		-0.23	[-0.52; 0.06]
v.d. Hulst 2008	TSK	No	SF36M	56		-0.12	[-0.41; 0.17]
v.d. Hulst 2008	TSK	No	SF36M	120		0.23	[-0.12; 0.58]

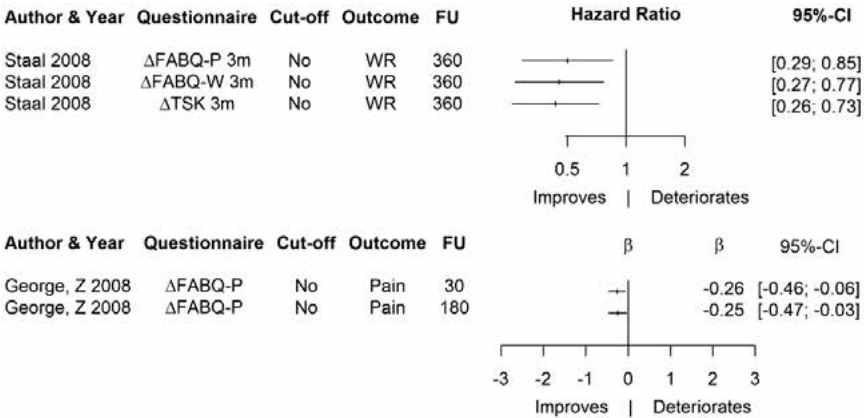
Author & Year	Questionnaire	Cut-off	Outcome	FU	Odds Ratio	OR	95%-CI
Rasmussen 2012	FABQ-P	≥ 14	Disability	360		0.80	[0.27; 2.35]
Rasmussen 2012	FABQ-P	≥ 14	Pain	360		1.21	[0.46; 3.17]
Rasmussen 2012	FABQ-P	≥ 14	Disability	1080		1.28	[0.38; 4.31]
Rasmussen 2012	FABQ-P	≥ 14	Pain	1080		0.97	[0.89; 1.06]
Magnussen 2007	FABQ-P	< 14.8	WR	360		0.77	[0.23; 2.61]
Magnussen 2007	FABQ-W	< 35	WR	360		0.33	[0.08; 1.40]

§ RMQ; # MVK; FU, follow-up given in days; WR, work-related; OR, odds ratio; β , Beta; SF-36D, SF-36 physical health sub-score: higher score indicates higher level of functioning; SF-36M, mental health sub-score: higher score indicates higher level of functioning

Moderating effect of fear avoidance beliefs on treatment efficacy in patients with LBP of up to 6 months

Ten RCTs investigated moderating effects of fear avoidance beliefs on specific treatments. Six studies found moderating effects on one or more treatments (Table 2, Figure 4). Baseline fear avoidance beliefs were associated with more pain or disability and less RTW in treatment arms based on biomedical concepts and fear avoidance strategies (Table 3, Figure 4). Higher scores in fear avoidance beliefs was associated with increased treatment efficacy of fear avoidance beliefs addressing treatment strategies. Interventions providing information to reduce fear avoidance beliefs (the Back Book, minimal intervention by GP’s) led to less reported disability in patients with high fear avoidance beliefs (FABQ-P) at short term follow-up [45, 47], Graded activity based on operant conditioning described by Vlaeyen et al. [48] was less effective in patients with high fear avoidance beliefs in two studies [40, 46]. Woods et al. compared graded activity alone to additional CBT based education and found an increased efficacy of combining the two in patients with high fear avoidance beliefs [46]. RCTs that combined fear avoidance addressing information, graded exposure and physical treatments found and increased treatment efficacy in patients with high fear avoidance beliefs with less disability [42, 43]. In treatment arms based on biomedical concepts eight out of nine (89%) found that baseline fear avoidance beliefs were associated with poor outcome for pain (n= 1), disability (n=4) and RTW (n=4) at short and long term follow-up (Table 3). Compared to the Back Book a booklet based on biomedical concepts (Handy Hints Booklet) was associated with less clinical meaningful important change (CMID) at follow-up compared to patients with the Back Book [47].

Figure 3. Mediating effect of change in FAB during treatment in subacute LBP



Δ change from baseline to follow-up; 3m, three months

* 1 day without pain

+ 7 days without pain

Table 3. Influence of fear avoidance beliefs on treatment arms addressing fear avoidance beliefs or based on biomedical concepts

Fear avoidance addressing treatments						
Treatment	FAB predictor	FAB moderator	Outcome	Standardized treatment	Adherence controlled	FU (months) Patients Drop-out (%)
The Back Book (BB) and GP care ID 2 [47]	∅	FABQ-P ↑: the BB efficacy ↑	CMID in disability (RMQ)	yes	no	3 162 22
Minimal intervention (MIS) by GP and GP care ID 12 [45]	∅	FABQ-P ↑: MIS efficacy ↑	CMID in disability (RMQ)	yes	no	1.5 314 19
Graded activity (GA, operant conditioning, Vlaeyen et al. 2002) ID 17 [40]	FABQ-P ↑ / -W ↑: less RTW	FABQ-P ↑ / -W ↑: GA efficacy ↑	RTW	yes	yes	12 134 11
Graded activity (GA, operant conditioning, Vlaeyen et al. 2002) ID 20 [46]	no (TSK <39 not included)	TSK >38 no pain reduction	Pain (PDI)	yes	yes	2 83 46
TBC-PT (treatment-based classification, Delitto et al. [68]) + graded activity (GA, George et al, 2003) ID 8 [33]	FABQ-P ↑ more pain and disability	∅	Pain and Disability	yes	yes	6 108 44
TBC-PT (treatment-based classification, [68]) + graded exposure (Give, based on Vlaeyen et al. 2001[48]) ID 8 [33]	FABQ-P ↑ more pain and disability	∅	Pain and Disability	yes	yes	6 108 44
Fear avoidance addressing PT (FAB-PT, Vlaeyen et al. 2000 [69]) using graded (GA) exercise and the Back Book (BB) ID 6 [42, 44]	FABQ-P ↑ more disability	FABQ-P ↑: FAB-PT efficacy ↑	Disability (ODI)	yes	yes	6 66 n.r.
Exercise with CBT approach (EG) ID 13 [43]	FABQ-P ↑ less CMID	FABQ-P ↑ EG efficacy ↑ with more CMID than UC	CMID in disability (RMQ)	yes	no	12 187 9
CBT and PT ID 18 [41]	FABQ-W ↑ less RTW	-	RTW	yes	yes	12 93

CBT based education (Psych. Student under supervision) on FAB (GivE, Vlaeyen et al., 2001 [48]) followed by GA ID 20 [46]	no (TSK <39 not included)	GivE + GA: pain reduction ↑ than GA	Pain (PDI)	yes	yes	6	83	46
Fear avoidance addressing PT (FAB-PT, only in 1 of the 3 trials) ID 7 [39]	FABQ-W ↑ less disability in men	∅	Disability (ODI)	n.r.		1	165	44
Exercise (PT provided back training, The Norwegian Aerobic Fitness Model) ID 18 [41]	FABQ-W ↑ less RTW	–	RTW	yes		12	93	39
NSAID vs. Placebo ID 9C [31, 32]	–	FABQ-P ↑ : NSAID efficacy ↑	Pain (days without pain)	yes	to recovery		239	2
Treatments based on biomedical concepts								
No FAB addressing therapy	FAB predictor	FAB moderator	Outcome	Standardized treatment	Adherence controlled	FU (months)	Patients	Drop-out (%)
UC by GPs ID 12 [45]	∅	–	CMID in disability (RMQ)	no	no	1.5	314	19
UC by GPs ID 13 [43]	FABQ-P ↑ less CMID	FABQ-P ↑ UC less CMID	CMID in disability (RMQ)	no	no	12	187	9
UC by GPs ID 18 [41]	FABQ-W ↑ less RTW	no	RTW	no	no	12	93	39
UC by occupational physician ID 17 [40]	FABQ-P ↑ / -W ↑ : less RTW	difference between UC and GA ↓	RTW	no	yes	12	134	11
Active exercise and / or manipulation ID 7 [39]	FABQ-W ↑ less RTW	–	Disability (ODI)	n.r.	n.r.	1	165	44
TBC (treatment-based classification by Delitto et al. 1995 [68]): specific exercise, manipulation or mobilization, lumbar stabilization or traction ID 8 [33]	FABQ-P ↑ + Catastroph ↑ : more pain and disability	–	Pain and Disability	yes	yes	6	108	44

Treatments based on biomedical concepts						
No FAB addressing therapy	FAB predictor	FAB moderator	Outcome	Standardized treatment	Adherence controlled (months)	FU Patients Drop-out (%)
PT based on recommendations of the Agency for Health Care Policy and Research Guidelines (AHCPR) ID 5 [17]	FABQ-W ↑ less RTW	–	RTW	yes	no	1 78 n.r.
TBC-PT(treatment-based classification [68]) manipulation, exercise, traction ID 5 [17]	FABQ-W ↑ less RTW	–	RTW	yes	no	1 78 n.r.
TBC-PT (treatment-based classification [68]) + Handy Hints (HH) educational pamphlet ID 6 [42, 44]	FABQ-P ↑ more disability	–	Disability (ODI)	yes	yes	6 66 n.r.
Handy Hints Booklet (HH) and GP care ID 2 [47]	∅	FABQ-P ↑: HH less likely	CMID in disability (RMQ)	yes	no	3 162 22
Spinal Manipulation 9B [31]	–	∅	Pain (days without pain)	yes	yes	to recovery 239 2

∅, not investigated; –, no effect FU, follow-up duration in months; TBC, treatment-based classification; GA: graded activity; PT, physical therapy; UC, usual care; MIS, minimal intervention strategy; AM, active management; CBT, cognitive behavioral therapy; GivE, Graded in vivo exposure, ODI, Oswestry Disability Index; RTW, return to work; CPG, chronic pain grade (von Korf) questionnaire: higher score indicates more pain; GCPS, graded chronic pain scale; PDI, pain disability index (Pollard, 1984); RMQ, Roland Morris Questionnaire; SF-36D function physical health sub-score: higher score indicates higher level of functioning

Influence of baseline fear avoidance beliefs on treatment outcome in patients with LBP of 6 months and longer

Out of five RCTs two found high baseline fear avoidance beliefs to be associated with increased pain (TSK [34]) and RTW (FABQ-W [35]) (Table 4, Figure 2). No association of fear avoidance beliefs with disability at follow-up was found in three studies (FABQ-P [36], TSK [49], FABQ-P [38]).

Mediating effect of change in fear avoidance beliefs on treatment outcome in patients with LBP of 6 months and longer

One study investigated the influence of a change of fear avoidance beliefs on treatment outcome [37]. While changes in FABQ-P and FABQ-W scores were non-predictive for disability, a small effect was found for change in FABQ-W on pain reduction (Table 4).

Moderating effect of fear avoidance beliefs on treatment efficacy in patients with LBP of 6 months and longer

In patients with chronic low back pain, only one study found a multidisciplinary rehabilitation in patients with high fear avoidance beliefs (TSK) to lead to less disability and an increased chance for RTW (Table 5, Figure 4) [35]. No moderating effect was found on all other treatments investigated [34-38]. Underwood et al. compared active advice including the Back Book to active advice with additional group therapy session of CBT and found no influence of fear avoidance beliefs on treatment efficacy [38]. A summary of the description of the interventions extracted from the publications is provided in Appendix 2.

Table 4. Summary of influence of fear avoidance beliefs on treatment outcomes and efficacy in patients with LBP of six months and more

ID	Study, Year	Scale: mean (SD)	Predictor	Mediator	Moderator
1	Ang, 2010	Modified 10-item TSK (mean value N.R.)	High TSK associated with an increased pain [GPCS pain intensity, beta 0.46, (CI 0.0092-0.91)] and GPCS activity interference [beta 0.57 (CI 0.119-1.02)] at 3months	+	∅
3B	Rasmussen-Barr, 2012	FABQ-P, (median 13, range 0-24) cut-off high ≥14	Baseline FABQ-P ≥ 14 does not predict pain (VAS) and disability (ODI) after 12 (pain OR 1.21 (CI 0.46-3.17), disability OR 0.8 (CI 0.27-2.35) and 36 months [pain OR 0.97 (CI 0.89-1.06)], disability OR 1.28 (0.38-4.31)	-	∅
16	Mannion, 1999	FABQ-P (mean 13.8, SD 5.3) FABQ-W (mean 15.7, SD 11.3)	∅	(+) Decrease of FABQ-W associated with pain (VAS) reduction at 6 months [10% variance in combination with reduction of disability (RMQ)]	- No difference between “modern active PT”, muscle conditioning on training devices and low-impact aerobic
				-	- Decrease in FABQ-P and FABQ-W non-predictive for change in RMQ.

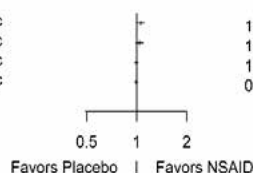
ID	Study, Year	Scale: mean (SD)	Predictor	Mediator	Moderator
19	van der Huilst, 2008	TSK (mean 39, SD 7)	TSK not associated with RMQ, SF-36D and SF-36M (end of treatment) and RMQ at 4 months [Beta- 0.13, (CI -0.29 to 0.027)]. Tendency to more decrease in SF-36D [Beta 0.23 (CI -0.06 to 0.52)] and worsening of SF-36M [Beta -0.23 (CI- 0.58 to 0.123)] at 4 months	∅	<p>+</p> <p>Multidisciplinary rehab (MR) better than UC: Higher TSK leads to less disability (more increase in SF-36D (beta 0.16 (CI 0.018 – 0.762) in MR after 4 weeks.</p> <p>–</p> <p>No influence on treatment efficacy for RMQ and SF-36M at end of treatment and 4 weeks FU or SF-36D at end of treatment.</p> <p>–</p> <p>No difference between AM vs. AM+CBT</p> <p>–</p> <p>No difference between control group and intervention with FAB reducing information</p>
14C	Underwood, 2011	FABQ-P (mean 13.5, SD 6.4) cut-off high: ≥14	FABQ-P non-predictive for RMQ, MVK Pain and Disability at 1 year	–	–
15	Magnussen, 2007	FABQ-P (cut-off m/s < 14.8, mean 14.6, SD 6.1) FABQ-W (m/s < 35, mean 30.8, SD 10.7)	<p>FABQ-W < 35: 3 times more likely to have entered RTW process at 1 year (self-reported 3-point scale: "Have you succeeded in returning to work?": "Yes", "No" and "Still in the process"), but small sample size and wide CI (0.4–4.6), FABQ-P non-predictive</p>	+	–

∅, not investigated LBP: N.R., not reported; m/s, median split; low back pain; NSLBP, non-specific low back pain; SLBP, specific low back pain; FU, Follow-up; FAB, fear avoidance beliefs; FABQ, fear avoidance questionnaire; FABQ-P, FABQ physical activity sub-scale; FABQ-W, FABQ work sub-scale; TSK, Tampa Scale of Kinesiophobia; TBC, treatment-based classification; GA: graded activity; GX, graded exposure; PT, physical therapy; UC, usual care; MIS, minimal intervention strategy; AM, active management; CBT, cognitive behavioral therapy; GiVE, Graded in vivo exposure; OR, odds ratio; RR, relative risk CMID, clinical meaningful important difference; ODI, Oswestry Disability Index: higher score indicates more disability, range 0-100% (Fairbank 1980): CMID = ≥12 points reduction; RMQ, Roland Morris Questionnaire: higher score indicates more disability; CMID = ≥2-3 point reduction, ≥30% change, range 0-24 (Morris 1984) RTW, return to work; CPG, chronic pain grade (von Korf 1992) questionnaire: higher score indicates more pain; MVK, modified von Korf (Underwood 1999), scale 0-100%, SF-36D, physical health sub-score: higher score indicates higher level of functioning; SF-36M, mental health sub-score: higher score indicates higher level of functioning (Aaronson 1998)

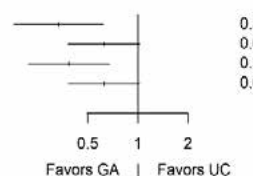
Figure 4. Moderating effect of fear avoidance beliefs on treatment efficacy

Low Back Pain of up to 6 months

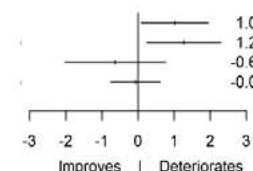
Author & Year	Treatment	Questionnaire	Cut-off	Outcome	FU	Hazard Ratio	HR	95%-CI
Hancock, D 2009	NSAID vs P	FABQ-P	No	Pain(*)	to rec		1.06	[1.00; 1.12]
Hancock, D 2009	NSAID vs P	FABQ-P	No	Pain(+)	to rec		1.04	[0.99; 1.10]
Hancock, D 2009	NSAID vs P	FABQ-W	No	Pain(*)	to rec		1.00	[0.97; 1.02]
Hancock, D 2009	NSAID vs P	FABQ-W	No	Pain(+)	to rec		0.99	[0.97; 1.02]



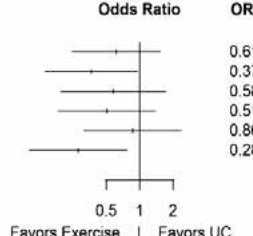
Author & Year	Treatment	Questionnaire	Cut-off	Outcome	FU	Hazard Ratio	HR	95%-CI
Staal 2008	GA vs UC	FABQ-P	<16	WR	360		0.33	[0.18; 0.62]
Staal 2008	GA vs UC	FABQ-P	>=16	WR	360		0.62	[0.38; 1.03]
Staal 2008	GA vs UC	FABQ-W short	<26	WR	360		0.38	[0.22; 0.67]
Staal 2008	GA vs UC	FABQ-W short	>=26	WR	360		0.62	[0.38; 1.03]



Author & Year	Treatment	Questionnaire	Cut-off	Outcome	FU	β	β	95%-CI
George, D 2003	PT	FABQ-P	No	Disability	28		1.02	[0.09; 1.95]
George, D 2003	PT	FABQ-P	No	Disability	180		1.27	[0.24; 2.30]
George, D 2003	FAB-PT	FABQ-P	No	Disability	28		-0.63	[-2.03; 0.77]
George, D 2003	FAB-PT	FABQ-P	No	Disability	180		-0.07	[-0.76; 0.62]

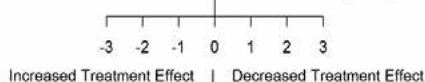


Author & Year	Treatment	Questionnaire	Cut-off	Outcome	FU	Odds Ratio	OR	95%-CI
Klaber 2004	Ex vs UC	FABQ-P	<14	Disability	42		0.61	[0.24; 1.54]
Klaber 2004	Ex vs UC	FABQ-P	>=14	Disability	42		0.37	[0.14; 0.96]
Klaber 2004	Ex vs UC	FABQ-P	<14	Disability	180		0.58	[0.19; 1.72]
Klaber 2004	Ex vs UC	FABQ-P	>=14	Disability	180		0.51	[0.18; 1.39]
Klaber 2004	Ex vs UC	FABQ-P	<14	Disability	360		0.86	[0.31; 2.37]
Klaber 2004	Ex vs UC	FABQ-P	>=14	Disability	360		0.28	[0.10; 0.77]



Low Back Pain of 6 months and longer

Author & Year	Treatment	Questionnaire	Cut-off	Outcome	FU	β	95%-CI
v.d. Hulst 2008	MR	TSK	No	Disability	56		8 [-0.12; 0.28]
v.d. Hulst 2008	MR	TSK	No	Disability	120		8 [-0.02; 0.38]
v.d. Hulst 2008	MR	TSK	No	SF36D	56		4 [-0.57; 0.09]
v.d. Hulst 2008	MR	TSK	No	SF36D	120		9 [-0.76; -0.02]
v.d. Hulst 2008	MR	TSK	No	SF36M	56		1 [-0.40; 0.42]
v.d. Hulst 2008	MR	TSK	No	SF36M	120		4 [-0.19; 0.67]



MIS, minimal intervention strategy; NSAID, non-steroidal anti-inflammatory drug; P, Placebo; Ex, exercise; UC, usual care; MR, multidisciplinary rehabilitation

Table 5. Effect of fear avoidance beliefs on treatment efficacy for fear addressing treatments

Fear avoidance addressing treatments						
Treatment	FAB predictor	FAB moderator	Outcome	Standardized treatment	Adherence controlled	FU (months) Patients Drop-out (%)
Active management advice (primary care nurse) + the Back Book ID 14 [38]	–	–	Pain and Disability	Yes	yes	12 598 15
Active management advice + the Back Book + 6 group therapy sessions with CBT ID 14 [38]	–	–	Pain and Disability	Yes	yes	12 598 15
Individual guidance how to improve physical fitness + 2 group sessions with lectures aiming to reduce FAB and motivational interviewing ID 15 [35]	FABQ-W ↑ less RTW	–	RTW	Yes	yes	12 89 36
Graded stabilizing exercises (protocol by Richardson et al., 1999 [70]) ID 3 [36]	–	–	Pain (VAS) and disability (ODI)	Yes	yes	36 71 23
Pain self-management program and antidepressant therapy ID 1 [34]	TSK ↑ more pain and disability	–	Pain (GCPS), Disability (GCPS)	Yes	no	3 250 0
Multidisciplinary rehabilitation (MR) ID 19 [49]	–	TSK ↑ MR better than waiting list (SF-36 function)	Disability (SF-36 function, RMQ)	Yes	yes	4 163 13
Treatments based on biomedical concepts						
No FAB addressing therapy						
Waiting list ID 19 [49]	–	∅	Disability (SF-36 function, RMQ)	No	no	4 163 13
Information about daily walks ID 3 [36]	–	∅	Pain (VAS) and disability (ODI)	Yes	yes	36 71 23

Fear avoidance addressing treatments

Treatment	FAB predictor	FAB moderator	Outcome	Standardized treatment	Adherence controlled	FU (months)	Patients Drop-out (%)
Advice should seek care for depressive symptoms ID 1 [34]	TSK ↑ more pain and disability	–	Pain (GCPs), Disability (GCPs)	No	no	3	250 0
Individual guidance how to improve physical fitness ID 15 [35]	FABQ-W ↑ less RTW	–	RTW	Yes	no	12	148 36
PT instruction for therapy ID 16 [37]	∅	–	Pain	Yes	yes	6	148 11
Muscle reconditioning group ID 16 [37]	∅	–	Pain	Yes	yes	6	148 11
Aerobics/stretching group ID 16 [37]	∅	–	Pain	Yes	yes	6	148 11

∅, not investigated; –, no effect FU, follow-up duration in months; TBC, treatment-based classification; GA: graded activity; PT, physical therapy; UC, usual care; MIS, minimal intervention strategy; AM, active management; CBT, cognitive behavioral therapy; GivE, Graded in vivo exposure, ODI, Oswestry Disability Index; RTW, return to work; CPG, chronic pain grade (von Korff) questionnaire: higher score indicates more pain; GCPs, graded chronic pain scale; PDI, pain disability index (Pollard, 1984); RMQ, Roland Morris Questionnaire; SF-36D function physical health sub-score: higher score indicates higher level of functioning

Differences between studies with non-predictive and predictive findings

Studies with non-predictive findings included on average more patients than studies with predictive findings (Table 6). This may be a spurious finding since it is mainly due to one study that included more than 580 chronic LBP patients [38]. The reporting of previously identified prognostic factors for chronic LBP could be improved in all studies [22]. Out of 16 domains predictive studies reported between 5 and 12 and non-predictive studies between 5 and 10 domains. Thus, reporting of prognostic factors did not differ between predictive and non-predictive studies.

Most often FABQ-P was investigated (15 times) followed by FABQ-W (10 times) and TSK (4 times). The most consistent predictor was the baseline FABQ-W value. All non-predictive studies had low baseline FABQ-W values (range of baseline mean values 14.3-14.5, Table 6) compared to the predictive studies (range of baseline mean values 15.1 - 30.8), indicating that non-predictive studies included a low risk population with little variability in FABQ-W. In non-predictive studies using FABQ-P, low baseline values could be responsible for non-predictive findings in two studies with chronic LBP (mean baseline value 13.5 [38], 13.0 [36]). The difference in the FABQ-P baseline values was not as pronounced as in the FABQ-W values. Eight studies (45%) used cut-off values with mixed findings. Most often the cut-off of ≥ 14 points for FABQ-P ($n = 5$, once median split) was used. Only two studies used FABQ-W cut-off values (≥ 26 [40] and median split of ≥ 35 [35]).

Table 6. Differences between predictive and non-predictive studies

	All		<6 months LBP		≥ 6 months LBP	
	+	-	+	-	+	-
Publications: n	9	5	7	2	2	3
n: range	78-250	71-598	78-187	83-239	89 - 250	71-598
Age: mean (range)	36.2-55	40.7-54	36.2-41	40.7-46.5	49 - 55	38-54
Disease duration: range of mean disease duration (days)	5 days - 11 years	9 days - 13 years	5.5-120	9-35	9-11 years	5-13 years
Follow-up: range (months)	1-12	2-36	1-12	to recovery - 2 months	3-12	4-36
Prognostic Domain [#] : range	5-12	5-10	5-12	5-6	7-8	7-10
FABQ-W: range baseline mean values	15.1-30.8	14.3-14.5	15.1-27.9	14.3-14.5	30.8	-
FABQ-P: range baseline mean values	13.3-16.9	13.5-18.9	13.3-16.9	17.0-18.9	-	13.5-14.6
TSK: range baseline mean values	39.4	39	39.4	-	N.R.	39
SIGN Quality	+(n=7), ++(n=2)	+(n=3), ++(n=2)	+(n=6), ++(n=1)	+(n=1), ++(n=1)	+(n=1), ++ (n=1)	+(n=2), ++(n=1)

[#]Number of prognostic domains reported (modified from Hayden et al. [22])

FABQ, fear avoidance beliefs questionnaire; FABQ-W, work sub-scale of the FABQ; FABQ-P, physical subscale of the FABQ; TSK, Tampa Scale of Kinesiophobia; SIGN, risk of bias: (++) , high quality: most of the criteria have been fulfilled; (+), moderate quality: some criteria fulfilled.

Discussion

Main findings

In this systematic review of 17 randomized controlled trials (RCTs), we found convincing evidence that the presence of fear avoidance beliefs (FAB) assessed by the most frequently used questionnaires, the FABQ and the TSK, influences treatment effects in patients with low back pain (LBP) of less than six months duration (GRADE high quality evidence, 831 patients vs. 322 in non-predictive studies). There was moderate evidence that a decrease in FAB values during treatment was associated with less pain and disability at follow-up (two RCTs with moderate quality, 242 patients). Interventions that addressed fear avoidance beliefs (GRADE moderate evidence, 1051 vs. 227 patients in studies without moderating effects) were more effective than treatments based on biomedical concepts. In chronic patients with NSLBP, the findings were less consistent. Two studies found baseline fear avoidance beliefs to be associated with more pain and disability and less RTW (339 patients) while three others (832 patients) found none (GRADE low evidence).

Results in light of existing literature

To our knowledge, this is the first systematic review summarizing the current evidence on the role of fear avoidance beliefs identified by the FABQ or the TSK as a moderator of treatment efficacy in RCTs for acute, sub-acute, and chronic NLSBP. Several non-systematic reviews highlighted the importance of the fear avoidance model [50-52] while others questioned the impact [14, 53, 54]. Other systematic reviews [55, 56] addressed the efficacy of specific treatment interventions, but to date none have systematically addressed the influence of baseline level of fear avoidance beliefs on treatment efficacy. Therefore it is unknown if treatments addressing fear avoidance beliefs are suitable for all patients with LBP and further if treatments based on biomedical concepts are less effective in patients with high fear avoidance beliefs. Our findings suggest that interventions addressing fear avoidance beliefs in patients with high fear avoidance beliefs score values at baseline are effective in patients with LBP of less than six months duration. In patients without or with low fear avoidance beliefs, such treatments are no more effective than other treatment strategies and likely unnecessary. To verify this, future studies should include a large enough sample of patients to allow for subgroup analyses.

A recent systematic review [65] of observational studies highlighted the importance of fear avoidance beliefs for poor work-related outcomes in subacute LBP patients and therefore recommended systematic assessment of fear avoidance beliefs in patients with persistent low-back pain. The current study further supports these findings and expands on the influence of fear avoidance beliefs on treatment response

and treatment outcome. The influence of fear avoidance beliefs in patients with persisting pain is less apparent. It is reasonable to believe that a chronic population has already solidified their beliefs and there may be less variation in beliefs about the effect of activity on pain in this group. It therefore makes sense that these beliefs are more important in guiding the course of LBP in the sub-acute phase when individuals are at risk of becoming chronic and beliefs may be more disparate as shown in a recent study [65]. Since there was an absence of assessment of additional psychological constructs with potential overlap in the studies reviewed here, this is a topic for future study.

Strength and limitations

The strengths of this systematic review are the assessment of fear avoidance beliefs in light of disease duration and the comprehensive evaluation of the currently available literature. The search was inclusive, no language limitations were applied, and a thorough bibliographic search was conducted in order to include all relevant studies. The extraction process was done in accordance with current guidelines and with the help of an experienced statistician. Potential factors of influence were identified by a multidisciplinary team (a psychologist, internist, physical therapist, statistician, rheumatologist, and methodologist).

