Fear Avoidance and Prognosis in Back Pain

A Systematic Review and Synthesis of Current Evidence

Tamar Pincus,1 Steven Vogel,2 A. Kim Burton,3 Rita Santos,1 and Andy P. Field4

Objective. Fear of pain, which is hypothesized to result in avoidance behavior, has been described as an obstacle to recovery in populations of patients with low back pain. However, the evidence to support the link between high levels of fear at early stages of pain and poor prognosis has yet to be systematically assessed. We undertook this review to explore current evidence and to propose further development of theoretical models.

Methods. We performed a systematic literature review of all prospective inception cohorts of patients with acute low back pain that measured fear of pain (often described as fear avoidance) at baseline.

Results. We reviewed 9 studies reported between 2001 and 2006. Several of these had acceptable/good methodology. Three studies, of which at least 1 had excellent methodology, showed no link between measures of fear at baseline and poor prognosis in the short term (3 months) or the long term (12 months). Three studies with acceptable methodology showed weak evidence for such a link, but the effect sizes were small. The only study with acceptable methodology to find a clear link suggested that fear of movement was linked to long-term pain.

Conclusion. Despite the prevalent focus on fear of pain at early stages of back pain, there is little evidence to link such fear states with poor prognosis. There is some evidence to suggest that fear may play a role when pain has become persistent. There is a growing consensus that distress/depression plays an important role at early stages, and clinicians should focus on these factors.

Individual psychological factors, along with related beliefs and behaviors, are accepted as having an important role in the experience of back pain. Throughout this article, we loosely use the term “fear avoidance” to mean fear of pain or movement. Fear of pain, postulated to result in self limitation of movement and activity (often labeled as fear avoidance) has become a focus for research (1) and intervention (2–5). It has been suggested that patients displaying high levels of fear avoidance can benefit from interventions targeted at the underlying unhelpful beliefs, yet the inherent mechanisms remain somewhat unclear. This model has been updated recently and reviewed by Asmundson and colleagues (6). We have considerably simplified it in our presentation (Figure 1) to emphasize the main components.

The model suggests that pain perception, in some people, can be imbued with catastrophic interpretation (which might result from beliefs about pain, emotional states, or other predisposing factors). This results in a fear-based state designed to protect the individual from the perceived catastrophic threat. In turn, pain-related anxiety develops, which is distinguished from fear by being focused on the future. Anxiety is associated with hypervigilance for evidence of harm, with high arousal, and ultimately with avoidance behavior. The model perceives catastrophic thinking as a prerequisite and elemental factor in the acquisition of avoidance behaviors. To date, however, there is no evidence to back the assumption of this causal path (7).

Other theories have challenged the role of cata-
strophizing and have suggested alternative mechanisms by which fear avoidance affects behavior, pain, and disability. An advance on the catastrophizing hypothesis has been proposed by Crombez and colleagues (8) and reviewed by Goubert and colleagues (9). Using modern learning theory as a basis for the model, Crombez and colleagues suggest that back pain becomes associated with movement, resulting in a conditioned response of fear and anxiety (Figure 2). The model is able to incorporate social, cultural, and individual factors as mediators in this relationship by stipulating that any factors that influence the strength of the movement–pain association will also influence the magnitude of the conditioned fear response. These include verbally and culturally transmitted information about the relationship (10), existing beliefs, expectations (11), and current emotional states (12).

In addition, there is evidence from learning models to suggest that the way in which people evaluate their pain will affect the level of acquired fear and, therefore, subsequent avoidance of movement. The model suggests that catastrophizing about pain inflates the aversiveness of the unconditioned stimulus (the pain) and thus increases the conditioned response to movement (fear) through its association with the unconditioned stimulus. Catastrophizing is thus included as one factor among many and not necessarily as a prerequisite influencing the acquisition of fear. Some interventions based on this model have been developed. In these studies, patients are physically exposed to the movement(s) they fear. Some interventions have been shown to be successful in reducing fear and influencing avoidance behavior, although they are based on small samples (13–16).

In both models, fear avoidance is conceptualized as a risk factor that can and should be addressed in clinical practice. The question, however, is whether fear avoidance is demonstrably a risk factor, and if so, for what?

The most appropriate method to investigate the role of risk factors is the prospective cohort study. Our previous systematic review (17) analyzed and reviewed reports from prospective cohort studies up to the end of 1999 that included measurement of psychological factors in groups of patients consulting for acute and subacute back pain. We concluded that there was robust evidence for the role of negative mood, in the form of distress or depression, in the transition to chronic states, along with limited evidence for catastrophizing and somatization. Perhaps surprisingly, however, there was no evidence for or against the role of fear avoidance. In fact, one commonly cited report did not actually measure fear avoidance, using instead a composite measure that included disability (18), while the only cohort study at that time to include fear avoidance as a baseline predictor demonstrated no significant relationship to outcome at 12 months (19). In the current review, we therefore set out to systematically assess additional evidence on psychological risk factors for disadvantageous outcomes in low back pain, with a focus on fear avoidance (defined as self report of cognition, without evidence of avoidance behavior), and to synthesize the findings with current theories and evidence from sources beyond the strict confines of the review.

