SPINE Volume 30, Number 7, pp 813–825

A Systematic Review of Sociodemographic, Physical, and Psychological Predictors of Multidisciplinary Rehabilitation— or, Back School Treatment Outcome in Patients With Chronic Low Back Pain

Marije van der Hulst, MD,*,† Miriam M.R. Vollenbroek-Hutten, PhD,* and Maarten J. IJzerman, PhD*

Study Design. A systematic review.

Objective. To determine predictors of outcome of multidisciplinary rehabilitation–or back school treatment for patients with chronic low back pain.

Summary of Background Data. Numerous reviews have been performed to gain insight into which patients benefit from which treatment. However, no review has systematically focused on predictors from multiple domains (*i.e.*, sociodemographic, physical, and psychological), or on treatment outcome measured as activity limitation or participation restriction.

Methods. Studies were found by searching medical and psychological databases, and screening references. Two reviewers independently assessed the methodological quality using standard criteria. Studies were only included if they met a predefined level of internal validity. A qualitative analysis was performed.

Results. Heterogeneity among studies in patient characteristics, predictors, treatment, and outcomes limited evidence. All reviewed studies were descriptive or exploratory in nature. Consistent evidence was found for the predictive value of pain intensity (more pain→ worse outcome), several work-related parameters (*e.g.*, high satisfaction→ better outcome), and coping style (less active coping→ better outcome). Other sociodemographic and physical variables consistently lacked predictive value. No consistent evidence was found for other psychological variables.

Conclusions. It is impossible to define a generic set of predictors of outcome of multidisciplinary rehabilitation and back schools for patients with chronic low back pain because the reviewed studies were descriptive or exploratory in nature, and most predictors were only studied once. Nevertheless, for several predictors, consistent evidence was found. Large confirmatory studies are needed to test the value of these predictors.

Key words: chronic low back pain, predictor, rehabilitation, treatment, outcome. **Spine 2005;30:813–825**

Chronic low back pain (CLBP) is a complex problem, and multiple authors have emphasized the biopsychosocial influences on the development of chronicity.[1–7] The multidimensional approach of CLBP has now been widely recognized. A variety of multidisciplinary treatments have been developed that focus on restoration of functional activity. Several systematic reviews[8–13] and meta-analyses[14,15] have been published. The conclusions are not uniform, and the efficacy of multidisciplinary treatment of CLBP is not yet clearly proven.

One of the explanations for this limited evidence could be the heterogeneity of the CLBP population, which makes it unlikely that one treatment benefits all.[16] Because of this result, it is important to understand which subtypes of patients benefit from which treatment module. Unfortunately, there is insufficient knowledge about the prognosis of different subgroups of patients.[2,15] To improve this insight, several reviews have been performed that study predictive factors of treatment outcome in patients with chronic (low back) pain.[2,3,17–25]

Regarding the *non-systematic* reviews, a great variability is found in study population, type of treatment, outcome measures, or duration of follow-up.[2,3,19–22,24,25] First, the patient characteristics differ. Some describe the heterogeneous pain population and do not focus specifically on low back pain (LBP).[20,22–25] Others do not confine themselves to either acute or chronic LBP[3,20,24] or specific or nonspecific CLBP.[18,20,24] Second, most studies investigate a variety of outcome measures (*e.g.*, pain reduction, return to work). Third, several studies include different and often poorly defined treatments (*e.g.*, conservative, multimodal, surgical).[2,3,17,19,20,23] Fourth, studies differ in duration of follow-up. Finally, the studies include different potential predictors in the analyses, thus making comparison difficult.[19] Based on this result, it is difficult to draw a final conclusion about prognostic factors of treatment outcome, and systematic reviews are necessary.

Moreover, 3 of the published reviews are *systematic* reviews[17,18,23] and study predictive factors of multidisciplinary treatment outcome of patients with CLBP. Only one of these reviews[18] addressed the concept of multidimensionality by including prognostic factors from different domains (*i.e.*, sociodemographic, physical, and psychological). However, this review did not focus on outcome measures as disability or handicap but only on

From the *Roessingh Research and Development, and †Rehabilitation Center "Het Roessingh," Roessinghsbleekweg, Enschede, The Netherlands.

Acknowledgment date: October 1, 2003. First revision date: May 18, 2004. Acceptance date: May 19, 2004.

Financial support provided by Joris Stichting.

The manuscript submitted does not contain information about medical device(s)/drug(s).

Foundation funds were received in support of this work. No benefits in any form have been or will be received from a commercial party related directly or indirectly to the subject of this manuscript.

Address correspondence and reprint requests to Marije van der Hulst, MD, Roessingh Research and Development, Roessinghsbleekweg 33b, 7522 AH Enschede, The Netherlands; E-mail: m.vanderhulst@rrd.nl

return-to-work rate. Therefore, there is a need for a systematic review focusing on prognostic factors from multiple domains, and the outcome measures disability and handicap. It is expected that gained insight from this review will facilitate patient classification into more homogeneous subgroups, which are likely to benefit from rehabilitation treatment.[16,21,26–32]

The objective of this systematic review is to determine which factors (*i.e.*, sociodemographic, physical, and psychological) predict outcome of rehabilitation treatment (*i.e.*, multidisciplinary treatment or back schools) of patients with nonspecific CLBP. Outcome is defined as activity limitation (*i.e.*, difficulties an individual may have in executing activities) and participation restriction (*i.e.*, problems an individual may have in life situations).[33]

■ Methods

The Review Process. In the first stage of the review process, 2 reviewers (M.v.d.H. and M.V.-H.) selected the studies to be included in the systematic review.[34,35] In the second stage, both reviewers independently assessed the methodological quality of the studies and excluded studies that were not internally valid from the final review. Disagreements concerning inclusion and quality assessment of studies were resolved by consensus, and a third independent reviewer (M.IJ.) could be consulted to make the final decision. From a practical point of view, articles were not blinded for authors, institution, journal, results, or conclusions.

Search Strategy

Appropriate Studies Were Traced by:

- A computer-aided search of the Medline, Psychinfo, Picarta, Web of Science, The Cochrane Library databases up to August 2003, and the Embase and Cinahl up to September 2003.
- Screening references given in relevant, identified publications (reviews, included articles).
- Manual search of relevant journals: *Spine* (to August 2003) and *Pain* (to August 2003), American Pain Society bulletin to August 2003 (www.ampainsoc.org).
- Recommended literature by experts in the field.

The most relevant used key words were: LBP, chronic, predictor, prognosis, treatment, therapy, rehabilitation, multidisciplinary, functional restoration, outcome, and effect. Articles published in English, German, French, or Dutch were included.

Inclusion and Exclusion Criteria

Types of Studies. (Non) randomized controlled trials (RCT) and prospective cohort studies were included. RCT were included if data concerning prognostic factors for treatment outcome could be extracted from the study cohort.

Types of Participants. Subjects between 18 and 65 years of age, with as primary complaint chronic nonspecific LBP (more than 12 weeks continual or recurrent episodes of LBP).[15,36] LBP is defined as pain under the scapulas, above the cleft of the buttocks, with or without radiation to the lower extremities.[36,37] Excluded were subjects with specific causes of LBP (*e.g.*, inflammatory disease, radicular syndrome), back surgery in the last 6 months, or a medical contraindication for active rehabilitation.

Types of Interventions. Multidisciplinary treatments and back schools were included. Multidisciplinary treatment was defined as physician consultation in addition to psychological, social, or vocational intervention, or a combination of these interventions.[10] Back schools at least consisted of an education and skills program, and included an exercise regimen. Instructions were given in groups, supervised by a physiotherapist or other (para)medical therapist.[11] Excluded were all other treatments or if nerve blocks were an additional component of the intervention.

Types of Baseline Measures. Only *baseline* measures of predictive factors were included because the time of assessment of the potential predictor (*i.e.*, at baseline or during therapy) may influence the prognostic value for treatment outcome.[19]

Types of Outcome Measures. Studies were included if at least one of the outcome measures was a measurement of activity limitation (*i.e.*, difficulties an individual may have in executing activities) and/or participation restriction (*i.e.*, problems an individual may experience in daily life situations).[33]

Criteria for Methodological Quality. There are no widely accepted quality criteria for assessing the methodological quality of prognostic studies.[38,39] Therefore, we used criteria as proposed by the Cochrane Collaboration for observational studies,[40] Altman,[38] and van der Windt *et al*[41] completed with criteria used in other systematic reviews of prognostic factors (E. Beeks and J. van Limbeek, unpublished data, December 1999)[1,42–44] (Appendix 1). Each criterion was graded as: yes, no, partially, not applicable, or can't tell (*i.e.*, insufficient information provided). Internal validity was assessed by a subset of the quality criteria, adapted from Côté *et al*[42] (Appendix 2). If any of these criteria was scored as "no," the study was rejected from the analysis.