The study's main limitation is a possibility of bias due to the moderate quality of most studies. Some studies did not meet the required sample size calculated for the primary outcome. None of the studies provided a power analysis for moderator effect. For reliable subgroup analysis even more patients are needed. We are aware of some limitations with regards to generalizability by limiting our search to studies that used FABQ or TSK only to assess fear avoidance beliefs. Other available questionnaires include questions that address fear avoidance beliefs but were not included into this analysis. The strength of our approach is that FABQ and TSK are widely used questionnaires and assess fear avoidance beliefs exclusively which allowed us to make comparisons among studies. Further, some of the recent questionnaires that assess different coping strategies including the STarT Back Tool or the Orebro questionnaire included questions that were derived from the FABQ or TSK [6, 57, 58]. Therefore it is likely that our findings would apply to studies that use such questionnaires with regards to the importance of fear avoidance beliefs. The heterogeneity of studies and the methodological limitations impeded us from conducting a meta-analysis [20]. We have tried to balance these limitations by providing a comprehensive comparative description of all the studies included and the prognostic factors reported.

Implications for research

Addressing fear avoidance beliefs leads to better outcome compared to treatments that are based on biomedical concepts in particular in patients at risk for chronic LBP.

While our review might suggest that distributing educational information will reduce fear avoidance beliefs equally compared to more intensive interventions, there is some evidence from one study that interventions that expose the patient to the feared activities are more effective than education alone [59]. However, this study population was small and therefore omitted from this review. Future research should further investigate how treatment interventions designed to address fear avoidance beliefs work. Further, the mechanism by which fear avoidance beliefs affect the outcome should be investigated. It has been suggested that fear of activities may be motivated by different aspects of injury, such as pain or likelihood of reinjury [14, 60]. It has also been suggested that pain intensity may influence fear avoidance beliefs separate from cognitive appraisal [61]. However, these constructs are difficult to disentangle and need to be clarified for future study. Because we used results from multivariate analyses where the main confounders including baseline pain are included in the model, our findings suggest that fear avoidance beliefs have an effect regardless of the influence of baseline pain. In addition, the process delineated in the fear avoidance model has been questioned. It has been suggested that one may show fear avoidance beliefs with or without catastrophizing as a prerequisite [62]. Therefore, the interaction between fear avoidance beliefs and the presence of catastrophizing or other coping strategies and their effects on outcome should be investigated further. One way to approach this is to include tools that assess different aspects of coping such as fear avoidance beliefs, catastrophizing, self-efficacy, and positive outcome expectations to test their relative utility. Another approach proposed by researchers [15, 60] is to further categorize patients in different types of avoiders (e.g. misinformed avoiders, learned pain avoiders, affective avoiders) [50, 60]. Those subjects who fall into these sub-groups could then be assigned to different treatment strategies.

The various outcome measures reported impeded us from conducting meta-analysis. In future studies a core set of information should be collected as proposed by the Multinational Musculoskeletal Inception Cohort Study Collaboration (MMICS) [64]. This will allow researchers to conduct sub-group analyses for specific prognostic risk factors and ultimately to pool study populations in future systematic reviews.

The current available literature does not support one scale over the other. TSK was used less often than the FABQ. It seems that FABQ-W and FABQ-P are complementary and, depending on the patient population, study participants express more fear of physical activity or fear of work. The independents of these constructs argue for the use of both, regardless of how they are measured. This information is particularly important for researchers that strive for including different coping strategies in one questionnaire by using single item questions. Whether the proposed cut-off values are valid for identifying patients at risk for poor outcome has not been shown to date. In the current review, few studies used cut-off values and the results were inconsistent. For clinicians it is important to know at what point treatments addressing fear avoidance beliefs should be initiated. Whether the proposed cut-off values are valid for identifying

patients at risk for poor outcome has not been shown to date. In the current review, few studies used cut-off values and the results were inconsistent. Recently the use of the following cut-off values for high fear avoidance beliefs has been proposed: FABQ-P of 16 and more points, FABQ-W of 25 and more points, TSK of 38 and more points [65]. For low fear avoidance the following cut-off values have been proposed: FABQ-P 14 or less points, FABQ-W 20 or less, and TSK score 17 and less points [65]. Future research should aim at validating these proposed cut-off values in the randomized clinical trial setting and clarifying the role of cut-off values.

Implication for practice

Treatment recommendations for LBP vary widely and often reflect the personal beliefs of physicians and other health care providers [66-68] without taking patient preference or need into consideration. In addition, the clinical judgment of therapists in the studies reviewed generally failed to correlate with patients' self-reported FAB [69]. In clinical practice, a screening instrument can allow clinicians to accurately identify risk factors like fear avoidance beliefs that have been shown to be modifiable [42, 45]. Therefore, standardized measurement methods are needed to improve outcomes. This systematic review supports a proactive standardized assessment of psychological risk factors when LBP persists. Further, it also supports the recommendation that patients with high fear avoidance beliefs should be treated with interventions targeting fear avoidance beliefs. The distribution of information material addressing fear avoidance beliefs has been shown to be effective in patients with high fear avoidance beliefs. It is an inexpensive and widely available intervention. Further physical therapy addressing fear avoidance beliefs also led to better outcome when compared to physical therapy based on bio-medical concepts. Fear avoidance beliefs should be taken into account in the treatment of patients with LBP. The comparison of the efficacy of different treatment interventions was beyond the scope of this review.

Conclusion

Evidence suggests that fear avoidance beliefs are associated with poor treatment outcome in patients with LBP of less than six months and thus early treatment, including interventions to reduce fear avoidance beliefs, may avoid delayed recovery and chronicity. Patients with high fear avoidance beliefs are more likely to improve when fear avoidance beliefs are addressed in treatments than when these beliefs are ignored, and treatment strategies should be modified if fear avoidance beliefs are present.

Acknowledgements

This study was performed in collaboration with many experienced researchers and clinicians. We thank the research team of the Occupational and Industrial Orthopaedic Center (OIOC), New York University (NYU) Langone Medical Center in New York for their very valuable discussions and input. Thanks are extended to Reto Kofmehl from Horten Centre of patient oriented research, University of Zurich, Switzerland for compiling the figures.

Disclosures

Conflict of interest: The authors of this work have no conflict of interest to declare.

Funding: This work was not funded.

Ethics: Not applicable, literature review only

References

1. Vos T, Flaxman AD, Naghavi M, et al. Years lived with disability (YLDs) for 1160 sequelae of 289 diseases and injuries 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. *The Lancet*. 2012;380(9859):2163-96.
2. Balague F, Mannion AF, Pellise F, Cedraschi C. Non-specific low back pain. *Lancet*. 2012;379(9814):482-91.
3. Fourney D, Andersson G, Arnold P, et al. Chronic low back pain: a heterogeneous condition with challenges for an evidence-based approach. *Spine (Philadelphia, Pa 1976)*. 2011;36(21 Suppl):S1-S9.
4. van Tulder M, Becker A, Bekkering T, et al. Chapter 3 European guidelines for the management of acute nonspecific low back pain in primary care. *Eur Spine J*. 2006;15(0):s169-s91.
5. Krismer M, van Tulder M. Strategies for prevention and management of musculoskeletal conditions. Low back pain (non-specific). *Best practice & research Clinical rheumatology*. 2007;21(1):77-91.
6. Hill J, Whitehurst DGT, Lewis M, et al. Comparison of stratified primary care management for low back pain with current best practice (STarT Back): a randomised controlled trial. *Lancet*. 2011;378(9802):1560-71.
7. Linton SJ, Shaw WS. Impact of Psychological Factors in the Experience of Pain. *Phys Ther*. 2011;91(5):700-11.
8. Chou R, Huffman L. Nonpharmacologic therapies for acute and chronic low back pain: a review of the evidence for an American Pain Society/American College of Physicians clinical practice guideline. *Ann Intern Med*. 2007;147(7):492-504.
9. Airaksinen O, Brox J, Cedraschi C, et al. Chapter 4 European guidelines for the management of chronic nonspecific low back pain. *Eur Spine J*. 2006;15(0):s192-s300.
10. Waddell G, Newton M, Henderson I, Somerville D, Main CJ. A Fear-Avoidance Beliefs Questionnaire (FABQ) and the role of fear-avoidance beliefs in chronic low back pain and disability. *Pain*. 1993;52(2):157-68.
11. Miller RP, Kori SH, Todd DD. Development of the Tampa Scale of Kinesiophobia (TSK) - unpublished report. 1991.
12. Vlaeyen JW, Kole-Snijders AM, Boeren RG, van Eek H. Fear of movement/(re)injury in chronic low back pain and its relation to behavioral performance. *Pain*. 1995;62(3):363-72.
13. Pincus T, Burton AK, Vogel S, Field AP. A systematic review of psychological factors as predictors of chronicity/disability in prospective cohorts of low back pain. *Spine (Phila Pa 1976)*. 2002;27(5):E109-20.
14. Pincus T, Vogel S, Burton AK, Santos R, Field AP. Fear avoidance and prognosis in back pain: a systematic review and synthesis of current evidence. *Arthritis Rheum*. 2006;54(12):3999-4010.
15. Rainville J, Smeets RJEM, Bendix T, Tveito TH, Poiraudreau S, Indahl AJ. Fear-avoidance beliefs and pain avoidance in low back pain-translating research into clinical practice. *The spine journal: official journal of the North American Spine Society*. 2011;11(9):895-903.
16. Westman AE, Boersma K, Leppert J, Linton SJ. Fear-Avoidance Beliefs, Catastrophizing, and Distress A Longitudinal Subgroup Analysis on Patients With Musculoskeletal Pain. *Clin J Pain*. 2011;27(7):567-77.
17. Fritz JM, George SZ. Identifying psychosocial variables in patients with acute work-related low back pain: the importance of fear-avoidance beliefs. *Phys Ther*. 2002;82(10):973-83.
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. *Ann Intern Med*. 2009;151(4):W-65.
19. Harbour R, Lowe G, Twaddle S. Scottish Intercollegiate Guidelines Network: the first 15 years (1993-2008). *J R Coll Physicians Edinb*. 2011;41(2):163-8.
20. Pincus T, Miles C, Froud R, Underwood M, Carnes D, Taylor S. Methodological criteria for the assessment of moderators in systematic reviews of randomised controlled trials: a consensus study. *BMC Medical Research Methodology*. 2011;11(1):14.

21. Wertli MM, Schob M, Brunner F, Steurer J. Incomplete reporting of baseline characteristics in clinical trials: an analysis of randomized controlled trials and systematic reviews involving patients with chronic low back pain. *PloS one*. 2013;8(3):e58512.
22. Hayden JA, Dunn KM, van der Windt DA, Shaw WS. What is the prognosis of back pain? Best practice & Research Clinical Rheumatology. 2010;24(2):167-79.
23. Swinkels-Meewisse EJCM, Swinkels RAHM, Verbeek ALM, Vlaeyen JWS, Oostendorp RAB. Psychometric properties of the Tampa Scale for kinesiophobia and the fear-avoidance beliefs questionnaire in acute low back pain. *Manual therapy*. 2003;8(1):29-36.
24. Vlaeyen JWS, Kole-Snijders AMJ, Rotteveel AM, Ruesink R, Heuts PHT. The role of fear of movement/(re)injury in pain disability. *Journal of Occupational Rehabilitation*. 1995;5(4):235-52.
25. Roelofs J, Goubert L, Peters ML, Vlaeyen JWS, Crombez G. The Tampa Scale for Kinesiophobia: further examination of psychometric properties in patients with chronic low back pain and fibromyalgia. *European Journal of Pain*. 2004;8(5):495-502.
26. Ostelo RWJG, Swinkels-Meewisse IJCM, Knol D, Vlaeyen JWS, de Vet HCW. Assessing pain and pain-related fear in acute low back pain: what is the smallest detectable change? *International Journal of Behavioral Medicine*. 2007;14(4):242-8.
27. Crombez G, Vlaeyen JW, Heuts PH, Lysens R. Pain-related fear is more disabling than pain itself: evidence on the role of pain-related fear in chronic back pain disability. *Pain*. 1999;80(1-2):329-39.
28. R Development Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2011; Available from: <http://www.R-project.org/>.
29. Guyatt G, Oxman A, Vist G, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ British medical journal*. 2008;336(7650):924-6.
30. Furlan AD, Pennick V, Bombardier C, van Tulder M. 2009 updated method guidelines for systematic reviews in the Cochrane Back Review Group. *Spine (Phila Pa 1976)*. 2009;34(18):1929-41.
31. Hancock MJ, Maher CG, Latimer J, Herbert RD, McAuley JH. Can rate of recovery be predicted in patients with acute low back pain? Development of a clinical prediction rule. *European Journal of Pain*. 2009;13(1):51-5.
32. Hancock MJ, Maher CG, Latimer J, McLachlan AJ, Day RO, Davies RA. Can predictors of response to NSAIDs be identified in patients with acute low back pain? *Clin J Pain*. 2009;25(8):659-65.
33. George SZ, Zeppieri G, Jr., Cere AL, et al. A randomized trial of behavioral physical therapy interventions for acute and sub-acute low back pain (NCT00373867). *Pain*. 2008;140(1):145-57.
34. Ang DC, Bair MJ, Damush TM, Wu J, Tu W, Kroenke K. Predictors of pain outcomes in patients with chronic musculoskeletal pain co-morbid with depression: results from a randomized controlled trial. *Pain medicine (Malden, Mass)*. 2010;11(4):482-91.
35. Magnussen L, Strand LI, Skouen JS, Eriksen HR. Motivating disability pensioners with back pain to return to work - a randomized controlled trial. *Journal of rehabilitation medicine : official journal of the UEMS European Board of Physical and Rehabilitation Medicine*. 2007;39(1):81-7.
36. Rasmussen-Barr E, Campello M, Arvidsson I, Nilsson-Wikmar L, Ång B-O. Factors predicting clinical outcome 12 and 36 months after an exercise intervention for recurrent low-back pain. *Disabil Rehabil*. 2012;34(2):136-44.
37. Mannion AF, Muntener M, Taimela S, Dvorak J. A randomized clinical trial of three active therapies for chronic low back pain. *Spine (Phila Pa 1976)*. 1999;24(23):2435-48.
38. Underwood M, Mistry D, Lall R, Lamb S. Predicting response to a cognitive-behavioral approach to treating low back pain: Secondary analysis of the BeST data set. *Arthritis Care Res*. 2011;63(9):1271-9.
39. George SZ, Fritz JM, Childs JD, Brennan GP. Sex differences in predictors of outcome in selected physical therapy interventions for acute low back pain. *J Orthop Sports Phys Ther*. 2006;36(6):354-63.
40. Staal JB, Hlobil H, Koke AJ, Twisk JW, Smid T, van Mechelen W. Graded activity for workers with low back pain: who benefits most and how does it work? *Arthritis Rheum*. 2008;59(5):642-9.
41. Storheim K, Brox JL, Holm I, Bo K. Predictors of return to work in patients sick listed for sub-acute low back pain: a 12-month follow-up study. *Journal of rehabilitation medicine : official journal of the UEMS European Board of Physical and Rehabilitation Medicine*. 2005;37(6):365-71.

42. George SZ, Fritz JM, Bialosky JE, Donald DA. The effect of a fear-avoidance-based physical therapy intervention for patients with acute low back pain: results of a randomized clinical trial. *Spine (Phila Pa 1976)*. 2003;28(23):2551-60.
43. Klaber Moffett JA, Carr J, Howarth E. High fear-avoiders of physical activity benefit from an exercise program for patients with back pain. *Spine (Phila Pa 1976)*. 2004;29(11):1167-72; discussion 73.
44. Beneciuk JM, Robinson ME, George SZ. Low back pain subgroups using fear-avoidance model measures: results of a cluster analysis. *Clin J Pain*. 2012;28(8):658-66.
45. Jellema P, van der Horst HE, Vlaeyen JW, Stalman WA, Bouter LM, van der Windt DA. Predictors of outcome in patients with (sub)acute low back pain differ across treatment groups. *Spine (Phila Pa 1976)*. 2006;31(15):1699-705.
46. Woods MP, Asmundson GJ. Evaluating the efficacy of graded in vivo exposure for the treatment of fear in patients with chronic back pain: a randomized controlled clinical trial. *Pain*. 2008;136(3):271-80.
47. Burton AK, Waddell G, Tillotson KM, Summerton N. Information and advice to patients with back pain can have a positive effect. A randomized controlled trial of a novel educational booklet in primary care. *Spine (Phila Pa 1976)*. 1999;24(23):2484-91.
48. Vlaeyen JW, de Jong J, Geilen M, Heuts PH, van Breukelen G. Graded exposure in vivo in the treatment of pain-related fear: a replicated single-case experimental design in four patients with chronic low back pain. *Behav Res Ther*. 2001;39(2):151-66.
49. van der Hulst M, Vollenbroek-Hutten MM, Groothuis-Oudshoorn KG, Hermens HJ. Multidisciplinary rehabilitation treatment of patients with chronic low back pain: a prognostic model for its outcome. *Clin J Pain*. 2008;24(5):421-30.
50. Rainville J, Smeets RJ, Bendix T, Tveito TH, Poiraudau S, Indahl AJ. Fear-avoidance beliefs and pain avoidance in low back pain-translating research into clinical practice. *Spine Journal*. 2011;11(9):895-903.
51. Vlaeyen JWS. Fear-avoidance model of chronic musculoskeletal pain: 12 years on. *Pain*. 2012;153(6):1144-7.
52. Nicholas MK, Linton SJ, Watson PJ, Main CJ, Decade Flags Working G. Early Identification and Management of Psychological Risk Factors ("Yellow Flags") in Patients With Low Back Pain: A Reappraisal. *Phys Ther*. 2011;91(5):737-53.
53. Kovacs FM, Abraira V, Moix J, et al. Fear avoidance beliefs and low back pain: "practical reviews" from expert panel discussions versus comprehensive systematic reviews . . . *Spine J*. 2011 Sep;11(9):895-903. *Spine Journal*. 2012;12(2):174-5.
54. Hasenbring M. Fear-avoidance and endurance-related responses to pain: New models of behavior and their consequences for clinical practice. *The Clinical Journal of Pain*. 2010;26(9):747-53.
55. Brox JI, Storheim K, Grotle M, Tveito TH, Indahl A, Eriksen HR. Systematic review of back schools, brief education, and fear-avoidance training for chronic low back pain. *The Spine Journal*. 2008;8(6):948-58.
56. Bunzli S, Gillham D, Esterman A. Physiotherapy-provided operant conditioning in the management of low back pain disability: A systematic review. *Physiother Res Int*. 16(1):4-19.
57. Hill JC, Fritz JM. Psychosocial Influences on Low Back Pain, Disability, and Response to Treatment. *Phys Ther*. 2011;91(5):712-21.
58. Linton SJ, Halld  n K. Can we screen for problematic back pain? A screening questionnaire for predicting outcome in acute and subacute back pain. *The Clinical Journal of Pain*. 1998;14(3):209-15.
59. de Jong J, Vlaeyen JWS, Onghena P, Goossens MIEJB, Geilen M, Mulder H. Fear of movement/(re)injury in chronic low back pain: education or exposure in vivo as mediator to fear reduction? *The Clinical Journal of Pain*. 2005;21(1):9-17.
60. Pincus T, Smeets RJEM, Simmonds MJ, Sullivan MJL. The Fear Avoidance Model Disentangled: Improving the Clinical Utility of the Fear Avoidance Model. *Clin J Pain*. 2010;26(9):739-46.
61. Eccleston C, Crombez G. Pain demands attention: a cognitive-affective model of the interruptive function of pain. *Psychol Bull*. 1999;125(3):356-66.
62. Crombez G, Eccleston C, Van Damme S, Vlaeyen JW, Karoly P. Fear-avoidance model of chronic pain: the next generation. *Clin J Pain*. 2012;28(6):475-83.

63. Leeuw M, Goossens MEJB, van Breukelen GJP, et al. Exposure in vivo versus operant graded activity in chronic low back pain patients: Results of a randomized controlled trial. *Pain*. 2008;138(1):192-207.
64. Pincus T, Santos R, Breen A, Burton AK, Underwood M. A review and proposal for a core set of factors for prospective cohorts in low back pain: a consensus statement. *Arthritis Rheum*. 2008;59(1):14-24.
65. Wertli MM, Rasmussen-Barr E, Weiser S, Bachmann LM, F B. The Role of Fear Avoidance Beliefs as a Prognostic Factor for Outcome in Patients with Non-Specific Low Back Pain – A Systematic Review. *The Spine Journal* 2013.
66. Rainville J, Carlson N, Polatin P, Gatchel RJ, Indahl A. Exploration of physicians' recommendations for activities in chronic low back pain . . . including commentary by Volinn E. *Spine*. 2000;25(17):2210-20.
67. Werner EL, Cote P, Fullen BM, Hayden JA. Physicians' determinants for sick-listing LBP patients: a systematic review. *Clin J Pain*. 2012;28(4):364-71.
68. Buchbinder R, Staples M, Jolley D. Doctors with a special interest in back pain have poorer knowledge about how to treat back pain. *Spine*. 2009;34(11):1218-27.
69. Calley DQ, Jackson S, Collins H, George SZ. Identifying patient fear-avoidance beliefs by physical therapists managing patients with low back pain. *J Orthop Sports Phys Ther*. 2010;40(12):774-83.
70. Delitto A, Erhard RE, Bowling RW. A treatment-based classification approach to low back syndrome: identifying and staging patients for conservative treatment. *Phys Ther*. 1995;75(6):470-85; discussion 85-9.
71. Vlaeyen JW, Linton SJ. Fear-avoidance and its consequences in chronic musculoskeletal pain: a state of the art. *Pain*. 2000;85(3):317-32.
72. Jull GA, Richardson CA. Motor control problems in patients with spinal pain: A new direction for therapeutic exercise. *J Manipulative Physiol Ther*. 2000;23(2):115-7.

Appendix 1:

Search history for PubMed and CINAHL January week 5 2013

PubMed

#	Query	Results
1	("low back pain"[MeSH Terms] OR ("low"[All Fields] AND "back"[All Fields] AND "pain"[All Fields]) OR "low back pain"[All Fields]) OR ("low back pain"[MeSH Terms] OR ("low"[All Fields] AND "back"[All Fields] AND "pain"[All Fields]) OR "low back pain"[All Fields] OR "lumbago"[All Fields]) OR ("low back pain"[MeSH Terms] OR ("low"[All Fields] AND "back"[All Fields] AND "pain"[All Fields]) OR "low back pain"[All Fields] OR ("low"[All Fields] AND "back"[All Fields] AND "ache"[All Fields]) OR "low back ache"[All Fields])	<u>23223</u>
2	("fear"[MeSH Terms] OR "fear"[All Fields]) AND avoidance[All Fields]	<u>3808</u>
3	#3 AND #4	<u>342</u>
4	Search #3 AND #4 Limits: Humans, Publication Date from 1990 to 2011	<u>305</u>

CINAHL (EBSCOhost)

#	Query	Results
1	(MH "Low Back Pain") OR "low back pain"	12490
2	"lumbago"	42
3	"low back ache" or "low backache"	13
4	S1 or S2	12521
5	"fear avoidance"	476
6	(MH "Avoidant Personality Disorder") OR "avoidance"	6265
7	"avoidance"	6219
8	S5 or S6 or S7	6265
9	S3 or S4	10977
10	S8 and S9	321

Appendix 2:**Internal validity (SIGN methodology checklist) [19] and methodology for moderator analysis [20]**

		1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.9	1.10	2.1.	MC	IA	ED
Fritz	2002	WC	PA	PA	PA	WC	AA	AA	PA	n/a	+	-	+	NA
Staal	2008	WC	WC	WC	WC*	WC*	AA	WC	WC*	N/A*	++ *	NA	+	+
Storheim	2003	WC	WC	AA	AA	AA	AA	AA	AA	AA	+	-	+	+
Ang	2010	WC	WC	WC*	AA	WC	WC*	WC	WC*	PA	++ *	-	+	
Mannion	1999	AA	AA	AA	AA	AA	AA	AA	AA	AA	+	-**	+	+
v.d. Hulst	2008	WC*	WC*	AA	AA*	WC*	AA	AA	AA	AA	++ *	-	+	+
Burton	1999	WC	AA	PA	WC*	WC	AA	WC	WC	NA	+	N/A	+	+
Jellema	2006	WC	AA	PA	PA	WC	AA	AA	AA	PA	+	-	+	+
Hancock	2009	WC	WC	WC	WC	WC	AA	AA	WC	AA	++	-	-	+
Georg, D	2003	AA	WC	AA	AA	AA	AA	AA	WC	NR	+	-	+	+
George, B	2006	AA	WC	AA	AA	PA	PA	WC	PA	PA	+	-	+	+
Georg, Z	2008	AA	AA	WC	AA	WC	AA	AA	WC	PA	+	-	+	+
Rasmussen	2012	WC	WC	AA	AA	AA	AA	AA	WC	AA	+	-	+	+
Underwood	2011	WC*	WC*	WC*	WC*	WC*	WC*	WC*	WC*	WC*	++ *	-	+	+
Klarber	2004	WC	WC	AA	AA	AA	AA	AA	AA	PA	+	N/A	+	+
Mangussen	2007	WC	WC	WC	WC	AA	AA	AA	WC*	N/A*	+	-	+	+
Woods	2008	AA	AA	PA	PA	AA	AA	WC	WC	AA	+	N/A	-	+

WC, well covered; AA, adequately addressed; PA, poorly addressed; NA, not addressed; NR, not reported; N/A, not applicable

1.1: appropriate and clearly focused question; 1.2, assignment is randomized; 1.3, adequate concealment; 1.4, blinding; 1.5; treatment and control groups are similar at baseline; 1.6, The treatment under investigation is the only difference; 1.7, all relevant outcomes are measured; 1.9, intention to treat analysis; 1.10, various sites are comparable; 2.1, risk of bias: (++), high quality: most of the criteria have been fulfilled. If not fulfilled, the conclusions of the study are very unlikely to alter; (+), moderate quality: some criteria fulfilled. Criteria not adequately described are unlikely to alter the conclusions; (-), low quality: few or no criteria fulfilled. The conclusions are likely to alter.

Additional methodological criteria not yet addressed:[20] MC, multiple comparison: adjustment of significance level from $p < 0.05$ to < 0.001 : +, yes; -, no; IA, interaction test between treatment and moderator; ED, equal distribution of the moderators between groups at baseline.

** problem of multiple testing mentioned

*Ratings the two reviewers (MW) agreed on after discussing additional published material. No third party arbitration was needed to achieve consensus.

