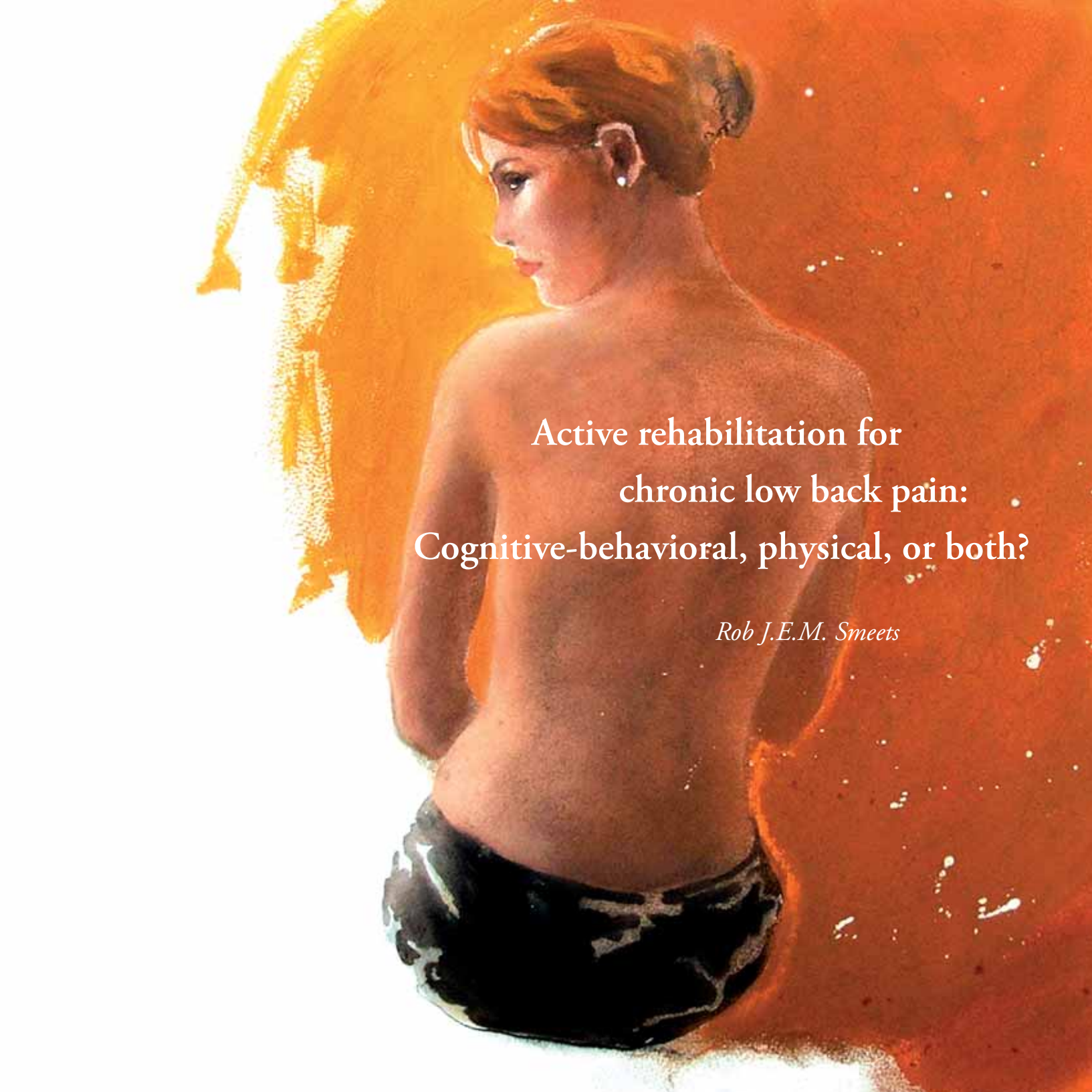

Active rehabilitation for chronic low back pain: Cognitive-behavioral, physical, or both?

Rob J.E.M. Smeets

Stellingen

behorende bij het proefschrift

Active rehabilitation for chronic low back pain:

Cognitive-behavioral, physical, or both?

Rob Johannes Elise Marie Smeets

Maastricht, 8 december 2006

Een behandeling die bestaat uit een combinatie van intensieve fysieke en cognitieve gedragstherapie is een jaar na het einde van de behandeling niet effectiever dan de afzonderlijke behandelonderdelen (dit proefschrift).

De effectiviteit van de fysieke training is deels verklaarbaar door een afname van pijn-catastroferen (dit proefschrift).

De geloofwaardigheid en verwachting ten aanzien van een behandeling voorspellen de uitkomst van behandeling van chronische lage rugklachten (dit proefschrift).

Gezien de diversiteit van de door lage rugpijnpatiënten geuite beperkingen verdient het aanbeveling meer patiënt relevante capaciteitstaken te ontwikkelen voor de evaluatie van de behandeling (dit proefschrift).

De wetenschappelijke impact factor van pijntijdschriften staat in geen verhouding tot de persoonlijke en maatschappelijke impact van pijn.

De betekenis en acceptatie (voor publicatie) van een wetenschappelijk artikel dient niet te worden afgemeten aan de verkregen uitkomsten maar aan de importantie van de getoetste hypothesen.