Data Extraction. Prognostic determinants were classified into 3 main domains: sociodemographic, physical, and psychological variables. Outcomes were classified as activity limitation or participation restriction. Studies were classified according to the phase of investigation (Phase I–III).[42,45] Phase I studies are descriptive, exploratory studies that seek an association between a prognostic marker and a certain outcome variable. Phase II studies are exploratory studies that value a set of prognostic variables to discriminate between high and low risk patients or to indicate which patients are likely to benefit from therapy.[42,45] Phase III studies are confirmatory studies that attempt to confirm *a priori* stated hypotheses of the value of a set of prognostic markers in predicting outcome.[42,45] The study population will be classified into patients recruited from a population of employees and patients seeking treatment at a rehabilitation center because these might differ with respect to prognosis.[42]

Data Analysis. If possible, statistical pooling will be performed. Otherwise, the results of the internal valid studies will be described qualitatively, with the overall conclusion of best evidence defined as "two or more studies reporting consistent results on the finding, or 75% of the studies reporting similar conclusions."[42] Results are statistically significant if $P \leq 0.05$.

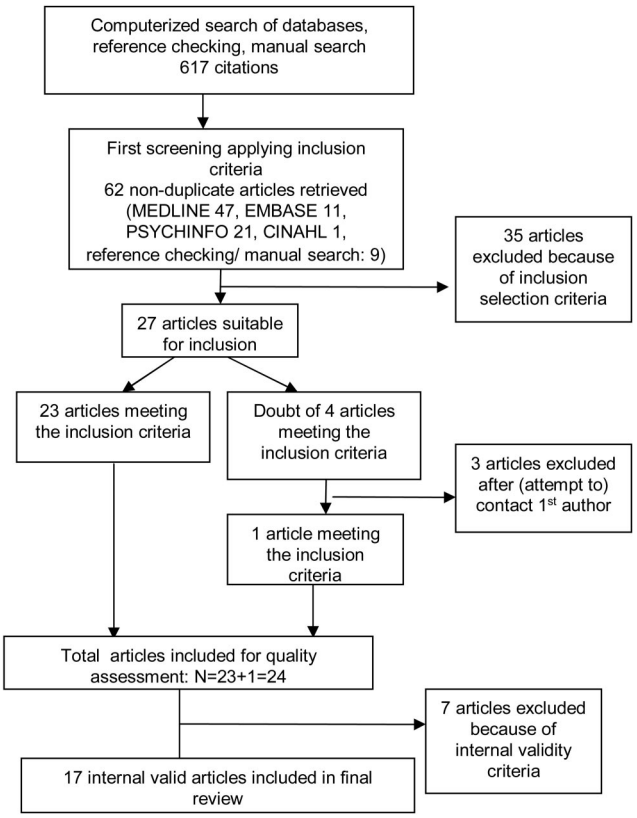

Figure 1. Flow diagram of articles accepted and rejected during the selection process.

■ Results

Selection of Studies

The flow chart of the total search and study selection is shown in Figure 1. In the first stage of the review process, the inclusion criteria were applied. Four articles[46–49] did not provide sufficient information about the studied population to apply the inclusion criteria. An attempt was made to contact the first authors for clarification, which was successful in 3 cases.[46–48] Of these 4 articles, 3 were excluded. One was excluded because "chronicity" was not defined, and the author could not be contacted.[49] The other 2 were excluded because they studied a mixed population of patients with other primary locations of pain than the back and did not perform a separate analysis for patients with LBP.[46,47] The fourth article was included because the author confirmed that "chronicity" was defined as pain duration longer than 3 months.[48] The first stage yielded 24 articles.

The second stage consisted of applying the methodological quality criteria to these 24 studies. Of these studies, 7 did not meet the quality cutoff point for internal validity,[50–56] leaving 17 internal valid studies (24 studies–7) for inclusion in the final review. Of the 7 studies, 3 were excluded because the inclusion and exclusion criteria were, although at first sight appropriate, not well defined.[50,53,55] Two studies were excluded because the participation rate and percentage follow-up were insufficient.[51,52] Invalid or unreliable outcome measurement was the reason for excluding the other 2 studies.[54,56] The quality assessment of all studies is available from the author. Final consensus was reached without needing to consult the third reviewer.

Study Characteristics

The main characteristics of the study populations of the 17 included internal valid studies are shown in Table 1.

Types of Studies. Of the 17 studies, 8 were RCT,[27,32,57–61,65] and 9 were prognostic cohorts.[28,48,62–64,66–69] Ten studies were classified as Phase I,[27,28,32,60–63,66,67,69] and 7 as Phase II studies.[48,57–59,64,65,68] No studies were classified as Phase III.

Types of Participants. Four studies included patients recruited from a population of employees,[27,59–61] and 13 studies included patients seeking treatment at a rehabilitation center. Sample sizes varied from 58 to 476 cases, with most studies including approximately 100 cases. The mean age of studied patients was approximately 40–45 years. The ratio of male-to-female differed per study, with one study including only females.[60] The duration of LBP varied from 3 months to a maximum of 26 years.

Types of Interventions. Six articles studied back schools,[27,32,48,60,62,64] 3 studied back schools *versus* multidisciplinary treatment,[28,57,60] and 8 studied multidisciplinary treatment alone.[58,59,61,63,65–69] Although the basic principles of multidisciplinary treatment and back schools are comparable, there is a large variety in duration, setting (inpatient or outpatient), and implementation between the studied interventions. For example, multidisciplinary treatment at one center may be based on cognitive behavioral concepts but in another, on operant behavioral ones. Also, several back schools offer consultation of a psychologist if needed, and others did not.

Follow-up. Outcome was measured after different periods of follow-up. The shortest follow-up was set at discharge,[66] and the longest at 30 months.[61] The percentage loss to follow-up varied from 0%[66,67] to 27%,[64] in 4 articles the percentage was unclear.[58,61,63,68]

Types of Baseline Measures. In total, 79 prognostic factors were studied. The number of relevant predictors differed substantially per author (Figure 2). Most studies focused on 1 to 3 prognostic variables. Five authors studied more than 9 variables,[57–60,65] with a maximum of 19 variables studied by Bendix *et al*.[57]

Sociodemographic variables were studied in 8 articles,[57–60,64–66,68] physical variables in 7,[32,48,57,60–63] and psychological variables in 10.[27,28,32,58,59,62,65,67–69] Four articles studied "other " predictors, which includes baseline measurements of activity and participation limitation.[57–59,69]

After classifying predictors into 3 main domains (*i.e.*, sociodemographic, physical, and psychological), it was clear that none of the authors studied predictors from these 3 domains simultaneously. Nine articles studied

Table 1. Summary of Design Characteristics of Studies on the Prognosis of Treatment Outcome for Patients with CLBP