Appendix 3:
Summary of study results

Acute low back pain

ID	Study	Sample size*	N	Drop-out %	DD days	FU days	QA	Cut-off used?	MA?	Domain inclusion in MA	FAB predictive for Outcome	FAB Influences treatment effect
5	Fritz 2002	N.R.	78	5.5	30	FABQ-W FABQ-P	No	Yes	Stepwise entry with p=0.05, out with p=0.15	High FABQ-W score predict non-RTW at 4 weeks (OR 1.2 (CI 1.07-1.34), FABQ-P non-predictive (N.R.)	No differences between Control PT (AHCPR clinical practice guidelines) and Intervention (PT Classification System by Delitto)	
7	George, Brennan, 2006	N.R.	165	44	13.3	30	FABQ-P and FABQ-W	No	Yes	Stepwise, variables in/out p=0.05.	FABQ-W predictive for Reduction of ODI in men (Beta 0.43 (CI 0.03-0.83)), not in all men in all treatments(manipulation + exercise; stabilization exercise; direction-specific exercise	More FABQ-W at baseline, less ODI reduction at 4 weeks in men in all
12	Jellema, 2006	290	314	19	12	42	FABQ-P	No	Yes	Univariate logistic regression, inclusion p=<0.20)	FABQ-P in UC non-predictive (N.R.), in MIS predictive for favourable outcome (RMQ improvement ≥ 30% = CMID) after 6 weeks	Patients with higher FABQ-P are more likely to have a CMID in MIS treatment (OR 1.07, CI 0.99-1.16) compared to UC

Acute low back pain

ID	Study	Sample size*	N	Drop-out %	DD days	FU days	QA	Cut-off used?	MA?	Domain inclusion in MA	FAB predictive for Outcome	FAB Influences treatment effect
9B	Hancock, James, 2009	240	239	2	9.13	To recovery	FABQ-W and FABQ-P	No	Yes	Univariate cox regression, values $p < 0.1$ included	FABQ-P non-predictive. FABQ-W univariate associated. Did not remain in the MA model: Non-predictive for days to recovery from pain (1 day without pain; VAS 0 or 1) (N.R.)	No specific information given for the treatments Diclophenac vs. Placebo + sham manipulation vs. spinal manipulative therapy
9C	Hancock, Davies, 2009	240	239	2	9.13	To recovery	FABQ-W and FABQ-P	No	Yes	Univariate cox regression, $p < 0.25$ included	FABQ-P predictive for response to NSAID treatment compared to placebo with regards to days to recovery from pain; FABQ-W non-predictive (HR for 1 day 0.996 (CI 0.972-1.02), HR for 7 days 0.992 (CI 0.967 – 1.018))	NSAID HR to have a day without pain compared to placebo: 1 day HR 1.059 (1.002-1.118) and 7 days HR 1.042 (0.988-1.10)

Subacute low back pain

ID	Study	Sample size*	N	Drop-out %	DD days	FU days	QA	Cut-off used?	MA?	Domain inclusion in MA	FAB predictive for outcome	FAB influences treatment effect
17	Staal, 2008	70/ group	134	11	67	360	FABQ-P and modified FABQ-W (0-36) TSK	FABQ-P ≥16, FABQ-W ≥26	Yes	N.R.	FABQ-P ≥16 and FABQ-W ≥26 predictive: less likely to RTW (N.R.). Reduction of FABQ-P, FABQ-W and TSK at 3 months increased chance of RTW (FABQ-P HR 2.0 (CI 1.2-3.5), FABQ-W 2.2 (CI 1.3-3.7), TSK 2.3 (CI 1.4-3.9)	Behavioural oriented GA in comparison to UC increased chance for RTW in patients with FABQ-P <16 [HR 3.0 (CI 1.6-5.5)], FABQ-P ≥16 [HR 1.6 (1.0-2.7)], FABQ-W <26 [HR 2.6 (CI 1.5-4.6)], and FABQ-W ≥26 [HR 1.6 (CI 1.0-2.7)]. Increased FABQ-W/P decreased treatment effect difference between GA and UC
8	George, Zeppieri, 2008	80	108	44	52	180	Δ FABQ-P Δ FABQ-W	FABQ-P >14	Yes	-	Change in FABQ-P was predictive for change in pain at 1 month [beta= 0.26 (CI -0.46 to-0.06)], 6 months [beta= 0.25 (-0.47 to-0.03)] Change in FABQ-W non-predictive. Cluster of FABQ-P high and Catastrophizing more change in pain and disability.	No differences between treatments, all treatments ↓ (TBC PT vs. TBC + GA vs. TBX + GX)
18A	Storheim, 2005	50/ group	93	39	70	360	FABQ-W FABQ-P	No	Yes	Univariate analysis, variables included p= <0.25	Higher FABQ-W predictive for longer time to RTW at 1 year [HR 1.082 (CI N.R.)] FABQ-P non-predictive (N.R.)	No differences between treatments, all treatments ↓ (UC, PT, CBT)

Acute to subacute low back pain

ID	Study	Sample size*	N	Drop-out %	DD out days	FU days	QA	Cut-off used?	MA?	Domain/inclusion in MA	FAB predictive for outcome	FAB influences treatment effect
6	Georg Donald, 2003	54	66	N.R.	27.6	180	FABQ-P FABQ-W	No	Yes, multiple L.R.	Pre-defined effects	FABQ-P increased at baseline predictive for more disability (ODI) at 1 month (beta 0.292, p= 0.042) and 6 months (beta 0.469, p= 0.002) in both groups. FABQ-W non-predictive (N.R.) (Graf 12)	FAB-treatment reduces FABQ-P effect at 1 month (beta -0.13) and 6 months (beta -0.15) compared to standard PT (Graf 14). Subgroup treatment effects show improvement of disability in FAB-PT compared to PT (Graf 15)
2	Burton, 1999	N.R.	162	22	<3 months	90	FABQ-P	Median split: >14	Repeated ANOVA	Only predefined parameters investigated	Patients with FABQ-P at baseline of >14 and improvement (FABQ-reduction of > 4 points) at 2 weeks and 3 months are more likely to improve in RMQ (≥ 3 points) at 3 months [RR to deteriorate at 0.56 (CI 0.31-1.03) and 0.46 (CI 0.26-0.8) respectively]	Experimental booklet more likely to have CMID in FABQ-P (>4 points): at 2 weeks (RR 2.72), 3 months (RR 1.53), 1 year (RR 1.47) compared to control booklet
208	Woods, 2008	N.R.	83	46	<7 weeks	60	TSK	≥38 = inclusion criteria	ANCOVA	Only 44 finished the study	TKS non-predictive for PDI: Pain Disability Index (Pollard 1984) at 2 months	Mean effect GivE for PDI pre- to 6 month 20.4 -> 12.2, GA 23.3 -> 23.7, Control 24 ->18.8. GivE better than GA

Subacute to chronic low back pain

ID	Study	Sample size*	N	Drop-out %	DD days	FU days	QA	Cut-off used?	MA?	Domain inclusion in MA	FAB predictive for outcome	FAB influences treatment effect
13A	Klauer Moffett, 2004	300	187	9	1- 6 months	360 days	FABQ-P	High: ≥14	Binary L.R. within sub-groups	Adjustment for post-randomization differences	FABQ-P <14 non-predictive and no differences between the groups at 1 year [Exercise group (EG) and UC]	FABQ-P ≥14: EG more likely to have CMID RMQ (≥2 points) compared to UC. The effect increases over time [OR 6 months 1.98 (CI 0.72-5.48), 12 months 3.58 (CI 1.30-9.84)].

Chronic low back pain

ID	Study	Sample size* N	Drop- out %	DD days	FU days	QA	Cut-off used?	MA?	Domain inclusion in MA	FAB predictive for outcome	FAB influences treatment effect
1	Ang, 2010	97/group	250 0	3240	90	modified 10-item TSK	No	Yes	Univariate variables with $p \leq 0.2$ included	High TSK associated with an increased pain [GCPS pain intensity, beta 0.46, CI 0.0092-0.91]) and GCPS activity interference [beta 0.57 (CI 0.119-1.02)] at 3months	Effect of both, the optimized antidepressant therapy + pain self-management program and UC ↓
38	Rasmussen- Barr, 2012	72	71 23	3600	1080	FABQ-P	High ≥ 14	Yes	Univariate variables with $p \leq 0.10$ included	FABQ-P ≥ 14 at baseline does not predict pain (VAS) and disability (ODI) after 12 (pain exercise group OR 1.21 (CI 0.46-3.17), disability OR 0.8 (CI 0.27- 2.35) and 36 months [pain OR 0.97 (CI 0.89-1.06)], disability OR 1.28 (0.38-4.31)	No differences between the reference group (advice) and reference group
16	Mannion, 1999	50-55	148 11	3924	180	change in FABQ-P and FABQ- W	No	Univ. stepwise correlation R.	Simple analyses with association in change of RMQ	Change in FABQ-P and FABQ- W non-predictive for change in RMQ. Change in FABQ-W small effect for pain reduction at 6 months (10% variance in combination with reduction of disability)	No difference between "modern active PT", muscle conditioning on training devices and low-impact aerobic

Chronic low back pain

ID	Study	Sample size*	N	Drop-out %	DD days	FU days	QA	Cut-off used?	MA?	Domain inclusion in MA	FAB predictive for outcome	FAB influences treatment effect
19	van der Hulst, 2008	N.R.	163	13	7043	120	TSK	No	Yes	Not specified	No influence of TSK on RMQ, SF-36D and SF-36M (end of treatment) and RMQ at 4 months [Beta - 0.13, (CI -0.29 to 0.027)]. Tendency to more decrease in SF-36D [Beta 0.23 (CI -0.06 to 0.52)] and worsening of SF-36M [Beta -0.23 (CI -0.58 to 0.123)] at 4 months	Multidisciplinary rehab (MR) better than UC: TSK does not significantly influence the effect on; RMQ and SF-36M at end of treatment and 4 weeks FU, SF-36D at end of treatment. More TSK leads to more favourable change in SF-36D (beta 0.16 (CI 0.018 – 0.762) in MR
14C	Underwood, 2011	AM 233, AM+CBT 467	598	15	4680	360	FABQ-P	High: ≥ 14	Yes	Forward selection, $p=0.025$	FABQ-P non-predictive for RMQ, MVK Pain and Disability at 1 year	AM vs. AM+CBT co-variant
15	Magnussen, 2007	Retrospective power 19%	89	36	3852	360	FABQ-P and FABQ-W	Median split: FABQ-P low < 14.8; FABQ-W <35	U.L.	-	FABQ-W <35: 3 times more likely to have entered RTW process at 1 year, but small sample size and wide CI (0.4-4.6), FABQ-P non-predictive	Control group vs. Intervention with FAB reducing information co-variant

low back pain; NSLBP, non-specific low back pain; SLBP, specific low back pain; DD, Disease Duration; FU, Follow-up; FAB, Fear Avoidance Beliefs; FABQ, fear avoidance questionnaire; FABQ-P, FABQ physical activity sub-scale; FABQ-W, FABQ work sub-scale; TSK, Tampa Scale of Kinesiophobia; TBC, treatment-based classification; GA: graded activity; GX, graded exposure; PT, physical therapy; UC, usual care; MIS, minimal intervention strategy; AM, active management; CBT, cognitive behavioral therapy; GiVE, Graded in vivo exposure; QA, Questionnaire used; OR, odds ratio; MA, multiple regression analysis; IQR: interquartile range; CMID, clinical meaningful important difference; Log. Reg., logistic regression; U.L., univariate logistic regression; L.R., linear regression; ODI, Oswestry Disability Index: higher score indicates more disability; CMID = ≥ 12 points reduction; RTW, return to work; CPG, chronic pain grade (von Korf, 1984): higher score indicates more disability; RMQ, Roland Morris Questionnaire: higher score indicates more disability; CMID = ≥ 2.3 point reduction, $\geq 30\%$ change; SF-36D, physical health sub-score: higher score indicates higher level of functioning; SF-36M, mental health sub-score: higher score indicates higher level of functioning. * Power analysis for sample size in all studies for main outcome not for sub-group analysis; RR, relative risk

Appendix 4:

Detailed description of interventions provided

Treatment better in high FAB

ID 12: FABQ-P high: MIS more effective than UC: UC provided by the GPs. MIS provided by the GPs: The MIS consultation took place within about 5 days for 20 minutes (three phases: exploration, information, and self care). Aim identification and discussion of psychosocial prognostic factors, such as the patient's own ideas on the cause of their LBP, FAB, worries regarding the pain, pain catastrophizing, pain behavior, and reactions from the social environment (family, friends, work) regarding LBP. GPs were asked not to send the intervention group to PT in the first 6 weeks.[45]

ID 9C: FABQ-P high, Patients with Diclophenac sooner without pain than the patients with placebo.[32]

ID 13: FABQ-P high Exercise Group (EG) more effective than UC: EG-group 1 our 8 times over 4 weeks: aim: to increase confidence in using the spine normally, including low-impact aerobic exercises and strengthening and stretching exercises for the main muscle groups. A CBT approach underpinned the program. Control group: usual GP care.[43]

ID 6: FABQ-P high, TBC+FAB-PT more effective the TBC-PT so FAB has not effect on treatment outcome in the FAB-PT treated group compared to patients with low FAB. FAB-TBC-PT group: GA + education (de-emphasizes anatomic findings, encourages to take an active role in the recovery, and educates to view back pain as a common condition, not a serious disease) + Back Book educational pamphlet.[42]

ID: 2: FABQ-P high: Experimental booklet more effective than control booklet: The Back Book, was developed to accompany the U.K.'s 1996 Royal College of General Practitioners (RCGP) guidelines: main message: 1) No sign of serious disease. 2) Spine is strong, no permanent damage even if it is very painful. 3) Back pain is a symptom that your back is not moving quite as it should. 4) There are many treatments that help control the pain but lasting relief depends on your effort. 5) Recovery depends on getting your back moving. The sooner you get active the sooner your back will feel better. Positive attitude are important. Compared to the Handy Hints Booklet based on traditional biomedical concepts.[47]

ID 19: High TSK increases treatment effect of the MR for Disability (SF-25 physical) but has a tendency for worse mental health (SF-36 mental). 8 patients 3 hours conditional training + sport, 0.5 hours swimming, 1.5 hours occupational therapy, 4 hours PT per week for 7 weeks. In work related to back pain, also individual occupational rehabilitation after the program. Supervision of a specialist in physical and rehabilitation medicine, conducted by a team consisting of a PT, occupational therapist, a sport therapist, and, if necessary, a psychologist and a dietician.[49]

ID 20: TSK \geq 38 inclusion criteria: GivE potentially more effective (?) than GA: GivE: Psych. student provided (under supervision) education about CBT perspective on FAB and its consequences, followed by GA techniques. 1. Session: assessment, educating (FAM of chronic pain), formulation patient's problems within this context + establishment of individualized hierarchy of fear-eliciting movements (Photograph Series of Daily Activities, PHODA; Kugler et al., 1999). Sessions 2-5: focus on exposure to the activities identified in the fear hierarchy. Behavioral tests to challenge patient expectations. Assessment of expected before and the effectively experienced FAB level after each exposure level of pain, the level of anxiety related to the activity (0–10 scales, 0 = no pain/fear, 10 = most pain/anxiety). Final session review the treatment process, provide direction on relapse prevention. Graded activity (GA) group: principles of operant conditioning, involves the shaping of healthy behaviours through positive reinforcement of predefined activity quotas (Vlaeyen et al., 2002).[46]

Treatment effect reduced by FAB

ID 17: FABQ-W ≥ 26 , FABQ-P ≥ 16 : Graded activity efficacy (GA) reduced by high FAB but still better than UC. Usual guidance by occupational physician for all, GA provided by PT and included 1-hour exercise sessions that participants attended twice per week for max 3 months. Two 1-hour sessions per week, supervised by the physiotherapist, until full return to regular work. The intervention has a maximal duration of 3 months. Education message: pain does hurt, but this does not mean it harms. Exercise, baseline values during 3 sessions. After the third session the worker proposes date for RTW.[40]

ID 6: FABQ-P high, reduced efficacy of standard PT (TBC, Treatment-based classification) + Handy Hints educational pamphlet (traditional biomedical concepts).[42]

FAB no effect

ID 3: FABQ-P high no effect on treatment outcome: Exercise group: progression of the graded stabilizing exercises. The subjects were informed of how the stabilizing muscles activate. Reference group: informed of the benefits of daily walks as physical activity. They were instructed to take a 30-minute walk every day.[36]

ID 7: FABQ-W high, reduced efficacy of all interventions: trials received randomly assigned PT interventions, including various forms of active exercise, and/or manipulation, with 1 trial including randomization to a fear avoidance supplement to PT.[39]

ID 1: TSK high: in clinically significant depression (PHQ-9 scores ≥ 10) both treatments are less effective: UC: informed that they have depressive symptoms and that they should seek advice about treatment. There are no other attempts by study personnel to influence depression management unless a psychiatric emergency. Intervention group 1. step 12 weeks optimized antidepressant th. 2 step: 12 weeks of pain self-management program: six clinical contacts (baseline and 1, 3, 6, 9, and 12 weeks) with the study nurse care manager. Following a rational algorithmic approach to antidepressant selection and dosing, the study nurse care manager assessed antidepressant adherence, adverse effects, and depression response. Depression rather than pain response dictated antidepressant adjustments.[34]

ID 18: High FABQ-W, reduced treatment efficacy of all treatments: Exercise (PT provided ongoing group back training, 15 weeks 2-3 sessions per week for 1 hour: modification of The Norwegian Aerobic Fitness Model: exercise physiology + ergonomic principles to increase fitness. CBT: 2 consultations 30 to 60 min with the PT: Explanation of pain mechanisms, discussion of baseline question in more depth. Functional examination with individual feedback and advice. Instruction in activation of deep stabilizing muscles + advice how to use it in functional and demanding tasks of daily life. Instruction in the squat technique when lifting is required. How to cope with new attacks. Reassure and emphasize that it is safe to move and to use the back without restriction. UC: GP based treatment.[41]

ID 8: high FABQ-P reduced efficacy: Treatment based classification by Delitto (TBC: provides framework for corresponding treatment category of specific exercise, manipulation or mobilization, lumbar stabilization, or traction) vs. TBC + GA (from baseline GA principles were used to progress exercise during subsequent treatment sessions) vs. TBC + graded exposure (GX, PT selected two items ranked as most fearful for implementation in the GX protocol. Lifting more than 20 pounds, carrying more than 20 pounds, and reaching to floor were commonly rated as highly fearful by this sample).[33]

ID 15: FABQ-W high decreases the efficacy of both interventions. All participants received individual guidance on how to improve their physical fitness. Intervention group (only 29 patients completed): 2 3-hour session within 2 to 3 days (groups of 5–11 persons): 2 hours lecture (spinal problems, pain mechanisms, aiming to reduce FAB related to activity and work) + 3 hours of motivational interviewing (help to focus strength and capacity), encouraging to identify barriers for RTW and find solutions. 1 hour information by counsellors from social insurance office and work office (options for combining health-adjusted work and disability pension). After the group sessions: individual follow-up by physician and nurse offered.[35]

ID 5: High FABQ-W reduced efficacy of both treatment: PT based on the recommendations of the Agency for Health Care Policy and Research (AHCPR) clinical practice guidelines. PT based on the classification system by Delitto et al. 4 classifications were: (1) manipulation followed by active range of motion exercises, (2) repeated end-range exercises (i.e., flexion or extension exercises), (3) spinal stabilization exercises, and (4) traction.[17]

ID 9B: FABQ-W and P no effect on outcome in the treatment groups: Manipulation / Sham Manipulation + Diclophenac / Placebo.[31]

ID 16: FABQ-W most likely no relevant effect on all treatments (FABQ-P no effect). PT group: 0.5 h individual therapy sessions (improving functional capacity, ergonomic principles), isometric exercises and Therabands, strength-training devices. Advice for home exercises. Muscle Reconditioning Group: Muscle strengthening devices 1 hour sessions for 12 weeks (David Back Clinic program), defined, progressive program with controlled, progressive exercises on training devices within pain-free range of motion. Aerobics/Stretching group: 12 person groups 1 hour stretching and aerobic and muscle toning exercises carried out to music with an appropriate tempo and rhythm to promote the desired level of exertion.[37]

ID 14: High FABQ-P no effect on AM and AM + CBT: All participants 15-minute session of active management advice, including the benefit of and how to remain active, avoidance of bed rest, appropriate use of pain medication and symptom management, and a copy of The Back Book (9). Intervention group: individual assessment lasting up to 1.5 hours and 6 sessions of group therapy using a cognitive-behavioral approach lasting 1.5 hours per session.[38]

CHAPTER 5

Concordance of qualitative bone scintigraphy results with presence of clinical complex regional pain syndrome 1: Meta-analysis of test accuracy studies

R. Ringer, M. Wertli, L.M. Bachmann, F.M. Buck, F. Brunner
Eur J Pain 16 (2012) 1347–1356

Abstract

Background: To date, no attempt has been made to investigate the agreement between qualitative bone scintigraphy (BS) and the presence of Complex Regional Pain Syndrome 1 (CRPS 1) and the agreement between a negative bone scintigraphy in the absence of CRPS 1.

Aims: To summarize the existing evidence quantifying the concordance of qualitative BS in the presence or absence of clinical CRPS 1.

Data sources: We searched Medline, Embase, Dare and the Cochrane Library and screened bibliographies of all included studies.

Study eligibility criteria: We selected diagnostic studies investigating the association between qualitative BS results and the clinical diagnosis of CRPS 1. The minimum requirement for inclusion was enough information to fill the two-by-two tables.

Results: Twelve studies met our inclusion criteria and were included in the meta-analysis. The pooled mean sensitivity of 12 two-by-two tables was 0.87 (95% CI, 0.68-0.97) and specificity was 0.69 (95% CI, 0.47-0.85). The pooled mean sensitivity for the subgroup with clearly defined diagnostic criteria (7 two-by-two tables) was 0.80 (95% CI, 0.44-0.95) and specificity was 0.73 (95% CI, 0.40-0.91).

Conclusions: Based on this study, clinicians must be advised that a positive BS is not necessarily concordant with presence or absence of CRPS 1. Given the moderate level of concordance between a positive BS in the absence of clinical CRPS 1, discordant results potentially impede the diagnosis of CRPS 1.

Introduction

The International Association for the Study of Pain (IASP) defined the presence of Complex regional pain syndrome 1 (CRPS 1) as a pain state following injury which exceeds in magnitude and duration the expected clinical course and that is usually associated with significant impairment of motor function [1]. Clinical manifestations include a broad spectrum of sensory, autonomic, motor and trophic changes [2]. Despite increasing research in this field the exact underlying processes of CRPS 1 remain unclear.

In clinical practice, the diagnosis of CRPS 1 is often challenging, because there are no unique and characteristic features that help to encompass CRPS 1 from other similar conditions. In particular, the distinction between CRPS 1 and expected posttraumatic changes is often difficult at early stages. Because of the broad spectrum of clinical manifestations the differential diagnosis is widespread and includes for example other neuropathic pain states, inflammation, vascular diseases and myofascial pain syndromes.

In the literature, there is a broad consensus that CRPS 1 is a clinical diagnosis based on clinical manifestations (e.g. [3]) and, given the unknown somatic anomaly of CRPS, many experts agree that definite diagnosis cannot be set with any kind of diagnostic test. Instead, clinicians and researchers must rely on various sets of diagnostic criteria [4].

Currently, the diagnostic criteria introduced by IASP [1] and the Budapest criteria [5] are widely used in clinical practice and represent an expert consensus on a specific set of clinical manifestations.

An academic discourse debating the role of qualitative bone scintigraphy (BS) to diagnose CRPS 1 began in the early 1980's [6-8]. Proponents argued that one of the underlying pathomechanisms of CRPS 1 was local osteopenia as a result of altered sympathetic activity, neurogenic inflammation and immobilization, and claimed that a typical pattern of findings in BS (an increased activity of the affected extremity with periarthicular accentuation in all three phases of scan and a decreased activity through all phases in later stages [9, 10]) could measure these bone alterations. Subsequently, BS gained popularity in clinical practice and some authors even recommended the use of bone scintigraphy as a gold standard for the diagnosis of CRPS 1 (e.g. [11, 12]). However, in the literature there is conflicting evidence about the diagnostic accuracy of BS in the diagnosis of CRPS 1 (e.g. [9, 13]).

To date no attempt has been made to assess the concordance of qualitative BS findings with the presence or absence of CRPS 1 in a systematic manner and applying up-to-date meta-analytic methods. Therefore, the aim of this study was to summarize the existing evidence quantifying the concordance of qualitative BS in the presence or absence of clinical CRPS 1.

Methods

The study was conducted in accordance to the PRISMA statement on the conduct of systematic reviews (Preferred Reporting Items for Systematic reviews and Meta-Analyses) [14].

Literature search

We searched Medline (PubMed), Embase, DARE and the Cochrane Library from inception to February 2011 using the following search terms in titles and abstracts: 'Complex regional pain syndrome' or 'CRPS' or 'reflex sympathetic dystrophy' or 'RSD' or 'algoneurodystrophy' or 'algodystrophy' or 'Sudeck' AND 'scintigraphy' or 'radionuclide' or 'radioisotope' or 'isotope'. We also screened bibliographies of all included studies for relevant articles. The search was conducted without restrictions to language or year of publication. Two independent reviewers (RR and FB) performed the literature review.

Inclusion criteria

We selected diagnostic studies investigating the association between bone scintigraphy result and the diagnosis of CRPS 1. The diagnosis of CRPS 1 was defined as the reference standard (presence or absence of CRPS 1) and BS served as the index test (positive or negative BS). The minimum requirement for inclusion was enough information to fill the two-by-two tables.

Search selection

All papers considered potentially eligible by one of the two reviewers (RR and FB) were ordered. Each reviewer independently selected studies for inclusion in the study by using a predefined checklist (see Appendix 1). Disagreement between the two extracting authors was resolved by consensus. If consensus between the two reviewers could not be reached, a third author (LMB) was deferred to arbitration and consensus.

Data extraction and quality assessment

The bibliographic details of all retrieved articles were stored in an Endnote file. We removed duplicate records resulting from the various database searches. We extracted the following variables from each study: Author, publication year, country of origin of the study, study population demographics, reference standard (presence of CRPS 1), index test (positive scintigraphy), and absolute numbers of true-positive (TP), false-positive (FP), false-negative (FN), true-negative (TN), and observations.

Two reviewers (FB, FMB) independently evaluated the methodological quality of included studies by using the QUADAS tool for assessing the quality of diagnostic accuracy studies [15]. Answers were dichotomized as 'yes' or 'no/unclear'. Disagreement between the two reviewers was resolved by a third author (LMB).

Data synthesis and interpretation

Descriptive statistical methods were used to calculate medians and standard deviation (SD). For each study, we constructed a two-by-two contingency table consisting of TP, FP, FN and TN results. For analysis, we called a result TP if the clinical diagnosis of CRPS 1 was in agreement with the pathological findings in the scintigraphy. We calculated sensitivity as $TP/(TP + FN)$ and specificity as $TN/(FP + TN)$. We plotted results from individual studies in receiver operating characteristic (ROC) space and estimated summary ROC curves using a unified model for meta-analysis of diagnostic test accuracy studies [16]. We also indicated on the ROC figures the confidence and prediction regions. This method provides estimates of average sensitivity and specificity across studies and it can be used to provide a 95% confidence region (CI) for this summary point and prediction regions within which we expect the sensitivity and specificity of 95% of future studies to lie. We performed stratified analyses for studies which did or did not report a gold standard for the clinical diagnosis of CRPS 1. All analyses were done using the Stata statistical software package v.11.1 (College Station, TX, USA).

Results

Study selection

The selection process of the studies is presented in Figure 1. Our search retrieved a total of 232 records. After reading titles and abstracts we excluded 189 papers. From the remaining 43 papers we excluded 31 papers based on the full text assessment, because they did not include a control group or they did not allow the construction of a two-by-two table. Finally, twelve studies met our inclusion criteria and were included in our analysis [10, 17-27].

Description of studies

The included studies originated from nine countries: Canada (n=2), Germany (n=2), South Korea (n=2), Austria (n=1), Belgium (n=1), China (n=1), USA (n=1), Turkey (n=1), Serbia (n=1). In total 882 participants (range 30-120) with a median age of 54.7 years (range 44-63 years) were included in the study. Median duration of symptoms was 21.7 weeks (range 0.3-75.9 weeks). In the majority (n=9) of the included studies the upper

extremity was the affected site [10, 17, 20-24, 26, 27]. A total of 6 studies [17, 18, 21, 22, 24, 27] used 3 established diagnostic criteria (Kozin [28], IASP [1], Bruehl [29]). Four studies used a self designed set of clinical manifestations to confirm the diagnosis and two studies did not report how the diagnosis was made. All except one study [19] used 3-phase Technetium 99m mono-methylene disphosphonate (Tc-99m MDP) or Technetium 99m disphosphono propan-dicarbonat (Tc-99m DPD) as the index test. The dose of the radioisotope ranged between 550-750 mega becquerel (MBq). The definition of the criteria that were used to define a positive did not vary greatly between the studies. Most studies defined an increased and asymmetric periarticular uptake in the affected extremity as a positive index test. The study characteristics are summarized in table 1.

Table 1. Study characteristics (n=12)

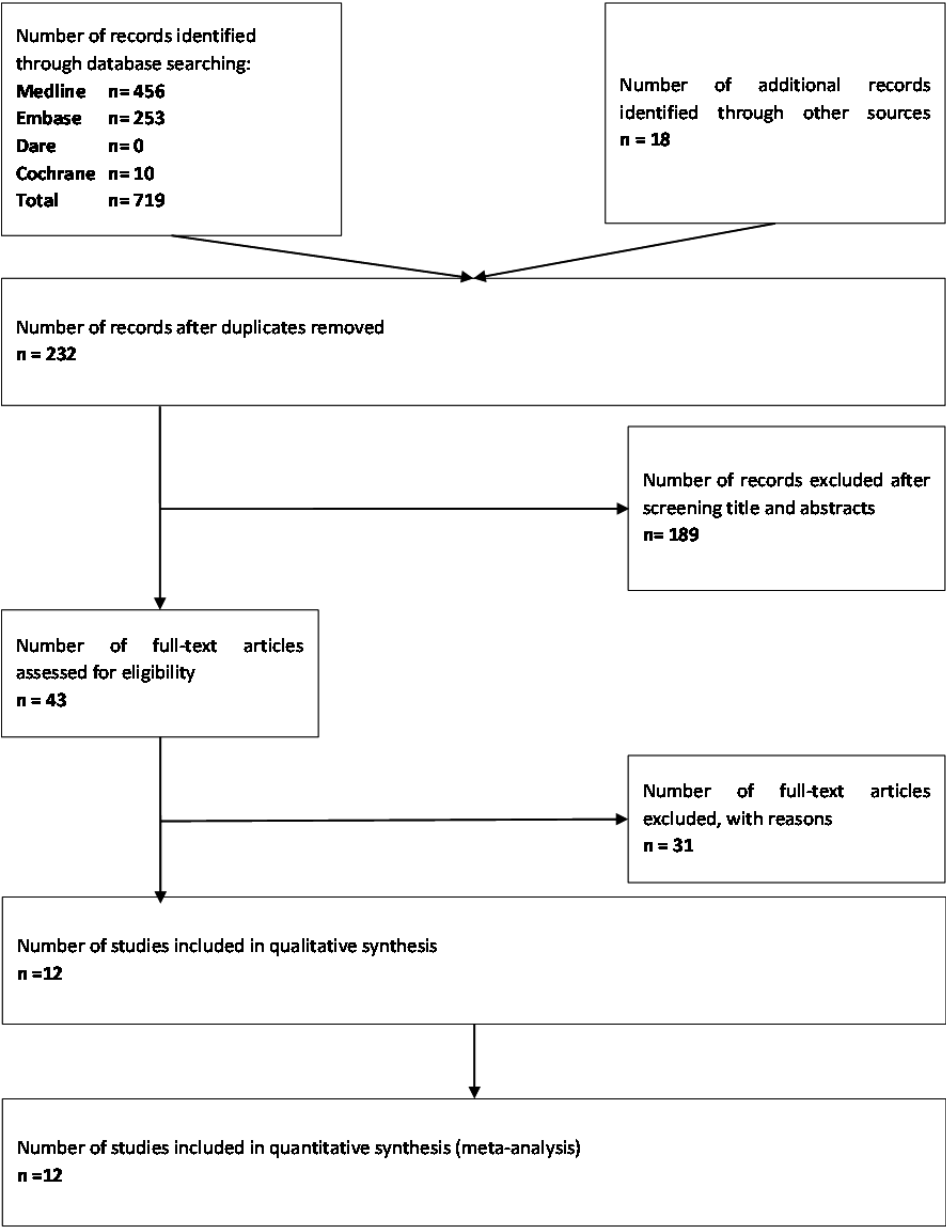
Author	Country	No. of patients	Age (median)	Duration symptoms (weeks)	Initiating event	Localization	Reference standard	Index test	Radioisotope dose	Interpretation of a positive index test
Greyson, 1984	Canada	85	60	11.0	CVE	UE	Kozin	3-phase Tc-99m MDP	550-740 MBq	Increased uptake in blood pool and delayed phase.
Kozin, 1981	USA	64	48.3	75.9	Trauma, CVE	UE, LE	Kozin	3-phase Tc-99m MDP	560 MBq	Increased periarticular activity involving multiple joints on the affected distal extremity and asymmetrical blood flow in the extremities.
Leitha, 1996	Austria	120	50.1	36.8	Trauma, CVE	UE, LE	Pain, swelling, physical force, range of motion, temperature	5-phase Tc-99m MDP	600 MBq	Diffusely increased periarticular uptake in the carpus or tarsus. Diffuse uptake in all small joints in the involved extremity.
O'Donoghue, 1993	Canada	78	NR	NR	NR	UE	Pain, trophic skin changes, vasomotor instability	3-phase Tc-99m MDP	550-740 MBq	Asymmetrical uptake, consistent with CRPS 1.
Okudan, 2005	Turkey	34	61	6.5	CVE	UE	Pain, swelling, vasomotor changes, tenderness in palpation	3-phase Tc-99m MDP	600 MBq	Diffuse asymmetric uptake on the blood flow, early and late blood pool images and increased periarticular uptakes on the delayed bone scan.
Park, 2007	Korea	38	52.2	0.3	CVE	UE	IASP	3-phase Tc-99m MDP	750 MBq	NR

Author	Country	No. of patients	Age (median)	Duration symptoms (weeks)	Initiating event	Localization	Reference standard	Index test	Radioisotope dose	Interpretation of a positive index test
Park, 2009	Korea	50	56	13.0	CVE, TBI	UE	IASP	3-phase Tc-99m DPD	580-620 MBq	All 3-phase images showed diffuse asymmetric periarticular uptake (early stage), blood flow and blood pool images showed decreased uptake or increased asymmetric periarticular accentuation (late stage)
Schierpers, 1998	Belgium	50	44	NR	Trauma, surgery	UE	NR	3-phase Tc-99m MDP	740 MBq	Diffusely increased periarticular uptake on the third phase of the scan.
Schurmann, 2007	Germany	118	59.9	8.0	Trauma ^a	UE	IASP, Bruehl	3-phase Tc-99m DPD	580-620 MBq	Arterial hyperperfusion of the affected hand, increased periarticular uptake on blood pool and delayed images.
Todorovic, 1995	Serbia	44	51	8.8	Trauma ^a	UE, LE	NR	3-phase Tc-99m DPD	555 MBq	Early phase: Increased circulation in the distal part of the affected limb in at least one of the images. Blood pool: Diffusely elevated radioactivity in the metacarpal and phalangeal areas of the hand or the corresponding areas of the foot. Delayed phase: Increased accumulation in all bones of the distal part of the affected limb.