**PATIENTS AND METHODS**

**Inclusion criteria.** The population reviewed included people consulting for musculoskeletal back pain. Settings included specialist clinics and primary care, secondary care, and occupational health care providers. This review focuses exclusively on factors associated with fear, including fear avoidance, kinesiophobia, pain-related anxiety, health anxiety, and general anxiety.

In line with the previous review, and concordant with recommendations for investigations into risk and prognosis (20), the review included only prospective cohort studies. We excluded retrospective studies, population-based studies, studies based on secondary analysis of data sets from clinical trials, and studies based in groups with chronic pain. However, evidence from excluded studies is addressed in the Discussion. We also excluded studies that measured and analyzed psycho-
logical items within a composite measure with other factors, such that extracting the information unique to fear factors was not possible (21,22).

We implemented several changes to the criteria of assessment from our previous systematic review (17) to reflect changes in the conceptualization of back pain. There is a growing consensus that the descriptors “acute” and “chronic” do not reflect the natural history of back pain. In their place, back pain is conceptualized as an intermittent lifetime problem for many among the general population. The complaint manifests as an untidy pattern of symptomatic periods interspersed with less-troublesome periods, although for some the symptoms (and associated disability) may become persistent (23,24). The recurrent nature of back pain stimulated us to review the definition of “new cases” for the purposes of this review. In the absence of precise data, an arbitrary but clinically reasonable criterion was chosen to define a new episode as symptoms starting after an absence of back pain reported in the previous 3 months. However, studies that defined new cases in other terms have also been included (see Table 1). As in the previous review, the focus was on identifying predictors at an early stage (defined as <3 weeks). This criterion aimed to include the entire population of persons with back pain, which is fundamental to the utility of early assessment of risk. The criterion reflects reports that measures of pain and disability decrease rapidly in the first 4 weeks but stabilize after this period (25).

While our previous review focused on long-term followup (12 months), the current review also aimed to consider a short-term outcome time point (4–6 weeks from new onset), again reflecting the variable nature of back symptoms. We considered that information on both long- and short-term outcomes would provide more comprehensive evidence for understanding the impact of psychological factors on back pain. As before, we decided not to exclude any studies from the review on the basis of methodologic weakness; instead, they would be graded accordingly. The findings from all studies, regardless of their grade, are considered in the Discussion.

Search strategy. A search strategy based on that from the original review was developed by 2 of the coauthors (TP and AKB). This strategy was used to search the following databases: PubMed/Medline, psycINFO, AMED (Allied & Complementary Medicine), CINAHL (Cumulative Index to Nursing and Allied Health Literature), Social Science Citation Index, and Science Citation Index. Despite the focus on fear avoidance, we widened the search to include all prospective cohort studies that measured psychological factors at baseline and that were reported since the end of 1999 (which is when our previous review ended). Our current search ended in January 2006. The search strategy was adapted slightly when each database was searched to ensure the greatest yield. The results of each search were downloaded into a Reference Manager database (Thomson Scientific, Americas, Philadelphia, PA), which enabled duplicates to be removed. Titles and abstracts were screened to remove obviously irrelevant articles. The search was limited to humans and combined strands containing key words associated with psychological risk (e.g., fear avoidance, fear of pain, pain anxiety, catastrophizing, psychosocial, somatization) with strands containing key words associated with the target population (e.g., musculoskeletal, back pain). Abstracts of potentially suitable titles were scrutinized by 2 coauthors (TP and SV) to identify relevant articles according to predetermined inclusion criteria.

Assessment protocol. Two reviewers (TP and SV), each blinded to the other’s assessment, coded every accepted article on the basis of predetermined criteria. (One coauthor specializes in research methods/epidemiology and psychological measurement, while the other is a clinician researcher specializing in back pain.) They then met to discuss and reach agreement on any differences in coding. A third experienced reviewer (AKB) assessed samples of the articles in a blinded manner. All 3 reviewers then reconciled remaining minor differences. An independent statistician (APF) performed the statistical conversion of reported results to effect sizes (r).

Assessment criteria. The criteria for assessment, initially derived from general evidence-based medicine guidelines (20), guidelines specific to back pain research (26), and issues specific to psychological measurement in pain (27,28), were a modification of those in the previous review (17). For consistency, the same 3 main quality-rating criteria for methodology were retained, but the additional criteria were coded to provide more detailed information. The criteria were divided into 3 sets and enabled “yes/no” coding according to the presence of each criterion in the reports.

Set 1 focused on methodology (based mainly on Sackett and colleagues [20]) and included the following:

1. Recruited subjects <3 weeks after current onset (main criterion for quality rating).
3. Dropout rate <20% (main criterion for quality rating).
4. Subjects reported absence of back pain in previous 3 months.
5. Homogeneity of sample (e.g., patients with neck pain or leg pain tested separately).
6. Adequate measurement of pain and disability at baseline.
8. Short-term and long-term followup (defined as 4–6 weeks and 6–12 months).