Pijnrevalidatie staat op pijnlijk gespannen voet met de daadwerkelijke vraag van de patiënt: pijnvermindering.

Het merendeel van de patiënten met chronische lage rugklachten kan meer fysieke belasting aan dan de meeste behandelaars voor mogelijk houden.

Het veelvuldig niet honoreren van zittend ziekenvervoer werpt voor de poliklinische revalidatie onoverbrugbare drempels op.

“Het afmatten van het lichaam heeft een bevrijdende werking op de geest” (In Joe Speedboot, Tommy Wieringa), of “Training the body is relaxing the mind” (Rob Smeets)

Een TomTom voor promovendi; een gat in de markt!

Active rehabilitation for chronic low back pain: Cognitive-behavioral, physical, or both?

Rob Johannes Elise Marie Smeets

Active rehabilitation for chronic low back pain: Cognitive-behavioral, physical or both?
Rob Smeets – Thesis Maastricht University – With references – With summary in Dutch.

ISBN: 90-9021283-3

Cover: Wil van der Laan

Financial support for the printing of this theses has been kindly provided by Care and Public Health Research Institute (CAPHRI), Rehabilitation Centre Blixembosch.

Book preparation, cover design: Reclamebureau KoolenVanWordragen,Valkenswaard, www.k-v-w.com.

Printing: Drukkerij Offsetservice, Valkenswaard.

Active rehabilitation for chronic low back pain: Cognitive-behavioral, physical, or both?

PROEFSCHRIFT

Ter verkrijging van de graad van doctor
aan de Universiteit Maastricht,
op gezag van de Rector Magnificus
Prof. mr. G.P.M.F. Mols
volgens het besluit van het College van Decanen,
in het openbaar te verdedigen
op vrijdag 8 december 2006 om 14.00 uur

door

Robert Johannes Elise Marie Smeets

Promotores

Prof. dr. J.A. Knottnerus
Prof. dr. J.W.S. Vlaeyen

Copromotor

Dr. A. Hidding (Atrium Medical Centre, Heerlen)

Beoordelingscommissie

Prof. dr. R.A. de Bie (voorzitter)
Prof. dr. J.H.B. Geertzen (Rijksuniversiteit Groningen)
Prof. dr. S.J. Morley (University of Leeds, UK)
Mevr. dr. A.M.C.F. Verbunt
Mevr. dr. T. van der Weijden

The studies presented in this thesis have been supported by the Zorgonderzoek Nederland and the Dutch Council for Medical and Health Research (ZON/MW), grant nr. 014-32-007, Rehabilitation Centre Blixembosch Eindhoven, and were conducted at the Care and Public Health Research Institute (CAPHRI), which participates in the Netherlands School of Primary Care research (CaRe), acknowledged in 1995 by the Royal Dutch Academy of Science (KNAW).

Contents

Contents

Introduction.

Chapter 1

Introduction

Low back pain disability: Defining the problem

Low back pain (LBP) is defined as pain, muscle tension or stiffness, localized in the back below the costal margin and above the inferior gluteal folds, with or without leg pain (sciatica). Nonspecific LBP is LBP not attributable to recognizable pathology (e.g. tumour, osteoporotic fracture, infection, rheumatoid arthritis, inflammatory process, traumatic fracture, spondylolisthesis or prolapsed intervertebral disc) (Deyo et al. 1992). LBP is called chronic when the complaints last for more than 3 months (Faas et al. 1996; Kendall et al. 1997; van der Heijden et al. 1995; Waddell 1998).

Nonspecific LBP is a very common disorder in the western industrialized world. The lifetime prevalence is reported as over 70% in developed countries. The adult incidence is 5% per year, the one-year prevalence is 15% to 45% (Andersson 1999). For the Dutch population, the one year prevalence in 1998 was 44%, the point prevalence 27% and the prevalence of chronic low back pain (CLBP) 21%. There is a slight decline in prevalence with increasing age. One third reported limitation in daily life due to their LBP and 24% reported sick leave due to LBP in the past year (Picavet and Schouten 2003). With a recovery rate of 75% to 90% within 4 to 6 weeks, the natural course of LBP is quite good. Only a small proportion will develop long lasting or frequently recurring LBP (Waddell 1998).

This group with CLBP accounts for 75% to 85% of the societal (direct and indirect) costs (Frymoyer 1988). In The Netherlands it is estimated that in 1991 (total population: 15.0 million), LBP accounted for 1.7% of the Gross National Product. The indirect costs (work absenteeism and disablement) constituted 93% of the total costs of back pain (US$ 4.8 billion). Sick leave accounted for US$ 3.1 billion and disablement for US$ 1.5 billion. The direct medical costs contributed only 7% (US$ 367 million) (van Tulder et al. 1995). These last findings were confirmed by Meerding et al. (1998) who calculated that musculoskeletal diseases of which LBP was the most important disorder, accounted for 6% to 7.3% of all Dutch health care costs in 1994 (total population: 15.3 million).

The total health care costs, representing 9.7% of the Dutch Gross National Product, were US$ 1381 per capita. Musculoskeletal diseases together with psychological/psychiatric diseases are the most important causes of sick leave. In the calculation of the direct medical costs, the costs of non-professional care and expenses paid by the patient himself (over the counter medication and so on) are not included, but seem to contribute 2/3 of all direct medical costs (Goossens et al. 1998; Goossens et al. 1996).