Reference	Design	Phase Study/Analysis	Source Population (No., workers/referrals, age, gender, duration of pain)	Follow- up	% Loss to Follow-up
Bendix *et al* [57]	RCT	II/Multiple regression	816 Pts. referred to Copenhagen Back Center (621 to functional restoration program ("multidisciplinary") and 195 to control program ("back school"). Age 40 yrs. 67% female. First LBP episode at 27 yrs (median).	1 yr	15%
Haazen *et al* [58]	RCT	II/Multiple regression	58 Pts. referred to Pain Clinic of University Hospital and a rehabilitation center (19 to operant-, 18 to operant-cognitive, 21 to operant-respondent treatment). M age 40.5 yrs. 75.9% female. Duration of LBP >6 mos (53.4% >10 yrs).	Discharge, 6, 12 mos	Unclear
Härkäpää *et al* [59]	RCT	II/Multiple regression	476 Pts. (156 inpts., 150 outpts., 153 controls) selected from blue-collar workers employed by Finnish railways, post, and telecommunications, enterprises in Helsinki area and farmers from southern Finland. All approached by mail. M age 45 yrs. 63% male. Duration of LBP: ± 14 yrs.	3 mos	4%
Hurri [60]	RCT	I/Discrimination analysis	188 Pts. (95 to "back school", 95 controls) selected among the employees of a major Finnish cooperative. Age 46.1 ± 9.5 yrs (M/SD), 100% Female. Duration of LBP: 11.6 ± 9.4 yrs.	12 mos	13%
Hurri *et al* [61]	RCT	I/One-way ANOVA	245 Pts. (81 inpts., 88 outpts., 76 controls) selected from blue-collar workers in Helsinki. Primary selection by mail. M age 44.5 yrs, 71% male. M duration of LBP ± 12 yr.	3 and 30 mos	Unclear
Hutten *et al* [62]	Prognostic cohort	I/Two-way ANOVA	84 Pts. referred to a Rehabilitation Center. Age 38.4 ± 8.6 yrs (M/SD), 49 males. Duration of LBP 6.6 ± 8.1 yrs (M/SD).	1 wk	10.5%
Julkunen *et al* [27]	RCT	I/Multiple discrimination analysis	188 (95 for treatment and 93 controls). Employees from a large commercial enterprise in Finland. M age 46.1 ± 9.5 yrs, M duration of LBP 11.6 ± 9.4 yrs.	12 mos	14%
Long [63]	Prognostic cohort	I/ANOVA	223 Pts. referred to Columbia Rehabilitation Center. Age 38.2 ± 10.4 yrs (centralizers); 39.3 ± 9.9 yrs (non centralizers). 64%–74% male. Duration of LBP 7.2 ± 6.4 mos (centralizers); 8.8 ± 9.2 mos (noncentralizers).	Discharge, 2 yrs	Unclear
Luoto *et al* [48]	Prognostic cohort	II/Multiple regression	68 Pts. admitted to rehabilitation centre Finland. Selected by Social Insurance Institution of Finland. Age 43.7 ± 8.8 yr (M/SD). 33 males. Duration of LBP 7.8 ± 8 yrs (women); 12.1 ± 7.8 yrs (men).	6 mo	4%
Rainville *et al* [64]	Prognostic cohort	II/Multiple regression	117 Pts. referred to rehabilitation program. M age 39 yrs (compensation); 43 yrs (no compensation). 40–58% Female. Duration of LBP 45 mos.	3, 12 mos	3 mos: 12% 12 mos: 27%
Talo *et al* [28]	Prognostic cohort	I/ANOVA	173 Pts. referred to Social Insurance Institution-financed rehabilitation program, Finland, Rehabilitation Research Center (63 to functional activation program ("multidisciplinary"), 107 to spa resort program ("back school")) Median age 40.4 yrs. 101 females. Duration of LBP 6–317 mos.	12 mos	2%
Talo *et al* [65]	RCT	II/Multiple regression	173 Pts. referred to rehabilitation Center in Finland financed by social insurance institution (60 to functional activation program ("multidisciplinary"), 105 to spa resort program ("back school"). Median age 40.4 yrs. 101 females. Duration of LBP 6–317 mos.	12 mos	2%
Trief and Stein [66]	Prognostic cohort	I/Multivariate ANOVA	81 Pts. referred to back rehabilitation program. M age 39.1 yrs (litigation)/42.2 yrs (non litigation). 47 females. Duration of LBP 3.4 yrs (litigation); 3.9 yrs (nonlitigation).	Discharge	0%
Vendrig *et al* [67]	Prognostic cohort	I/Partial correlations	120 Pts. referred to the Netherlands Back Advice Center. Age 41.3 ± 9.0 yrs (M/SD). 78 males. Duration of LBP 47.6 ± 37.6 mos (M/SD).	6 mos	0%
Vendrig *et al* [68]	Prognostic cohort	II/Hierarchical regression	120 Pts. referred to the Netherlands Back Advice Center. Age 41.3 ± 9.0 yrs (M/SD). 78 males. Duration of LBP 47.6 ± 37.6 mos (M/SD).	6 mos	Unclear
Vollenbroek-Hutten *et al* [32]	RCT	I/ANOVA	142 Pts. (69 for treatment and 73 controls) referred to a Rehabilitation Center. Age 38.5 ± 9.8 yr (M/SD). Median duration of LBP 72 mos.	Discharge, 6 mos	6 mos: 13%
Walsh and Radcliffe [69]	Prognostic cohort	I/Univariate analysis	84 Pts. referred to back pain unit (referred from another hospital department at King's Mill Center for Health Care Services) Median age 47 yrs (22–70). 45 Males. Duration of LBP min.12 mos.	Discharge, 3 mos	3 mos: 12%

Phase I are exploratory studies (hypothesis generating) that seek an association between a prognostic marker and a certain outcome variable. Phase II are extensive exploratory studies (hypothesis generating) that value one or more prognostic variables. Phase III are confirmatory studies of *a priori* stated hypotheses of the value of a set of prognostic markers.
ANOVA = analysis of variance; M = mean; Pts = patients.

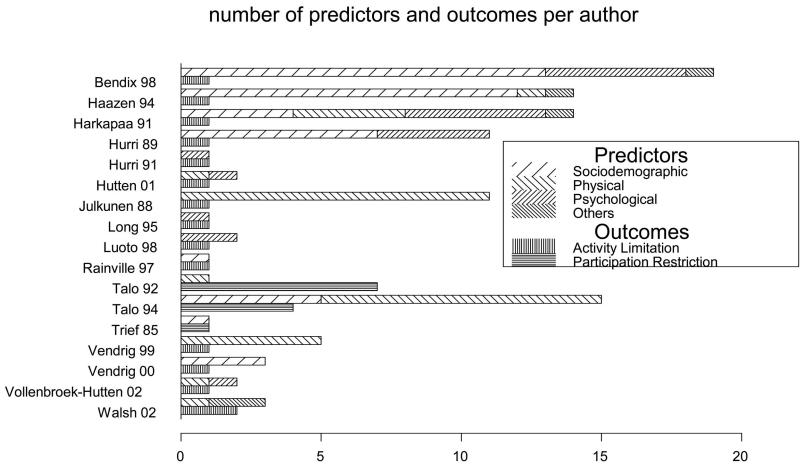

Figure 2. Number of predictors and outcome measures per study, classified by type of predictor (*i.e.,* sociodemographic, physical, psychological, others) and type of outcome measure (*i.e.,* activity limitation, participation restriction).

predictors from one domain,[27,28,48,61,63,64,66,67,69] and 8 predictors from one domain.[32,57–60,62,65,68]

Types of Outcome Measures. In total, 19 outcome measures were used to measure the domains activity limitation or participation restriction. Eight different measures were used to measure activity limitation and 11 to measure participation restriction. Only 3 authors studied more than one outcome measure. Talo *et al* studied 7[28] and 4[65] different outcome measures, respectively, and Walsh and Radcliffe[69] 2. Figure 2 gives an overview of the number of prognostic factors and outcome measures per study. Three authors studied measures of participation restriction,[28,65,66] and the others studied measures of activity limitation.[27,32,48,57–64,67–69]

Overall Level of Evidence

Table 2 provides an overview of the available evidence for the different prognostic factors per treatment (back schools/multidisciplinary) and outcome measure (activity limitation/participation restriction). From the table it is clear that none of the articles, with the exception of 2,[32,63] studied the relation of a specific predictor, treatment, and outcome measure more than once. This result means that evidence in this study is limited. Conclusions can only be drawn if predictors, treatments, and outcomes are grouped together in comparable domains. The heterogeneity of the study population, prognostic factors, and outcome measures precluded statistical pooling of the results and necessitated a qualitative summary of the results.

Sociodemographic Predictors

Consistent evidence was found that personal characteristics like age[57–59,68] and gender[57–59] were not predictive. Height and weight also lacked predictive value, although evidence was weak.[57]

For health related variables (*e.g.*, smoking), different variables were studied, most lacking predictive value.[57–60] Different results were found for the use of medication at baseline. One article studied back schools and found no predictive value at 12 months.[60] Another article studied multidisciplinary treatment and found a negative predictive value for outcome at discharge and at 6 months.[58] However, in this study, the explained variance in outcome that could be attributed to medication was only 10% at 6-month and 0% at 12-month follow-up.[58] Thus, both articles showed that use of medication had no predictive value at 1 year, but no conclusion could be drawn for shorter follow-ups.

Pain related variables were studied by 6 authors.[57–60,65,68] Pain duration consistently lacked predictive value.[57–59,68] Consistent results were also found for the negative predictive value of pain intensity,[57,65] although not for pain intensity in the leg.[57] Higher pain intensity at baseline predicted worse outcome. Talo *et al*[65] drew the same conclusion for pain interference (*i.e.*, if patients have more interference with activities, the outcome was worse). It is noteworthy that Talo *et al*[65] studied different outcome measures and patient groups (*e.g.*, "fit" and "unfit" patients with CLBP), and these results were only found for specific outcome measures and patient groups.

Not predictive for outcome are social status related variables.[57,58,60,65,68] There was one study that found that "better functioning in leisure time" was, in combination with other prognostic factors, predictive for better outcome.[65]

Concerning work related variables,[57,58,60,64–66] evidence was found for corresponding variables measuring subjective work capacity and experience. The "ability to,"[57] "functioning at,"[65] "adjustment at,"[65] and "satisfaction of" work[60] were all positive predictors for outcome of both treatment methods. The only exception was the variable "ability to work," which showed different predictive values for different treatments or outcome measures.[57,58] "Physical strenuousness of the job" consistently lacked predictive value,[57,60] and "vibrations in the job" showed inconsistent results for different treatments.[57]

Sick leave was an inconsistent predictor: a negative predictive value in one study[57] and was not confirmed in another.[60] The variable "compensation" showed comparable results: varying from a negative predictive value,[64] no predictive value,[66] to a positive predictive value[58] for treatment effect.