Set 2 focused on the quality of the measurement of fear-related factors and included the following:

1. Valid measurement (developed for populations of persons with pain or validated on them) (main criterion for quality rating).
3. Complete measurement instrument used (not single items extracted)—if a composite measure was used, the psychological component was tested separately (main criterion for quality rating).

Set 3 focused on the statistical criteria and included the following:

1. Multivariate analysis.
2. Adjustment for baseline variables.
3. More than 200 subjects in analyzed sample (based on recommendation from Tabachnick and Fidell [29]).
Table 1. Systematic review of prospective studies of fear avoidance, 2001–2006*

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;3 weeks since onset</td>
<td>No (73 acute patients but analyzed with others)</td>
<td>No (between 7 and 180 days)</td>
<td>Yes (64 acute patients, analyzed with others, but included chronicity as variable)</td>
<td>No (acute/subacute and chronic)</td>
<td>No (mean of 6 weeks duration)</td>
<td>No (subacute and chronic)</td>
<td>Yes</td>
<td>Yes</td>
<td>No (but &lt;6 weeks)</td>
</tr>
<tr>
<td>No back pain in previous 3 months</td>
<td>Not reported</td>
<td>No (53% had previous episode of LBP within 2 years)</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Yes</td>
<td>Not reported</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Stated exclusion criteria</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Homogeneous sample</td>
<td>Yes (recorded leg pain)</td>
<td>Yes</td>
<td>Not reported</td>
<td>Yes</td>
<td>No (but recorded and analyzed separately)</td>
<td>Yes (for subacute group)</td>
<td>No</td>
<td>Yes (80% at 3 months, 77% at 12 months)</td>
<td>Yes</td>
</tr>
<tr>
<td>Dropout rate &lt;20%</td>
<td>No</td>
<td>Yes</td>
<td>No (53% had previous episode of LBP within 2 years)</td>
<td>Not reported, but appears to be nil</td>
<td>Not reported</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Measured pain and disability at baseline</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Compared completed and dropout groups on baseline variables</td>
<td>Yes</td>
<td>NA</td>
<td>NA</td>
<td>Yes</td>
<td>NA</td>
<td>Yes</td>
<td>Not reported</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Short-term and long-term followup</td>
<td>Yes</td>
<td>No (1 year only)</td>
<td>No (14 days only)</td>
<td>No (6 months only)</td>
<td>No (3 months only)</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Methodologic quality†</td>
<td>*</td>
<td>*</td>
<td>**</td>
<td>*</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>***</td>
<td>**</td>
</tr>
<tr>
<td>Psychological criteria</td>
<td>Valid for population</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Compare factors</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Complete</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Psychological quality†</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
</tr>
<tr>
<td>Statistical criteria</td>
<td>Multivariate</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Sample size &gt;200</td>
<td>Yes (n = 131)</td>
<td>Yes (n = 90)</td>
<td>Yes (n = 411)</td>
<td>Yes (n = 440)</td>
<td>Yes (n = 126 subacute, n = 56 chronic)</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Adjustment for baseline</td>
<td>No</td>
<td>Yes</td>
<td>Yes (LBP and PCS-12)</td>
<td>No (although work-related back pain and sex were controlled)</td>
<td>Yes (in final model)</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Statistical quality†</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
</tbody>
</table>

* LBP = lower back pain; NA = not available; PCS-12 = Pain Catastrophizing Scale 12.
† A summary rating for each domain (methodology, psychological measurement, and statistical analysis) was constructed and is presented under the following system of asterisks: *** = good (meets all main criteria); ** = acceptable (meets >1 main criterion); * = unacceptable (meets ≤1 main criterion). The studies were scored for overall quality on the basis of the number of asterisks awarded across the 3 domains (maximum score of 9), as follows: 8–9 asterisks = high quality; 6–7 asterisks = acceptable quality; 0–5 asterisks = unacceptable quality. The studies were graded for overall quality as follows: Burton et al, 5; Koleck et al, 6; Kovacs et al, 6; Picavet et al, 6; Poiraudeau et al, 7; Sieben et al (2002), 6; Sieben et al (2005), 8; Schultz et al, 7; Werneke and Hart, 7.
A summary rating for each domain (methodology, psychological measurement, and statistical analysis) was then constructed and is presented under a system of asterisks (see Table 1), in which 3 asterisks indicate “good” (meets all main criteria), 2 asterisks indicate “acceptable” (meets >1 main criterion), and 1 asterisk indicates “unacceptable” (meets ≤1 main criterion). The studies were finally scored for overall quality on the basis of the number of asterisks awarded across the 3 domains (maximum score of 9), as follows: 8–9 asterisks = high quality; 6–7 asterisks = acceptable quality; 0–5 asterisks = unacceptable quality.