These findings confirm that CLBP is not only a tremendous health problem, but also a huge socio-economic problem. Despite the increasing amount of resources and the enormous variety of CLBP treatments being offered to CLBP sufferers and the fact that the incidence of nonspecific LBP is not increasing, the LBP disability is still getting steadily worse (Waddell 1998). This further underscores the pressing need to evaluate the effectiveness of interventions aimed at management of this chronic condition.

Active rehabilitation for CLBP: Are model-based treatments effective?

The studies presented in this thesis are part of the third research project of the LOBADIS-research program. This LOw BAck pain DISability and rehabilitation strategy-program aims at further refinement of knowledge of model-based treatments for patients with chronic disabling nonspecific LBP. The ultimate goal of the LOBADIS-program is to reduce CLBP disability by matching rehabilitation treatment to individual patient needs.

In the literature, three frequently used models regarding the development and maintenance of CLBP functional limitations are described:

1) The physical deconditioning model assuming that loss of muscle strength and endurance including aerobic capacity is responsible for reduced activity levels and hence functional limitations (Mayer et al. 1998; Verbunt et al. 2003).

2) The cognitive-behavioral model postulating that functional limitations results from maladaptive beliefs and avoidance behaviors that are maintained by learning processes (Fordyce 1976; Turk et al. 1983; Turk and Okifuji 2002; Vlaeyen et al. 1995).

3) The biopsychosocial model assuming that loss of functional abilities results from both the deconditioning and the cognitive-behavioral model (Waddell 1998).

Similarly, three major rehabilitation strategies for the reduction of CLBP disability can be distinguished: a biomedical approach aimed at restoring physical abilities, a cognitive-behavioral one aiming at the modification of maladaptive beliefs about pain and the acquisition of cognitive and behavioral coping skills, and finally, a combination of both approaches.

Based on the latest systematic reviews and meta-analyses, there is growing evidence that strengthening exercises combined with aerobic exercises as well as cognitive-behavioral treatment (CBT) are worth the effort when compared with no treatment or waiting list control. There is insufficient evidence for the effectiveness of strengthening and aerobic exercises versus other active therapies (Hayden et al. 2005a; Liddle et al. 2004; van Tulder et al. 2000). However, just recently, Hayden et al. (2005b) showed that exercise therapy consisting of individually designed programs, including stretching and high intensity strengthening exercises, improves pain and function. Controversy exists regarding the effectiveness of CBT when compared with alternative active treatments (Morley et al. 1999; Ostelo et al. 2005; van Tulder et al. 2000). Multidisciplinary treatment of at least 100 hours, combining exercise therapy, functional restoration and CBT appeared promising in comparison to other non-multidisciplinary treatments, whereas multidisciplinary rehabilitation programs of less than 30 hours failed to show improvements on several relevant

outcome measures. It should be taken into account though, that there is no consensus about the content, intensity and frequency of the different training sessions and the results are based on a relative low number of studies (Guzman et al. 2002).

However, by taking a closer look at the studies included in the systematic reviews and meta-analyses, it appeared that many therapies were not solely based on one of the three models mentioned above. For example in CBT, exercise therapy was used to increase a patient's level of activity while applying the operant learning principles. As a result, it is not clear whether CBT, exercise therapy or the combination of both was responsible for the improvement. Furthermore, many exercise therapies were not of sufficient intensity, frequency and duration to fulfill the physiologic training principles and therefore should not have been classified as reconditioning or strengthening therapies (ACSM 1998; Haskell 1994; Hilde and Bo 1998; Liddle et al. 2004; McArdle et al. 1996). At the start of our research project in 2001 and even now, the scientific evidence for the effectiveness of model-based treatments still is scarce and we do not know what treatment elements or combinations are really necessary to reach positive treatment results. Nor do we know how the offered treatments exert their effect and what treatment works best for whom. The treatments offered resemble a black box; the patient is offered several treatment modalities but the only thing that was really assessed is the final outcome. We do not know what really happens, how the results are obtained, who benefits most, and whether the effectiveness can be improved be adding, changing or removing treatment elements.

The main aim of this third research project of LOBADIS was to study the differential effect of model-based treatment approaches on disability by performing a comprehensive trial to assess the effectiveness of three treatments, each based on one of the abovementioned models. Furthermore, the trial was designed to get more insight in how a treatment works (mediation), and to get more insight in variables that are associated with the outcome of these model-based treatments.

Outline of this thesis

At the start of this research project, there was an ongoing debate regarding the presence of physical deconditioning and more specifically, loss of aerobic capacity and loss of strength/endurance of the lower back muscles in patients with CLBP. *Chapter 2* provides a systematic review regarding the association of physical deconditioning and CLBP. The physical deconditioning theory is described and features of physical deconditioning are defined. Next an overview of the evidence whether CLBP patients are physical deconditioned, and the effectiveness of treatments specifically aimed at physical reconditioning are presented and discussed.