Author	Country	No. of patients	Age (median)	Duration symptoms (weeks)	Initiating event	Localization	Reference standard	Index test	Radioisotope dose	Interpretation of a positive index test
Wang, 1998	China	30	63	6.1	CVE	UE	Pain at rest or with mobilization, swelling, vasomotor changes, tenderness	3-phase Tc-99m DPD	740 MBq	Diffusely increased uptake, in particular increased periarticular uptake involving multiple joints of the affected extremity.
Wüppenhorst, 2010	Germany	78	50.7	56.8	Trauma, surgery, spontaneous onset	UE	Bruehl	3-phase Tc-99m MDP	500-700 MBq	Increased tracer uptake on the affected side in each of the three phases.

Abbreviations: NR= not reported, CVE= cerebrovascular event, TBI= traumatic brain injury, UE= upper extremity, LE=lower extremity, IASP=International Association for the Study of Pain, MBq=mega Becquerel, Tc=technetium, MDP=methylene disphosphonate, DPD=disphosphono propan dicarbonat
^a Fracture

Figure 1. Study flow



Methodological quality

Table 2 shows a summary of the quality assessment of the included studies. Overall, the methodological quality was moderate. None of the included studies fulfilled all QUADAS criteria [15]. The main limitation was inadequate reporting. For example, in half of the included studies the reference standard was not likely to classify the target condition correctly [10, 19, 20, 23, 25, 26]. Furthermore, none of the included studies reported whether the index test results were interpreted without knowledge of the results of the reference standard.

Table 2. Methodological quality of the included studies (Quadas tool [15])

Author	Was the spectrum of participants representative of the patients who will receive the test in practice?	Were selection criteria clearly described?	Was the reference standard likely to classify the target condition correctly?	Was the period between performance of the reference standard and the index test short enough to be reasonably sure that the target condition did not change between the two?	Did the whole sample or a random selection of the sample receive verification using the reference standard?	Did participants receive the same reference standard regardless of the index test result?	Was the reference standard independent of the index test? (that is, the index test did not form part of the reference standard)	Was the execution of the index test described in sufficient detail to permit its replication?	Was the execution of the reference standard described in sufficient detail to permit its replication?	Were the index test results interpreted without knowledge of the results of the reference standard?	Were the reference standard results interpreted without knowledge of the results of the index test?	Were the same clinical data available when the test results were interpreted as would be available when the test is used in practice?	Were uninterpretable, indeterminate or intermediate test results reported?	Were withdrawals from the study explained?
Greyson, 1984	Yes	Yes	Yes	Unclear	Yes	Yes	Yes	Yes	Yes	Unclear	Yes	Yes	Yes	Yes
Kozin, 1981	Yes	Yes	Yes	Unclear	Yes	Yes	Yes	Yes	Yes	Unclear	Yes	Yes	Yes	No
Leitha, 1996	Yes	Yes	No	Unclear	Yes	Yes	Yes	Yes	Unclear	Unclear	Yes	Yes	No	Yes
O'Donoghue, 1993	No	Unclear	No	Unclear	Yes	Yes	Yes	Yes	Unclear	Unclear	Yes	Yes	No	Yes
Okudan, 2005	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	Unclear	Unclear	Yes	Yes	No	Yes
Park, 2007	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Unclear	Yes	Yes	No	Yes
Park, 2009	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Unclear	Yes	Yes	No	Yes
Schleppers, 1998	Yes	Yes	No	Unclear	Yes	Yes	Yes	Yes	No	Unclear	Yes	Yes	Yes	Yes
Schurmann, 2007	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Unclear	Yes	Yes	Yes	Yes
Todorovic, 1995	Yes	No	No	Yes	Yes	Yes	Yes	Yes	No	Unclear	Yes	Yes	No	Yes
Wang, 1998	Yes	Unclear	No	Yes	Yes	Yes	Yes	Yes	Unclear	Unclear	Yes	Yes	No	Yes
Wüppenhorst, 2010	Yes	Yes	Yes	Unclear	Yes	Yes	Yes	Yes	Yes	Unclear	Yes	Yes	No	No

Pooled sensitivity and specificity

Table 3 shows the summary estimates of the included studies. The pooled mean sensitivity of 12 two-by-two tables was 0.87 (95% CI, 0.68-0.97) and specificity was 0.69 (95% CI, 0.47-0.85) (Figure 2). The pooled mean sensitivity for the subgroup with clearly defined diagnostic criteria (6 two-by-two tables) was 0.80 (95% CI, 0.44-0.95) and specificity was 0.73 (95% CI, 0.40-0.91).

Table 3. Summary estimates for all included studies and the subgroup with a specified definition of CRPS 1

Summary estimates	Overall	Subgroup ^a
Sensitivity (95% CI ^b)	0.87 (0.68 to 0.97)	0.80 (0.44 to 0.95)
Specificity (95% CI)	0.69 (0.47 to 0.85)	0.73 (0.40 to 0.91)
DOR ^c (95% CI)	17.23 (7.71 to 38.52)	10.6 (3.76 to 29.72)
LR+ ^d (95% CI)	2.83 (1.68 to 4.79)	2.92 (1.33 to 6.43)
LR- ^e (95% CI)	0.16 (0.06 to 0.42)	0.28 (0.10 to 0.76)
Prevalence	0.51	0.48
CRPS probability negative test	0.14 (0.06-0.31)	0.20 (0.08-0.41)
CRPS probability positive test	0.75 (0.64-0.83)	0.73(0.55-0.85)

a Specifying CRPS definition

b Confidence interval

c Receiver operating characteristic

d pooled positive likelihood ratio

e pooled negative likelihood ratio

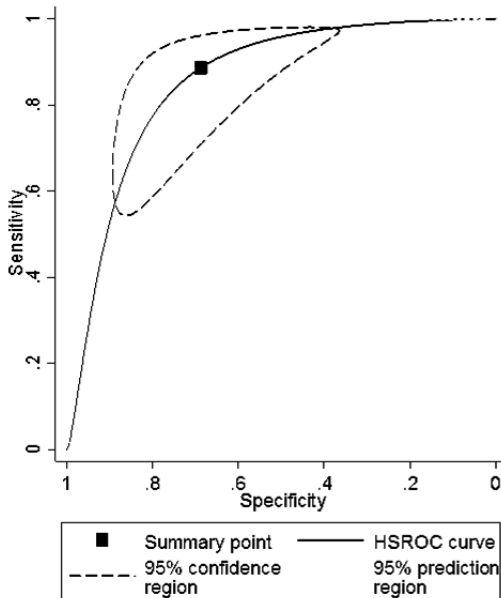


Figure 2. Diagnosed CRPS 1 (overall) – pathologic correlative

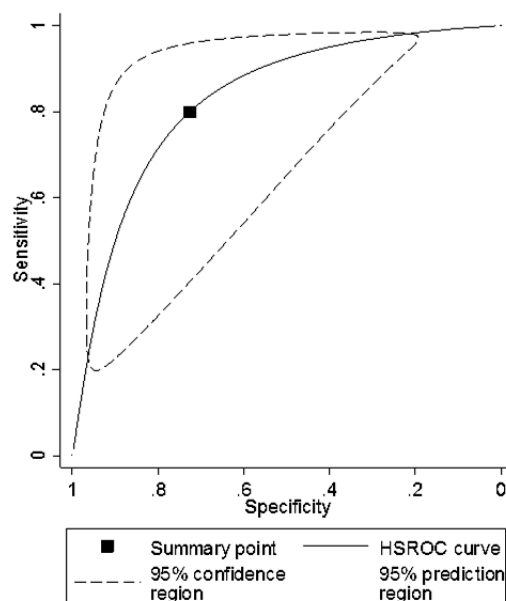
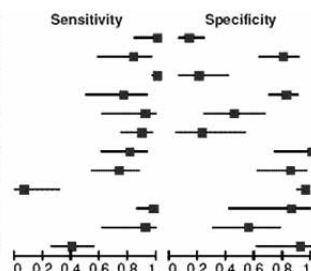


Figure 3. Diagnosed CRPS 1 (with reported diagnostic criteria) – pathologic correlative

HSOC= hierarchical summary receiver operating characteristic

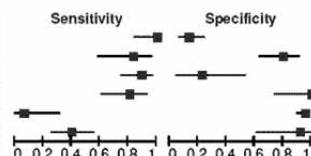
Figure 4. True positive (TP) and false positive (FP), true negative (TN) and false negative (FN), sensitivity and specificity for each study (grouped with and without a specified definition of CRPS 1 [95% confidence interval])

Study	TP	FP	FN	TN	Sensitivity	Specificity
Greyson 1984	21	55	0	9	1.00 [0.84, 1.00]	0.14 [0.07, 0.25]
Kozin 1981	15	7	3	28	0.83 [0.59, 0.96]	0.80 [0.63, 0.92]
Leitha 1996	96	19	0	5	1.00 [0.96, 1.00]	0.21 [0.07, 0.42]
O'Donoghue 1993	13	11	4	50	0.76 [0.50, 0.93]	0.82 [0.70, 0.91]
Okudan 2005	11	12	1	10	0.92 [0.62, 1.00]	0.45 [0.24, 0.68]
Park 2007	33	10	4	3	0.89 [0.75, 0.97]	0.23 [0.05, 0.54]
Park 2009	21	0	5	12	0.81 [0.61, 0.93]	1.00 [0.74, 1.00]
Schiepers 1998	22	3	8	17	0.73 [0.54, 0.88]	0.85 [0.62, 0.97]
Schurmann 2007	1	4	14	88	0.07 [0.00, 0.32]	0.96 [0.89, 0.99]
Todorovic 1995	36	1	1	6	0.97 [0.86, 1.00]	0.86 [0.42, 1.00]
Wang 1998	11	8	1	10	0.92 [0.62, 1.00]	0.56 [0.31, 0.78]
Wuppenhorst 2010	18	1	27	11	0.40 [0.26, 0.56]	0.92 [0.62, 1.00]



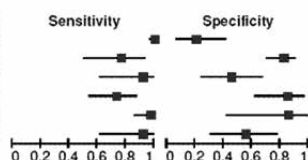
With specified diagnostic criteria

Study	TP	FP	FN	TN	Sensitivity	Specificity
Greyson 1984	21	55	0	9	1.00 [0.84, 1.00]	0.14 [0.07, 0.25]
Kozin 1981	15	7	3	28	0.83 [0.59, 0.96]	0.80 [0.63, 0.92]
Park 2007	33	10	4	3	0.89 [0.75, 0.97]	0.23 [0.05, 0.54]
Park 2009	21	0	5	12	0.81 [0.61, 0.93]	1.00 [0.74, 1.00]
Schurmann 2007	1	4	14	88	0.07 [0.00, 0.32]	0.96 [0.89, 0.99]
Wuppenhorst 2010	18	1	27	11	0.40 [0.26, 0.56]	0.92 [0.62, 1.00]



Without specified diagnostic criteria

Study	TP	FP	FN	TN	Sensitivity	Specificity
Leitha 1996	96	19	0	5	1.00 [0.96, 1.00]	0.21 [0.07, 0.42]
O'Donoghue 1993	13	11	4	50	0.76 [0.50, 0.93]	0.82 [0.70, 0.91]
Okudan 2005	11	12	1	10	0.92 [0.62, 1.00]	0.45 [0.24, 0.68]
Schiepers 1998	22	3	8	17	0.73 [0.54, 0.88]	0.85 [0.62, 0.97]
Todorovic 1995	36	1	1	6	0.97 [0.86, 1.00]	0.86 [0.42, 1.00]
Wang 1998	11	8	1	10	0.92 [0.62, 1.00]	0.56 [0.31, 0.78]



Discussion

Main findings

The pooled data revealed an adequate agreement between a positive qualitative BS and the presence of CRPS 1 (sensitivity) and only moderate agreement between a negative BS in the absence of CRPS 1 (specificity). The considerable number of instances of positive BS scans in the absence of clinical CRPS 1 might be due to other clinical conditions mimicking similar biological processes other than CRPS 1 such as posttraumatic or postoperative bone affections. However, these results did not change greatly if the clinical diagnosis was based on established diagnostic criteria.

Results in light of the existing literature

To the best to our knowledge, this is the first attempt to assess the concordance of BS findings with the presence or absence of CRPS 1 in a systematic manner.

In 1995, Lee et al summarized the validity of scintigraphy in diagnosing CRPS of the upper extremity [9]. Their data showed a wide variability in scintigraphic accuracy in patients with clinically obvious CRPS 1. The results of BS correlated best with the clinical diagnosis of CRPS within the first 20-26 weeks of onset but the sensitivity was only about 50%. After 26 weeks, there was a poor correlation between BS and CRPS 1. The authors recommended that TPBS should not be used as a major criterion in diagnosing CRPS and that the diagnosis remained a clinical diagnosis made by an experienced clinician. The methodological quality of this review is moderate because the authors did not report the detailed search strategy and failed to perform a meta-analysis.

Our results are in line with several available guidelines. For example, the current evidence based guideline published by the Netherlands Society of Rehabilitation Specialists and the Netherlands Society of Anaesthesiologists in 2006 does not recommend the use of BS for the diagnosis of CRPS 1 [13]. The multidisciplinary committee argues that because of methodological shortcomings of the available literature, a sufficient discriminatory capacity and an additional diagnostic value of BS in diagnosing CRPS 1 could not be established. In a recent Delphi survey, an international expert panel agreed that CRPS 1 is a clinical diagnosis and no further examinations are necessary to confirm the diagnosis [30].

Strength/limitation

The strengths of this meta-analysis include the broad literature searches and careful data extraction, and the assessments of methodological quality of diagnostic accuracy tests. We modeled sensitivity and specificity by applying a valid meta-analytic method, which incorporates both within and between-study variance [16].

In some of the trials there were major methodological limitations that may have introduced bias in our results. For example, only seven studies used established diagnostic criteria for the diagnosis of CRPS 1. Two studies did not even report the detailed clinical manifestations on which the diagnosis was based. However, our meta-analysis of the subgroup having used well established diagnostic criteria did not reveal a significant improvement of the results. One major limitation of our study is the lack of a valid reference test for the diagnosis of CRPS 1. As already stated in the introduction the diagnosis of CRPS 1 remains a challenge in clinical practice because it is solely based on clinical manifestations.

Implication for practice

Based on the results of our meta-analysis, BS does not add any benefits to the clinical diagnosis of CRPS 1 and should not be used for confirmatory purposes. The diagnosis should be still be based on presenting signs and symptoms according to the current diagnostic criteria [5].

Implication for research

Given the unknown somatic anomaly underlying CRPS 1, further studies evaluating diagnostic accuracy are questionable. Therefore, future research should rather focus on prognostic aspects of CRPS 1 and for example identify potential subgroups with different prognostic outcome. For example, we are unaware of any study investigating whether the subgroup of CRPS 1 patients with a positive BS is related with a different disease severity or prognosis. On the other hand, sustained efforts should flow into a deeper understanding of the somatic anomaly underlying CRPS coming from basic scientific inquiry. Further research should also focus on investigating the underlying mechanisms of CRPS 1. To date, three major pathophysiological pathways have been identified for CRPS 1 [2]: Aberrant inflammatory mechanisms, vasomotor dysfunction, and maladaptive neuroplasticity. It's still a matter of debate which of these pathways is responsible for the bony alterations in CRPS 1 and why these changes only occur in a subset of patients. It has been hypothesized that the increased tracer uptake in a positive BS results from a neurogenic inflammation, a decreased sympathetic activity or a demineralization of the bone because of an activated bone metabolism [27].

Conclusions

Based on this study, clinicians must be advised that a positive BS is not necessarily concordant with presence of absence of CRPS 1. Given the moderate level of concordance between a positive BS in the absence of clinical CRPS 1, discordant results potentially impede the diagnosis of CRPS 1.

References

1. Merskey, H. and Bogduk, N., Classification of chronic pain: description of chronic pain syndrome and definitions of pain terms. 2nd. ed1994, Seattle: IASP Press.
2. Marinus, J., Moseley, G.L., Birklein, F., et al., Clinical features and pathophysiology of complex regional pain syndrome. *Lancet Neurol*, 2011. 10(7): p. 637-48.
3. van Eijs, F., Stanton-Hicks, M., Van Zundert, J., et al., Evidence-based interventional pain medicine according to clinical diagnoses. 16. Complex regional pain syndrome. *Pain Pract*, 2011. 11(1): p. 70-87.
4. Steurer, J., Bachmann, L.M., and Miettinen, O.S., Etiology in a taxonomy of illnesses. *Eur J Epidemiol*, 2006. 21(2): p. 85-9.
5. Harden, R.N., Bruehl, S., Stanton-Hicks, M., et al., Proposed new diagnostic criteria for complex regional pain syndrome. *Pain Med*, 2007. 8(4): p. 326-31.
6. Holder, L.E. and Mackinnon, S.E., Reflex sympathetic dystrophy in the hands: clinical and scintigraphic criteria. *Radiology*, 1984. 152(2): p. 517-22.
7. Mackinnon, S.E. and Holder, L.E., The use of three-phase radionuclide bone scanning in the diagnosis of reflex sympathetic dystrophy. *J Hand Surg Am*, 1984. 9(4): p. 556-63.
8. Simon, H. and Carlson, D.H., The use of bone scanning in the diagnosis of reflex sympathetic dystrophy. *Clin Nucl Med*, 1980. 5(3): p. 116-21.
9. Lee, G.W. and Weeks, P.M., The role of bone scintigraphy in diagnosing reflex sympathetic dystrophy. *J Hand Surg Am*, 1995. 20(3): p. 458-63.
10. O'Donoghue, J.P., Powe, J.E., Mattar, A.G., et al., Three-phase bone scintigraphy. Asymmetric patterns in the upper extremities of asymptomatic normals and reflex sympathetic dystrophy patients. *Clin Nucl Med*, 1993. 18(10): p. 829-36.
11. Steinert, H. and Hahn, K., [The value of 3-phase skeletal scintigraphy for early diagnosis of Sudeck disease]. *Rofo*, 1996. 164(4): p. 318-23.
12. Demangeat, J.L., Constantinesco, A., Brunot, B., et al., Three-phase bone scanning in reflex sympathetic dystrophy of the hand. *J Nucl Med*, 1988. 29(1): p. 26-32.
13. Geertzen, J.H.B., Perez, R.S.G.M., Dijkstra, P.U., et al., Complex regional pain syndrome 1 guidelines., N.S.o.R. Specialists and Anaesthesiologists, N.S.o., Editors. 2006, Van Zuiden Communications B.V.: Alphen an den Rijn.
14. Moher, D., Altman, D.G., Liberati, A., et al., PRISMA statement. *Epidemiology*, 2011. 22(1): p. 128; author reply 128.
15. Whiting, P., Rutjes, A.W., Reitsma, J.B., et al., The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Med Res Methodol*, 2003. 3: p. 25.
16. Harbord, R.M., Deeks, J.J., Egger, M., et al., A unification of models for meta-analysis of diagnostic accuracy studies. *Biostatistics*, 2007. 8(2): p. 239-51.
17. Greyson, N.D. and Tepperman, P.S., Three-phase bone studies in hemiplegia with reflex sympathetic dystrophy and the effect of disuse. *J Nucl Med*, 1984. 25(4): p. 423-9.
18. Kozin, F., Soin, J.S., Ryan, L.M., et al., Bone scintigraphy in the reflex sympathetic dystrophy syndrome. *Radiology*, 1981. 138(2): p. 437-43.
19. Leitha, T., Staudenherz, A., Korpan, M., et al., Pattern recognition in five-phase bone scintigraphy: diagnostic patterns of reflex sympathetic dystrophy in adults. *Eur J Nucl Med*, 1996. 23(3): p. 256-62.
20. Okudan, B., Celik, C., Serttas, S., et al., The predictive value of additional late blood pool imaging to the three-phase bone scan in the diagnosis of reflex sympathetic dystrophy in hemiplegic patients. *Rheumatol Int*, 2005. 26(2): p. 126-31.
21. Park, S.A., Yang, C.Y., Kim, C.G., et al., Patterns of three-phase bone scintigraphy according to the time course of complex regional pain syndrome type I after a stroke or traumatic brain injury. *Clin Nucl Med*, 2009. 34(11): p. 773-6.
22. Park, S.G., Hyun, J.K., Lee, S.J., et al., Quantitative evaluation of very acute stage of complex regional pain syndrome after stroke using three-phase bone scintigraphy. *Nucl Med Commun*, 2007. 28(10): p. 766-70.

23. Schiepers, C., Bormans, I., and De Roo, M., Three-phase bone scan and dynamic vascular scintigraphy in algoneurodystrophy of the upper extremity. *Acta Orthop Belg*, 1998. 64(3): p. 322-7.
24. Schurmann, M., Zaspel, J., Lohr, P., et al., Imaging in early posttraumatic complex regional pain syndrome: a comparison of diagnostic methods. *Clin J Pain*, 2007. 23(5): p. 449-57.
25. Todorovic-Tirnanic, M., Obradovic, V., Han, R., et al., Diagnostic approach to reflex sympathetic dystrophy after fracture: radiography or bone scintigraphy? *Eur J Nucl Med*, 1995. 22(10): p. 1187-93.
26. Wang, Y.L., Tsau, J.C., Huang, M.H., et al., Reflex sympathetic dystrophy syndrome in stroke patients with hemiplegia-three phase bone scintigraphy and clinical characteristics. *Kaohsiung J Med Sci*, 1998. 14(1): p. 40-7.
27. Wuppenhorst, N., Maier, C., Frettlow, J., et al., Sensitivity and specificity of 3-phase bone scintigraphy in the diagnosis of complex regional pain syndrome of the upper extremity. *Clin J Pain*, 2010. 26(3): p. 182-9.
28. Kozin, F. The painful shoulder and the reflex sympathetic dystrophy syndrome., in *Arthritis and Allied Conditions*, D.J. McCarthy, Editor 1979, Lea & Febinger: Philadelphia. p. 1097-1120.
29. Bruhl, S., Harden, R.N., Galer, B.S., et al., External validation of IASP diagnostic criteria for Complex Regional Pain Syndrome and proposed research diagnostic criteria. *International Association for the Study of Pain. Pain*, 1999. 81(1-2): p. 147-54.
30. Brunner, F., Lienhardt, S.B., Kissling, R.O., et al., Diagnostic criteria and follow-up parameters in complex regional pain syndrome type I - a Delphi survey. *Eur J Pain*, 2008. 12(1): p. 48-52.

Appendix 1:
Checklist on study in-/exclusion

<hr/>		
Autor, Year:		
Inclusion criteria	Yes	No
<hr/>		
Diagnostic study		
Clinical diagnosis of CRPS 1 = Index test		
Bone scintigraphy = Reference standard		
Enough information to fill the two by two table (absolute number of true-positive, false-positive, false-negative , true-negative observations)		
Definitive inclusion		
<hr/>		

CHAPTER 6

Prognostic factors in complex regional pain syndrome 1: A systematic review

Maria M. Wertli, MD, Lucas M. Bachmann, MD, PhD, Shira Schechter Weiner, PT, PhD, CIE, Florian Brunner, MD, PhD
J Rehabil Med 2013; 45: 225–231

Abstract

Objective: The aim of this systematic review is to merge and summarize the current evidence of prognostic factors relevant for the course of Complex Regional Pain Syndrome 1.

Methods: We searched Medline, Embase, PsychINFO, CENTRAL and screened reference lists of included studies. We considered any type of study investigating parameters associated with the prognosis of the condition.

Results: Searches retrieved 2577 references of which 14 articles were included. The preferred diagnostic criteria were the Veldman and the International Association for the Study of Pain criteria. The general level of study quality was insufficient. We found 34 parameters for the following seven different clusters: clinical characteristics, type of initiating event, localization, gender, older age, contextual factors and time point of diagnosis. Most consistency was found for delayed recovery of patients with sensory disturbances and cold Complex Regional Pain Syndrome 1. For many parameters the evidence is contradictory.

Conclusion: Evidence is scarce and contradictory and impedes drawing firm conclusions for practice. Further high quality research in this field is needed and we propose concerted action in etiologic and clinical research.

Introduction

Complex Regional Pain Syndrome (CRPS) Type 1 is a syndrome with a significant morbidity and loss of quality of life [1, 2]. It usually appears after a noxious event such as trauma or surgery [3] and the clinical manifestations includes sensory, autonomic, motor and trophic changes [4]. Despite increasing research in this field, the exact underlying mechanisms are still unknown. In a recent review article, Marinus et al. concluded that compelling evidence implicates biological pathways that underlie aberrant inflammation, vasomotor dysfunction, and maladaptive neuroplasticity in the clinical features of CRPS [5]. In contrast to CRPS Type 2, which is characterized by a definable nerve lesion, CRPS Type 1, formerly known as reflex sympathetic dystrophy or algodystrophy, appears without definable nerve lesion [4].

There is broad consensus in the literature that once diagnosed, every patient should be treated early and aggressively in the hope of preventing chronicity [6, 7]. Treatment is based on the pathomechanisms mentioned above and usually includes a multidisciplinary approach using a combination of pharmacologic, physical, occupational and psychological therapies [6]. However, clinical observation reveals that in a substantial proportion of patients, resolution occurs spontaneously or the natural course is benign [8] sometimes even without treatment. A subgroup of patients suffering from CRPS 1 will experience an unfavorable course of the disease and will contribute to high healthcare costs. If this subgroup of patients could be identified at an early stage i.e. with prognostic instruments, treatment activities could be focused and specifically tailored to fit the needs of these patients.

Until now, evidence regarding prognostic aspects of CRPS 1 has not been assessed systematically. The literature is scattered and not easy to access. Therefore, the purpose of this systematic review was to merge and summarize the current evidence of prognostic factors relevant for the course of CRPS 1. We defined prognostic factors as all clinical and non-clinical parameters with relevant impact on clinical course and treatment response, reflected by persisting impairment, disease duration and long term disability.

Methods

Literature search

The search method was used according to the MOOSE statement (figure 1) on conducting a meta-analysis of observational studies in epidemiology [9]. We identified all observational studies investigating prognostic factors of CRPS 1, published between 1990 and July 2011, by searching the following databases: MEDLINE (OvidSP), MEDLINE In-Process Citations (OvidSP), Embase (OvidSP), Embase (OvidSP), PsychINFO and the

Cochrane Database of Systematic Reviews (CDSR). We restricted our searches to 1990 onwards, because the current definition of Complex Regional Pain Syndrome was introduced in the early nineties. The search was conducted with the help of an experienced information specialist working in the field of systematic reviews. Search terms included in addition to medical subject headings (MeSH terms) all commonly used terms for CRPS (e.g. complex regional pain syndrome(s), reflex sympathetic dystrophy, sudeck atrophy, algodystrophy, shoulder hand syndrome). A detailed search strategy is depicted in Table 1. To ensure the completeness of the literature search, the reviewers, experienced researchers in the field of CRPS, screened bibliographies of all included studies, retrieved review articles and current treatment guidelines in an additional hand search and all (inclusion and exclusion criteria applied) potential studies were additionally included.