Table 2. Overall Level of Evidence for Prognostic Factors and Their Association with Outcome (activity limitation/ participation restriction)

Domain	Group	Prognostic Factor	Treatment	Outcome	Studies Assessed	Association*	No Association†	Evidence‡
Sociodemographic	Personal	Age	Bs	A	1		Bendix *et al*[57]	W
		Age	Mu	A	4		Bendix,[57] Haazen,[58] Härkäpää[59] *et al*, Vendrig[68]	C
		Gender	Bs	A	1		Bendix *et al*[57]	W
		Gender	Mu	A	3		Bendix,[57] Haazen,[58] Härkäpää[59] *et al*	C
		Height	Bs	A	1		Bendix *et al*[57]	W
		Height	Mu	A	1		Bendix *et al*[57]	W
		Weight	Bs	A	1		Bendix *et al*[57]	W
		Weight	Mu	A	1		Bendix *et al*[57]	W
	Health	Smoking	Bs	A	1		Bendix *et al*[57]	W
		Smoking	Mu	A	1		Bendix *et al*[57]	W
		No. of back surgeries	Mu	A	2		Haazen,[58] Härkäpää[59] *et al*	C
		Use of nerve blocks	Mu	A	1		Haazen *et al*[58]	W
		Use of supportive equipment	Mu	A	1		Haazen *et al*[58]	W
		TENS	Mu	A	1		Haazen *et al*[58]	W
	Health	Analgesics	Bs	A	1		Hurri[60]	W
		Analgesics	Mu	A	1	Haazen *et al*[58]		W
	Pain	Age first pain	Bs	A	1		Bendix *et al*[57]	W
		Age first pain	Mu	A	1		Bendix *et al*[57]	W
		Duration	Mu	A	3		Haazen,[58] Härkäpää[59] *et al*, Vendrig[68]	C
		Intensity back	Bs	A	1	Bendix *et al*[57]		C
		Intensity back	Mu	A	1	Bendix *et al*[57]		
		Intensity	Bs	P	1	Talo *et al*[65]§		
		Intensity	Mu	P	1	Talo *et al*[65]§		
		Intensity leg	Bs	A	1		Bendix *et al*[57]	W
		Intensity leg	Mu	A	1		Bendix *et al*[57]	W
		Interference	Bs	P	1	Talo *et al*[65]§		W
		Interference	Mu	P	1	Talo *et al*[65]§		W
		No. of painful spots	Bs	A	1		Hurri[60]	W
	Social	Education	Bs	A	1		Hurri[60]	W
		Education	Mu	A	2		Haazen *et al*[58], Vendrig[68]	C
		Social status (blue *vs* white collar)	Bs	A	1		Bendix *et al*[57]	W
		Social status (blue *vs* white collar)	Mu	A	1		Bendix *et al*[57]	W
		Civil status	Mu	A	1		Haazen *et al*[58]	W
		Functioning (leisure)	Bs	P	1	Talo *et al*[65]§		W
		Functioning (leisure)	Mu	P	1	Talo *et al*[65]§		W
	Work	Work ability	Bs	A	1		Bendix *et al*[57]	W
		Work ability	Mu	A	2	Bendix *et al*[57]	Haazen *et al*[58]	NC
		Work satisfaction	Bs	A	1	Hurri[60]		W
		Functioning (work)	Bs	P	1	Talo *et al*[65]§		W
		Functioning (work)	Mu	P	1	Talo *et al*[65]§		W
		Work adjustment	Bs	P	1	Talo *et al*[65]§		W
		Work adjustment	Mu	P	1	Talo *et al*[65]§		W
	Work	Vibrations in job	Bs	A	1		Bendix *et al*[57]	W
		Vibrations in job	Mu	A	1	Bendix *et al*[57]		W
		Physical strenuousness	Bs	A	2		Bendix *et al*[57], Hurri[60]	C
		Physical strenuousness	Mu	A	1		Bendix *et al*[57]	W
		Sick leave	Bs	A	2	Bendix *et al*[57]	Hurri[60]	NC
		Sick leave	Mu	A	1		Bendix *et al*[57]	W
		Compensation	Bs	A	1	Rainville *et al*[64]		W
		Compensation	Mu	A	1	Haazen *et al*[58]		W
		Litigation	Mu	P	1		Trief and Stein[66]	W

(*Table continues*)

Table 2. (*Continued*)

Domain	Group	Prognostic Factor	Treatment	Outcome	Studies Assessed	Association*	No Association†	Evidence‡
Physical	-	Aerobic capacity	Bs	A	1		Bendix *et al*[57]	W
		Aerobic capacity	Mu	A	2		Bendix *et al*[57], Hurri[60]	C
		Muscle endurance	Bs	A	1		Bendix *et al*[57]	W
		Muscle endurance	Mu	A	1		Bendix *et al*[57]	W
		Trunk muscle strength	Bs	A	1		Hurri[60]	W
	-	Mobility	Bs	A	2		Bendix *et al*[57], Hurri[60]	C
		Mobility	Mu	A	1		Bendix *et al*[57]	W
		Ability to do squats	Bs	A	1		Hurri[60]	W
		Sport activities	Bs	A	1		Bendix *et al*[57]	W
		Sport activities	Mu	A	1	Bendix *et al*[57]		W
		Postural control	Bs	A	1		Luoto *et al*[48]	W
		Psychomotor speed	Bs	A	1		Luoto *et al*[48]	W
		Centralization phenomenon	Mu	A	1		Long[63]	W
		Dynamometry	Bs	A	2	Hutten *et al*[62]	Vollenbroek-Hutten *et al*[32]	NC¶
Psycho-logical	Psychic	Symptom Checklist-90 (SCL-90)	Bs	A	1	Hutten *et al*[62]		W
	Health	Rorschach test‖	Bs	A	1	Julkunen *et al* [27]		W
		Sentence Completion Test (SCT)	Bs	A	1		Julkunen *et al* [27]	W
		Distress scale	Bs	P	1		Talo *et al*[65]§	W
		Distress Scale	Mu	P	1	Talo *et al*[65]§		W
	Psychic	Stress Appraisal Questionnaire	Bs	P	1	Talo *et al*[65]§		W
	Health	Stress Appraisal Questionnaire	Mu	P	1		Talo *et al*[65]§	W
		Severity Scale of Mental Disorders	Bs	P	1	Talo *et al*[65]§		W
		Severity Scale of Mental Disorders	Mu	P	1	Talo *et al*[65]§		W
		General Health Questionnaire (GHQ)	Mu	A	1		Härkäpää *et al*[59]	W
		Middlesex Hospital Questionnaire (MHQ)**	Bs	A	1		Julkunen *et al*[27]	W
		Minnesota Multiphasic Personality Inventory (MMPI)	Mu	A	1		Haazen *et al*[58]	W
		MMPI-2 selected scales††	Mu	A	1		Vendrig *et al*[67]	W
		MMPI-2 Personality Psychopathology Five (Psy-5)	Mu	A	1		Vendrig[68]	W
		Beck Depression Inventory	Bs	P	1	Talo *et al*[65]§		W
		Beck Depression Inventory	Mu	P	1		Talo *et al*[65]§	W
	Psychic	Hypochondria scale	Bs	P	1	Talo *et al*[65]§		W
	Health	Hypochondria scale	Mu	P	1		Talo *et al*[65]§	W
		Denial scale	Bs	P	1	Talo *et al*[65]§		W
		Denial scale	Mu	P	1	Talo *et al*[65]§		W
	Cognitive variables	Multidimensional Pain Inventory (MPI)	Bs	A	1		Vollenbroek-Hutten *et al*[32]	W§§
		MPI	Bs	P	1	Talo *et al*[28]§		W§§
		MPI	Mu	P	1	Talo *et al*[28]§		W§§
		Cognitive Self-Statements scale	Bs	P	1	Talo *et al*[65]§		W
		Cognitive Self-Statements scale	Mu	P	1		Talo *et al*[65]§	
		Increased Activity scale	Bs	P	1	Talo *et al*[65]§		W
		Increased Activity scale	Mu	P	1		Talo *et al*[65]§	W
		Belief in control by others	Mu	A	1		Härkäpää *et al*[59]	W
		Internal Locus of Control	Mu	A	1		Härkäpää *et al*[59]	W
		Pain Beliefs Questionnaire (PBQ)-organic beliefs	Mu	A	1	Walsh and Radcliffe[69]		W
	Cognitive variables	Belief in chance control of disease scale	Mu	A	1		Härkäpää *et al*[59]	W
		Belief in chance control of disease scale	Bs	P	1	Talo *et al*[65]§		W

(*Table continues*)

Table 2. (*Continued*)