Because decisions concerning quality are somewhat subjective even with the use of explicit coding criteria, none of the studies were excluded from the analysis, the presentation of the results, or the Discussion. This approach presents a complete picture of current evidence and permits readers to assess independently the weighting they might wish to attribute to each study.

RESULTS

Of the 9 studies that we reviewed, 8 were of acceptable quality overall, and 7 contributed information about the link between fear and outcome. Only 1 study (30) achieved a score of “good” for methodology and for overall quality. Factors that commonly reduced quality scores were late recruitment, small sample size, and high dropout rates. Psychological measurement in the majority of studies used validated, reliable complete questionnaires. In terms of statistical analysis, 7 studies were acceptable according to our preset criteria. The most common problems (Table 2) were as follows: 1) failure to use multivariate statistics to control baseline variables, 2) failure to report sufficient information that would allow effect sizes to be computed (especially for nonsignificant results), and 3) failure to use analytic strategies that allow comparisons between competing predictor variables (e.g., analyzing catastrophizing and fear avoidance separately, converting continuous variables into grouping variables based on arbitrary criteria, and failing to control for other measured variables). In general, the statistical analyses made it difficult to assess the unique contribution of fear avoidance as a risk factor.

Summary of results. None of the studies that measured fear avoidance provided convincing evidence that fear-avoidance beliefs are a risk factor for poor outcomes. The highest-scoring study (30) included good measurement of fear, using not only the Tampa Scale of Kinesiophobia (TSK) but also a measure of avoidance of physical activity and the Pain Catastrophizing Scale (31). Only negative affect (depression, measured by the Beck Depression Inventory [32], minus somatic items) predicted outcome (measured on the Graded Chronic Pain Scale [33]) at 3, 6, and 12 months. Despite the final sample size being <200 (n = 158, with 22% loss to followup), thorough history-taking of current and previous back pain episodes made this the most informative of the studies.

Another high-scoring study (34) showed that fear of work activities significantly predicted pain intensity and (delayed) return to work in the univariate analysis, but not in the multivariate analysis. Sieben and colleagues (35) used a sophisticated design to test relationships across time between scores on the TSK (36), catastrophizing, and pain severity. Their results indicated that peaks on all 3 measures occurred together, but they were unable to provide evidence for a causal path between the variables. Two other studies that focused on short-term outcome failed to demonstrate that fear-avoidance beliefs affect outcome. Poiraud and colleagues (37) measured short-term outcome at 3 months in a large sample with almost no loss to followup. The Fear Avoidance Beliefs Questionnaire (FABQ) (38) score and anxiety measured by the Hospital Anxiety and Depression Scale (39) did significantly predict persistence of back pain; however, the FABQ score had almost a zero effect size (r ≈ 0.005), and the effect size for anxiety was small to medium (r = 0.22).

Kovacs and colleagues (40) studied patients with acute, subacute, and chronic pain, but included chronicity in the analysis and found no interaction. However, only very short-term outcome was measured (14 days). The FABQ score was not a significant predictor of the mental health component score of the Short Form 12 (SF-12) health survey (41), but it did predict the physical component score of the SF-12 health survey on days 1 and 15 (r = 0.221 and r = 0.210, respectively). The FABQ score also predicted disability on days 1 and 15, but the medium-size effect observed on day 1 (r = 0.25) had almost halved by day 15 (r = 0.15). In contrast, the ability of the SF-12 health survey mental component score to predict disability almost doubled from day 1 (r = 0.18) to day 15 (r = 0.29). It is also noteworthy that the ability of the FABQ score to predict disability was 2–3-fold less than that of low back pain (42). Reported scores on the TSK and on an instrument measuring catastrophizing independently predicted pain, but the analysis was not appropriate (see comments in Table 2). Finally, Burton and colleagues (43) analyzed a 4-year followup of an original prospective cohort study in which fear-avoidance beliefs were not found to predict disability at 12 months (44). Yet again, FABQ scores did not predict disability at 4 years.

To estimate the approximate strength of effect, measures of fear-avoidance beliefs as predictors of poor outcomes were pooled. Effect sizes from Table 2 were
**Table 2. Statistical assessment**