As appeared from the review there were still many uncertainties regarding the existence of aerobic deconditioning in CLBP patients. Many researchers postulated that the activity level, especially regarding sport might be a very important contributing factor. Surprisingly, a sound comparison of CLBP patients and healthy controls matched for their sport activity had never been carried out. Furthermore, several other variables were mentioned as putative influential for the reduction of aerobic capacity, but most studies only calculated univariate correlations. In *chapter 3*, a study is presented in which the aerobic capacity of CLBP patients assessed with a submaximal Åstrand bicycle test is compared with the aerobic capacity of healthy controls matched for age, gender and level of sport activity. For this comparison an extensive data base regarding the cardiovascular capacity (also assessed with the Åstrand submaximal bicycle test) of Dutch healthy people was used. Furthermore, by using multiple linear regression techniques we analysed whether the difference in aerobic capacity between CLBP patients and healthy controls was associated with the putative influential variables such as pain intensity, pain-related fear, level of disability and working activity.

The main aim of this research was to test the effectiveness of model-based treatments for CLBP-disability. First, three different model-based treatments were developed based on scientific literature and in close collaboration with experts in the field of each of the abovementioned models. Next, a randomized controlled trial (RCT) was performed. In *chapter 4*, the short term results of this RCT in which three model-based active treatments are compared with a waiting list control group, and the combination treatment is compared with the single component treatments, are presented. In *chapter 5*, the one year follow-up results of this RCT are described by comparing the combination treatment with the two single treatments.

In order to improve the effectiveness/efficacy of treatment we need insight into what mechanisms are responsible for the desired outcomes and what kind of treatment is the best for a particular patient. These questions can be approached by using the moderator-mediator distinction (Vlaeyen and Morley 2005). Moderators provide the answer to the question "in what circumstances does the treatment work?", whereas mediators concern the question "how does the treatment work?" The RCT was designed to study the influence of potential mediating factors such as pain catastrophizing and internal control of pain. In *chapter 6*, the results on the mediating role of pain catastrophizing and internal control of pain are presented. In *chapter 7*, the question "which variables are associated with the outcome of these model-based treatments?" is addressed by studying the role of treatment and credibility assessed immediately after the patient has been told the treatment rationale.

At the start of this research, treatment evaluation in scientific research was mainly based on patient's reports by using questionnaires despite the fact that there was growing scientific evidence that the levels of perceived disability and physical activity are significantly lower than those observed or measured (Lee et al. 2001; Simmonds et al. 1998; Verbunt et al. 2001; Wittink 2005). Therefore it was decided that 'more objective' performance/capacity tasks should also be used in the evaluation of effectiveness of the treatments offered in our RCT (Simmonds et al. 1998; Waddell

1998). From the available scientific literature regarding cheap and easy-to-administer performance/capacity tasks in CLBP, six tasks were selected. In order to improve objectivity and reliability of the assessment within the RCT, the testing protocols of the tasks had to be adjusted. As a result, the existing psychometric data were not valid anymore. Furthermore, no data were available on the limits of agreement, defining the minimal change needed to be confident that the observed change between two measurements reflects real change in an individual patient and not just measurement error. *Chapter 8* describes the results of a test-retest reliability study in a subgroup of the trial population, addressing test-retest reliability, the limits of agreement and the clinical usefulness of each task.

Although the performance/capacity tasks are thought to be more objective than questionnaires and tend to measure the real capacity of patients and not the experienced disability, these tasks are also influenced by cognitive and emotional processes. *Chapter 9* describes the cross-sectional relationship assessed by multiple linear regression analysis, between both the physical and psychological factors and each of the six performance/capacity tasks.

Finally, *chapter 10* provides a general discussion evaluating the methods, findings and conclusions of all these studies together. Implications for daily, clinical practice and recommendations for future research are addressed.