Domain	Group	Prognostic Factor	Treatment	Outcome	Studies Assessed	Association*	No Association†	Evidence‡
	Cognitive variables	Belief in chance control of disease scale	Mu	A	1		Härkäpää *et al*[59]	W
		Belief in chance control of disease scale	Bs	P	1	Talo *et al*[65]§		W
		Belief in chance control of disease scale	Mu	P	1	Talo *et al*[65]§		W
		Wechsler Adult Intelligence Scale	Bs	P	1	Talo *et al*[65]§		W
		Wechsler Adult Intelligence Scale	Mu	P	1	Talo *et al*[65]§		W
Others	-	Activity of Daily Living Scale	Bs	A	1	Bendix *et al*[57]		W
		Activity of Daily Living Scale	Mu	A	1	Bendix *et al*[57]		W
		Activity change during baseline	Mu	A	1	Haazen *et al*[58]		W
		Low Back Pain Disability Index (LBPDI)	Mu	A	1	Härkäpää *et al*[59]		W
		Modified Roland and Morris Disability Questionnaire (RMDQ)	Mu	A	1	Walsh and Radcliffe[69]		W
		SF-36 physical functioning	Mu	A	1	Walsh and Radcliffe[69]		W

* Only significant associations are noted (*i.e.*, $P < 0.05$).
† No significant associations ($P \geq 0.05$)
‡ Consistent (C) evidence is ≥2 studies reporting consistent results or 75% of the studies reporting similar conclusions. No consistent (NC) evidence is ≥2 studies reporting inconsistent results. Weak (W) evidence is from 1 study.
§ Talo *et al*[65] found different predictive values for different patient groups (*i.e.*, "fit" and "unfit" patients) and for different outcome measures of participation restriction.
¶ Noted as inconsistent, although both studies find the *same* association but Vollenbroek-Hutten *et al*[32] slightly failed to reach statistical significance.
‖ Rorschach variables significantly associated with outcome: number of form-color responses (FC), total number of answers (R), reality index by Neiger (Ri), modified genetic level index (GL).
** MHQ consists of 6 scales: anxiety, phobic anxiety, obsessional, somatic, depressive, hysterical. The subscales and total score have no predictive value.
†† MMPI-2 scales: Obsessiveness, Lassitude-Malaise, Somatic Complaints, Depression, Hypochondriasis.
§§ MPI clusters significantly associated with better outcome (depending on outcome measure used): "dysfunctionals" and "interpersonally distressed." Vollenbroek-Hutten *et al*[32] find the same association as Talo *et al*[65] but slightly failed to reach statistical significance.
A = activity limitation; Bs = back school; Mu = multidisciplinary treatment; P = participation restriction.

Physical Predictors

Physical variables were studied by 7 authors,[32,48,57,60–63] and, overall, these were not predictive. Physical variables as strength, endurance, or mobility had no prognostic value.[57,60,61] A trend was found for performance on a dynamometer.[32,62] A dynamometer measures angular position, velocity and torque of the 3 primary movement axes of the back. Patients with an "expected" performance (*i.e.*, lower than healthy subjects but with consistent test behavior) have better outcome following a back school treatment.[62] The same trend was found in another study, although the results were not significant.[32] Postural control, psychomotor speed,[48] or the centralization phenomenon according to Long[63] were not predictive either. One study found that participation in sports predicted better outcome.[57]

Psychological Predictors

A variety of measures of psychic health and cognitive variables were studied.[27,28,32,58,59,62,65,67–69] No consistent results were found for the predictive value of measures for the overall level of psychic health. Psychic health was measured with a variety of scales measuring overall psycho neuroticism. All scales were only studied once, except for the Minnesota Multiphasic Personality Inventory (MMPI), which was included in 3 studies.[58,67,68] The MMPI provides patient profiles of psychopathology by combining scales[21] and was not associated with treatment success.[58] However, 2 articles used an alternative approach by examining the predictive value of *individual* scales of the MMPI-2, the successor of the MMPI. Scores on the MMPI-2 "Personality Psychopathology Five scales" were not associated with outcome either.[68] In another article concerning the same study, the author found that high scores on several other scales of the MMPI-2 (*i.e.*, obsessiveness, depression, hypochondriasis, lassitude-malaise, somatic complaints) were associated with worse outcome,[67] although the explained variance was low, and the results slightly failed to reach statistical significance ($P < 0.06$).

Other scales, including the Middlesex Hospital Questionnaire (MHQ),[27] the Sentence Completion Test,[27] and the General Health Questionnaire,[59] were not predictive either, with exception of the Symptom Check List-90)[62] and the Rorschach inkblot test.[27] A total score more than 198 on the Symptom Check List-90 was a negative predictor for outcome.[62] The Rorschach inkblot is used as a diagnostic tool for psychiatric diagnoses and for particular psychological symptoms.[70] Several Rorschach variables were associated with better outcome. As the author explained, patients with "good cog-

nitive, intellectual capacity with undisturbed reality testing" were more likely to respond well to treatment.[27]

Although measures of overall psycho neuroticism were studied only once, several aspects of psychic health (*i.e.*, depression, hypochondriasis, and obsessiveness) were studied more frequently. However, consistent results were not found for any of these variables. Depression was measured with the Beck Depression Inventory[65] but also with the MMPI-2 scale depression[67] and the MHQ.[27] One study found that high levels of depression were associated with worse outcome,[67] although the correlations were low and slightly failed to reach statistical significance ($P < 0.06$). Another study found no association,[27] and Talo *et al*[65] showed different results for multidisciplinary or back school treatment. Also, no consistent results were found for the variable "obsessiveness," studied by the MMPI-2[67] and the MHQ.[27] The same was true for hypochondriasis; 2 authors found a negative predictive value[65,67] and again, Talo *et al*[65] found different results for multidisciplinary or back school treatment.

Cognitive related variables, such as coping and beliefs, were studied in 5 articles.[28,32,59,65,69] The Multidimensional Pain Inventory (MPI) was classified as cognitive variable because one of the measured aspects is coping behavior. However, it may also be classified as a measure of psychic health because it also measures pain-relevant psychosocial aspects.[71]

Although studies included different measures of coping variables, all showed that low levels of active coping skills at baseline were predictive of better outcome.[28,32,65] Talo[28] and Vollenbroek[32] *et al* used the MPI which showed that patients characterized as "dysfunctional " and "interpersonally distressed" had better outcomes after back school treatment than "adaptive copers."[28] The results by Vollenbroek-Hutten *et al*[32] failed to reach significance due to small subgroups. Patients who are "dysfunctional" and "interpersonally distressed" are both characterized by high psychological distress, pain intensity, and interference with daily activities and low levels of life control. In contrast, "adaptive copers" have relatively low levels of psychological distress, pain intensity, and interference and higher life control. Active coping was also studied by the Cognitive Self-Statements scale and the Increased Activity Scale, with higher scores referring to more active coping.[65] In accordance with aforementioned results, Talo *et al*[65] found that lower active coping skills, measured with the Cognitive Self-Statements scale, were associated with better outcome after back school treatment, although no predictive value was found for multidisciplinary treatment. In general, the Increased Activity Scale had no predictive value, although it was positively associated with the outcome measure "panel assessment" after back school treatment.

Disease related beliefs were studied in 3 articles,[59,65,69] but no consistent results were found, and most variables were only studied once. Different beliefs were measured, like the belief that the disease can be controlled by chance, or by others, or by oneself (internal locus of control). The first 2 variables could also be classified as passive coping strategies and the last as an active coping strategy. Concerning belief in chance, Härkäpää *et al*[59] found no predictive value in contrast to Talo *et al*[65] who showed that belief in chance was associated with worse outcome after both multidisciplinary and back school treatment. Belief in control by others had no value either.[59] One study found that belief in control of back pain was a positive predictor for better outcome,[59] although it slightly failed to reach statistical significance ($P < 0.059$). Finally, strong beliefs in an organic cause of pain (and not a psychological one) were associated with better outcome, although the correlation was low.[69]

Intelligence, which was also classified as a cognitive variable, was studied only once[65] and showed contradictory results within that study. Intelligence was either positively or negatively associated with outcome, or not associated with outcome at all, depending on the patient groups (those who are "fit" or "unfit") or outcome measure used.

Other Predictors

Different authors included baseline scores of different measurements of activity limitation as predictors of outcome.[57–59,69] All showed that high baseline scores predicted higher reduction in scores after treatment. In general, this result implies that patients who have more limitations at baseline will benefit most from treatment. One exception is the physical functioning scale of the Short Form-36. High baseline scores on this scale Short Form-36 (*i.e.*, better physical functioning) predicted higher reduction after treatment (*i.e.*, *worse* outcome).[69]

Sub Analyses and the Influence on the Level of Evidence

The level of evidence hardly changed if the results were analyzed for Phase I and II studies, or for different subpopulations. In general, evidence became weaker. If only positive findings from Phase II studies[48,57–59,64,65,68] were included (there were no Phase III studies), evidence was limited because less prognostic factors were studied. The prognostic value of "dynamometry" and the "MPI" was not found because the articles that studied these variables were excluded from the analysis.[28,32,62] Furthermore, the inconsistent predictive value of psycho neuroticism scales changed into consistent evidence having no predictive value.