<table>
<thead>
<tr>
<th>Author, year (ref.)</th>
<th>n†</th>
<th>Outcome</th>
<th>Psychological factors</th>
<th>Medical or demographic factors‡</th>
<th>Multivariate method</th>
<th>Results and effect sizes (r)§</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Burton et al, 2004 (43)</td>
<td>131</td>
<td>RMD (at 4-year followup)</td>
<td>DRAM, MSPQ, Zung, FABQ, CSQ</td>
<td>Demographic characteristics, clinical history, previous health care usage, clinical examination findings, back function tests, pain intensity, previous RDO score</td>
<td>Yes (stepwise multiple regression)</td>
<td>The only predictors were depression (r = 0.47) and baseline pain intensity (r = 0.20)</td>
<td>Adjustment for baseline failed to meet criterion for inclusion in a stepwise analysis</td>
</tr>
<tr>
<td>Koleck et al, 2006 (59)</td>
<td>90</td>
<td>FNA and ENA based on PCA of the MOS and NHP</td>
<td>STAI, CES-D, LCS (state and trait in all cases), CSQ, PSSS (state only), VAS pain intensity; composite state variables based on PCA: distraction-praying (a composite of LCS and CSQ), helplessness-hopelessness (a composite based on STAI and CES-D), cognitive restructuring (sic) (a composite based on CSQ subscales), perceived control (a composite based on LCS and PSSS)</td>
<td>Sex, number of children, education, income, excessive weight, reduced activity, poor life comfort, previous LBP episode, history of trauma &gt;1 year, job satisfaction</td>
<td>Hierarchical multiple regression</td>
<td>Trait depression predicted ENA; in a separate analysis, helplessness-hopelessness predicted ENA; FNA was significantly predicted by sex, history of trauma &gt;1 year, reduced activity, and, in a separate analysis, distraction-praying</td>
<td>Insufficient statistical details to compute effect sizes; a very conservative estimate (i.e., taking (P = 0.05)) would give rise to (r = 0.17)</td>
</tr>
<tr>
<td>Kovacs et al, 2005 (40)</td>
<td>163–172 (based on estimate of missing data in the multivariate analysis)</td>
<td>RMD and QOL (1 day and 15 days, respectively)</td>
<td>FABQ, QOL based on mental factors (SF-12 mental health component score)</td>
<td>Sex, age, chronicity, sick leave, presence of leg pain, SLR test, LBP, QOL based on physical factors (SF-12 physical component score), fear-avoidance beliefs–chronicity interaction</td>
<td>Yes (stepwise multiple regression)</td>
<td>Disability at 1 day was predicted by LBP ((r = 0.58)), QOL (physical) ((r = 0.29)), fear-avoidance beliefs ((r = 0.25)), and QOL (mental) ((r = 0.29)); disability at 15 days was predicted by QOL (physical) ((r = 0.61)), LBP ((r = 0.45)), QOL (mental) ((r = 0.29)), fear-avoidance beliefs ((r = 0.15)), SLR ((r = 0.14)), and sex ((r = 0.12)); QOL (mental) at 1 day and 15 days was predicted by RMD ((r = 0.342) and (r = 0.442), respectively) and chronicity ((r = 0.197) and (r = 0.176), respectively); QOL (physical) at 1 day and 15 days was predicted by RMD ((r = 0.488) and (r = 0.582), respectively), fear-avoidance beliefs ((r = 0.221) and (r = 0.210), respectively), and also at day 1 by the presence of leg pain ((r = 0.167))</td>
<td>The use of stepwise analysis meant that only LBP and QOL (physical) were controlled in both models and, additionally, sex and SLR in the 15-day model; the influence of fear-avoidance beliefs on disability halved over 14 days; conversely, the effect of QOL (mental) doubled over the same period; fear-avoidance beliefs had 2–3-fold less impact than QOL (physical) and LBP in the models</td>
</tr>
<tr>
<td>Author, year (ref.)</td>
<td>n†</td>
<td>Outcome</td>
<td>Psychological factors</td>
<td>Medical or demographic factors‡</td>
<td>Multivariate method</td>
<td>Results and effect sizes (r)§</td>
<td>Comments</td>
</tr>
<tr>
<td>---------------------</td>
<td>-----</td>
<td>---------</td>
<td>----------------------</td>
<td>---------------------------------</td>
<td>--------------------</td>
<td>--------------------------</td>
<td>----------</td>
</tr>
<tr>
<td>Picavet et al, 2002 (42)</td>
<td>411</td>
<td>Current LBP, LBP limitation, severe LBP, chronic LBP, LBP and disability (all at 6-month followup)</td>
<td>PCS and kinesiophobia (modified TSK)</td>
<td>None in the first analysis, but a second analysis corrected for severity and disability at baseline</td>
<td>Yes (logistic regression)</td>
<td>Medium and high levels of pain catastrophizing and kinesiophobia independently predicted LBP relative to low levels of pain catastrophizing and kinesiophobia, in terms of odds ratios</td>
<td>Regression statistics were not reported; people with back pain at baseline were analyzed separately from those without back pain at baseline, rather than included as a predictor; pain catastrophizing and kinesiophobia were entered into separate analyses, so shared variance is not known; pain catastrophizing and kinesiophobia were both converted from continuous variables into tertiles</td>
</tr>
<tr>
<td>Poiraudeau et al, 2006 (37)</td>
<td>440</td>
<td>Persistence of back pain at 3 months (single question with yes/no response)</td>
<td>FABQ, HADS</td>
<td>Occupational activities, education level, LBP in parents, duration, work-related back pain, sports activities, medication in past week, pain intensity in last 48 hours, level of handicap, disability (QBPDS)</td>
<td>Yes (stepwise logistic regression)</td>
<td>Work-related back pain (r = 0.29), anxiety (HADS) (r = 0.22), sex (r = 0.18), and FABQ for work-related activities (r = 0.005) predicted the outcome</td>
<td>The use of a stepwise procedure meant that only work-related back pain and sex were controlled in the final model; FABQ-work (although statistically significant) was not of practical importance as shown by an effect size only slightly &gt;0</td>
</tr>
<tr>
<td>Schultz et al, 2002 (58)</td>
<td>159 subacute, 56 chronic</td>
<td>Return to work</td>
<td>Anxiety (STAI), depression (CES-D), SF-36 (social functioning, emotional functioning, mental health)</td>
<td>Group (subacute or chronic), union membership, total time in current job, total time employed by employer, percentage of current body pain, SF-36 (physical functioning, vitality, and health transition), skill discretion, pain behavior (guarding), injury intensity, perception of severity of disability, Waddell nonorganic signs, T12 extension, worst pain during examination, time to complete walk, right leg typical sciatica</td>
<td>Yes (logistic regression)</td>
<td>Anxiety (STAI), depression (CES-D), and SF-36 (social functioning, emotional functioning, mental health) were not significant predictors of return to work</td>
<td>Effect sizes cannot be computed because no statistics are reported for nonsignificant predictors</td>
</tr>
</tbody>
</table>
Effect sizes cannot be computed because no statistics are reported for nonsignificant predictors.