References

- ACSM. ACSM position stand on the recommended quantity and quality of exercise for developing and maintaining cardiorespiratory and muscular fitness, and flexibility in adults. Med Sci Sports Exerc 1998;30:975-91.
- Andersson GB. Epidemiological features of chronic low back pain. Lancet 1999;354:581-5.
- Deyo RA, Rainville J, Kent DL. What can the history and physical examination tell us about low back pain? JAMA 1992;268:760-5.
- Faas A, Chavannes AW, Koes BW, van den Hoogen JMM, Mens JMA, Smeele LJM, Romeijnders ACM, van der Laan JR. NHG-standaard lage rugpijn. Huisarts en Wetenschap 1996;39:18-31.
- Fordyce WE. Behavioral methods for chronic pain and illness. St. Louis: Mosby, 1976.
- Frymoyer JW. Back pain and sciatica. N Engl J Med 1988;318:291-300.
- Goossens ME, Rutten-van Molken MP, Kole-Snijders AM, Vlaeyen JW, van Breukelen G, Leidl R. Health economic assessment of behavioural rehabilitation in chronic low back pain: a randomised clinical trial. Health Econ 1998;7:39-51.
- Goossens ME, Rutten-van Molken MP, Leidl RM, Bos SG, Vlaeyen JW, Teeken-Gruben NJ. Cognitive-educational treatment of fibromyalgia: a randomized clinical trial. II. Economic evaluation. J Rheumatol 1996;23:1246-54.
- Guzman J, Esmail R, Karjalainen K, Malmivaara A, Irvin E, Bombardier C. Multidisciplinary bio-psycho-social rehabilitation for chronic low back pain. Cochrane Database Syst Rev 2002:CD000963.
- Haskell WL. Dose-response issues from a biological perspective. In: C Bouchard, RJ Shephard, T Stephens. Physical activity, fitness, and health: international proceedings and consensus statement. Champaign: Human Kinetics, 1994, pp. 1030-9.
- Hayden JA, van Tulder MW, Malmivaara AV, Koes BW. Meta-analysis: exercise therapy for nonspecific low back pain. Ann Intern Med 2005a;142:765-75.
- Hayden JA, van Tulder MW, Tomlinson G. Systematic review: strategies for using exercise therapy to improve outcomes in chronic low back pain. Ann Intern Med 2005b;142:776-85.
- Hilde G, Bo K. Effect of exercise in the treatment of chronic low back pain: a systematic review, emphasising type and dose of exercise. Phys Ther Rev 1998;3:107-17.
- Kendall NAS, Linton SJ, Main CJ. Guide to assessing psychosocial yellow flags in acute low back pain: Risk factors for long-term disability and work loss. Wellington: AAC, 1997.
- Lee CE, Simmonds MJ, Novy DM, Jones S. Self-reports and clinician-measured physical function among patients with low back pain: a comparison. Arch Phys Med Rehabil 2001;82:227-31.
- Liddle SD, Baxter GD, Gracey JH. Exercise and chronic low back pain: what works? Pain 2004;107:176-90.
- Mayer TG, Polatin PB, Gatchel RJ. Functional restoration and other rehabilitation approaches to chronic musculoskeletal pain disability syndromes. Crit Rev Physical Rehab Med 1998;10:209-21.
- McArdle W, Catch FI, Katch VL. Training for anaerobic and aerobic power. In: W McArdle, FI Catch, VL Katch. Exercise physiology. Baltimore: Williams & Wilkins, 1996, pp. 393-415.
- Meerding WJ, Bonneux L, Polder JJ, Koopmanschap MA, van der Maas PJ. Demographic and epidemiological determinants of healthcare costs in Netherlands: cost of illness study. BMJ 1998;317:111-5.
- Morley S, Eccleston C, Williams A. Systematic review and meta-analysis of randomized controlled trials of cognitive behaviour therapy and behaviour therapy for chronic pain in adults, excluding headache. Pain 1999;80:1-13.
- Ostelo RW, Tulder M, Vlaeyen J, Linton S, Morley S, Assendelft W. Behavioural treatment for chronic low-back pain. Cochrane Database Syst Rev 2005:Cd002014.

References

- Picavet HS, Schouten JS. Musculoskeletal pain in The Netherlands: prevalences, consequences and risk groups, the DMC(3) study. Pain 2003;102:167-78.
- Simmonds MJ, Olson SL, Jones S, Hussein T, Lee CE, Novy D, Radwan H. Psychometric characteristics and clinical usefulness of physical performance tests in patients with low back pain. Spine 1998;23:2412-21.
- Turk DC, Meichenbaum D, Genest M. Pain and behavioral medicine: A cognitive behavioral perspective. New York: Guilford Press, 1983.
- Turk DC, Okifuji A. Psychological factors in chronic pain: evolution and revolution. J Consult Clin Psychol 2002;70:678-90.
- van der Heijden GJ, Beurskens AJ, Koes BW, Assendelft WJ, de Vet HC, Bouter LM. The efficacy of traction for back and neck pain: a systematic, blinded review of randomized clinical trial methods. Phys Ther 1995;75:93-104.
- van Tulder M, Malmivaara A, Esmail R, Koes B. Exercise therapy for low back pain. A systematic review within the framework of the Cochrane Collaboration Back Review Group. Spine 2000;25:2784-96.
- van Tulder MW, Koes BW, Bouter LM. A cost-of-illness study of back pain in The Netherlands. Pain 1995;62:233-40.
- Verbunt JA, Seelen HA, Vlaeyen JWS, van der Heijden GJ, Heuts PH, Pons K, Knottnerus JA. Disuse and deconditioning in chronic low back pain: concepts and hypothesis on contributing mechanisms. Eur J Pain 2003;7:9-21.
- Verbunt JA, Westerterp KR, van der Heijden GJ, Seelen HA, Vlaeyen JW, Knottnerus JA. Physical activity in daily life in patients with chronic low back pain. Arch Phys Med Rehabil 2001;82:726-30.
- Vlaeyen JWS, Kole-Snijders AM, Boeren RG, van Eek H. Fear of movement/(re)injury in chronic low back pain and its relation to behavioral performance. Pain 1995;62:363-72.
- Vlaeyen JWS, Morley S. Cognitive-behavioral treatments for chronic pain: what works for whom? Clin J Pain 2005;21:1-8.
- Waddell G. A new clinical model of low back pain and disability. The back pain revolution. London: Churchill Livingstone, 1998.
- Wittink H. Functional capacity testing in patients with chronic pain. Clin J Pain 2005;21:197-9.

Chapter 2

The association of physical deconditioning and chronic low back pain:
A hypothesis-oriented systematic review.

Rob J.E.M. Smeets, Derick Wade, Alita Hidding, Peter J.C.M. van Leeuwen, Johan W.S. Vlaeyen, J. André Knottnerus

Published as:
Smeets RJEM, Wade D, Hidding A, van Leeuwen PJCM, Vlaeyen JWS, Knottnerus JA.
The association of physical deconditioning and chronic low back pain: A hypothesis-oriented systematic review. Disability and Rehabilitation 2006;28:673-693.