Table 1. Search strategy Medline (OvidSP) (1990-2011/07/wk 27)

	Search	Results
1	complex regional pain syndromes/ or reflex sympathetic dystrophy/	3520
2	(CRPS1 or CRPS or complex regional pain syndrome\$ or RND).ti,ab.	1823
3	(reflex\$ sympathetic dystroph\$ or sudeck\$ atroph\$ or algodystroph\$ or algoneurodystroph\$).ti,ab.	2065
4	(algo dystroph\$ or algo neurodystroph\$).ti,ab.	13
5	(shoulder hand syndrom\$ or shoulder hand dystroph\$).ti,ab.	259
6	cervical sympathetic dystroph\$.ti,ab.	0
7	or/1-7	5017
8	animals/ not (animals/ and humans/)	10061
9	8 not 9	5015
10	prognosis/ or exp treatment outcome/	769448
11	(outcome\$ or predict\$ or prognosis or recover\$ or remission or relaps\$ or deteriorat\$ or exacerbat\$ or worsen\$ or course\$).ti,ab.	2146668
12	(Cure\$ or curative\$ or resolv\$ or resolution\$ or heal\$ or improv\$ or recuperat\$).ti,ab.	2498217
13	(convales\$ or alleviat\$ or decreas\$ or lessen\$).ti,ab.	1408991
14	or/11-14	5273995
15	10 and 15	1869
16	limit 16 to yr="1990 -Current"	1648

Study selection, data extraction and synthesis

The bibliographic details of all retrieved articles were stored in an Endnote file. Two reviewers (MW and FB) independently screened all references by title and abstract. We selected observational studies investigating prognostic parameters of CRPS 1. We did not apply any language restrictions. All included references were independently reviewed in full text (MW and FB). During the screening and inclusion process all disa-

agreements were discussed between the two reviewers and resolved by consensus. A designated third author (LMB) arbitrated any disagreement and facilitated consensus. Based on this review we extracted and catalogued all reported prognostic factors and data on salient clinical features. For any abstract where the full text was not available, the author was contacted. Alternative researchers with specific language proficiencies were used for non-English language references. For descriptive purposes and to weight the included studies, study quality was assessed according to the proposed guidelines for assessing Quality in Prognostic Studies [10] in a two step approach. Two reviewers (FB and MW) first individually compiled the fully operationalized, prognostic factors (correlation between the prognostic factor and the outcome). Second, these prognostic factor responses were tested for each of six potential biases: representative study population, drop out, adequate measurement of the prognostic factor, outcome measurement, confounding measurement and account, and analysis. Reviewers discussed the independent ratings and sought consensus about the overall risk of bias. A summary is reported in Table 2. As the included studies did not allow a statistical synthesis of outcome measures, quality criteria were used for descriptive purposes only and not for exclusion criteria [10]. Synthesis of quality was categorized as good (good for all six potential biases), acceptable (at least partly fulfilling all six criteria) and poor.

Results

Study selection

Figure 1 shows the study selection process and agreement on study inclusion. Our search retrieved 2577 records from which 64 were identified for full review based on title and abstract. Full text assessment utilizing inclusion and exclusion criteria resulted in exclusion of 50 studies. The main reasons for exclusion were study design (clinical trials) and outcome measures (no prognostic factors investigated). In total 14 studies fulfilled the inclusion criteria [8, 11-23].

Study characteristics

The study characteristics are summarized in Table 3. In total, 2298 subjects (median number of subjects 60, range 16-106) were investigated. Symptom duration ranged from less than three months [12-14] to more than eight years [21].

Seven studies followed a prospective study design. With the exception of one study [17], the prospective studies, included fewer patients ($n=16-47$, mean 28) when compared to retrospective design ($n=42-199$, mean 98).

Based on the quality assessment, quality was good in four, acceptable in one and poor in nine studies (Table 2).

The preferred diagnostic criteria were the Veldman criteria (n=5) [12, 19-22] and the International Association for the Study of Pain (IASP) criteria (n=2) [8, 15]. In one prospective and two retrospective studies no diagnostic criteria were reported [11, 17, 23].

Figure 1. Study flow

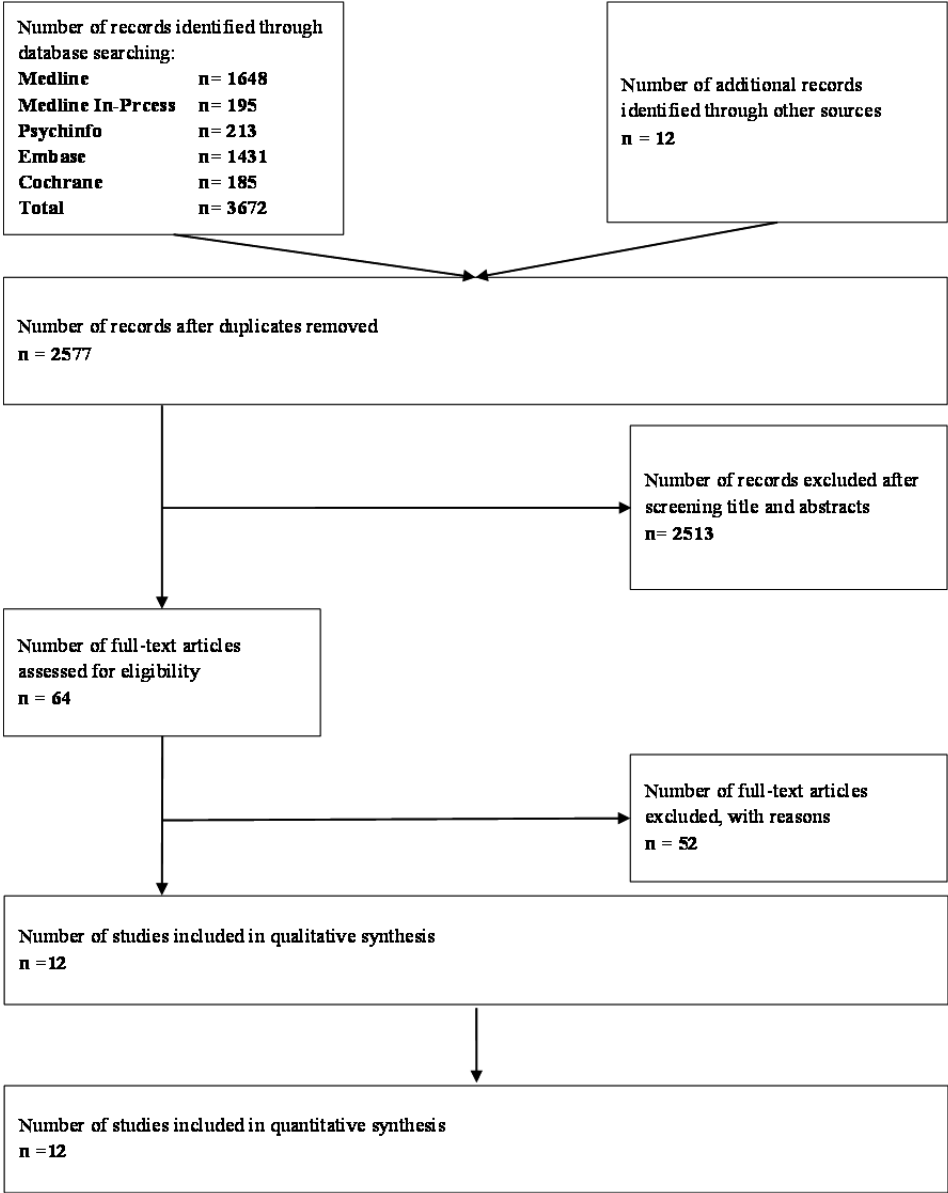


Table 2. Methodological quality of the included studies

Author	Year	Study participation ¹⁾	Study attrition ²⁾	Prognostic factor measurement ³⁾	Outcome measurement ⁴⁾	Confounding ⁵⁾	Analysis ⁶⁾	Synthesis of Quality
Bejja et al. [11]	2005	Unsure	Unsure	Yes	Yes	Unsure	Yes	Poor
Dauty et al. [12]	2001	Partly	Yes	Yes	No	Unsure	Unsure	Poor
Daviet et al. [13]	2001	Yes	Unsure	Yes	Unsure	Unsure	No	Poor
Daviet et al. [14]	2002	Yes	Yes	Yes	Yes	Yes	Yes	Good
De Mos et al. [15]	2009	Yes	Yes	Yes	Yes	Yes	Yes	Good
Eulry et al. [16]	1990	Partly	Yes	Unsure	Unsure	Unsure	No	Poor
Goris et al. [17]	1990	Yes	Unsure	Yes	Yes	Unsure	Yes	poor
Laulan et al. [18]	1997	Yes	Yes	Yes	Yes	Unsure	Yes	Poor
Sandroni et al. [8]	2003	Partly	Yes	Yes	Yes	Yes	Yes	Acceptable
Tan et al. [19]	2009	Yes	Yes	Unsure	Yes	Unsure	Yes	Poor
Van der Laan et al. [20]	1998	Unsure	No	Unsure	Yes	Unsure	Yes	Poor
Vaneker et al. [21]	2005	Yes	Yes	Yes	Yes	Yes	Yes	Good
Vaneker et al. [22]	2006	Yes	Yes	Yes	Yes	Yes	Yes	Good
Zyluk [23]	1998	Yes	Unsure	Yes	Yes	Yes	Yes	Poor

¹⁾ Study participation: study sample represents population of interest: yes, partly, no, unsure

²⁾ Study attrition: Loss to follow-up is not associated with key characteristics

³⁾ Prognostic factor measurement: prognostic factor of interest is adequately measured

⁴⁾ Outcome measurement: outcome of interest is adequately measured in study participants

⁵⁾ Confounding measurement and account: important potential confounders are appropriately accounted for?

⁶⁾ Analysis Statistical analysis is appropriate for the desing of the study

Prognostic factors

A wide spectrum of outcome parameters were investigated including self reported symptoms [8, 19, 21, 23], clinical severity scores [11, 14, 18, 22], development of complications [20], duration of disease [17] and return to work [12, 15]. The included studies revealed a total of 34 prognostic factors. A variety of clinical manifestations of CRPS 1 were linked to both positive (n=5) and negative prognosis (n=13) and are in detail

summarized in Table 4. In particular, cold CRPS 1 (primary or secondary cold skin temperature, n=5) and the presence of sensory disturbances (n=4) seem to represent parameters associated with a poor prognosis of CRPS 1. One study reported that with the onset of CRPS 1 after fracture, factors associated with a high resolution rate include the presence of swelling and the absence of sensory changes [8].

We found no study developing or validating a prognostic instrument to identify patients with poor CRPS 1 prognosis. Moreover, the evidence at hand did not allow for developing a diagnostic algorithm.

Table 3. Characteristics of the included studies (n=14)

Author	Year	Study design	Diagnostic criteria	Disease duration	Initiating event	Location	N (f/m)	Age (years)	Positive prognostic factors	Negative prognostic factors	Outcome
Beija et al. [11]	2005	Retro-spective	NR	16.7±25.5 months	Traumatic, atraumatic	UE, LE	60 (40/20)	51.6 (16-81)	Simple CRPS: result satisfactory 91 % (vs. 55% in secondary)	Male gender Age <40 years Cold CRPS Secondary diagnosis Delayed diagnosis (>2months)	Outcome criteria for treatment response of the society of rheumatology France. Very good, good, medium, bad.
Dauty et al. [12]	2001	Pros-pective	Veldman	35±10.2 days	Traumatic	UE, LE	16 (1/15)	34±11 (22-59)		Distal articular location (wrist, hand, ankle foot) Comorbidities (e.g. alcoholism) Polytrauma	Return to work
Daviet et al. [13]	2001	Pros-pective	Severity scale of Labrousse	<3 months	Stroke	UE	28 (NR)	70,7 ± 10,9	Perrigot score 0-1	Motor deficit Spasticity Perrigot score 2-5	Severity of CRPS 1
Daviet et al. [14]	2002	Pros-pective	Severity scale of Labrousse	<3 months	Stroke	UE	24 (NR)	NR	Perrigot score 0-1	Motor deficit Spasticity Sensory deficit Initial coma Perrigot score: 2-5	Severity scale of Labrousse
De Mos et al. [15]	2009	Retro-spective	IASP	5.8 years (2.1-10.8)	NR	NR	102 (81/21)	51 (12-86)		Upper extremity Event other than fracture Cold CRPS	Self reported recovery and work status
Eulry et al. [16]	1990	Retro-spective	Doury	Atraumatic: 5.2 months (0.5-22) Traumatic: 3.7	Traumatic, atraumatic	LE	199 (60/139)	34.2±17.4 (11-84)		Psychological background in non traumatic CRPS	Improvement or cure

Author	Year	Study design	Diagnostic criteria	Disease duration (0.1-13)	Initiating event	Location	N (f/m)	Age (years)	Positive prognostic factors	Negative prognostic factors	Outcome
Goris et al. [17]	1990	Prospective	NR	NR	Traumatic, atraumatic	UE, LE	441 (316/125)	10-83 (44)		Cold CRPS Exacerbation of the symptoms under influence of muscular exercise	Duration of disease
Laulan et al. [18]	1997	Prospective	Own criteria	NR	Wrist fracture	UE	102 (68/34)	54.8		Severe initial injury Pain at distal radioulnar joint Clinical algodystrophy score >7 at 6 weeks	Only 26% with CRPS 1. Clinical score with 5 parameters (severity 1-4: range 5-20)
Sandroni et al. [8]	2003	Retrospective	IASP	11.6±12.4 months	Traumatic, atraumatic	UE, LE	74 (NR)	46.9±16 (15-86)	Fracture RR 91% Absence of sensory symptoms Presence of objective swelling	Sprain RR 78% Miscellaneous RR 55%	Resolution rate (RR in %)
Tan et al. [19]	2009	Retrospective	Veldman	NR	Traumatic, atraumatic (onset of CRPS in childhood, age <16 years)	UE, LE	42 (37/5)	25.5 (16-34)	Disease duration on physical functioning on SF36	Male gender Older age, low score on SF36 on general health and physical functioning)	SF-36
Van der Laan et al. [20]	1998	Retrospective	Veldman	NR	Traumatic, atraumatic	UE, LE	1006 (763/243)	43.3		Young, female Lower extremity Initial cold CRPS Complications (infection, ulcers, chronic edema, dystonia, myoclonus)	Development of complications and development of CRPS in another limb
Vaneker et al. [21]	2005	Prospective	Veldman	>8 years	NR	UE	47 (33/14)	58±15		Cold CRPS 1 (pain outcome, signs of central sensitization	Self reporting: symptomatic or not,

Author	Year	Study design	Diagnostic criteria	Disease duration	Initiating event	Location	N (f/m)	Age (years)	Positive prognostic factors	Negative prognostic factors	Outcome
Vaneker et al. [22]	2006	Retro-spective	Veldman	<12 months	NR	UE	45 (32/13)	51±15	No significant differences between warm and cold subgroups regarding function	with disease progression) Increased pressure hyperalgesia on affected limb with disease progression Sub score VAS best predictive power over time and provides best prediction of long-term functional impairment.	VAS: Warm or cold at diagnosis Impairment Level SumScore (ISS)
Zyluk [23]	1998	Retro-spective	NR	NR	Traumatic, atraumatic	UE	112 (75/37)	58.2 (23-84)		Symptoms > 12 months, 2° and 3° stage, coexistence of misdiagnosed nerve injuries or compression	Good outcome (little or no pain, normal range of motion), moderate, poor (=persistence of spontaneous pain +/- malfunction of the hand)

UE= Upper extremity, LE= Lower extremity, NR= Not reported

Table 4. Prognostic factors grouped within seven clinical clusters

Cluster	Positive prognostic factors	Negative prognostic factors
Gender (n=3)		Female [20] Male [11, 19]
Age (n=3)		Age <40 years [11] Population with complication, “younger” [20] median 35 years (range 15-65) vs. 44 years (range 10-84) Age [19]
Diagnosis (n=1)		Delayed diagnosis (>2 months) [11]
Initiating event (n=5)	Fracture [8] Swelling [8]	Polytrauma [12] Event other than fracture [15] Severe initial injury [18]
Localization (n=4)		Distal articular location [12, 18] Upper extremity [15] Lower extremity [20]
Clinical features (n=29)	Primitive CRPS [11] Absence of sensory changes [8] Swelling [8] Perrigot score [13] Disease duration [19]	Exercise induced pain [17] Sensory disturbances [14, 18, 21, 22] Primary cold CRPS [20] Cold CRPS [11, 15, 17, 20, 21] Motor disturbances [13, 14] Spasticity [13, 14] Complications (infection, ulcers, chronic edema, dystonia, myoclonus) [20, 21] Initial coma [14] High Perrigot score [13, 14] Clinical algodystrophy score >7 [8] General health [19] Disease duration [23] Coexistence of misdiagnosed nerve injury or compression [23]
Contextual factors (n=2)		Comorbidities [12] Psychological background in non traumatic CRPS [16]

Discussion

Main findings

This systematic review revealed a wide scatter of general prognostic factors in CRPS 1 that were sometimes contradictory. Most consistency was found for clinical manifestations such as the presence of sensory disturbances [14, 18, 21, 22] and cold CRPS 1 [11, 15, 17, 20, 21]. Only few studies used reliable and validated measures to assess prognostic factors and co-factors that might influence the course of the condition. Therefore, no assumption can be made regarding the causality of these findings. We failed to

quantify and rank order prognostic parameters and we were unable to derive an algorithm allowing clinicians to assess patients' prognosis at early stages of disease.

Results in light of existing Literature

To the best of our knowledge, this is the first attempt to systematically identify and consolidate prognostic factors influencing the course of CRPS 1. Despite broad inclusion criteria only a small number of studies fulfilled our inclusion criteria. This suggests that prognostic aspects of CRPS 1 have only received little attention in methodological high level research.

Our findings are in line with the results of a recently published Delphi survey [24]. In the absence of evidence-based prognostic factors, the authors performed a survey aimed at reaching an expert consensus on poor prognostic factors in CRPS 1. The expert panel agreed on 49 items which in their opinion are associated with poor prognosis in CRPS 1. These factors consisted primarily of clinical manifestations such as sensory disturbances and primary cold CRPS 1. For many of the factors, we found only weak (disease duration, comorbidities) or conflicting evidence (localization, initiating event). For example the consensus was reached that the onset after a fracture will likely result in a poor prognosis whereas Sandroni et al. found a higher resolution rate if CRPS 1 appeared after a fracture than after other triggering events [8].

Strength and limitation

The strength of this study is the application of a robust systematic review methodology. A comprehensive review of the literature based on a broad search of all relevant databases and the additional search of bibliographies of included studies, retrieved review articles and current treatment guidelines reduces the risk of a selection bias. All relevant reports were searched systematically and without language restriction.

The main limitation of our study was the quality of the included studies. The average number of subjects in the included studies was small and the quality of retrospective and prospective reporting was often limited. Moreover, different diagnostic criteria for CRPS 1 were used and three studies did not report any definition of CRPS1. We found no systematic difference in the patterns of prognostic factors reported and therefore decided to report the results jointly. Variability in study characteristics, diagnostic criteria, outcomes and follow-up periods impeded us thus from providing a quantitative analysis, resulting in only a systematic description and an inventory of prognostic parameters.

Implication for research

The results of this review suggest the need for large prospective cohort studies aimed at predictors of outcome and effective interventions. CRPS 1 is a syndromal condition and all currently used diagnostic criteria show relatively good negative but only moderate positive predictive values (increased likelihood of false positives) [25]. Until diagnostic criteria improve, future research should focus on identification of relevant prognostic factors that influence the course of disease. A defined prognostic profile may compensate for the current lack of diagnostic precision and will allow clinicians to guide treatment according to risk profile. This will further enhance treatment outcomes in at-risk patients and may prevent overtreatment of patients with a favorable course of the disease. We are aware of one ongoing study in Switzerland following patients with suspected CRPS 1 of the hand or the foot through their observational study over a period of 1–2 years [26] and we encourage similar research in other countries.

Second, whereas these patients may be identified early, there still is a relative lack of evidence about the utility of various modalities that patients may be treated with. We require sufficiently powered intervention studies investigating the efficacy of various treatment options in CRPS 1 management. These studies should focus on classifying subgroups of patients and risk profiles that are associated with an unfavorable or favorable treatment response.

Recently, a multidisciplinary taskforce from the Netherlands completed a systematic review of the CRPS literature, and concluded that to date, due to insufficient data, no strong treatment recommendations could be made in treatment guidelines for CRPS 1 [27]. This highlights the need for further and conjoint research in all fields of medicine, to broaden our knowledge about this syndromal condition.

Implication for practice

To date, practitioners assess the prognosis of a CRPS 1 patients based on the individual clinical experience. Although there is broad consensus in clinical practice [6, 7] that early treatment of CRPS 1 is associated with a more favorable outcome, our findings indicate that little is known about prognostic relevant factors predictive of an unfavorable course of the disease. As epidemiological studies suggest that many mild forms of CRPS 1 resolve spontaneously [8], it is clinically important and relevant to recognize prognostic factors associated with unfavorable outcome. Timely identification of negative predictors will most likely lead to early referral for current best practice treatment and therefore better treatment outcome. One promising way could be a severity score [28], that may help guide treatment intensity. However, to date, the impact of such a score on treatment outcome is unknown and requires further investigation.

In addition, as our systematic review highlights, nearly all prognostic factors identified the Delphi Survey [24] are currently not or only weakly supported by the existing evidence, and therefore further research is needed. Based on our review, patients with

clinical features including sensory disturbances and cold CRPS 1 are likely to have an unfavourable course of disease and should be treated aggressively. Physicians, treating patients with CRPS 1, are therefore strongly encouraged to include patients in a registry to further enhance our knowledge about the disease.

Conclusions

In conclusion, this systematic review highlights the need for studies investigating prognostic factors for the course of CRPS 1 that allow clinicians to define specific risk profile in patients. The current evidence is weak and consistency was only found for negative predictive features such as sensory disturbances and cold CRPS 1. These findings are in agreement with a consensus on poor prognostic factors reached by an expert panel. However, for many other proposed relevant factors, we were unable to identify evidence to support their influence on the course of the disease. We presume that this finding is due to lack of evidence regarding the etiologic and prognostic understanding of CRPS1. Further research should therefore aim to investigate the clinical value of factors believed to be of importance both for the development and course of the disease.

Acknowledgments

The authors would like thank Kleijnen Systematic Reviews Ltd., York, U.K. for their support in the literature search.

References

1. Allen, G., Galer, B.S., and Schwartz, L., Epidemiology of complex regional pain syndrome: a retrospective chart review of 134 patients. *Pain*, 1999. 80(3): p. 539-44.
2. Duman, I., Dincer, U., Taskaynatan, M.A., et al., Reflex sympathetic dystrophy: a retrospective epidemiological study of 168 patients. *Clin Rheumatol*, 2007. 26(9): p. 1433-7.
3. Schwartzman, R.J., Erwin, K.L., and Alexander, G.M., The natural history of complex regional pain syndrome. *Clin J Pain*, 2009. 25(4): p. 273-80.
4. Merskey, H. and Bogduk, N., Classification of chronic pain: description of chronic pain syndrome and definitions of pain terms. 2nd. ed1994, Seattle: IASP Press.
5. Marinus, J., Moseley, G.L., Birklein, F., et al., Clinical features and pathophysiology of complex regional pain syndrome. *Lancet Neurol*, 2011. 10(7): p. 637-48.
6. Perez, R.S., Zuurmond, W.W., Bezemer, P.D., et al., The treatment of complex regional pain syndrome type I with free radical scavengers: a randomized controlled study. *Pain*, 2003. 102(3): p. 297-307.
7. Poplawski, Z.J., Wiley, A.M., and Murray, J.F., Post-traumatic dystrophy of the extremities. *J Bone Joint Surg Am*, 1983. 65(5): p. 642-55.
8. Sandroni, P., Benrud-Larson, L.M., McClelland, R.L., et al., Complex regional pain syndrome type I: incidence and prevalence in Olmsted county, a population-based study. *Pain*, 2003. 103(1-2): p. 199-207.
9. Stroup, D.F., Berlin, J.A., Morton, S.C., et al., Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. *JAMA*, 2000. 283(15): p. 2008-12.
10. Hayden, J.A., Cote, P., and Bombardier, C., Evaluation of the quality of prognosis studies in systematic reviews. *Ann Intern Med*, 2006. 144(6): p. 427-37.
11. Bejia, I., Khalifa, A.B., Salah, Z.B., et al., Predictifs factors of the algodystrophy evolution in a rheumatology department among 60 cases. *Tunis Med*, 2005. 83(3): p. 163-167.
12. Dauty, M., Renaud, P., Deniaud, C., et al., Professional outcome of reflex sympathetic dystrophy. [French]. *Ann Readapt Med Phys*, 2001. 44(2): p. 89-94.
13. Daviet, J.C., Preux, P.M., Salle, J.Y., et al., The shoulder-hand syndrome after stroke: clinical factors of severity and value of prognostic score of Perrigot. *Ann Readapt Med Phys*, 2001. 44(6): p. 326-32.
14. Daviet, J.C., Preux, P.M., Salle, J.Y., et al., Clinical factors in the prognosis of complex regional pain syndrome type I after stroke: A prospective study. *Am J Phys Med*, 2002. 81(1): p. 34-39.
15. de Mos, M., Huygen, F.J.P.M., van der Hoeven-Borgman, M., et al., Outcome of the complex regional pain syndrome. *Clin J Pain*, 2009. 25(7): p. 590-7.
16. Eulry, F., Aczel, F., Vasseur, P., et al., Treatment and evolution of algodystrophy of the foot. Retrospective study of 199 cases. *Ann Med Interne*, 1990. 141(1): p. 20-5.
17. Goris, R.J., Reynen, J.A., and Veldman, P., The clinical symptoms in post-traumatic dystrophy. *Ned Tijdschr Geneeskd*, 1990. 134(44): p. 2138-41.
18. Lulan, J., Bismuth, J.P., Sicre, G., et al., The different types of algodystrophy after fracture of the distal radius. Predictive criteria of outcome after 1 year. *Journal of Hand Surgery*, 1997. 22 B(4): p. 441-447.
19. Tan, E.C.T.H., van de Sandt-Renkema, N., Krabbe, P.F.M., et al., Quality of life in adults with childhood-onset of Complex Regional Pain Syndrome type I. *Injury*, 2009. 40(8): p. 901-4.
20. van der Laan, L., Veldman, P.H., and Goris, R.J., Severe complications of reflex sympathetic dystrophy: infection, ulcers, chronic edema, dystonia, and myoclonus. *Arch Phys Med Rehabil*, 1998. 79(4): p. 424-9.
21. Vaneker, M., Wilder-Smith, O.H., Schrombges, P., et al., Patients initially diagnosed as 'warm' or 'cold' CRPS 1 show differences in central sensory processing some eight years after diagnosis: A quantitative sensory testing study. *Pain*, 2005. 115(1-2): p. 204-211.
22. Vaneker, M., Wilder-Smith, O.H.G., Schrombges, P., et al., Impairments as measured by ISS do not greatly change between one and eight years after CRPS 1 diagnosis. *Eur J Pain*, 2006. 10(7): p. 639-44.
23. Zyluk, A., The natural history of post-traumatic reflex sympathetic dystrophy. *J Hand Surg Br*, 1998. 23(1): p. 20-3.

24. Brunner, F., Nauer, M., and Bachmann, L.M., Poor prognostic factors in complex regional pain syndrome 1: A Delphi survey. *J Rehabil Med*, 2011. 43(9): p. 783-6.
25. Quisel, A., Gill, J.M., and Witherell, P., Complex regional pain syndrome underdiagnosed. *J Fam Pract*, 2005. 54(6): p. 524-32.
26. Brunner, F., Bachmann, L.M., Weber, U., et al., Complex regional pain syndrome 1--the Swiss cohort study. *BMC Musculoskelet Disord*, 2008. 9: p. 92.
27. Perez, R.S., Zollinger, P.E., Dijkstra, P.U., et al., Evidence based guidelines for complex regional pain syndrome type 1. *BMC Neurol*, 2010. 10: p. 20.
28. Harden, R.N., Bruehl, S., Perez, R.S., et al., Validation of proposed diagnostic criteria (the "Budapest Criteria") for Complex Regional Pain Syndrome. *Pain*, 2010.

CHAPTER 7

Rational pain management in complex regional pain syndrome 1 (CRPS 1):

A network meta-analysis

M. M. Wertli M.D., A.G.H. Kessels, R.S.G.M. Perez Ph.D., L.M. Bachmann M.D. Ph.D., F. Brunner M.D. PhD
Submitted

Abstract

Background and Objective: Guidelines for CRPS 1 advocate several substance classes to reduce pain and support physical rehabilitation, but guidance about which agent should be prioritized when designing a therapeutic regimen is not provided. Using a network meta-analytic approach, we examined the efficacy of all agent classes investigated in randomized clinical trials of CRPS1, and provide a rank-order of various substances stratified by length of illness duration.

Method and data treatment: Network-Meta-analysis. Searches in electronic, previous systematic reviews, conference abstracts, book chapters, and the reference lists of relevant articles were performed. Eligible studies were randomized trials comparing at least one analgesic agent in patients with CRPS 1 with placebo or with another analgesic, and reporting efficacy in reducing pain. Summary efficacy stratified by symptom duration and length of follow-up was computed across all substance classes. Two authors independently extracted data.

Results: In total, 16 studies were included in the analysis. Bisphosphonates appear to be the treatment of choice in early stages of CRPS 1. The effects of calcitonin overpass that of bisphosphonates and other substances as a short term medication in more chronic stages of the illness. While most medications showed some efficacy on short term follow-up, only bisphosphonates, NMDA analogues and vasodilators showed better long term pain reduction than placebo.

Conclusion: This network meta-analysis indicates that a rational therapeutic strategy of pain management in CRPS 1 should consider bisphosphonates in early CRPS 1 and a short term course of calcitonin in later stages. Other available treatments were less effective.