Although nonorganic signs (r = 0.17) and fear of work activities (r = 0.14) significantly predicted pain intensity and return to work, respectively, in the univariate analysis, they were not significant predictors in the multivariate analysis.

Pain catastrophizing and pain-related fear were not analyzed as predictors of back pain, and the study focused on subgroups of people with increasing or decreasing pain-related fear/catastrophizing: peaks in pain catastrophizing, pain-related fear, and pain itself tended to co-occur, and so it is difficult to determine which variables, if any, caused the others.

There was insufficient detail to compute effect sizes for BDI, PCS, and NEM; all TSK variables were not significant predictors of GCPS, but, more important, they yielded, at best, only small effect sizes.

Table 2.

<table>
<thead>
<tr>
<th>Author, year (ref.)</th>
<th>n†</th>
<th>Outcome</th>
<th>Psychological factors</th>
<th>Medical or demographic factors‡</th>
<th>Multivariate method</th>
<th>Results and effect sizes (r)§</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sieben et al, 2002 (35)</td>
<td>34 (2-week followup), 33 (3 months), 30 (1 year)</td>
<td>RMD</td>
<td>Dutch PCS, Dutch TSK</td>
<td>None in analysis</td>
<td>Yes/no (time series analysis, MANOVA, Kruskal-Wallis, Wilcoxon)</td>
<td>Those with rising pain-related fear showed significantly lower RMD than those with descending pain-related fear at baseline (r = 0.46), but significantly higher RMD than those with descending pain-related fear at 3 months (r = 0.25) and 12 months (r = 0.44); this trend was also present at 2 weeks (r = 0.32) but not significantly so (despite the medium effect size); increases in pain were associated with a rising level of pain-related fear (r = 0.49) and pain catastrophizing (r = 0.35) In the final model, age (r = 0.005), number of previous episodes (r = 0.004) predicted end-of-study GCPS; nonsignificant predictors were disability (r = 0.006), followup TSK (somatic) (r = 0.042), baseline TSK (somatic) (r = 0.064), followup pain intensity (r = 0.075), followup TSK (avoidance) (r = 0.107), days between baseline and end of study (r = 0.041), and baseline TSK (avoidance) (r = 0.114)</td>
<td>Pain catastrophizing and pain-related fear were not analyzed as predictors of back pain, and the study focused on subgroups of people with increasing or decreasing pain-related fear/catastrophizing: peaks in pain catastrophizing, pain-related fear, and pain itself tended to co-occur, and so it is difficult to determine which variables, if any, caused the others.</td>
</tr>
<tr>
<td>Sieben et al, 2005 (30)</td>
<td>158</td>
<td>GCPS</td>
<td>TSK, NEM, PCS, BDI</td>
<td>Age, sex, education, employment status, LBP history, previous LBP treatment, time since LBP onset, QBPDS, PARS</td>
<td>Yes (backward eliminating ordinal regression)</td>
<td></td>
<td>There was insufficient detail to compute effect sizes for BDI, PCS, and NEM; all TSK variables were not significant predictors of GCPS, but, more important, they yielded, at best, only small effect sizes.</td>
</tr>
<tr>
<td>Wernoke and Hart, 2001 (34)</td>
<td>187</td>
<td>Pain intensity, return to work, sick leave from work, activity interference at home, continued use of health care (all at 12-month followup)</td>
<td>Nonorganic physical signs, depressive symptoms, somatization, fear-avoidance beliefs (work), fear-avoidance beliefs (physical)</td>
<td>Multiple sites of pain, leg pain at intake, pain at intake, overt pain behaviors, perceived disability at discharge, pain pattern classification</td>
<td>Yes/no (logistic regression including only significant univariate variables)</td>
<td>Although nonorganic signs (r = 0.17) and fear of work activities (r = 0.14) significantly predicted pain intensity and return to work, respectively, in the univariate analysis, they were not significant predictors in the multivariate analysis.</td>
<td>Effect sizes cannot be computed because no statistics are reported for nonsignificant predictors.</td>
</tr>
</tbody>
</table>