Abstract

Purpose. Does physical deconditioning (loss of cardiovascular capacity and strength/endurance of paraspinal muscles) exist in patients with chronic low back pain (CLBP) and are treatments specifically aimed to reduce these signs effective?
Method. Systematic literature search in PUBMED, MEDLINE, EMBASE and PsycINFO until December 2004 to identify observational studies regarding deconditioning signs and high quality RCTs regarding the effectiveness of cardiovascular and/or muscle strengthening exercises. Internal validity of the RCTs was assessed by using a checklist of nine methodology criteria in accordance with the Cochrane Collaboration.
Results. There is conflicting evidence that cardiovascular deconditioning is present in CLBP and limited evidence for wasting of the multifidus muscle. No study examined the effectiveness of cardiovascular training specifically. General and lumbar muscle strengthening are equally effective as other active treatments. Only moderate evidence is available for the effectiveness of intensive low back extensor muscle strengthening compared to less intensive strengthening.
Conclusion. Probably reactivation caused by active treatment and not the reconditioning itself is the important factor in the reduction of disability. Further prospective and evaluative research into the role of physical deconditioning is necessary. It seems more promising to further explore the interplay between biological, social and psychological factors.

Introduction

Chronic low back pain (CLBP) is a common health complaint of which the societal burden is enormous (Frymoyer 1988; Meerding et al. 1998; van Tulder et al. 1995). People with low back pain frequently limit their work and leisure time activity. There is circumstantial evidence that the social and personal context of the person influences at least the medical presentation with back pain and also the nature and extent of activity limitation (Abenhaim et al. 2000; Flor et al. 1992; Fordyce 1976; International Association for the Study of Pain 1986; Morley et al. 1999; Pincus et al. 2002; Vlaeyen and Linton 2000; Waddell 1998). It is also postulated that the presence of persistent low back pain causes patients to avoid daily activities, which may lead to physical deconditioning, both generally (e.g. loss of cardiovascular capacity) as specifically (e.g. loss of strength and endurance of paraspinal muscles). Such signs of physical deconditioning may result in even more pain and disability and so contribute to the chronicity of low back pain (Jette and Jette 1996; Mayer et al. 1998). This pattern of signs and symptoms is called the 'deconditioning syndrome' and was first described in the mid 1980s as a factor contributing to the intolerance to physical activities and subsequent further loss of function and disability in patients with CLBP.
Over the last decade much research has been done on the specific role of the low back muscles, especially the paraspinal and in particular the multifidus muscle, in the segmental stability of the lower spine as an important factor in the persistence of low back pain. When the stability at the segmental level is not sufficient, even little disturbances at this level might result in high forces at the region of the discus and surrounding tissues. This results in repetitive high strains and possibly re-injury of these tissues and eventually causing pain (Danneels et al. 2001a; Panjabi 1992a; Panjabi 1992b). It has been suggested that specific training of this paraspinal muscles is worth the effort in nonspecific CLBP (O'Sullivan et al. 1997; Richardson et al. 1999; Richardson et al. 2002).

In order to find out whether there is enough evidence for the existence of physical deconditioning and its role in the persistence of low back pain disability, this paper reviews the available evidence critically.
This review is particularly interested in features of physical deconditioning that are generally considered to be the most important, being loss of cardiovascular fitness and loss of paraspinal muscle strength and endurance (Mayer and Gatchel 1988; Verbunt et al. 2003).
Furthermore, this review wants to study whether there is evidence that treatments specifically aiming to increase the cardiovascular capacity or the strength/endurance of the paraspinal muscles are effective.

The hypotheses being tested in this review are:
People with nonspecific CLBP:

- have cardiovascular deconditioning,
- have 'deconditioning' of their back musculature, comprising weakness and/or a lack of endurance,

- will respond to a treatment that increases cardiovascular capacity by having less pain and/or less activity limitation, and
- will respond to a treatment that reverses local back muscle deconditioning by having less pain and/or less activity limitation.

Methods

For his review the following definitions were used:

- Nonspecific CLBP is any pain felt in the lower back, or lower back and leg, present for more than 3 months without any particular cause such as lumbar disc herniation with neurological symptoms, inflammatory or neoplastic disease, fracture of the spine and spondylolisthesis or spondylolysis (AAC Committee 1997; Frymoyer 1988).
- Deconditioning is a general term that encompasses several features, of which the most important ones are:
    - Reduced cardiovascular capacity which can be measured by:
        - VO_2max (directly or estimated), or work capacity at a certain percentage of heart rate.
    - Muscle weakness defined as less force generated in a short burst, and it can be measured by:
        - Reduced maximum force produced.
        - Reduced muscle bulk or increased fatty infiltration. The quantification of wasting of muscle can be done by using CT scanning or MR-imaging.
    - Reduced muscular endurance defined as less work delivered in a fixed time. It can be measured by:
        - Reduced power produced.
        - Change in muscle fibre types and the quantification of the fibre type proportions, which can be done by muscle biopsy.