The results were analyzed for the 2 predefined subpopulations, "employees" and "referrals to a rehabilitation center." Only 4 of the 17 studies studied a population of employees,[27,59–61] so it was not possible to draw conclusions about the overall evidence for this particular subgroup. For referrals to a rehabilitation center, the overall conclusion did not change.

■ Discussion

This systematic review summarizes the results of 17 internal valid studies focusing on the prognostic value of various factors for (back school or multidisciplinary) treatment outcome in patients with CLBP. All studies

were classified as Phase I or II (exploring the value of one or more prognostic factors) and none as Phase III (testing a prognostic model). Outcome was measured as activity limitation or participation restriction. Because of heterogeneity in study population, number of prognostic factors, treatment and outcomes, a meta-analysis could not be performed,[38,45,72–75] and a qualitative analysis was done to support our final conclusion.

Sociodemographic Predictors

For sociodemographic variables, this study showed no prognostic value for personal, health, and social status related variables. Other reviews confirm that in general, these variables have no value[22] or will only explain a small portion of the variance in outcome measures.[19] Some reviews show weak evidence of these prognostic variables because conclusions are drawn from only one or 2 included studies.[2,3,18,20]

The predictive value of pain intensity or interference shown in this study is confirmed by other reviews as well.[2,3,18,19] Not only is higher pain intensity related with worse outcome,[3,18,19] but higher interference of pain with activities is also associated with reduced treatment success.[2,3,19]

Concerning work related variables, this study showed no consistent evidence for sick leave or receiving compensation. However, other reviews find that both variables are negative predictors for return to work.[2,19,20,23,25] Two reviews[23,24] show that compensation is also associated with reduced treatment response. However, McCracken and Turk[22] find inconsistent results that they explain by the often unclear definitions used for compensation, what obscures the prognostic role in treatment outcome. This result could also be an explanation for the inconsistency found in this study.

On the other hand, the present study showed a consistent trend that parameters measuring subjective work capacity, like the ability to or adjustment at work, predicted better outcome. This trend is in line with other reviews.[3,19,22,25] In these reviews, the variables "less work disability,"[2] "availability of the job at return,"[18,19,22] and "longer time in the job"[18] are all related with better outcome.

Physical Predictors

Considering physical variables, this study found no predictive value. This finding is confirmed by other reviews who show that these variables are of minor importance[19,22] or find only weak evidence.[17,18]

Psychological Predictors

For psychological variables, this study showed no consistent evidence for the overall level of psychic health. Only 2 other reviews looked at overall psycho neuroticism level,[21,22] and both also concluded that it shows inconsistent predictive value. It is striking that in contrast to current practice, included in 3 studies,[27,67,68] the prognostic value of individual scales of psycho neuroticism (*e.g.*, MMPI-2) is described instead of the overall level of psychopathology.

Looking at the individual aspects of psychic health, our study showed that depression was an inconsistent predictor. This result is in contrast with other reviews showing that depression is a negative predictor. High levels of depression at baseline are associated with worse outcome,[2,3,22] and reduction of depressive symptoms is related to better outcome.[19,22] This effect suggests that depressed patients have more to gain from treatment. No clear explanation can be given for the discrepancy between our study and the other reviews.

For cognitive variables, this study concluded that low levels of active coping at baseline were related to better outcome. This result was reflected by the fact that patients, classified by the MPI as "dysfunctionals" or "interpersonally distressed," benefit more from treatment than "adaptive copers." These findings are in accordance with Turk and Okifuji,[16] who reviewed studies using the MPI classification and found comparable results. It could be that patients with poor functional profiles ("dysfunctionals" and "interpersonally distressed") benefit more from treatment than "adaptive copers," who have less to gain. Treatment helps the first group reducing distress and improving adequate coping skills. This hypothesis agrees with another review, which shows that adoption of a more active self-management orientation to pain is associated with better treatment outcome.[22]

Our results that "dysfunctionals" and "interpersonally distressed" (both with high levels of pain and low perceived life control) are likely to benefit from treatment may seem contradictory with our finding that higher pain intensity and interference predicts worse outcome. However, the MPI defines subgroups based on multiple aspects of chronic pain.[3,16] It is possible that the predictive value of pain intensity and interference varies with coping ability. In that case, patients with low perceived life control may benefit from treatment despite the high level of pain intensity. It could be that coping with pain is a more important prognostic factor than pain intensity alone.

We could not draw conclusions regarding the value of disease related beliefs because most authors studied different measures of beliefs. Only "belief in chance control of disease" was studied twice,[59,65] showing inconsistent results. However, both articles differed in outcome measure and population studied, which may explain the inconsistency. Other reviews show consistent results that maladaptive beliefs (*e.g.*, catastrophisizing) are associated with poor outcome and stronger beliefs in control with better outcome.[19,22]

Other Predictors

Our results showed that a high level of perceived disability at baseline was related to better outcome, which is in contrast with other studies showing it to be associated with *worse* outcome.[2,3,19,22] An explanation for our results could be that it is a reflection of the phenomenon

"regression to the mean." Or, it is possible that persons with high levels of activity limitation "have more to gain" with treatment, which leads to better outcome.[76] It was shown before that a decrease in perceived disability *during* treatment is related with treatment success.[22]

Limitations and Recommendations

Several limitations of this review must be considered. First, publication bias cannot be excluded.[77] Second, the review process must be considered. It is known that the risk of missing prognostic studies because of difficulties searching the literature is higher than for randomized trials.[72] Third, the criteria and operationalization list we used for quality assessment is subject for debate because a generally accepted criteria list for assessing prognostic studies does not exist yet. However, the used criteria were based on frequently used checklists, thus, it is unlikely that relevant criteria would have been missed. General consensus of a methodological criteria list for prognostic studies is needed. We recommend, in accordance with Altman[38] and Scholten-Peeters *et al*,[78] to include the following criteria in this list: identified population (criterion 2), defined inclusion and exclusion criteria (criterion 3), valid and reliable measurements of prognostic (criterion 8) and outcome variables (criterion 15), explicitly described and standardized intervention (criterion 12), drop outs acceptable (criteria 5 + 17), follow-up sufficiently long (criterion 19), appropriate analysis (criterion 21), and adjustment for important prognostic factors. Finally, evidence was limited because there were only 2 prognostic cohorts studying the exact same prognostic factors, intervention and outcome measures.[32,62] However, we could draw general conclusions about prognostic variables from comparable domains for certain types of intervention and outcome measures.

To facilitate future prognostic studies of treatment outcome, 2 things are important. Future research is necessary to confirm the generated hypotheses derived from the descriptive (Phase I) and exploratory (Phase II) studies. In the current literature, there is a lack of confirmatory (Phase III) studies, which study *a priori* stated hypotheses of the value of a set of prognostic markers.[42] Besides understanding which prognostic factors predict outcome of treatment, insight in treatment process variables should be improved. This result will help to understand why and how specific prognostic factors are associated with treatment outcome. The knowledge of treatment process variables will enable the development of adequate treatment modules matched to specific patient characteristics, with different prognoses. It is assumed that tailoring different interventions to different subgroups of patients will enhance treatment effects.[16]

■ Conclusion and Clinical Implications

In addition to physician experience, knowledge of prognostic factors may be very useful (*i.e.*, patients with favorable prognostic factors are likely to benefit from treatment and those with unfavorable prognostic factors are not). It is likely that defining subgroups of patients may have to be based on the multiaxial assessment of functioning because it is shown that prognostic factors from several domains are of value for predicting outcome. However, a generic set of predictors of outcome in multidisciplinary rehabilitation and back schools for patients with CLBP cannot be defined.

With caution, several guidelines based on several consistent predictive factors of rehabilitation outcome in patients with nonspecific CLBP may be given. Physicians seeing patients with high pain intensity, problems at work (*e.g.*, functioning at work, dissatisfaction) should be aware that these patients are likely to have poor treatment outcome. In addition, the low use of active coping skills and high perceived limitations of activity at baseline may predict better treatment outcome. Other sociodemographic and physical variables probably should not play a role in the treatment decision because these consistently lacked predictive values. The value of other psychological variables is not clear because no consistent evidence was found.