* In studies using both univariate and multivariate analyses, only the multivariate analyses predicting disability/adjustment are summarized. RMD = Roland Morris Disability; DRAM = Distress and Risk Assessment Method; MSPQ = Modified Somatic Perceptions Questionnaire; Zung = Zung Depression Scale; FABQ = Fear Avoidance Beliefs Questionnaire; CSQ = Coping Strategies Questionnaire; RDQ = Roland Disability Questionnaire; FNA = functional nonadjustment; ENA = emotional nonadjustment; PCA = Principal Components Analysis; MOS = Medical Outcomes Study Short Form general health survey; NHP = Nottingham Health Profile; STAI = Spielberger State-Trait Anxiety Inventory; CES-D = Center for Epidemiologic Studies Depression Scale; LCS = Locus of Control Scale; PSSS = Perceived Social Support Scale; VAS = visual analog scale; LBP = lower back pain; QOL = quality of life; SF-12 = Short Form 12 health survey; SLR = straight leg raising; PCS = Pain Catastrophizing Scale; TSK = Tampa Scale of Kinesiophobia; HADS = Hospital Anxiety and Depression Scale; QBPDS = Quebec Back Pain Disability Scale; MANOVA = multivariate analysis of variance; GCPS = Graded Chronic Pain Scale; NEM = Negative Emotionality Scale; BDQ = Beck Depression Inventory; PARS = Physical Activity Rating Scale.
† Sample sizes quoted in this table reflect the sample on which statistics relating to psychological factors are based. As such, these values may differ from the total sample sizes reported for a given study (because of missing data).
‡ Only factors used in statistical analyses are listed.
§ In some cases effect sizes were based on the probability values quoted for a given effect (by first converting to z-values using the table in Field [47]). Probability values are often rounded up and so the resulting effect sizes are only approximations and reflect a conservative estimate. For comparisons between groups, the reported P value is assumed to be 2-tailed (unless it is stated that a 1-tailed test was performed); if this assumption is incorrect, the effect sizes in the table will overestimate the true effect.
used from studies that had used the FABQ and TSK (3 different studies in all, 4 effect sizes per study). Hedges and Vevea’s random effects meta-analytic method (45) was applied to these effect sizes (for technical details, see refs. 46 and 47) because random-effects methods are arguably better for real-world data and for trying to generalize beyond the studies included (46,47).

Hedges and Vevea’s estimate of between-study variance, $\hat{\tau}^2$, was 0.013 (SD 0.11). A chi-square test of homogeneity of effect sizes was highly significant ($\chi^2_{[11\text{df}]} = 26.88, P < 0.01$). These measures suggest considerable variation in effect sizes overall, which is not surprising given the number of different outcome measures and measures of pain avoidance used. The mean effect size based on Hedges and Vevea’s random-effects model was 0.173 (95% confidence interval 0.087–0.256). This overall effect is small to medium in size by Cohen’s criterion (48). In summary, the evidence from prospective cohort studies suggests that any causal link between fear avoidance and long-term measures of disadvantageous outcome is at best weak.

**DISCUSSION**

This systematic review of studies of prospective cohorts of subjects with low back pain has not provided evidence for fear avoidance as a strong risk factor for poor outcome; 6 of 9 studies failed to show a statistically significant link (or showed only a weak link) between measures of fear at baseline and a variety of both short- and long-term outcome measures. Only 3 studies succeeded at measuring truly early stages of back pain at baseline (30,35,40), and none of them showed a significant link between measures of fear at baseline and outcome.

This may have been due to lack of statistical power, since most studies had fewer than 200 subjects; unfortunately, pooling of data was not possible due to the use of different instruments. Another possible explanation is shared variance with other factors such as catastrophizing, which might obscure an independent relationship in the multivariate analysis (34,43). A common finding was the failure to measure distress/depression; the inclusion of negative affect arguably would improve explanatory power. With regard to outcome, the selection of one primary outcome measure for a given study is understandable, but testing several outcome measures would help clarify the picture and enable better pooling of data. However, at least 1 study (30) addressed the problems outlined above systematically and adequately and found evidence only for depression, but not fear, as a predictor of poor prognosis.

One possible explanation is that fear plays an important role only in later stages of pain, in which a negative cycle of cognition–emotion–behavior (as described below) is already established. There is some evidence to support this assumption. Linton and Boersma (22) found that beliefs that activity would result in injury or increased pain explained unique variance in both pain and function at 1-year followup. However, they later reported (49) that the relationship between fear of movement and function is mediated by chronicity, and that they had failed to observe a link between fear of movement and function among groups with duration of pain of <1 year.