All studies on muscle strength (maximum force or power produced) were excluded from further review because major methodological difficulties exist in knowing whether the observed force or power actually represents the maximum (or even a consistent proportion of the maximum) force or power the muscle can produce. Voluntary muscle strength is influenced by many factors such as motivation, fear that the action will cause pain or damage and emotional state (Crombez et al. 1996; Crossman et al. 2004; Gibbons et al. 1997a; Gibbons et al. 1997b; Kankaanpaa et al. 1998; Keller et al. 1999; Newton et al. 1993; Simmonds et al. 1998; Watson 1999). Also EMG-studies were excluded because reliable and valid assessment of back muscle weakness and muscle composition does not seem to be possible (Crossman et al. 2004; Lariviere et al. 2002).

Systematic reviews were identified in the Cochrane Collaboration database (up to December 2004). Observational studies were identified through systematic searches of PUBMED, MEDLINE and EMBASE (from the beginning of each data base up to December 2004) using the terms: low back pain, back pain, deconditioning, disuse, physical fitness, aerobic fitness, cardiovascular fitness/capacity, VO_2max, loss of muscle strength, muscle atrophy, muscular endurance, neuromuscular changes. One person (RS) screened all the titles, read the abstracts of possibly relevant articles and read the full text of all putative relevant articles. Recent books and theses which discuss deconditioning theory and evolving treatments were reviewed (Danneels 2001; McArdle et al. 1996; Richardson et al. 1999; Waddell 1998), and their reference lists and the reference lists of all selected articles were checked for additional relevant studies.

Relevant RCTs of interventions were identified through systematic searches of PUBMED, MEDLINE, EMBASE and PsycINFO (from the beginning of each data base up to December 2004) using the terms: randomized controlled trial, clinical trial, low back pain, back pain, exercise(s), aerobic exercise, strengthening exercise, strengthening, aerobic or cardiovascular fitness, reconditioning, physical therapy, physical exercise, stabilization/stabilizing exercises. The specific searches are available from the first author. Abstract books of relevant congresses were checked (e.g. ISPRM 2001, Low Back Pain Forum 2001, IASP World Congresses 1999, 2002) for relevant studies, and leading Dutch and Belgium experts were asked whether any relevant publication was missing. The reference lists of all articles and systematic reviews were checked for RCTs that might be relevant. Finally, for the already selected studies, the function 'related articles' in PUBMED was used for an additional check.

The methodological quality of all RCTs was assessed using the van Tulder criteria (2000b). Within this method, the nine criteria reflecting the internal validity of the RCTs were used (see table 1). The total of positive scores ranges from 0 to 9 and a RCT scoring five or more points is considered to be of sufficiently high quality. Where the quality had already been assessed within the review by van Tulder et al. (2000b), that review's measurement was used, and in other cases each paper was independently reviewed by two authors (RS and PvL), and any disagreement was discussed and a consensus reached. Only RCTs scoring five or more points were included in this review.

A rating system was used to summarize the strength of scientific evidence consisting of five levels:

- Strong evidence: generally consistent findings among more than two high quality RCTs.
- Moderate evidence: generally consistent findings among two high quality RCTs.
- Limited evidence: only findings of one high quality RCT.
- Conflicting evidence: inconsistent findings among multiple RCTs.
- No evidence: no high quality RCTs available.

Table 1: Quality assessment of randomized controlled trials

Study	Concealment of allocation	Withdrawal/ drop-out rate	Co-intervention avoided/ equal	Blinding of patients	Blinding of observer	Intention-to-treat analysis	Compliance	Similarity baseline characteristics	Blinding care-provider	Total score (max. 9)
Bendix (2000)[a]	+	-	+	-	+	-	+	+	-	5
Brontfort (1996)[c]	+	-	+	+	+	+	+	+	-	7
Danneels (2001; 2001a; 2001b)[a]	+	-	+	-	+	-	+	+	-	5
Frost (1995; 1998)[b]	+	+	+	-	+	-	+	+	-	6
Hansen (1993)[c]	+	+	+	-	+	-	-	+	-	5
Helmhout (2004)[a]	+	-	+	-	+	-	+	+	-	5
Kankaanpää (1999)[a]	+	+	+	-	-	-	+	+	-	5
Ljunggren (1997)[c]	+	+	-	-	-	+	+	+	-	5
Manniche (1988; 1991)[c]	+	+	+	-	+	+	-	+	-	6
Mannion (1999)[a]	+	+	+	-	+	+	+	+	-	7
Moffett (1999)[a]	+	+	+	-	+	+	+	+	-	7
Petersen (2002)[a]	+	-	+	-	-	+	+	+	-	5
Torstensen (1998)[b]	+	+	+	-	+	+	+	+	-	7
Vollenbroek (2004)[a]	+	+	+	-	+	+	-	+	-	6

[a] = studies not included in the Cochrane review on exercise therapy; [b] = studies included in the Cochrane review on exercise therapy but not on strengthening exercises; [c] = studies included in the Cochrane review on exercise therapy and strengthening exercises.

Results

The results will be presented for each hypothesis.