■ Key Points

- The evidence of prognostic factors for multidisciplinary and back school treatment outcome in patients with CLBP has been reviewed systematically by examination of prospective studies.
- Conclusions were based on 17 internally valid studies, which all explored the value of prognostic factors (Phase I or II), and none tested a prognostic model (Phase III).
- Study heterogeneity necessitated a qualitative summary of the results and limited evidence.
- Pain intensity, several work-related parameters, and coping style were consistently associated with outcome. Other sociodemographic and physical variables consistently showed no association with outcome. No consistent evidence was found for the predictive value of other psychological variables.
- Future confirmatory studies of prognostic factors and studies of treatment process variables may lead to improved interventions and higher treatment success.

References

1. Crook J, Milner R, Schultz IZ, et al. Determinants of occupational disability following a low back injury: A critical review of the literature. *J Occup Rehabil* 2002;12:277–95.
2. Feuerstein M, Beattie P. Biobehavioral factors affecting pain and disability in low back pain: Mechanisms and assessment. *Phys Ther* 1995;75:267–80.
3. Gatchel RJ, Gardea MA. Psychosocial issues: Their importance in predicting disability, response to treatment, and search for compensation. *Neurol Clin* 1999;17:149–66.
4. Gatchel RJ. A biopsychosocial overview of pretreatment screening of patients with pain. *Clin J Pain* 2001;17:192–9.
5. Truchon M. Determinants of chronic disability related to low back pain: Towards an integrative biopsychosocial model. *Disabil Rehabil* 2001;23:758–67.
6. Vlaeyen JW, Linton SJ. Fear-avoidance and its consequences in chronic musculoskeletal pain: A state of the art. *Pain* 2000;85:317–32.

7. Weiser S, Cedraschi C. Psychosocial issues in the prevention of chronic low back pain: A literature review. *Baillieres Clin Rheumatol* 1992;6:657–84.
8. Faas A. Exercises: Which ones are worth trying, for which patients, and when? *Spine* 1996;21:2874–8.
9. Guzman J, Esmail R, Karjalainen K, et al. Multidisciplinary rehabilitation for chronic low back pain: Systematic review. *BMJ* 2001;322:1511–6.
10. Karjalainen K, Malmivaara A, van Tulder M et al. Multidisciplinary biopsychosocial rehabilitation for subacute low back pain in working-age adults: A systematic review within the framework of the Cochrane Collaboration Back Review Group. *Spine* 2001;26:262–9.
11. Koes BW, van Tulder MW, van der Windt DAWM, et al. The efficacy of back schools: A review of randomized clinical trials. *J Clin Epidemiol* 1994;47:851–62.
12. Nentwig CG. Effectiveness of the back school. A review of the results of evidence-based evaluation. *Orthopade* 1999;28:958–65.
13. Schonstein E, Kenny DT, Keating J, et al. Work conditioning, work hardening and functional restoration for workers with back and neck pain. *Cochrane Database Syst Rev* 2003;CD001822.
14. Di Fabio RP. Efficacy of comprehensive rehabilitation programs and back school for patients with low back pain: A meta-analysis. *Phys Ther* 1995; 75:865–78.
15. van Tulder MW, Ostelo R, Vlaeyen JW, et al. Behavioral treatment for chronic low back pain: A systematic review within the framework of the Cochrane Back Review Group. *Spine* 2001;26:270–81.
16. Turk DC, Okifuji A. Directions in prescriptive chronic pain management based on diagnostic characteristics of the patient. *American Pain Society Bulletin* 1998. Available at: http://www.ampainsoc.org/pub/bulletin /sep98 /resupdate.htm. Accessed August 15, 2002.
17. Borge JA, Leboeuf-Yde C, Lothe J. Prognostic values of physical examination findings in patients with chronic low back pain treated conservatively: A systematic literature review. *J Manipulative Physiol Ther* 2001;24:292–5.
18. Feuerstein M, Menz L, Zastowney TR, et al. Chronic pain and disability: Vocational outcomes following multidisciplinary rehabilitation. *J Occup Rehabil* 1994;4:229–51.
19. Hasenbring M. Predictors of efficacy of treatment in chronic low back pain. *Curr Opin Anaesthesio*l 1998;11:553–8.
20. Hunter J. Medical history and chronic pain. *Clin J Pain* 2001;17:S20–5.
21. Love AW, Peck CL. The MMPI and psychological factors in chronic low back pain: A review. *Pain* 1987;28:1–12.
22. McCracken LM, Turk DC. Behavioral and cognitive-behavioral treatment for chronic pain: Outcome, predictors of outcome, and treatment process. *Spine* 2002;27:2564–73.
23. Rohling ML, Binder LM, Langinrichsen-Rohling J. Money matters: A meta-analytic review of the association between financial compensation and the experience and treatment of chronic pain. *Health Psychol* 1995;14:537–47.
24. Teasell RW. Compensation and chronic pain. *Clin J Pain* 2001;17:S46–64.
25. Walsh NE, Dumitru D. The influence of compensation on recovery from low back pain. *Occup Med* 1988;3:109–21.
26. Gatchel RJ, Noe CE, Pulliam C, et al. A preliminary study of multidimensional pain inventory profile differences in predicting treatment outcome in a heterogeneous cohort of patients with chronic pain. *Clin J Pain* 2002;18:139–43.
27. Julkunen J, Hurri H, Kankainen J. Psychological factors in the treatment of chronic low back pain. Follow-up study of a back school intervention. *Psychother Psychosom* 1988;50:173–81.
28. Talo S, Rytokoski U, Puukka P. Patient classification, a key to evaluate pain treatment: A psychological study in chronic low back pain patients. *Spine* 1992;17:998–1011.
29. Talo S, Forssell H, Heikkonen S, et al. Integrative group therapy outcome related to psychosocial characteristics in patients with chronic pain. *Int J Rehabil Res* 2001;24:25–33.
30. Turk DC, Rudy TE. Towards an empirically derived taxonomy of chronic pain patients: Integration of psychological assessment data. *J Consult Clin Psychol* 1988;56:233–8.
31. Turk DC. Customizing treatment for chronic pain patients: Who, what, and why. *Clin J Pain* 1990;6:255–70.
32. Vollenbroek-Hutten MMR, Hermens HJ, Wever D, et al. Main and subgroup specific effects of a multidisciplinary rehabilitation program for patients with chronic low back pain (submitted).
33. World Health Organization. International Classification of Functioning, Disability and Health. ICF checklist. 2001. [*World Health Organisation* web site]. Available at: http://www3.who.int/icf /icftemplate.cfm. Accessed August 1, 2002.
34. Assendelft WJ, van Tulder MW, Scholten RJ, et al. The practice of systematic reviews. II. Searching and selection of studies. *Ned Tijdschr Geneeskd* 1999; 143:656–61.
35. van Tulder MW, Assendelft WJ, Koes BW, et al. Method guidelines for systematic reviews in the Cochrane Collaboration Back Review Group for Spinal Disorders. *Spine* 1997;22:2323–30.
36. Faas A, Chavannes AW, Koes BW, et al. NHG-Standaard Lage-Rugpijn. *Huisarts Wet* 1996;39:18–31.
37. van Tulder M, Malmivaara A, Esmail R, et al. Exercise therapy for low back pain: A systematic review within the framework of the Cochrane Collaboration Back Review Group. *Spine* 2000;25:2784–96.
38. Altman DG. Systematic reviews of evaluations of prognostic variables. *BMJ* 2001;323:224–8.
39. Clarke M, Oxman AD. Collecting data. Cochrane Reviewers Handbook [updated April 2002]. Section 5. In: *The Cochrane Library*. Issue 2. Oxford: 2002. Update software. Updated quarterly.
40. Cochrane Collaboration. Etiology and prognosis: Critical appraisal form for observational studies. [*Cochrane Collaboration web site]*. Available at: http://www. cochrane.dk/nrsmg. Accessed August 12, 2002.
41. van der Windt DA, Zeegers MP, Kemper HC, et al. Practice of systematic reviews. VI. Searching, selection and methodological evaluation of etiological research. *Ned Tijdschr Geneeskd* 2000;144:1210–4.
42. Côté P, Cassidy JD, Carroll L, et al. A systematic review of the prognosis of acute whiplash and a new conceptual framework to synthesize the literature. *Spine* 2001;26:E445–58.
43. Kwakkel G, Wagenaar RC, Kollen BJ, et al. Predicting disability in stroke: A critical review of the literature. *Age Ageing* 1996;25:479–89.
44. Pincus T, Burton AK, Vogel S, et al. A systematic review of psychological factors as predictors of chronicity/disability in prospective cohorts of low back pain. *Spine* 2002;27:E109–20.
45. Altman DG, Lyman GH. Methodological challenges in the evaluation of prognostic factors in breast cancer. *Breast Cancer Res Treat* 1998;52:289–303.
46. Bergstrom G, Jensen IB, Bodin L, et al. The impact of psychologically different patient groups on outcome after a vocational rehabilitation program for long-term spinal pain patients. *Pain* 2001;93:229–37.
47. Jensen IB, Bodin L, Ljungqvist T, et al. Assessing the needs of patients in pain: A matter of opinion? *Spine* 2000;25:2816–23.
48. Luoto S, Taimela S, Alaranta H, et al. Psychomotor speed in chronic low-back pain patients and healthy controls: construct validity and clinical significance of the measure. *Percept Mot Skills* 1998;87:1283–96.
49. Strategier LD, Chwalisz K, Altmaier EM. Multidimensional assessment of chronic low back pain: Predicting treatment outcomes. *J Clin Psychol Med Settings* 1997;4:91–110.
50. Blake C, Garrett M. Impact of litigation on quality of life outcomes in patients with chronic low back pain. *Ir J Med Sci* 1997;166:124–6.
51. Carmody TP. Psychosocial subgroups, coping, and chronic low-back pain. *J Clin Psychol Med Settings* 2001;8:137–48.
52. Dozois DJA, Dobson KS, Wong M, et al. Predictive utility of the CSQ in low back pain: Individual vs. composite measures. *Pain* 1996;66:171–80.
53. Härkäpää, K, Järvikoski A, Estlander AM. Health optimism and control beliefs as predictors for treatment outcome of a multimodal back treatment program. *Psychol Health* 1996;12:123–34.
54. Mannion AF, Müntener M, Taimela S, et al. A randomized clinical trial of three active therapies for chronic low back pain. *Spine* 1999;24:2435–48.
55. O'Farrell V, Tate N, Aitken C. Attitudes and prognosis in chronic low back pain. *J Psychosom Res* 1993;37:415–22.
56. Peebles JE, Moore RJ. Illness schemata in patients with low back pain: Predictions of rehabilitation success. *Pain Clinic* 2000;12:237–46.
57. Bendix AF, Bendix T, Haestrup C. Can it be predicted which patients with chronic low back pain should be offered tertiary rehabilitation in a functional restoration program? A search for demographic, socioeconomic, and physical predictors. *Spine* 1998;23:1775–83.
58. Haazen IWCJ, Vlaeyen JWS, Kole-Snijders AMJ, et al. Behavioral rehabilitation of chronic low back pain: Searching for predictors of treatment outcome. *J Rehabil Sci* 1994;34–43.
59. Härkäpää K, Järvikoski A, Mellin G, et al. Health locus of control beliefs and psychological distress as predictors for treatment outcome in low-back pain patients: Results of a 3-month follow-up of a controlled intervention study. *Pain* 1991;46:35–41.
60. Hurri H. The Swedish back school on chronic low back pain II: Part II. Factors predicting the outcome. *Scand J Rehabil Med* 1989;21:41–4.
61. Hurri H, Mellin G, Korhonen O, et al. Aerobic capacity among chronic low-back-pain patients. *J Spinal Disord* 1991;4:34–8.
62. Hutten MM, Hermens HJ, Zilvold G. Differences in treatment outcome between subgroups of patients with chronic low back pain using lumbar dynamometry and psychological aspects. *Clin Rehabil* 2001;15:479–88.
63. Long AL. The centralization phenomenon. Its usefulness as a predictor or outcome in conservative treatment of chronic low back pain (a pilot study). *Spine* 1995;20:2513–20.
64. Rainville J, Sobel J, Hartigan C, et al. The effect of compensation involvement on the reporting of pain and disability by patients referred for rehabilitation of chronic low back pain. *Spine* 1997;22:2016–24.
65. Talo S, Puukka P, Rytokoski U, et al. Can treatment outcome of chronic low