The strict criteria for selection and evaluation in a systematic review can lead to the exclusion of articles containing important information. To redress the balance, a number of recent articles deserve mention here to broaden the discussion and help place our findings in perspective. Some other studies have shown somewhat contrary findings. Notably, Fritz and colleagues (50) reported that work-related fearful beliefs in patients undergoing a clinical trial (excluded from the current review due to secondary analysis of randomized controlled trial) predicted disability and work status at 4 weeks. The discrepancy between this and subsequent similar findings of that group and findings in our review could be explained by postulating that work-related fears are qualitatively different from other types of fear and avoidance behavior and may warrant separate investigation. Another finding contrary to our own (reported after the analysis of the present review) (51) was that baseline measures of fear (on the TSK), above and beyond pain at baseline, predicted future disability in patients attending primary care clinics. However, the
loss to followup (>30%) and incompatibility of respondents and nonresponders on baseline measures of fear avoidance might compromise interpretation. The apparent discrepancies serve to highlight the complexity of this field and the difficulties of interpretation across differing time frames, outcomes measures, and environments.

A consistent finding in people with acute low back pain is the role of distress/depression as an obstacle to recovery (17,30,52). The implication is that clinicians would do better to concentrate on eliciting and managing distress rather than fear, at least during the early stages. We address the postulated role of distress/depression below, in a proposed model for poor prognosis.

The present review focused on the cognitive component of fear avoidance, which has been hypothesized to result in reduced activity leading to poor outcome, in the dominant model described in the introduction. While we agree that many patients reduce their levels of daily activities, including work, and that this reduction in activity is an important factor in recovery and therefore warrants further investigation, the findings from the present review and from other work (53) suggest that there may be other pathways to reduced activity.

Alternative models might account for decreased levels of physical activity (disuse) as a risk factor for poor prognosis. These might or might not be a result of avoidance behavior. Such models need not include fear as a causal component; in fact, in their review of learning theory as applied to avoidance behavior, Goubert and colleagues (9) state explicitly that there is no strong relationship between fear and avoidance behavior.

We propose 2 possible models that describe disuse without fear (Figures 3 and 4). The first, which we label the social-beliefs approach (Figure 3), postulates that the macrosystem of health beliefs and health culture, combined with the microsystem of personal health beliefs, is sufficient to account for avoidance behavior without fear. These factors have been incorporated into the leading models as antecedents to the acquisition of fear. We argue that emotional processing, whether through fear or catastrophizing, is not necessary for the outcome of reduced activity. In fact, positive reinforcement from significant others for expressed beliefs and avoidance behavior would suffice to increase the likelihood of such behavior without fear (54). Indeed, there are examples of successful interventions for back pain that, while focusing on beliefs and behaviors, do not specifically target fear, avoidance, or catastrophizing (55).

Our second model (Figure 4) is based on the observation that clinical depression is often associated with general lethargy, social withdrawal, and increased passivity. It is plausible that a minority of patients with back pain have coexisting clinical depression (not necessarily as a response to pain, but as a coexisting health problem). We propose that at least in some people, this negative affect constitutes a long-term, trait-like vulnerability (56,57). There is some evidence, from a prospective cohort in which both depression and fear avoidance have been measured, to suggest that the factors are reasonably independent of each other (43). We also note that while there is agreement that negative affect, as measured by self-reported depression, is an indicator of poor outcome, little is known about which components of this concept constitute high risk.

Traditionally, depression is considered to have 3 major components: cognitions (including catastrophizing), negative mood, and somatic symptoms (fatigue and the like). The individual contribution of each of these to poor outcome has not been examined, and their unique contribution, together with their interaction with fear and anxiety, should be a focus for future research. Explicit hypotheses can be extracted from current theories of depression and from current theories of fear. For example, we anticipate that depression would result in more general reduced activity, while fear might be associated with particular movements and activities.

While social factors and depression might or might not affect disuse independently of fear, it seems sensible to attempt to classify patients accordingly. Thus, those high in maladaptive health beliefs and/or depression who also report high levels of fear might be considered complex fear avoidant. In contrast, in simple fear-avoidant groups there would not be an association either with depression or with strongly held maladaptive health beliefs.

Until avoidance behavior itself is measured in prospective cohorts, we cannot start to assess its impact on outcome. If we suppose that agents other than fear/anxiety alone can lead to such behavior, such agents warrant investigation. The implication for intervention is that at early stages of the pain experience, each of these factors could require a different approach. At later stages, perhaps it might be more useful to focus on changing the avoidance behavior itself rather than attempting to affect its antecedents.

REFERENCES
FEAR-AVOIDANCE BEHAVIOR AND PROGNOSIS IN BACK PAIN


27. Burton AK, Tillotson KM, Main CJ, Hollis S. Psychosocial pre-