Introduction

In Complex Regional Pain Syndrome (CRPS) Type 1, sensory changes including pain, allodynia and hyperalgesia represent cardinal symptoms that form a considerable health burden for the patient [1]. Several agents have been suggested to reduce the severity of these symptoms, but to date there is no clear guidance about which agent should be prioritized when tailoring a personalized therapeutic regimen for individual patients. Guidelines propose several substance classes to reduce pain or pain sensitization and support physical rehabilitation. For example, the two latest guidelines remain unspecific with respect to treatment alternatives for different illness stages [2, 3]. These guidelines only provide a summary of the available evidence without putting treatment options into clinically applicable context. Therefore, in clinical practice the view that a patient should be treated early and aggressively in the hope to prevent chronic stages still prevails.

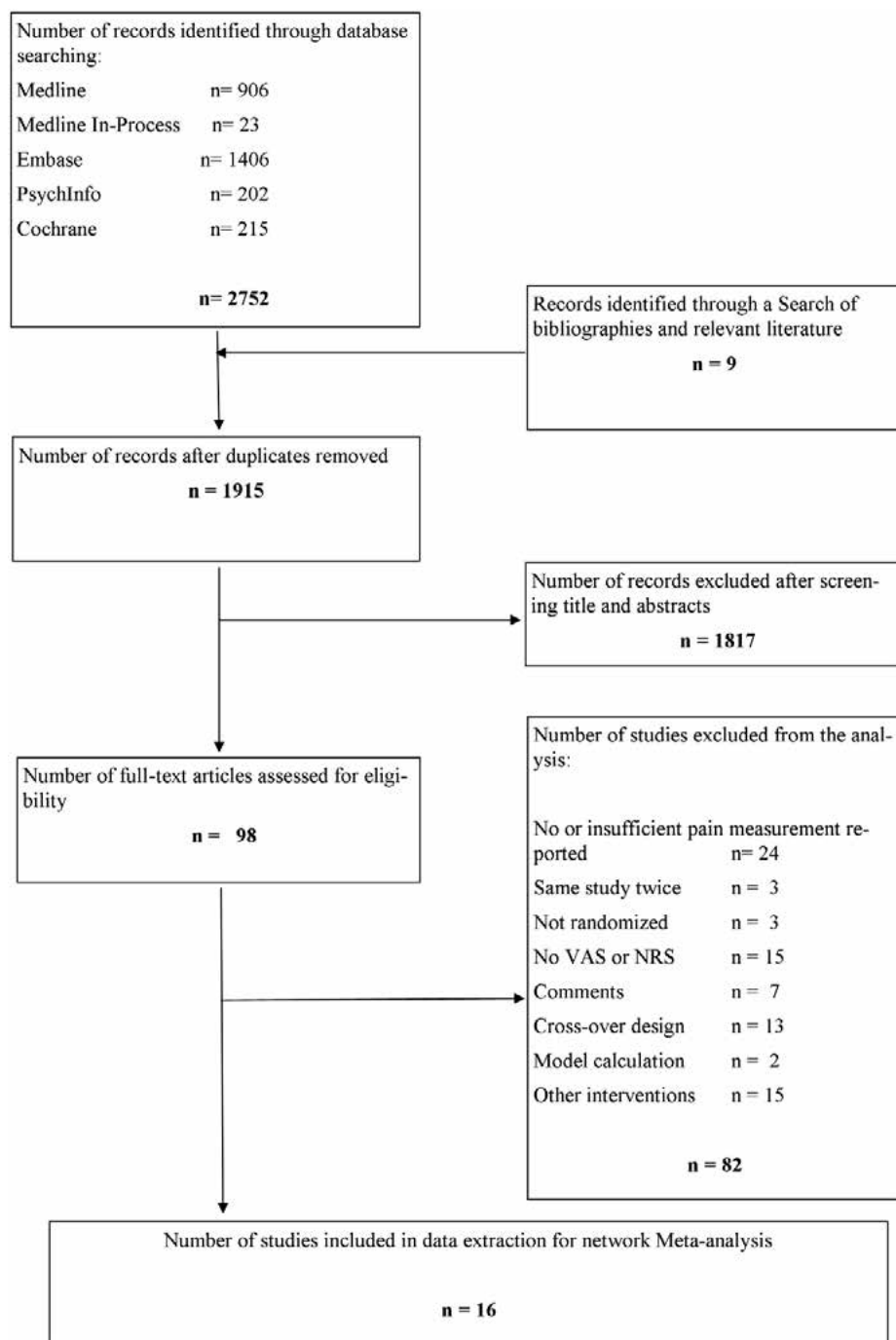
At present, treatment recommendations include conventional analgesics (paracetamol, non-steroidal anti-inflammatory drugs (NSAID) and opioids), anaesthetics, anti-convulsants, anti-depressants, free radical scavengers, oral muscle relaxants, corticosteroids, calcitonin, bisphosphonates and calcium channel blockers [4]. The magnitude of treatment options reflects the uncertainty and dilemma regarding the optimal choice. Uncertainty for both the patient and the clinician regarding the efficacy of a treatment prevails and could have a negative effect on the course of CRPS.

Recently, network meta-analysis, a new systematic review approach has become available, which allows for complete assessment across different drugs used for a specific indication [5, 6]. In particular, this method provides a rank-order that can be used for benchmarking purposes and decision-analytic modelling. Using such a network meta-analytic approach, this study examines the efficacy of all agent classes investigated within randomized clinical trials, and provides a rank-order of various substances stratified by length of illness duration.

Methods

A search method was used according to the PRISMA statement (Figure 1) for conducting meta-analyses of randomized controlled trials [7, 8].

Figure 1. Study flow



Literature search

We identified randomized controlled trials (RCTs) in patient with CRPS 1, published between 1990 and January 2013, by searching the following databases: MEDLINE (OvidSP), MEDLINE In-Process Citations (OvidSP), Embase (OvidSP), Cochrane Database of Systematic Reviews (CDSR), Cochrane Central Register of Controlled Trials (CENTRAL) and PsycINFO (OvidSP). The terms for the search strategies were identified through discussion between an information specialist and the review team, by scanning the background literature, and by browsing the MEDLINE Thesaurus (MeSH). Two detailed search strategies are described in Appendix I. To ensure the completeness of the literature search, the reviewers, experienced researchers in the field of CRPS 1, screened bibliographies of all included studies, retrieved review articles and current treatment guidelines in an additional hand search and all potentially eligible references were included in the full text review (inclusion and exclusion criteria applied).

Eligibility criteria

All RCTs were considered eligible for inclusion in this investigation that were published between January 1990 and January 2013 and met the PICO reporting system (patient, intervention, control, outcome) the following way: CRPS 1 patients, effect of pharmaceutical treatments, placebo controlled, pain and disability reduction. In order to reduce potential confounders caused by clinical, interventional and pharmacological heterogeneity, we decided to limit our analysis to medications administered orally or intravenously. No limits for the study setting or language of the publication were applied. Excluded were non-randomized studies, or conference proceedings.

Study selection

The bibliographic details of all retrieved articles were stored. Two reviewers (MW and FB) independently screened 1915 references by title and abstract. The full text was reviewed by both reviewers (MW and FB) independently in all studies meeting the pre-defined eligibility criteria (n= 98). Disagreements were discussed and resolved by consensus or by third party arbitration (LMB). Researchers with specific language proficiencies reviewed non-English language references. In the case of several publications for the same RCT without change in outcome or follow-up duration, the most recent publication was chosen and missing information from the previous publication added.

Outcome

Due to insufficient and heterogeneous reporting on other outcomes such as physical function or health-related quality of life, the analysis was limited to the outcome

“pain”. All studies reporting a valid and comparable pain measure on a visual analogue scale (VAS) or a numeric rating scale (NRS) were included in the analysis.

Data extraction and synthesis

Summary estimates per group (means, changes in means) with measures of variability (standard deviation [SD], 95% confidence interval [CI]), as available, were extracted. We summarized different formulations and routes of administration. Medications were categorized into the following groups: 1) calcitonin; 2) bisphosphonates; 3) traditional analgesics (acetaminophen, non-steroidal anti-inflammatory drugs (NSAID) and opioids); 4) radical scavengers 5) NMDA; 6) steroids; 7) NO-transmitted vasodilatation and 8) anticonvulsants.

Methodological quality

Two reviewers (MW, FB) independently assessed the methodological quality of each study as recommended [7] by using the Jadad-Score [9]. The Jadad-Score is a simple, short, reliable and valid 3-item scale (randomization procedure, blinding, drop out) developed to assess the quality of clinical reports in pain relief. Methodology was considered high when the score was 3 and more. A score of less than 3 comprises an increased risk for bias. The authors agreed in 96%, consensus was reached in a total 4% of the ratings. We did not exclude studies based on their quality rating.

Statistical analysis

Network meta-analysis allows for complete assessment across different drugs used for a specific indication [5, 6, 10-12]. The analysis was based on methodology described earlier by Kessels et al. [6]. A linear regression is used to determine the parameters describing the difference in effect between a specific intervention and the reference intervention and to check the assumptions needed to model the effect parameters. The method provides an easy and transparent way to estimate treatment effect parameters in meta-analyses involving studies with more than two arms [6].

Whenever available, we used results from the intention-to-treat analysis. If required, we imputed missing SDs of mean changes for each treatment using the largest SD reported in the set of included studies for this outcome. This procedure was necessary in 6 cases. For each participant, we simulated the outcome by sampling from a normal distribution with the mean and SD of the outcome in a specific treatment arm as described in the study report. Because of chance, the mean and SD parameters could be different from the original values. Therefore, these differences were corrected by a simple linear transformation. For all the treatment classes, a data set was generated in such a manner that it led to the same likelihood function as that from the original data.

To that new data set, a linear regression model was fitted. Drug classes, creating a unique code for each class, were entered as covariates. To preserve randomization within each trial, we included an indicator variate for each study. This variate adjusted for all differences in risk profiles and study setup among trials. From this regression model, we estimated an effect size and 95% CI's between placebo and all other treatment options.

Analyses were repeated for two a priori defined subgroups: i) studies with baseline mean disease duration of less than 12 months disease duration vs. studies with baseline mean disease duration of 12 months and more; and ii) follow-up duration of less than 2 months vs. follow-up duration of 2 months and more. All analyses were performed with Stata SE 11.2 (Copyright 1996–2010 StataCorp LP, 4905 Lakeway Drive, College Station, TX 77845 USA).

Ethics statement

For this study no ethical approval was required. No protocol was published or registered. All methods were determined a priori.

Results

Study selection

The search and inclusion process is summarized in Figure 1. Out of 1915 records, 98 were reviewed in full text. The full text assessment utilizing the inclusion and exclusion criteria resulted in the exclusion of 82 studies. The main reasons for exclusion are summarized in Figure 1 and included other non-randomized studies ($n=18$), other interventions not meeting the inclusion criteria ($n=15$) and insufficient or invalid outcome measures ($n=39$). In total 16 RCTs were included in the analysis.

Study characteristics

A detailed description of the studies, the treatment under investigation and additional treatments is summarized in Table 1. The diagnosis CRPS 1 was either based on the IASP-Orlando diagnostic criteria ($n=6$), the Bruehl/Harden criteria ($n=3$) or the Kozin classification ($n=2$), the Steinbrocker ($n=2$), the Atkins ($n=1$) or Budapest ($n=1$) criteria. One study reported no diagnostic criteria [13]. Mean disease duration was less than 12 months in eight studies [13–19]. Mean disease duration of 12 months or more was reported in seven studies [20–26]. One study did not report disease duration and was classified as mean duration less than 12 month [27]. The number of patients included in each treatment arm ranged from nine to 30 patients, the follow-up duration was be-

tween 14 and 127 days. Follow-up duration of two and more months was reported in eight studies [13, 15, 16, 18, 20, 24-26]. Study quality was moderate to good in 11 studies (Jadad score 3 and higher). Five studies had a Jadad score of less than 3 points [13, 17, 22, 25].

Randomized comparison against placebo or a control group was conducted for the groups as follows: 1) calcitonin in three studies (calcitonin 100 U i.m. [22], 200 [17] – 400 U i.n. [15]); 2) bisphosphonates in four studies (clodronate 300mg [19], pamidronate 60mg [25] and alendronate 7.5mg i.v. [14] alendronate 40mg p.o. [16]); 3) pain medications in 2 studies (paracetamol 1500mg [17], parecoxib 0.7 and 2.9 i.v. [27]) ; 4) radical scavengers in one study (mannitol 10% i.v. [24]); 5) NMDA in three studies (memantine 40mg p.o. [21], ketamin 100mg i.v. [18] and 7.2ug/kg/min. [26]), 6) corticosteroids 2 studies (methylprednisolone 5.7mg i.v. [28], prednisone 5mg p.o. [13]), 7) NO-transmitted vasodilatation in one study (tadalafil 16mg p.o. [20]); 8) anticonvulsants in one study (gabapentin 1800mg [23]).

Table 1. Baseline characteristics

Author	Treatment	N (f/m)	A (SD) years	Diagnostic criteria	Disease duration (SD) days	Additional treatment	Treatment duration (days)	Follow up (days)
Adami et al. 1997 [14]	Group 1: Placebo	10 (5/5)	60 (6.4)	Kozin	133 (133)	PT	3	14
Birkenstaff et al. 1991 [15]	Group 2: Alendronate 7.5mg iv	10 (7/3)	64 (8.0)		112 (119)	PT	3	14
	Group 1: Placebo	20 (15/5)	65.5 (8.0)	Atkins	41 (8.5)	None	28	84
	Group 2: Calcitonin 400 IU in	20 (19/1)	60.8 12.1		38 (8)	None	28	84
Frade et al. 2005 [27]	Group 1: Clonitrol	12 (6/6)	41 (8.0)	Budapest	n.r.	loco regional 1mg/kg lidocaine +30mg clonidine, Aminotryptilin 25mg and PT	7	7
Groeneweg et al. 2008 [20]	Group 2: iv. loco-regional 5mg Parecoxib 1x/week	10 (5/6)	41 (8.0)		n.r.	Same	7	7
	Group 3: 20mg Parecoxib i.v. 1/week	12 (6/6)	44 (10)		n.r.	Same	7	7
	Group 1: Placebo	12 (11/1)	36.5 (10.6)	Bruehl	1671 (1569)	Standardized graded EP (by Kempel), instruction 1x/w	84	84
Gustin et al. 2010 [21]	Group 2: Tadalafil 10mg/4w, 20mg/8w	12 (9/4)	39.8 (13.1)		1113 (735)	Same	84	84
	Group 1: Placebo	20 (12/8)	51 (12)	n.r.	492 (360)	Standardized PT: (manual, balneophysical, manual lymphatic drainage), OC therapy	49	56
	Group 2: Memantine (NMDA receptor antagonist, 40mg/day)				468 (269)	Same		56
Hamamci et al. 1996 [22]	Group 1: Placebo	16 (10/6)	59.5 (5.9)	Steinbrocker	756 (1008)	PT	28	28
	Group 2: calcitonin 100U i.m./d	25 (11/14)	57.8 (10.2)		648 (504)	Same	28	28

Author	Treatment	N (f/m)	A (SD) years	Diagnostic criteria	Disease duration (SD) days	Additional treatment	Treatment duration (days)	Follow up (days)
Lucovic et al. 2006 [13]	Group 1: Placebo	30 (21/9)	48.6 (13.8)	n.r.	Acute	Actively supported EP, PEMF+IC	127	127
Manicourt et al. 2004 [16]	Group 2: Prednison 5mg po	30 (21/9)	46.3 (12)		Acute	Same	119	119
	Group 1: Placebo	20 (12/8)	45.2 (12.5)	IASP	240 (90)	None	56	84
Perez et al. 2008 [24]	Group 2: Alendronate 40mg po/d	19 (9/10)	44.6 (12.3)		210 (60)	None	56	84
	Group 1: Placebo	19 (15/4)	43.9 (12.9)	Bruehl	525 (556)	PT, NSAID, antidepressants, anticonvulsants	5	63
Robinson et al. 2004 [25]	Group 2: Mannitol 10% iv	22 (18/4)	46.5 (11.5)		495 (556)	Same	5	63
	Group 1: Placebo	13 (4/9)	45 (6.0)	IASP	648 (414)	Paracetamol and Morphin at stable dosage if needed	1	90
Sahin et al. 2006 [17]	Group 2: Pamidronate 60mg iv. single infusion	14 (5/9)	45 (6.0)		648 (414)	Same	1	90
	Group 1: Paracetamol 1500mg/d	17 (13/4)	60 (12.3)	Steinbrocker	42 (25)	PT, EP, TENS	60	60
Schwartzmann et al. 2009 [18]	Group 2: Salmon calcitonin 200U/d nasal + calcium 500mg/d	18 (12/6)	60 (12.3)		38 (17.5)	Same	60	60
	Group 1: Placebo	10 (9/1)	45.5 (10.5)	IASP	219 (673.6)	Clonidine + Midazolam	10	84
Sigtermans et al. 2009 [26]	Group 2: Ketamin max 100mg iv	9 (9/0)	38 (7.6)		177 (369)	Same	10	84
	Group 1: Placebo	30 (26/4)	47.5 (13.1)	n.r.	2196 (1836)	Pre-existing medication: e.g. NSAID, SSRI	4.2	77
Taskaynatan et al. 2004 [28]	Group 2: Ketamin 7.2 ug/kg/min	30 (22/8)	43.7 (11.5)		3384 (2880)	Same	4.2	77
	Group 1: Placebo	10 (n.r.)	22.3 (1.6)	n.r.	93 (42)	ROM. Stretching, strengthening, bath 1x/d	21	45

Author	Treatment	N (f/m)	A (SD) years	Diagnostic criteria	Disease duration (SD) days	Additional treatment	Treatment duration (days)	Follow up (days)
van de Vusse et al. 2004 [23]	Group 2: Methylprednisolon 40mg, Lidocain 10ml 2% iv	12 (n.r.)	22.3 (1.6)		93 (42)	Same	21	45
	Group 1: Placebo	24 (21/3)	42 (13)	IASP	1290 (1080)	None	21	21
	Group 2: Gabapentin (escalating dosage 600mg to 3x/d from day 5- 21)	22 (18/4)	47 (14)		1320 (630)	None	21	21
Varenna et al. 2000 [30]	Group 1: Placebo	17 (10/7)	53.4 (9)	Kozin	126 (78)	None	10	40
	Group 2: Clodronate 300mg /d i.v.	15 (9/6)	58.1 (7.7)		111 (57)	None	10	40

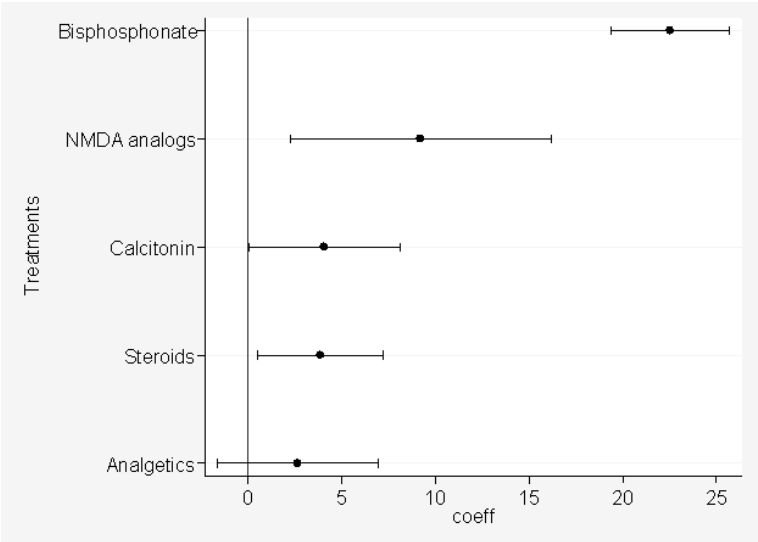
n.r.: not reported; NSAID, non-steroidal anti-inflammatory drugs; Sarpogrelate H: Sarpogrelate Hydrochloride (selective 5-HT₂ Antagonist); EP, Exercise program; PEMF Pulsed electromagnetic field; TENS: PT; physical therapy; OC, occupational therapy; ROM, range of motion

The overall treatment effect of all of the included pharmacological categories is summarized in Appendix III. Bisphosphonates and calcitonin were most effective followed by NMDA analogues, conventional analgetics, vasodilators and steroids. The radical scavengers mannitol and the anticonvulsants gabapentin were similarly effective then placebo in decreasing pain.

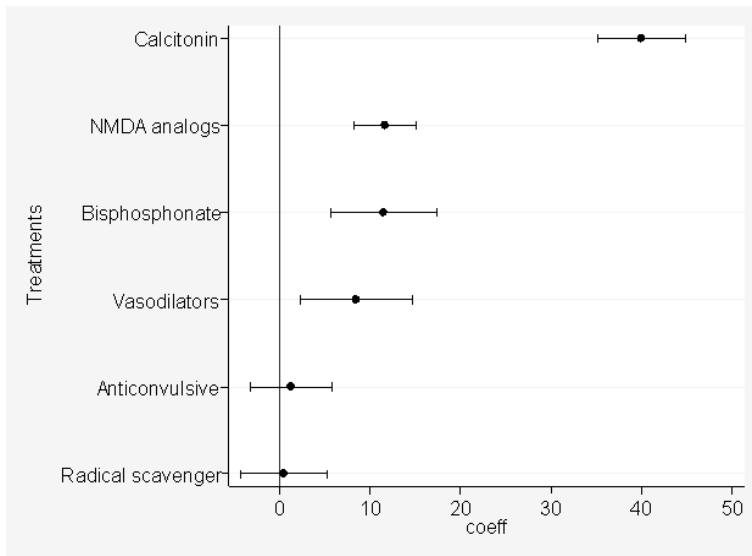
Results for symptom duration of less and more than 12 months

In patients with symptom duration of less than 12 months, bisphosphonates were most effective. Patients with symptom duration for more than 12 months responded best to calcitonin. For details please see figure 2 and figure 3.

Figure 2. Summary of Results in mean symptom duration less than 12 months



Coefficient, effect size and 95% CI's in the regression model between placebo and intervention

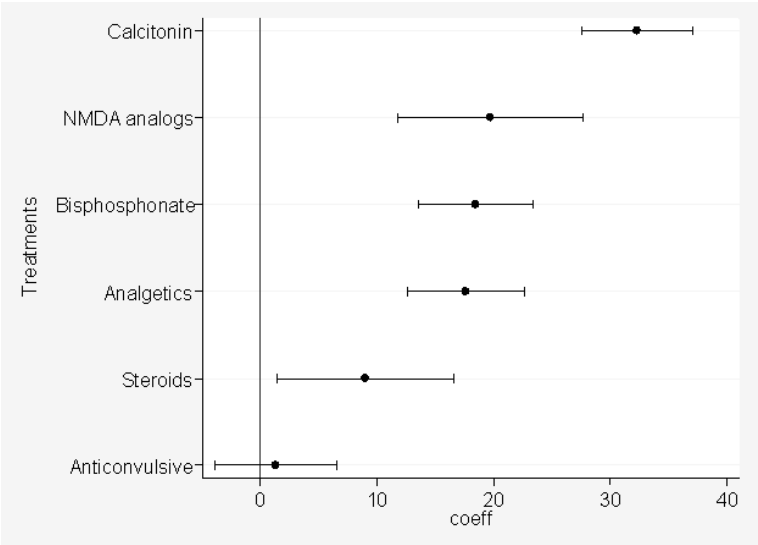
Figure 3. Summary results mean disease duration 12 months and more

Coefficient, effect size and 95% CI's in the regression model between placebo and intervention

Results for follow-up after less and more than 2 months

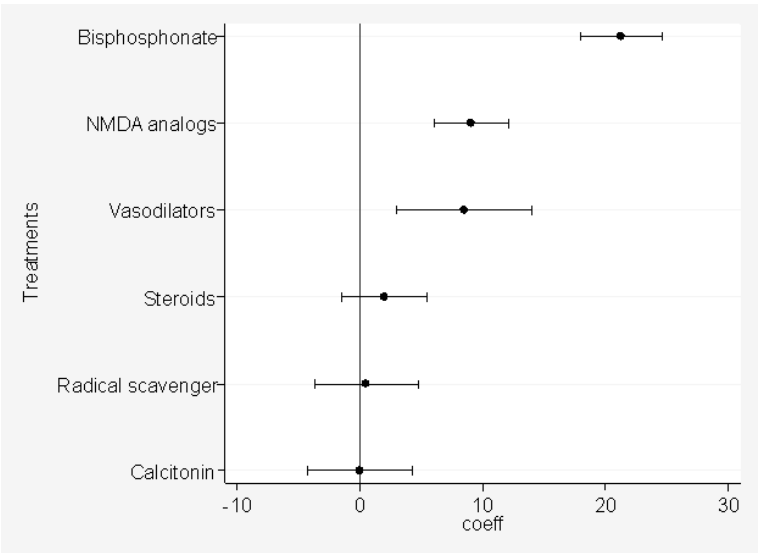
In studies with a follow-up of less than two months, calcitonin was most effective. More effective than placebo were NMDA analogs, bisphosphonates, analgesics and steroids (figure 4). The effect of steroids was more effective than placebo but less effective than bisphosphonates and analgesics. In studies with follow-up of two months and more bisphosphonates (n=14) were most effective (figure 5). NMDA analogs and the vasodilator tadalafil were also more effective than placebo. Steroids, the radical scavenger mannitol and calcitonin were not more effective than placebo.

Figure 4. Results follow-up less than 2 months



Coefficient, effect size and 95% CI's in the regression model between placebo and intervention

Figure 5. Results Follow-up 2 months and more



Coefficient, effect size and 95% CI's in the regression model between placebo and intervention

Discussion

Main findings

Whereas bisphosphonates appeared to be most beneficial in mean disease duration of CRPS of less than twelve months the effects of calcitonin overpass those of bisphosphonates and other substances in more chronic stages of the illness. Calcitonin appears to be the treatment of choice as a short term medication in advanced stages of CRPS. While most medications showed some efficacy on short term follow-up, only bisphosphonates, NMDA analogs and vasodilators showed better long term pain reduction than placebo.

Results in light with existing literature

In a recent systematic review bisphosphonates and oral tadalafil were identified to significantly reduce pain [29]. The current analysis expands our understanding of the efficacy in short and long term CRPS 1. Further, short term effects were seen in several medications including calcitonin, radical scavenger, NMDA analogs, bisphosphonates, analgesics and steroids while long term efficacy was only present in bisphosphonates and NMDA analogs. This is in agreement with the findings of a recently published RCT and therefore not yet included in this analysis that showed good efficacy for bisphosphonates in patients with CRPS 1 of less than four months duration [30]. Just recently, two updated guidelines for the diagnosis and treatment of CRPS were published [2, 3]. Although they both thoroughly review the existing literature, they do not provide a rank-order of the available substances stratified by the length of illness duration.

Strength and limitation

To our knowledge this is the first network meta-analysis providing a comprehensive evaluation of currently available pharmaceutical treatment strategies in patients with CRPS 1. In particular, the network approach allowed us to provide a rank-ordering of drug classes and allowed us to explore the temporal aspect of drug efficacy in more detail. The review was conducted according to modern methodological standards, searches were comprehensive, no language limitations were imposed, a thorough bibliographic search was conducted in order to include all relevant studies and the extraction process was done carefully.

The weaknesses of our study are twofold. First, the majority of studies included only a small group of patients and for some drug classes assessed, the number on included studies was small for most, and in a number of cases only one drug in a drug class was assessed. Second, although we aimed at investigating pharmaceutical efficacy on disability outcomes as well, presentation of results (heterogeneity on the level of

outcome measurement and reporting) impeded us from conducting efficacy analysis beyond pain outcome measures.

Implication for research

Further research should aim at clarifying the role of calcitonin and bisphosphonates in short and long term CRPS 1- in terms of short term and long term pain control and disability.

Moreover, the analgesic effects of bisphosphonates remains incompletely understood. We are aware of a few studies on animals and humans postulating potential pathophysiologic pathways of action [31]. However, further investigations should aim at exploring these further.

Quite similarly, the analgesic properties of calcitonin remain unclear. A few, often uncontrolled studies postulated a central nervous involvement which was independent from an effect on osteoclastic bone resorption [32-34].

Based on our current pathophysiologic understanding of CRPS [35], treatment regimens should distinguish between early (neurogenic inflammation, vasomotor dysfunction) and chronic (neuroplastic changes in the central nervous system) illness stages. Further research should take this aspect into account.

From a more methodological viewpoint initiatives aiming at reaching a consensus regarding relevant outcome measures and their quantification, particularly to assess function restoration should be encouraged [36].

Implication for practice

Our findings encourage using bisphosphonates as a first line treatment for pain control in early CRPS and calcitonin as a short term addition to bisphosphonates in later stages (after 12 month of symptom duration). Our findings endorse a recent warning to limit the calcitonin treatment period to six weeks in maximum, after finding a potential association with cancer incidence of various types [37].

Conclusion

Based on an efficacy network meta-analysis of 16 randomized controlled trials a rational therapeutic strategy should consider illness duration and should either contain bisphosphonates (disease duration less than 12 months, follow-up 2 months and more) or short-term calcitonin (for disease duration 12 months and more, follow-up less than 2 months). Further studies are warranted exploring the analgesic effects of bisphosphonates in more detail.

Acknowledgments

The authors would like to thank Kleijnen Systematic Reviews Ltd., York, U.K. for providing the literature searches.

Author contributions

All authors discussed the results and commented on the manuscript. F.B. A.K. and L.M.B. designed the study. M.W. and F.B. performed the literature search. M.W., F.B., A.K., R.P., and L.M.B. analysed the data and interpreted the results. M.W., F.B., R.P. and L.M.B. wrote the manuscript. F.B. oversaw the execution of the project.