back pain be predicted? Psychological disease consequences clarifying the issue. *Clin J Pain* 1994;10:107–21.
66. Trief P, Stein N. Pending litigation and rehabilitation outcome of chronic back pain. *Arch Phys Med Rehabil* 1985;66:95–9.
67. Vendrig AA, Derksen JJL, de Mey HR. Utility of selected MMPI-2 scales in the outcome prediction for patients with chronic back pain. *Psychol Assess* 1999;11:381–5.
68. Vendrig AA. MMPI-2 Personality Psychopathology Five (PSY-5) and prediction of treatment outcome for patients with chronic back pain. *J Pers Assess* 2000;74:423–38.
69. Walsh DA, Radcliffe JC. Pain beliefs and perceived physical disability of patients with chronic low back pain. *Pain* 2002;97:23–31.
70. Wood JM, Lilienfeld SO, Garb HN, et al. The Rorschach test in clinical diagnosis: A critical review, with a backward look at Garfield (1947). *J Clin Psychol* 2000;56:395–430.
71. Lousberg R, Van Breukelen GJ, Groenman NH, et al. Psychometric properties of the Multidimensional Pain Inventory, Dutch language version (MPI-DLV). *Behav Res Ther* 1999;37:167–82.
72. Altman DG. Systematic reviews of evaluations of prognostic variables. In: Egger M, Davey Smith G, Altman DG, eds. *Systematic Reviews in Health Care. Meta-analysis in Context*. London: BMJ Publishing Group, 2001: 228–247.
73. Buettner P, Garbe C, Guggenmoos-Holzmann I. Problems in defining cutoff points of continuous prognostic factors: example of tumor thickness in primary cutaneous melanoma. *J Clin Epidemiol* 1997;50:1201–10.
74. Greener J, Langhorne P. Systematic reviews in rehabilitation for stroke: Issues and approaches to addressing them. *Clin Rehabil* 2002;16:69–74.
75. Scholten RJ, Assendelft WJ, Kostense PJ, et al. The practice of systematic reviews. V. Heterogeneity between studies and subgroup analysis. *Ned Tijdschr Geneeskd* 1999;143:843–8.
76. Lofland KR, Burns JW, Tsoutsouris J, et al. Predictors of outcome following multidisciplinary treatment of chronic pain: Effects of changes in perceived disability and depression. *International Journal of Rehabilitation and Health* 1997;3:221–32.
77. Easterbrook PJ, Berlin JA, Gopalan R, et al. Publication bias in clinical research. *Lancet* 1991;337:867–72.
78. Scholten-Peeters GG, Verhagen AP, Bekkering GE, et al. Prognostic factors of whiplash-associated disorders: A systematic review of prospective cohort studies. *Pain* 2003;104:303–22.

■ Appendix 1

Criteria List for Methodological Quality Assessment[1,35,38,40–44]

1. The research question is well stated[40,42]

Patient selection:

2. The population is well identified (*)[38,40]
3. The inclusion and exclusion criteria are defined and appropriate (*)[1,35,38,41–44]
4. For RCT: treatment allocation[35,40]
   a) Was a method of randomization performed?
   b) Was the treatment allocation concealed?
5. Participation rate is reported and appropriate (*)[40,42]
6. Are all subjects representative of the same underlying population?[40,41]
7. Are the various groups comparable at baseline?[35,42]

Prognostic factors:

8. The methods used to measure the baseline prognostic variables are valid and reliable (*)[1,40–44]
9. The prognostic factor(s) is (are) measured in a standardized way[38,40,42]
10. Other relevant prognostic factors are measured[42,44]

Interventions:

11. Additional treatment effects during period of observation are avoided or comparable[35,43]
12. The intervention(s) is (are) explicitly described[35,38]
13. The compliance is acceptable in all groups[35]

Outcome measurement:

14. The same data collection is used for all members of the cohort[35,40,42]
15. The methods used to measure the outcome are defined and measurable (*)[1,38,40,42]
16. The methods used to measure the outcome are valid and reliable(*)[41–43]
17. % Follow-up is reported, explained, and reasonable (*)[1,35,40–44]
18. Loss to follow-up is equal in different groups[41,42]
19. The duration of follow-up is adequate[35,41–43]
    a) Was a short-term follow-up measurement performed?[35]
    b) Was a long-term follow-up measurement performed?[35,38]

Statistics

20. The sample size provides adequate statistical power [35,40–44]
21. Was the statistical methodology appropriate for the research question and study design?[1,38,40–44]
22. An intention-to-treat analysis is performed[35]
23. Control for statistical significance[43]
24. Control for multicollinearity[38,43,44]
25. The results are verifiable from the data[35,42]

General

26. Was bias or random error likely to have been avoided? (*)[40–42]

ad (*) criteria of internal validity

■ Appendix 2

Internal Validity Criteria

- The source population was well identified (criterion 2)
- Inclusion and exclusion criteria were defined and appropriate (criterion 3)
- The methods used to measure the prognostic factors were valid and reliable (criterion 8)
- The outcome was well defined and measurable (criterion 15)
- The measures of outcome were valid and reliable (criterion 16)
- The participation rate and percentage follow-up was reported and appropriate (together at least 60%), or a comparative analysis of participants and nonparticipants was presented (criteria 5 + 17)
- Was bias or random error likely to have been avoided? (criterion 26)