References

1. Schwartzman, R.J., Erwin, K.L., and Alexander, G.M., The natural history of complex regional pain syndrome. *CLINICAL JOURNAL OF PAIN*, 2009. 25(4): p. 273-80.
2. Harden, R.N., Oaklander, A.L., Burton, A.W., et al., Complex regional pain syndrome: practical diagnostic and treatment guidelines, 4th edition. *Pain Med*, 2013. 14(2): p. 180-229.
3. Goebel, A., Barker, C.H., and Turner-Stokes, L. Complex regional pain syndrome in adults: UK guidelines for diagnosis, referral and management in primary and secondary care. 2013 May 2012 [cited 2013 August 20]; Available from: <http://www.rcplondon.ac.uk/sites/default/files/documents/complex-regional-pain-full-guideline.pdf>.
4. Perez, R.S., Zollinger, P.E., Dijkstra, P.U., et al., Evidence based guidelines for complex regional pain syndrome type 1. *BMC Neurol*, 2010. 10: p. 20.
5. Lu, G. and Ades, A.E., Combination of direct and indirect evidence in mixed treatment comparisons. *Stat Med*, 2004. 23(20): p. 3105-24.
6. Kessels, A.G.H., ter Riet, G., Puhan, M.A., et al., A simple regression model for network meta-analysis. *OA Epidemiology*, 2013. 1(1): p. 7.
7. Liberati, A., Altman, D.G., Tetzlaff, J., et al., The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. *BMJ*, 2009. 339: p. b2700.
8. Moher, D., Liberati, A., Tetzlaff, J., et al., Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Ann Intern Med*, 2009. 151(4): p. 264-9, W64.
9. Jadad, A.R., Moore, R.A., Carroll, D., et al., Assessing the quality of reports of randomized clinical trials: is blinding necessary? *Control Clin Trials*, 1996. 17(1): p. 1-12.
10. Berlin, J.A., Santanna, J., Schmid, C.H., et al., Individual patient- versus group-level data meta-regressions for the investigation of treatment effect modifiers: ecological bias rears its ugly head. *Stat Med*, 2002. 21(3): p. 371-87.
11. Hasselblad, V., Meta-analysis of multitreatment studies. *Med Decis Making*, 1998. 18(1): p. 37-43.
12. Lumley, T., Network meta-analysis for indirect treatment comparisons. *Stat Med*, 2002. 21(16): p. 2313-24.
13. Lukovic, T.Z., Ilic, K.P., Jevtic, M., et al., Corticosteroids and physical agents in treatment of complex regional pain syndrome type I. *Medicus*, 2006. 7(2): p. 70-2.
14. Adami, S., Fossaluzza, V., Gatti, D., et al., Bisphosphonate therapy of reflex sympathetic dystrophy syndrome. *ANNALS OF THE RHEUMATIC DISEASES*, 1997. 56(3): p. 201-4.
15. Bickerstaff, D.R. and Kanis, J.A., The use of nasal calcitonin in the treatment of post-traumatic algodystrophy. *BRITISH JOURNAL OF RHEUMATOLOGY*, 1991. 30(4): p. 291-4.
16. Manicourt, D.H., Brasseur, J.P., Boutsen, Y., et al., Role of alendronate in therapy for posttraumatic complex regional pain syndrome type I of the lower extremity. *ARTHRITIS AND RHEUMATISM*, 2004. 50(11): p. 3690-7.
17. Sahin, F., Yilmaz, F., Kotevoglou, N., et al., Efficacy of salmon calcitonin in complex regional pain syndrome (type 1) in addition to physical therapy. *CLINICAL RHEUMATOLOGY*, 2006. 25(2): p. 143-8.
18. Schwartzman, R.J., Alexander, G.M., Grothausen, J.R., et al., Outpatient intravenous ketamine for the treatment of complex regional pain syndrome: a double-blind placebo controlled study. *PAIN*, 2009. 147(1-3): p. 107-15.
19. Varenna, M., Zucchi, F., Ghiringhelli, D., et al., Intravenous clodronate in the treatment of reflex sympathetic dystrophy syndrome. A randomized, double blind, placebo controlled study. *The Journal of rheumatology*, 2000. 27(6): p. 1477-83.
20. Groeneweg, G., Huygen, F.J., Niehof, S.P., et al., Effect of tadalafil on blood flow, pain, and function in chronic cold complex regional pain syndrome: a randomized controlled trial. *BMC musculoskeletal disorders*, 2008. 9: p. 143.
21. Gustin, S.M., Schwarz, A., Birbaumer, N., et al., NMDA-receptor antagonist and morphine decrease CRPS-pain and cerebral pain representation. *PAIN*, 2010. 151(1): p. 69-76.

22. Hamamci, N., Dursun, E., Ural, C., et al., Calcitonin treatment in reflex sympathetic dystrophy: a preliminary study. *The British journal of clinical practice*, 1996. 50(7): p. 373-5.
23. van de Vusse, A.C., Stomp-van den Berg, S.G., Kessels, A.H., et al., Randomised controlled trial of gabapentin in Complex Regional Pain Syndrome type 1 [ISRCTN84121379]. *BMC neurology*, 2004. 4: p. 13.
24. Perez, R.S., Pragt, E., Geurts, J., et al., Treatment of patients with complex regional pain syndrome type I with mannitol: a prospective, randomized, placebo-controlled, double-blinded study. *The journal of pain : official journal of the American Pain Society*, 2008. 9(8): p. 678-86.
25. Robinson, J.N., Sandom, J., and Chapman, P.T., Efficacy of pamidronate in complex regional pain syndrome type I. *Pain medicine (Malden, Mass.)*, 2004. 5(3): p. 276-80.
26. Sigtermans, M.J., Van Hilten, J.J., Bauer, M.C., et al., Ketamine produces effective and long-term pain relief in patients with Complex Regional Pain Syndrome Type 1. *PAIN*, 2009. 145(3): p. 304-311.
27. Frade, L.-C.P., Lauretti, G.R., Lima, I.C.P.R., et al., The antinociceptive effect of local or systemic parecoxib combined with lidocaine/clonidine intravenous regional analgesia for complex regional pain syndrome type I in the arm. *Anesthesia & Analgesia*, 2005. 101(3): p. 807-11.
28. Taskaynatan, M.A., Ozgul, A., Kenan Tan, A., et al., Bier block with methylprednisolone and lidocaine in CRPS type I: A randomized, double-blinded, placebo-controlled study. *Regional anesthesia and pain medicine*, 2004. 29(5): p. 408-412.
29. Cossins, L., Okell, R.W., Cameron, H., et al., Treatment of complex regional pain syndrome in adults: a systematic review of randomized controlled trials published from June 2000 to February 2012. *European Journal of Pain*: Ejp, 2013. 17(2): p. 158-73.
30. Varena, M., Adami, S., Rossini, M., et al., Treatment of complex regional pain syndrome type I with neridronate: a randomized, double-blind, placebo-controlled study. *Rheumatology*, 2013. 52(3): p. 534-542.
31. Rodan, G.A. and Reszka, A.A., Bisphosphonate Mechanism of Action *Current Molecular Medicine*, 2002. 2(6): p. 571-577.
32. Gennari, C., Analgesic effect of calcitonin in osteoporosis. *BONE*, 2002. 30(5, Supplement 1): p. 67-70.
33. Moreau, M.F., Guillet, C., Massin, P., et al., Comparative effects of five bisphosphonates on apoptosis of macrophage cells in vitro. *BIOCHEMICAL PHARMACOLOGY*, 2007. 73(5): p. 718-723.
34. Russell, R.G.G., Bisphosphonates: The first 40 years. *BONE*, 2011. 49(1): p. 2-19.
35. Marinus, J., Moseley, G.L., Birklein, F., et al., Clinical features and pathophysiology of complex regional pain syndrome. *Lancet Neurol*, 2011. 10(7): p. 637-48.
36. Tugwell, P., Boers, M., Brooks, P., et al., OMERACT: an international initiative to improve outcome measurement in rheumatology. *Trials*, 2007. 8: p. 38.
37. The Medical Letter on Drugs and Therapeutics, In Brief: Cancer Risk with Salmon Calcitonin. *The Medical Letter Online*, 2013. 29, <http://secure.medicalletter.org/w1414a>.

Appendix I:

Search strategies for Medline and Embase

(a) Medline (1990-2013/01/week 2) (OvidSP)

The Medline search was from 1990 to 2013/01/week 2 and identified 906 references.

No.	Search	Results
1	randomized controlled trial.pt	342334
2	controlled clinical trial.pt.	85694
3	random\$.ab	572469
4	placebo.ab	136550
5	drug therapy.fs.	1588363
6	random\$.ti.	94252
7	trial.ab.	253825
8	groups.ab.	1145730
9	or/1-8	3070837
10	animals/ not (animals/ and humans/)	3720385
11	9 not 10	2608039
12	complex regional pain syndromes/ or reflex sympathetic dystrophy/	3791
13	(CRPS or complex regional pain syndrome\$ or RND or CRPS1).ti,ab.	2140
14	posttrauma\$ dystroph\$ or post trauma\$ dystroph\$ or reflex\$ neurovascular dystroph\$).ti,ab.	58
15	(reflex\$ sympathetic dystroph\$ or sudeck\$ atroph\$ or algodystroph\$ or algoneurodystroph\$).ti,ab.	2099
16	(algo dystroph\$ or algo neurodystroph\$).ti,ab.	12
17	(shoulder hand syndrom\$ or shoulder hand dystroph\$).ti,ab.	175
18	cervical sympathetic dystroph\$.ti,ab.	0
19	or/12-18	5383
20	11 and 19	1167
21	limit 20 to yr="1990 -Current"	906

Based on the following Trials filter: Lefebvre C, Manheimer E, Glanville J. Chapter 6: searching for studies. Box 6.4.c: Cochrane Highly sensitive search strategy for identifying randomized controlled trials in Medline: Sensitivity-maximizing version (2008 version); OVID format. In: Higgins JPT, Green S (editors). Cochrane Handbook for Systematic Reviews of Interventions Version 5.0.1 [updated September 2008]. The Cochrane Collaboration, 2008. Available from www.cochrane-handbook.org. (b) Embase (1990-2013/week 2) (OvidSP).

The Embase search was from 1990 to 2013/week 2 and identified 1406 references.

No.	Search	Results
1	random.tw. or clinical trial.mp. or exp treatment outcome/	1814870
2	animal/	1802202
3	animal experiment/	1554445
4	(rat or rats or mouse or mice or murine or rodent or rodents or hamster or hamsters or pig or pigs or porcine or rabbit or rabbits or animal or animals or dogs or dog or cats or cow or bovine or sheep or ovine or monkey or monkeys).mp.	5230067
5	or/2-4	5230067
6	exp human/	14036759
7	human experiment/	307990
8	or/6-7	14038153
9	5 not (5 and 8)	4134417
10	1 not 9	1753891
11	complex regional pain syndrome/ or exp complex regional pain syndrome type i/	6240
12	(CRPS1 or CRPS or complex regional pain syndrome\$ or RND).ti,ab	3173
13	(posttrauma\$ dystroph\$ or post trauma\$ dystroph\$ or reflex\$ neurovascular dystroph\$).ti,ab.	62
14	(reflex\$ sympathetic dystroph\$ or sudeck\$ atroph\$ or algodystroph\$ or algoneurodystroph\$).ti,ab.	2767
15	(algo dystroph\$ or algo neurodystroph\$).ti,ab.	12
16	(shoulder hand syndrom\$ or shoulder hand dystroph\$).ti,ab.	231
17	cervical sympathetic dystroph\$.ti,ab.	0
18	or/11-17	8153
19	10 and 18	1423
20	limit 19 to yr="1990 -Current"	1406

Trials filter: Wong SS, Wilczynski NL, Haynes RB. Developing optimal search strategies for detecting clinically sound treatment studies in EMBASE. *Journal of the Medical Library Association* 2006;94(1):41-7.

Appendix II:

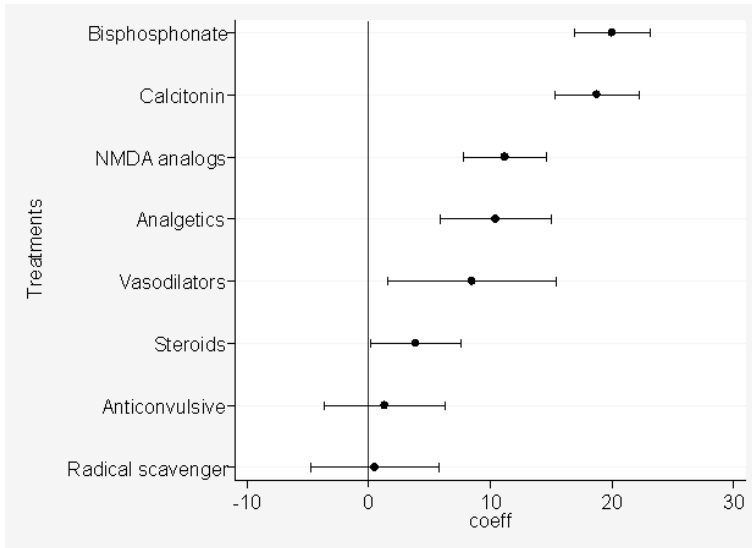
Summary of study quality according to the jadad score quality assessment [9]

Author	Year	Randomisation	Randomisation correct	double blind	blinding adequate?	drop out described?	Score
Adami	1997	1	-1	1	1	1	3
Birkenstaff	1991	1	1	1	1	1	5
Frade	2005	1	1	1	1	0	4
Groeneweg	2008	1	1	1	1	1	5
Gustin	2010	1	-1	1	1	1	3
Hamamci	1996	1	-1	0	-1	0	0
Lucovic	2006	1	-1	0	-1	0	0
Manicourt	2004	1	1	1	1	1	5
Ogawa	1998	1	-1	0	-1	1	0
Perez	2008	1	1	1	1	1	5
Robinson	2004	1	-1	1	-1	1	1
Sahin	2006	1	-1	0	-1	0	0
Schwartzmann	2009	1	1	1	1	1	5
Sigtermans	2009	1	1	1	1	1	5
Taskaynatan	2004	1	1	1	1	1	5
van de Vusse	2004	1	1	1	1	1	5
Varena	2000	1	1	1	1	1	5
Wallace	2002	1	1	1	1	1	5

Jadad Scoring system: methodological quality high ≥ 3 points; increased risk for bias in < 3 points.

Scoring: 1 point if adequately described, 0 points: not adequately described. Give 1 additional point if for question 1, the method to generate the sequence of randomization was described and it was appropriate (table of random numbers, computer generated, etc.) AND / OR if for question 2 the method of double blinding was described and it was appropriate (identical placebo, active placebo, dummy, etc.). deduct 1 point if for question 1, the method to generate the sequence of randomization was described and it was inappropriate (patients were allocated alternately, or according to date of birth, hospital number, etc.) AND / OR for question 2, the study was described as double blind but the method of blinding was inappropriate (e.g., comparison of tablet vs. injection with no double dummy)

Appendix III: Summary of treatment efficacy overall



Coefficient, effect size and 95% CI's in the regression model between placebo and intervention.

CHAPTER 8

General discussion

The research presented in this doctoral thesis examined a series of problems associated with two pain conditions—non-specific low back pain (LBP) and complex regional pain syndrome type 1 (CRPS 1)—where the diagnosis relies on patient-relevant symptoms and clinical findings. The work presented in this thesis addressed topics at the diagnostic-prognostic-therapeutic junction. It suggests future studies that are needed to investigate the influence of clinical information on the prognosis and treatment outcome in patients suffering from these pain disorders. The findings presented in this thesis may inspire future research in both LBP and CRPS 1.

Results in light of existing literature

Hayden et al. identified a set of factors associated with poor prognosis in observation studies of patients with LBP [5]. It is reasonable to believe that these factors potentially influence treatment outcome in RCTs. Research on disease prognosis is an emerging field and many prognostic factors have been identified recently [6]. The amount of indexed publications addressing prognosis in back pain and cited in PubMed increased from 5 percent in 1990 to 22 percent in 2012. We demonstrated that in studies of LBP, many important prognostic factors are incompletely assessed or reported. Even 10 years after the publication of the CONSORT statement [7] which aimed at improving the reporting in RCTs, on average only half of the prognostic factors identified by Hayden et al. are reported [8]. We therefore conclude that the influence of prognostic factors on the course of LBP and treatment efficacy is not fully understood. Future research should systematically assess and report prognostic factors. We also support current research initiatives aimed at conducting individual participant data meta-analysis based on high quality RCTs [9]. This will eventually allow researchers to understand the influence of prognostic factors and to develop individualized treatment strategies for specific subgroups.

Maladaptive cognitive behaviors are considered to play a significant role in the development of chronic LBP [10]. However, the importance of the fear avoidance model, a theoretical model used to explain how psychological factors affect the experience of pain, has been questioned [11, 12]. The two systematic analyses presented in this thesis support the importance of fear avoidance beliefs in LBP. We showed that high fear avoidance beliefs were associated with poor work-related outcome in patients with persisting LBP of less than six months duration. The review also highlights important gaps in the literature that should be considered in future research. Most importantly, the interaction of different coping strategies should be further studied. In the original Fear Avoidance Model, catastrophizing is a prerequisite for fear avoidance. However, there is evidence that fear avoidance can be present without catastrophizing [13]. In addition, other coping strategies might be equally or more important.

We analyzed how fear avoidance beliefs influenced treatment efficacy in various treatment strategies [14]. While the findings were consistent, the heterogeneity of the

outcomes measured, the various analyses conducted, and the different study protocols used impeded us from conducting meta-analyses. The results presented in this study should inspire future research to assess the influence of various prognostic factors on treatment outcome. It can be hypothesized that patients with different coping strategies respond better to different treatment interventions. Our results show that patients with LBP and fear avoidance beliefs respond well to treatments addressing those beliefs, while patients without fear avoidance beliefs have no benefit. This might explain negative study findings when a treatment strategy is used for all patients without risk stratification. Risk stratified treatment strategies have been shown to be cost effective [15]. The STarT Back tool used by Hill et al. [15] includes single item questions for fear avoidance beliefs, catastrophizing, and depression. Future studies should systematically assess coping strategies and their influence on treatment efficacy.

The diagnostic accuracy of bone scintigraphy for the diagnosis of CRPS 1 was systematically assessed. While in patients with non-specific LBP, the use of imaging studies is not recommended [1], the recommendations for those with CRPS 1 are less clear. The use of bone scintigraphy has been advocated by some as a gold standard test for the diagnosis [2]. Our analysis showed that a positive finding is not necessarily concordant with and a negative scan does not necessarily rule out CRPS 1 [3]. These findings are important for clinicians: bone scintigraphy should be only used in conjunction with clinical findings. Future research should aim at understanding when bone scintigraphy or other imaging techniques might be of importance in the diagnostic process. We agree with Hancock and colleagues that future research should further investigate the diagnostic and prognostic utility of findings in laboratory tests and imaging studies in non-specific pain disorders [4]. However, before recommending the use of any test in clinical practice, their diagnostic accuracy and clinical validity need to be established.

Whereas clinical LBP management involves early prognostic screening for delayed recovery [16, 17], these concepts are largely lacking in CRPS 1. The systematic analysis of prognostic factors in CRPS 1 identified 28 potentially relevant factors. Those most consistently associated with poor prognosis were cold skin temperature (5 studies) and the presence of sensory disturbances (3 studies) [18]. However, quality was poor in many of the studies. The measurement of prognostic factors or potential confounders and the statistical methods used were often inappropriate. Most factors only weakly correlated with 49 prognostic factors identified by an expert Delphi consensus study [19]. Future research should systematically assess the influence of the factors identified by this systematic analysis and the expert Delphi consensus study.

A recently published analysis showed that the presence of “allodynia” and “hypoesthesia” were negative predictors for treatment response to sympathetic blockade with local anesthetics [20]. This finding supports that research should systematically study the influence of prognostic factors on treatment response. Further, validated and comparable measurements and appropriate statistical methods should be used to further investigate the importance of prognostic factors for the course of CRPS.

While in patients with LBP, maladaptive cognitive behaviors are considered to play a significant role in the development of chronicity [10], in patients with CRPS their influence is not clear [19, 21]. In the systematic analysis of prognostic factors, only one study identified previous psychological problems associated with poor prognosis [22]. In a recently published cross-sectional analysis of patients with chronic CRPS, passive coping but not fear avoidance beliefs contributed to limitations in activity [23]. In the early 1990s, based on their comparative findings in patients with CRPS, LBP, or post-herpetic neuralgia, Rose and colleagues suggested that the fear avoidance model might explain psychological overlay in chronic pain conditions regardless of pathology [24]. More than 30 years later, there is little research on psychological factors in CRPS available. Future research should include psychological measures and investigate how coping strategies influence prognosis and treatment outcome in patients with CRPS.

Treatment recommendations for patients with CRPS 1 include a magnitude of treatment options from analgesics, anticonvulsants, anti-depressants, free radical scavengers, oral muscle relaxants, corticosteroids, calcitonin, and bisphosphonates to calcium channel blockers [25, 26]. The magnitude of treatment options reflects the uncertainty and dilemma regarding the optimal choice. By using a network-meta-analytic approach of 16 randomized controlled trials we intended to establish a rank order of treatment efficacy for pain management. Based on the results a rational therapeutic strategy should consider illness duration and should either contain bisphosphonates (disease duration less than 12 months, follow-up 2 months and more) or short-term calcitonin (for disease duration 12 months and more, follow-up less than 2 months). Consistent with these findings, a recently published RCT in patients with CRPS 1 of less than 3 months disease duration found a good treatment response to bisphosphonates [27]. Additional research should aim at clarifying the role of calcitonin and bisphosphonates in short- and long-term CRPS 1 in terms of short-term and long-term pain control and disability.

Strength and limitation

The work presented in this thesis was conducted according to modern methodological standards. Great care was taken to comply with current guidelines on how to conduct systematic reviews, meta-analysis and network meta-analysis.

There are some limitations to the findings presented in this thesis. First, the heterogeneity in studies, the differences in outcome assessment and non-reporting in many studies impeded us from conducting meta-analysis in most systematic reviews. Future studies need to address this by using thorough methodology, providing comprehensive description of study populations and using a core set of measures for outcomes and prognostic factors that will allow for comparison between studies. Second, while the network meta-analysis is an exciting and promising technique that allows for comparison between treatments where no head-to-head comparison is available, the findings

are as robust as the quality of the randomized controlled trials they are based on. In many studies presented in this thesis, study quality was poor and only limited data was available for clinically relevant outcomes. Therefore, the findings have to be interpreted with caution. Third, LBP and CRPS 1 are heterogeneous conditions and the findings presented in this thesis might not be generalizable to all patients with LBP and CRPS 1. Many studies included a small group of patients and for many treatment groups only one study was available.

Implication for research

We suggest the following implication on future research based on the various aspects discussed above: Large prospective cohort studies are needed that aim at predictors of outcome and effective interventions research in both, non-specific LBP and CRPS 1. Future research should focus on identification of relevant prognostic factors that influence the course of a disease. A defined prognostic profile may compensate for the current lack of diagnostic precision and will allow clinicians to guide treatment according to a specific risk profile. In addition to cohort studies, clinical registries could provide valuable information on all patients treated for LBP and CRPS 1. In comparison to clinical trials, such registries also contain information on patient populations not included in the clinical trials. This additional information would improve generalizability of findings. Further, for various treatment modalities there is still a relative lack of evidence on efficacy. We require sufficiently powered intervention studies investigating the efficacy of various treatment options in LBP and CRPS 1. These studies should focus on classifying subgroups of patients and risk profiles that are associated with an unfavorable or favorable treatment response. This will further enhance treatment outcomes in at-risk patients and may prevent overtreatment of patients with a favorable course of the disease.

Implication for practice

Our findings may have a substantial impact on health care cost in Switzerland. LBP and CRPS represent an important spectrum of pain disorders with substantial disability, loss of productivity, and health care costs [5, 28-32]. In LBP, the total economic burden to Swiss society was estimated between 1.6 and 2.3% of GDP [34]. According to the Swiss Accident Insurance (SUVA), the costs for one single CRPS case amount to the considerable sum of 152'000 Swiss Francs within a 10-year period whereas the average cost per case comes to 5'200 Swiss Francs [33]. The main drivers of these costs are long-term disability, compensation claims, and disability pension. Therefore, early effective and targeted treatment interventions are needed to prevent long-term disability and address this growing economic and health care problem.

Conclusion

With our research activities we expect to make a sustained contribution to the emerging recognition of the importance of prognostic factors in chronic pain. We hope that clinicians, researchers and patients affected with CRPS and LBP can benefit from our findings. The results presented in this thesis demonstrate that many questions need to be addressed until we know how prognostic factors influence the course of the disease and treatment efficacy. The findings for LBP might influence research strategies in CRPS 1. Future clinical trials should systematically assess and report important prognostic information and use of valid, reliable, standardized, and comparable measures. Further, we support the development and the use of core sets of factors to measure, as proposed for LBP research [35], all patients with pain conditions.

References

1. Airaksinen, O., et al., Chapter 4 European guidelines for the management of chronic nonspecific low back pain. *European Spine Journal*, 2006. 15(0): p. s192-s300.
2. O'Donoghue, J.P., et al., Three-phase bone scintigraphy. Asymmetric patterns in the upper extremities of asymptomatic normals and reflex sympathetic dystrophy patients. *Clin Nucl Med*, 1993. 18(10): p. 829-36.
3. Ringer, R., et al., Concordance of qualitative bone scintigraphy results with presence of clinical complex regional pain syndrome 1: meta-analysis of test accuracy studies. *Eur J Pain*, 2012. 16(10): p. 1347-56.
4. Hancock, M., et al., Discussion paper: what happened to the 'bio' in the bio-psycho-social model of low back pain? *European Spine Journal*, 2011. 20(12): p. 2105-2110.
5. Hayden, J.A., et al., What is the prognosis of back pain? *Best Practice & Research Clinical Rheumatology*, 2010. 24(2): p. 167-179.
6. Hayden, J.A., P. Côté, and C. Bombardier, Evaluation of the Quality of Prognosis Studies in Systematic Reviews. *Annals of Internal Medicine*, 2006. 144(6): p. 427-437.
7. Junker, C., et al., The CONSORT statement. *JAMA*, 1996. 276(23): p. 1876-7; author reply 1877.
8. Wertli, M.M., et al., Incomplete reporting of baseline characteristics in clinical trials: an analysis of randomized controlled trials and systematic reviews involving patients with chronic low back pain. *PLoS One*, 2013. 8(3): p. e58512.
9. Hayden, J.A., et al., Exercise therapy for chronic low back pain: protocol for an individual participant data meta-analysis. *Syst Rev*, 2012. 1: p. 64.
10. Borkum, J.M., Maladaptive cognitions and chronic pain: Epidemiology, neurobiology, and treatment. *Journal of Rational-Emotive & Cognitive Behavior Therapy*, 2010. 28(1): p. 4-24.
11. Pincus, T., et al., A systematic review of psychological factors as predictors of chronicity/disability in prospective cohorts of low back pain. *Spine (Phila Pa 1976)*, 2002. 27(5): p. E109-20.
12. Pincus, T., et al., Fear avoidance and prognosis in back pain: a systematic review and synthesis of current evidence. *Arthritis and Rheumatism*, 2006. 54(12): p. 3999-4010.
13. Westman, A.E., et al., Fear-Avoidance Beliefs, Catastrophizing, and Distress A Longitudinal Subgroup Analysis on Patients With Musculoskeletal Pain. *Clinical Journal of Pain*, 2011. 27(7): p. 567-577.
14. Wertli, M.M., et al., The Role of Fear Avoidance Beliefs as a Prognostic Factor for Outcome in Patients with Non-Specific Low Back Pain – A Systematic Review. *The Spine Journal*, 2013(0).
15. Hill, J., et al., Comparison of stratified primary care management for low back pain with current best practice (STarT Back): a randomised controlled trial. *Lancet*, 2011. 378(9802): p. 1560-1571.
16. Von Korf, M., et al., Comparison of Back Pain Prognostic Risk Stratification Item Sets. *The Journal of Pain*, 2014. 15(1): p. 81-89.
17. van Tulder, M., et al., Chapter 3 European guidelines for the management of acute nonspecific low back pain in primary care. *European Spine Journal*, 2006. 15(0): p. s169-s191.
18. Wertli, M., et al., Prognostic factors in complex regional pain syndrome 1: a systematic review. *J Rehabil Med*, 2013. 45(3): p. 225-31.
19. Brunner, F., M. Nauer, and L.M. Bachmann, Poor prognostic factors in complex regional pain syndrome 1: A Delphi survey. *Journal of Rehabilitation Medicine*, 2011. 43(9): p. 783-6.
20. van Eijs, F., et al., Predictors of pain relieving response to sympathetic blockade in complex regional pain syndrome type 1. *Anesthesiology*, 2012. 116(1): p. 113-121.
21. Hill, R., P. Chopra, and T. Richardi, Rethinking the Psychogenic Model of Complex Regional Pain Syndrome: Somatoform Disorders and Complex Regional Pain Syndrome. *Anesthesiology and Pain Medicine*, 2012. 2(2): p. 54-59.
22. Eulry, F., et al., [Treatment and evolution of algodystrophy of the foot. Retrospective study of 199 cases]. *Ann Med Interne (Paris)*, 1990. 141(1): p. 20-5.
23. Marinus, J., et al., The role of pain coping and kinesiophobia in patients with complex regional pain syndrome type 1 of the legs. *The Clinical journal of pain*, 2013. 29(7): p. 563-569.

24. Rose, M.J., et al., An application of the fear avoidance model to three chronic pain problems. *Behaviour Research and Therapy*, 1992. 30(4): p. 359-365.
25. Perez, R.S., et al., Evidence based guidelines for complex regional pain syndrome type 1. *BMC Neurol*, 2010. 10: p. 20.
26. Harden, R.N., et al., Complex regional pain syndrome: practical diagnostic and treatment guidelines, 4th edition. *Pain Med*, 2013. 14(2): p. 180-229.
27. Varenna, M., et al., Treatment of complex regional pain syndrome type I with neridronate: a randomized, double-blind, placebo-controlled study. *Rheumatology*, 2013. 52(3): p. 534-542.
28. Allen, G., B.S. Galer, and L. Schwartz, Epidemiology of complex regional pain syndrome: a retrospective chart review of 134 patients. *Pain*, 1999. 80(3): p. 539-44.
29. Duman, I., et al., Reflex sympathetic dystrophy: a retrospective epidemiological study of 168 patients. *Clin Rheumatol*, 2007. 26(9): p. 1433-7.
30. Schwartzman, R.J., K.L. Erwin, and G.M. Alexander, The natural history of complex regional pain syndrome. *Clin J Pain*, 2009. 25(4): p. 273-80.
31. Hayden, J.A., et al., Systematic reviews of low back pain prognosis had variable methods and results: guidance for future prognosis reviews. *J Clin Epidemiol*, 2009. 62(8): p. 781-796.e1.
32. Balagu, F., et al., Non-specific low back pain. *Lancet*, 2012. 379(9814): p. 482-491.
33. Jänig, W., R. Schaumann, and W. Vogt, CRPS Complex regional pain syndrome, SUVA, Editor 2013, SUVA: Luzern.
34. Wieser, S., et al., Cost of low back pain in Switzerland in 2005. *Eur J Health Econ*, 2011. 12(5): p. 455-67.
35. Pincus, T., et al., A review and proposal for a core set of factors for prospective cohorts in low back pain: a consensus statement. *Arthritis and Rheumatism*, 2008. 59(1): p. 14-24.

Summary

Non-specific low back pain (LBP) and complex regional pain syndrome 1 (CRPS 1) are two important examples of non-specific pain conditions. While non-specific LBP is common, CRPS 1 is less known but economically very important. Both illnesses share a high patient burden and lead to substantial healthcare expenditures for pain management.

Non-specific LBP is defined as lumbar pain without a specific morphological correlate and accounts for over 90 percent of all LBP cases. In the absence of a structural abnormality explaining the patient complaint, the diagnosis is solely based on the clinical findings. Further examinations—in particular imaging—do not give a causal explanation. CRPS 1 is defined as a pain state following injury, which exceeds in magnitude and duration the expected clinical course when other underlying diseases are excluded. Clinical manifestations include a broad spectrum of sensory, autonomic, motor, and trophic changes usually associated with significant impairment of motor function. For diagnosis, the use of the “Budapest criteria” is recommended. These criteria consider signs and symptoms of the clinical manifestation.

While in specific pain conditions a targeted therapy of the underlying somatic anomaly may be initiated, in non-specific LBP and CRPS 1, no causal derived therapies are possible and the therapy remains symptom-based. In order to avoid chronicity, it is important to identify early those patients at risk for delayed recovery. However, when no clear underlying cause can be found, the identification of patients at risk represents a major challenge in clinical practice.

The results presented in this thesis should be seen as a contribution to the problems outlined above. The research presented in this thesis comprises a series of projects that were performed at the Horten Centre, University of Zurich, Zurich, Switzerland, and at the Occupational and Industrial Orthopaedic Center (OIOC), New York University Hospital for Joint Diseases (NYU-HJD), New York, U.S.A, in collaboration with the University Hospital Balgrist, Zurich, Switzerland.

In Chapter 2 we demonstrated that in LBP studies many important prognostic factors are incompletely assessed or reported. Even 10 years after the publication of the CONSORT statement which aimed at improving the reporting in RCTs, on average only half of the prognostic factors identified by Hayden and colleagues are reported. We therefore concluded that the influence of prognostic factors on the course of LBP and

treatment efficacy is not fully understood. Future research should systematically assess and report prognostic factors.

In Chapter 3 we studied the influence of fear avoidance beliefs on the prognosis of LPB. High fear avoidance beliefs were associated with poor work-related outcome in patients with persisting LBP of less than three months duration. Further, important gaps in the literature are highlighted in this review.

In Chapter 4 we analyzed how fear avoidance beliefs influenced treatment efficacy in various treatment strategies. While the findings were consistent, the heterogeneity of outcomes measured, the various analyses conducted and the different study protocols used impeded us from conducting meta-analyses. The results presented in this study should inspire future research to assess the influence of various prognostic factors on treatment outcome.

In Chapter 5 we systematically assessed the diagnostic accuracy of bone scintigraphy for the Diagnosis of CRPS 1. Our analysis showed that a positive finding is not necessarily concordant with and a negative scan does not necessarily rule out CRPS 1.

Whereas clinical LBP management involves early prognostic screening for delayed recovery, these concepts are largely lacking in CRPS 1. The systematic analysis of prognostic factors in In Chapter 6 we identified 28 potentially relevant factors for the outcome of CRPS 1. Most consistently associated with poor prognosis were cold skin temperature (5 studies) and the presence of sensory disturbances (3 studies). However, the study quality was poor in many studies. The measurement of prognostic factors or potential confounders and the statistical methods used were often inappropriate. While in patients with LBP, maladaptive cognitive behaviors are considered to play a significant role in the development of chronicity, in CRPS their influence is not clear. In the systematic analysis of prognostic factors, only one study identified previous psychological problems associated with poor prognosis.

In patients with CRPS 1, treatment recommendations include a magnitude of options. The wide variety of treatment options reflects the uncertainty and dilemma regarding the optimal choice. In Chapter 7 we used a network-meta-analytic approach of 16 randomized controlled trials to establish a rank order of treatment efficacy for pain management. Based on the results, a rational therapeutic strategy should consider illness duration. Further bisphosphonates (disease duration less than 12 months, follow-up 2 months and more) or short-term calcitonin (for disease duration 12 months and more, follow-up less than 2 months) were most effective. However, there are some limitations that need to be addressed in future studies. For many treatments, only a few studies and small patient samples were available. Future research should further investigate the pathway of action of pharmacological agents. Further treatment regimens and administration routes should be studied. Finally, studies should use reliable and comparable outcome measures so clinicians and researchers will be able to compare study outcomes.

In conclusion, prognostic factors may play an important role in chronic non-specific pain conditions. Future research should aim at understanding the influence of prognostic factors and how they affect treatment efficacy and prognosis. With our research activities we expect to make a sustained contribution to the emerging recognition of the importance of prognostic factors in chronic pain. We hope that clinicians, researchers, and patients affected with LBP and CRPS can benefit from our findings.

Samenvatting

Aspecifieke lage rugpijn (LRP) en complex regionaal pijn syndroom 1 (CRPS 1) zijn twee belangrijke voorbeelden van non-specifieke pijnklachten. Terwijl LRP veel voorkomt, is CRPS 1 minder bekend, maar heeft een hoge economische impact. Voor beide aandoeningen geldt dat ze een hoge ziektelast voor de patiënt veroorzaken, en tot hoge gezondheidsuitgaven leiden.

Aspecifieke LRP wordt gedefinieerd als lumbale pijn zonder een specifiek morfoloogisch substraat, en omvat meer dan 90 procent van alle gevallen van lage rugpijn. Vanwege het ontbreken van een structurele afwijking die een verklaring kan bieden voor de klachten van de patiënt, wordt de diagnose volledig gebaseerd op klinische bevindingen. Aanvullende onderzoeken – in het bijzonder beeldvormende technieken – leveren geen causale verklaring voor de klachten. CRPS 1 wordt omschreven als een pijnlijke conditie na een trauma, die in ernst en duur het verwachte klinische beloop overtreft, en andere onderliggende aandoeningen zijn uitgesloten. Het klinisch beeld omvat een breed spectrum aan sensorische, autonome en trofische verschijnselen, die doorgaans gepaard gaan met substantiële motorische beperking. Voor het stellen van de diagnose worden de “Boedapest criteria” aanbevolen, welke gebaseerd zijn op door de patiënt gerapporteerde en door de arts waargenomen verschijnselen.

Terwijl bij specifieke pijncondities een gerichte causale behandeling van de onderliggende afwijking mogelijk is, is dit niet het geval bij aspecifieke LRP en CRPS 1, en is de behandeling doorgaans symptomatisch. Om het chronisch worden van deze aandoeningen te voorkomen moeten risicopatiënten vroeg worden geïdentificeerd. Echter, het ontbreken van duidelijke onderliggende oorzaak ontbreekt de identificatie van risicofactoren in de klinische praktijk.

De resultaten beschreven in dit proefschrift moeten worden gezien als een bijdrage aan de oplossing van de problemen zoals hierboven geschetst. Het onderzoek gepresenteerd in dit proefschrift omvat een serie onderzoeken uitgevoerd aan Horten Centre, University of Zurich, Zwitserland, en aan het Occupational and Industrial Orthopaedic Center (OIOC), New York University Hospital for Joint Diseases (NYU-HJD), Verenigde Staten, in samenwerking met het University Hospital Balgrist, Zurich, Zwitserland.

In Hoofdstuk 2 hebben we aangetoond dat in LRP studies belangrijke prognostische factoren onvoldoende worden onderzocht of gerapporteerd. Zelf tien jaar na de publicatie van de CONSORT statement gericht op het verbeteren van de verslaglegging van

RCTs, werd gemiddeld genomen slechts de helft van de prognostische factor zoals door Hayden voorgesteld gerapporteerd. Wij concludeerden derhalve dat de invloed van prognostische factoren voor beloop en behandeling van LRP nog onvoldoende duidelijk is. In toekomstig onderzoek moeten prognostische factoren voor deze aandoening systematisch worden onderzocht.

In Hoofdstuk 3 hebben we de invloed van angst-vermijdingsgedachten (“fear avoidance beliefs”) op de prognose van LBP onderzocht. Een hoge mate van angst-vermijdingsgedachten was geassocieerd met lage werk-gerelateerde uitkomst bij patiënten met persisterende LRP korter dan 3 maanden. Daarnaast werden belangrijke hiaten in de literatuur belicht in deze literatuurstudie.

In Hoofdstuk 4 hebben we geanalyseerd op welke wijze angst-vermijdingsgedachten de uitkomst van verschillende behandelstrategieën beïnvloedt. Hoewel de bevindingen consistent waren, was het verrichten van een meta-analyse niet mogelijk vanwege de heterogeniteit in uitkomstmaten, de verschillende wijzen van analyseren en de verschillen in studiemethodiek tussen de geëvalueerde studies. De resultaten uit dit onderzoek kunnen als inspiratie dienen voor toekomstig onderzoek naar de invloed van diverse prognostische factoren op de behandelingsuitkomst.

In Hoofdstuk 5 hebben we de diagnostische precisie van bot scintigrafie voor het diagnosticeren van CRPS 1 onderzocht. Onze resultaten laten zien dat een positieve scan niet noodzakelijk wijst op CRPS 1, en een negatieve scan deze aandoening niet altijd uitsluit.

Daar waar het bij LRP gebruikelijk is vroegtijdig te screenen op prognostische factoren voor vertraagd herstel, is dit niet het geval bij CRPS 1. In hoofdstuk 6 beschrijven we de identificatie van 28 potentieel relevante factoren voor beloop van CRPS. Koude huidtemperatuur (5 studies) en de aanwezigheid van sensorische stoornissen (3 studies) waren het meest consistent in verband gebracht met slechte prognose van CRPS 1. Echter, de methodologische kwaliteit van de meeste studie was beperkt. Het meten van prognostische factoren en het toepassen van statistische analyses was doorgaans onvoldoende. Hoewel bij LRP patiënten men het er over eens is dat maladaptieve cognities een rol spelen bij de ontwikkeling van chroniciteit, is de invloed hiervan bij CRPS onduidelijk. Bij de systematische analyse van prognostische factoren werd slechts één studie gevonden waarin voorgaande psychologische problematiek geassocieerd werd met slechte prognose.

Veel verschillende therapeutische mogelijkheden worden beschreven voor de behandeling van CRPS 1. Het grote aanbod aan behandelingsopties reflecteert de onzekerheid over de beste keuze voor de behandeling van deze klacht. In Hoofdstuk 7 hebben we met behulp van een netwerk meta-analytische methodiek een rangorde voor behandelingseffect bepaald op basis van 16 gerandomiseerde gecontroleerde studies. Voor een rationale therapeutische strategie zou de ziekteduur in ogenschouw moeten worden genomen. Bisfosfonaten (ziekteduur korter dan 12 maanden, follow-up 2 maanden of meer) of kortdurend calcitonine ziekteduur korter dan 12 maanden, fol-

low-up 2 korter dan 2 maanden) leken daarbij het meest effectief. Echter, ook hiervoor geldt dat methodologische tekortkomingen in de onderliggende artikelen werden gevonden. Voor veel behandelingen geldt dat er slechts een beperkt aantal studies werden gevonden, onderzocht bij kleine patiëntengroepen. Het werkingsmechanisme van verschillende farmaca zou verder moeten worden onderzocht, evenals de behandelingsprotocollen en toedieningswegen van medicamenten. Voorts zouden betrouwbare meetinstrumenten moeten worden gebruikt om behandelingseffecten op vaardigheidsniveau en andere gezondheid gerelateerde uitkomsten in kaart te brengen.

Concluderen kan worden gesteld dat prognostische factoren een belangrijke rol kunnen spelen in chronische specifieke condities, en toekomstig onderzoek zou erop gericht moeten zijn de rol van prognostische factoren verder te ontrafelen. Met ons onderzoek willen we een bijdrage leveren aan de erkenning van het belang van prognostisch onderzoek voor chronische pijn. We hopen dat klinici, onderzoekers en patiënten met LRP en CRPS kunnen profiteren van onze bevindingen.

Valorization addendum

Relevance

The work presented in this thesis addressed topics at the diagnostic-prognostic-therapeutic junction. It contributes important new evidence to three main areas of patient-centered care for non-specific low back pain and CRPS 1.

First: We showed that fear is a potentially modifiable psychological factor that influences both the prognosis and treatment success in patients with persistent low back pain. Fear may lead to fear of activity and avoidance, which leads to disuse and distress and results in a deleterious cycle. Stratified primary care management that included psychological factors for patients with low back pain improved the outcome and reduced the cost of care compared to standard care [1]. Our findings emphasized the importance of systematic risk stratification for patients with low back pain, and we demonstrated that fear reduced the efficacy of treatment strategies based solely on biomechanical principles.

These findings may have important social and economic impact. Because musculoskeletal pain results in distress, loss of productivity, and increased health care cost [2, 3], the Bone and Joint Decade recommended that musculoskeletal pain conditions should be a priority within the World Health Organization (WHO) non-communicable disease initiatives. This will attract the attention of policy makers and enable a collaborative, preventive, treatment and research approach [4]. Our findings also contributed to the discussion on the need for standardized disease definitions and outcome measures to facilitate comparisons of study outcomes.

Second: In this thesis, we presented the first systematic analysis of prognostic factors in patients with CRPS 1. More research is needed to fully identify early indicators of the course of CRPS 1. As shown in patients with low back pain, the identification of specific subgroups may lead to individualized treatment approaches, which may improve the outcomes. Our findings highlight the importance of systematic data collection from patient registries. In Switzerland, the Swiss Accident Insurance agency (SUVA) analyzes patients with CRPS 1 that are unable to work. However, little data are available for

patients that display a benign course of CRPS 1. To promote the identification of prognostic factors, clinicians and health care policy makers should provide a framework that facilitates the collection and analysis of data on all patients treated for CRPS 1.

Third: Systematic reviews and meta-analyses are considered the highest level of evidence in evidence-based medicine. However, it is important for clinicians, policy makers, and health insurance agencies to be aware of the limitations of meta-analyses. In meta-analyses, a summary estimate is calculated across different studies. In that process, study-specific information is potentially lost. For example, treatment A may be more effective in the early course of a disease and treatment B may be more effective during a chronic state. When disease duration is not considered in the analysis, the summary estimate in the meta-analysis shows that both treatments are equally effective. Moreover, the patient clinical profile is highly important in non-specific pain conditions, where no underlying cause can be identified. In randomized controlled trials for low back pain, we showed that, on average, only 50% of the known prognostic factors were reported; therefore, only limited information on the patient clinical profiles were known. Therefore, all stakeholders should support a collaborative effort to obtain a consensus on disease definitions and outcomes, in an effort to support the collection of high quality data for patients with chronic pain conditions.

Target groups

Clinicians

Our study findings are important for clinicians that care for patients with low back pain and CRPS 1. Our findings provided further support and guidance on the use of risk stratification in patients with low back pain, and we highlighted the potential for enhancing patient outcome and reducing treatment costs. We critically reviewed the use of bone scintigraphy for the diagnosis of CRPS 1. We provided a rank order of treatments to optimize efficacy and showed that disease duration should be considered in the treatment decision.

Health care policy makers

Risk stratification for patients with non-specific low back pain has been shown to improve patient outcomes and reduce health care costs. The findings in this study further supported those findings by addressing fear avoidance behavior in patients with persistent low back pain. We showed that this approach is important and might further improve outcome. We also found some evidence that strategies for addressing fear avoidance applied to patients without fear avoidance beliefs might be ineffective or

might even have negative effects. Therefore, a systematic risk assessment should be recognized as an important strategy that can reduce unnecessary diagnostic imaging studies and treatments in low risk patients and provide early targeted care in patients at risk for chronic low back pain.

Health care policy makers should also be aware of the limitations of systematic reviews and meta-analyses outlined above.

Stakeholders

Our findings support the recommendation that stakeholders, including medical, pharmaceutical, and surgical support industries, should obtain a consensus on definitions for conditions and outcomes [4]. In particular, for non-specific pain conditions, it is difficult or even impossible to compare results between studies without a consensus on the disease definition and outcome. Although the evidence has suggested that a risk-stratified approach for non-specific low back pain treatments improved patient outcome with lower cost, not all stakeholders are likely to welcome the focus on prevention and the evidence-based, cost-efficient management of low back pain [5]. However, we believe that standardized definitions and outcome measures and a systematic assessment of prognostic factors will eventually facilitate the identification of patient subgroups, and that some subgroups may benefit from treatments that are currently considered ineffective.

Activities

The results presented in this thesis have led to several activities in the field. Our results on the influence of fear avoidance beliefs were presented at the Public Health Conference in Switzerland, where stakeholders from various areas participated (policy makers, public health experts, clinicians, and researchers) [6]. Several presentations recommended the clinical implementation of systematic assessments of prognostic factors with a convenient screening tool, known as the STarT Back Tools [1]. This is currently being implemented or tested in several clinics. As a result of the work presented in this thesis, a standardized screening tool was developed – the patient attitude survey (PAS). The PAS includes in addition to the STarT Back questions additional questions derived from validated questionnaires for other psychological domains. The PAS was and tested by a multidisciplinary team at the New York University in New York. The PAS will be translated and cross-culturally validated for Swiss and Swedish populations. The goal of this process is to provide a rapid, generic screening tool for identifying potentially modifiable psychological risk factors that can be used in different musculoskeletal conditions.

Furthermore, a cohort study was conducted to identify early prognostic factors for the disease course in patients with CRPS 1 that had an otherwise unexplained painful swelling of the hand or foot eight weeks after trauma or surgery. The results from that study will soon be published [7]. The findings of this study will directly influence treatment of patients with CRPS 1.

Innovation

In recent years, increasing attention has been focused on data quality and reporting quality. In analyzing the current quality of reporting prognostic factors in low back pain, we found a need for improvement. We used an innovative approach for analyzing studies; we systematically extracted prognostic information and studied the influence of prognostic factors on the study outcome. With this approach, we could identify clinically relevant subgroups in an otherwise poorly-defined population of patients with non-specific low back pain.

In this thesis, we presented the first network meta-analysis of pharmaceutical therapies for treating CRPS 1. This newly-available technique allows comparisons of different drugs with no available head-to head comparison data. This approach provided a rank order assessment of treatment efficacy.

Implementation

In a collaborative, international effort, we will conduct a systematic assessment of potentially modifiable psychological factors that apply to various pain conditions. The aim is to validate a rapid, convenient tool that can be implemented by clinicians in a busy clinic and will enable early detection of maladaptive coping behavior. Several steps must be taken to achieve this goal:

- Translate and perform a cross-cultural validation of the PAS for patients with low back pain
- Validate the PAS for patients with other pain conditions
- Compare the predictive validity of the PAS across different pain conditions and across different cultural settings
- Develop an open source, web-based questionnaire and associated apps to facilitate its use and access in clinical practice
- Negotiate with stakeholders, including insurance agencies, health care policy makers, and clinicians
- Establish registries to facilitate data collection for systematic assessments of prognostic information and outcome measures for patients with back pain conditions

- Publicly advertise, present, and publish evidence to support the importance of coping strategies in pain conditions

References

1. Hill, J., et al., Comparison of stratified primary care management for low back pain with current best practice (STarT Back): a randomised controlled trial. *Lancet*, 2011. **378**(9802): p. 1560-1571.
2. Buchbinder, R., et al., Placing the global burden of low back pain in context. *Best Pract Res Clin Rheumatol*, 2013. **27**(5): p. 575-89.
3. Roehr, B., US needs new strategy to help 116 million patients in chronic pain. *BMJ*, 2011. **343**.
4. Haldeman, S., et al., Advancements in the Management of Spine Disorders. *Best Practice & Research Clinical Rheumatology*, 2012. **26**(2): p. 263-280.
5. Kmietowicz, Z., President of British Pain Society is forced from office after refusing to denounce NICE guidance on low back pain. *Vol.* 339. 2009.
6. Wertli, M.M., U. Held, and J. Steurer. Psychologische Faktoren bei lumbalen Rückenschmerzen – Hype oder wichtige prognostische Information? in *Swiss Public Health Conference 2013*. 2013. Zurich, Switzerland.
7. Brunner, F., et al., Complex regional pain syndrome 1--the Swiss cohort study. *BMC Musculoskelet Disord*, 2008. **9**: p. 92.

Dankwoord

I would like to thank Florian Brunner, a friend and mentor who supervised this thesis and contributed to all of the chapters. He organized my fellowship at the New York University (USA), which greatly influenced my work.

I also would like to thank Lucas M. Bachmann, whose most valuable advice inspired my research career and my work.

I am grateful to Johann Steurer, director of the Horten Center for Patient Oriented Research and Knowledge Transfer, for his support and for letting me be part of his research team.

I am also very grateful that Marten van Kleef agreed to act as the promoter for this thesis. It is indeed a great honor for me.

Thanks are extended to Margareta Nordin, Marco Campello, and Sherri Weiser who allowed me to be part of their research group at the Occupational and Industrial Orthopaedic Center (OIOC), New York University Hospital for Joint Diseases (NYU-HJD), U.S.A. Participating in their research projects influenced my research tremendously.

Research is a collaborative effort. I would therefore like to thank all of the remaining coauthors for their excellent work: Manuela Schöb (Chapter 2), Eva Rasmussen-Barr (Chapters 3 and 4), Ulrike Held (Chapter 4), Ralph Ringer (Chapter 5), Florian M. Buck (Chapter 5), Shira Schechter Weiner (Chapter 6), Alfons Kessels (Chapter 7), and Roberto Perez (Chapter 7).

Curriculum vitae

Maria Monika Wertli was born in Muri, Switzerland on February 6th, 1973. After obtaining her high school diploma from the Theresianum Ingenbohl, Brunnen, Switzerland in 1993, she read medicine at the University of Zurich, graduating in 2000. In 2005, she finished her MD thesis, "The predictive value of Protein Polymorphisms for weight loss and improvement of hypertension in severely obese patients." Her professional career included clinical training in emergency medicine, internal medicine, and intensive care. In 2006, she achieved the Swiss Board Certification in Internal Medicine, and in 2010 she completed a Certificate of Advanced Studies (CAS) in Health Science in health care science methodology, epidemiology, and evidence-based healthcare. Between September 2006 and June 2010, she worked as an attending physician at the Cantonal Hospital Winterthur in Winterthur, Switzerland.

Clinical questions have stimulated research activities throughout her entire career. With a broad clinical expertise, she joined the research group at the Horten Centre, University of Zurich and University Hospital Balgrist, Zurich in 2010. During a research fellowship at the New York University Hospital for Joint Diseases, Occupational and Industrial Orthopaedic Center (OIOC), New York, U.S.A., from July 2011 to July 2012, she developed her research skills in the field of systematic reviews and prognostic factors for pain conditions. The work undertaken for this thesis was carried out from 2010 to 2014.

List of publications

(last update Nov 2013)

Peer reviewed publications

- Wertli, M.M., Steurer J, Wildi L.M, Held U, Cross-cultural adaptation of the German version of the spinal stenosis measure, *European Spine Journal*, accepted for publication, 2014
- Wildi, L.M., Hensel, A., Wertli, M., et al., Relevant baseline characteristics for describing patients with knee osteoarthritis: results from a Delphi survey. *BMC Musculoskelet Disord*, 2013. 14(1): p. 369.
- Wertli, M.M., Rasmussen-Barr, E, Weiser Weiser, Bachmann, L.M., Brunner, F, The Role of Fear Avoidance Beliefs as a Prognostic Factor for Outcome in Patients with Non-Specific Low Back Pain – A Systematic Review. *The Spine Journal*, Epub. Oct. 2013
- Wertli, M.M., Ruchti, K.B., Steurer, J., et al., Diagnostic indicators of non-cardiovascular chest pain: a systematic review and meta-analysis. *BMC Med*, 2013. 11(1): p. 239
- Wertli, M.M., Burgstaller, J.M., Weiser, S., et al., The Influence of Catastrophizing on Treatment Outcome in Patients With Non-Specific Low Back Pain: A Systematic Review. *Spine (Phila Pa 1976)*, 2013.
- Andreisek, G., Imhof, M., Wertli, M., et al., A Systematic Review of Semiquantitative and Qualitative Radiologic Criteria for the Diagnosis of Lumbar Spinal Stenosis. *American Journal of Roentgenology*, 2013. 201(5): p. W735-W746
- Andreisek, G., Jenni, M., Klingler, D., et al., Access routes and reported decision criteria for lumbar epidural drug injections: a systematic literature review. *Skeletal Radiology*, 2013: p. 1-10.
- Wertli, M.M., Schöb, M., Brunner, F., et al., Incomplete Reporting of Baseline Characteristics in Clinical Trials: An Analysis of Randomized Controlled Trials and Systematic Reviews Involving Patients with Chronic Low Back Pain. *PLoS ONE*, 2013. 8(3): p. e58512.
- Wertli, M., Bachmann, L.M., Weiner, S.S., et al., Prognostic factors in complex regional pain syndrome 1: A systematic review. *J Rehabil Med*, 2013. 45(3): p. 225-31
- Ringer R, Wertli M, Bachmann LM, Brunner F: The accuracy of three-phase bone scintigraphy for the diagnosis of Complex Regional Pain Syndrome 1 – a systematic review, *European J. Pain* 2012 Apr 4.
- Potoczna N, Wertli M, Horber F: Protein Polymorphisms Do Not Predict Weight Loss and Improvement of Hypertension in Severely Obese Patients. *J. of Gastrointestinal Surgery*, 2004;8(7):862 – 868

Currently under peer review

- M. M. Wertli M.D., A.G.H. Kessels, R.S.G.M. Perez, L.M. Bachmann M.D. Ph.D., F. Brunner M.D. PhD: Rational pain management in Complex Regional Pain Syndrome 1 (CRPS 1) – a network meta-analysis
- Maria M. Wertli, M.D., Eva Rasmussen-Barr, RPT; Ph.D., Ulrike Held, Ph.D., Sherri Weiser, Ph.D., Lucas M. Bachmann, M.D.; Ph.D., Florian Brunner, M.D.; Ph.D.:The Influence of Fear Avoidance Beliefs on Treatment Outcome in Patients with Low Back Pain – a Systematic Review

Maria M. Wertli M.D., Rebekka Eugster, Ulrike Held Ph.D., Johann Steurer M.D., Reto Kofmehl B.Sc., Sherri Weiser Ph.D.: Catastrophizing - A Prognostic Factor for Outcome in Patients with Low Back Pain – A Systematic Review

Reviews not peer reviewed

Heim, M. and Wertli, M., [Severe hypertriglyceridemia]. Praxis (Bern 1994), 2013. 102(17): p. 1023-32.
Kuhn, N. and Wertli, M., [Needlestick-injuries among health care professionals: guidelines and reality]. Praxis (Bern 1994), 2011. 100(3): p. 151-8.
Franzini C, Wertli M: Vitamin D – the less well known effects. Swiss Med Forum 2009;9(13):260–264
Wertli, M., Schulthess, G., Kolyvanos, N., et al., [Amyloidosis]. Praxis (Bern 1994), 2006. 95(34): p. 1257-62; quiz 1263
Wertli, M. and Suter, P.M., [Water--the forgotten nutrient]. Praxis (Bern 1994), 2006. 95(39): p. 1489-95

Dissertations

Rebekka Eugster, 2013, Katastrophisierende Gedanken als Prädiktor für die Entwicklung chronischer lumbaler Rückenschmerzen
Katrin B. Ruchti, 2013, Nicht kardiale Thoraxschmerzen – Bedeutung und Genauigkeit durchgeführter diagnostischer Tests: ein systematischer Review der Literatur
Boris Jenni, 2013, Prognose und Effektivität bisheriger Therapieansätze bei Patienten mit nicht-kardialen Thoraxschmerzen: eine systematische Analyse der Literatur