Cardiovascular deconditioning – observational studies.
Nine studies in which at least 50% of the patients had CLBP were identified (see table 2). Two studies concerned mainly patients without CLBP (Battie et al. 1989; Kellett et al. 1991). There are great differences in the measurement and calculation of the cardiovascular capacity and sometimes no adjustments for gender are made. When no appropriate controls are available, less relevant normative data are usually used for comparison. Since the identified studies on cardiovascular deconditioning are all cross-sectional, they do not provide any evidence that patients with CLBP develop cardiovascular deconditioning after the onset of low back pain. And the evidence that patients with CLBP do have clinically significant cardiovascular deconditioning relative to age-matched controls is itself conflicting. There is limited evidence that people with CLBP who are in work are fitter than those not in work (Verbunt et al. 2003).

Low back muscle deconditioning – observational studies.
Two variables have been studied: (1) muscle bulk using CT-scanning or MR-imaging to quantify the cross-sectional area and amount of fat (density) of paraspinal muscles, and (2) muscle endurance using muscle biopsies to identify changes in muscle fibre types and size.

The reliability of CT-scanning or MR-imaging for the measurement of muscle bulk is acceptable (Keller et al. 2003). Studies on muscle bulk are shown in table 3. The evidence suggests that people with CLBP may have muscle wasting and some fatty infiltration of paraspinal muscles in the back but often also of the psoas muscle in both operated and non-operated CLBP patients. This might to be caused, at least partly, by inactivity/disuse and aging. One study on patients with unilateral back pain shows clear atrophy of both the multifidus and psoas muscle at the symptomatic side suggesting that not only disuse but also neuromuscular changes play a role (Barker et al. 2004). However, there are inconsistencies between studies. When one corrects for age, gender, body mass index and activity level there is limited evidence that once LBP originates, atrophy of the multifidus muscle might play a role in the complex of chronic low back pain and its resulting disability.

Table 2: Study characteristics of observational studies concerning cardiovascular capacity

Author	Patient population	Controls	Exercise test procedure	Compared to controls	Notes
Schmidt (1985)	39 male NSLBP patients (> 6 months), referred to orthopaedic clinic, 25-55 yr, duration of complaints/level of disability/work status: ?	39 males from general population, age matched	Symptom limited treadmill test	Lower total testing time, equal heart rate and respiratory quotient	
Battie (1989)	690 aircraft plant workers, current or previous LBP < 10 yr, 80% male, 36 (21-67) yr, 10 months complaints, level of disability (Os): 29.7 (± 15.18) for 124 patients; 100 % working	1744 workers without LBP	Submaximal treadmill test	No difference in VO_2max	% CLBP patients very low or zero
Hurri (1989; 1991)	245 blue collar workers, CLBP or ≥ 2 yr recurrent LBP and sick leave < 2 yr, 71 % male, 44 yr and 46 yr, time since first symptoms of LBP: 13.3 yr and 9.9 yr, level of disability (LBPDI): 15.6 and 19.3, 100% working	Normative Czechoslovakia population (1968-1974)	Symptom limited maximal graded bicycle test	No difference in VO_2max	
Kellett (1991)	111 blue collar workers, current or previous LBP and < 50 days of sick leave < 1½ yr, ? % male, 41.5 (25-62) yr, duration of complaints/level of disability/work status: ?	Normative Swedish population (1960)	Submaximal bicycle test (Åstrand-Rhyming)	Average to above-average VO_2max	% CLBP patients very low or zero
Nielens (1991)	56 patients, 64% lumbo-radicular pain, 16% LBP ≥ 6 months, 48% male, referred to pain clinic, age: 42.9 (± 10) yr and 47.2 (± 13) yr, duration of complaints/level of disability: ?, 26% and 17% working	42 healthy members of medical staff, 50 % male	Discontinuous submaximal step test	PWC 65%/kg lower for men, no difference for women	Healthy controls: mean age 5 yr lower
Davis (1992)	46 patients, predominant CLBP, 59% male, referred to pain and rehabilitation centre, age: 38.4 (± 8.7) yr, duration of complaints: 41.2 (± 48) months, level of disability/work status: ?	Normative Brazilian working population (1986)	Symptom limited bicycle test	VO_2max observed/ VO_2max level: equal to sedentary individuals	VO_2max is not presented for gender, % back pain unknown
Nielens (1994)	42 patients, 57% lumbo-radicular pain, 14% LBP ≥ 1 year, 55% male, referred to pain clinic, age: 41 (± 8.2) yr, duration of complaints/level of disability/work status: ?	34 healthy members of medical staff, 50% male	Submaximal bicycle test	PWC 65%/kg lower for men, no difference for women	
Van der Velde (2000)	285 CLBP patients, 21% also neck or thoracic pain, 50% male, referred for fitness treatment, age: 34.2 (20-49) yr, 10 months complaints, level of disability (Os): 29.7 (± 15.18) for 124 patients, ? working	I. 1001 persons starting fitness program (29.1 yr, 41% male) II. Normative Canadian population (1981)	Submaximal step test (Canadian Aerobic Fitness test)	Lower versus I and II (CLBP = 19.64 mean percentile rank)	
Wittink (2000a)	50 CLBP patients, 44% male, referred to pain clinic, age: 39.8 (± 8.3) yr, 40.2 (± 50) months complaints, level of disability: ?, 36% and 54% working	Normative U.S. population (1973)	Symptom limited modified Bruce treadmill test	Men: level of healthy sedentary Women: level of healthy active	
Wittink (2000b)	30 CLBP patients, 43% radicular pain, 26% radiculopathy, 31% LBP, 47% male, referred to physical therapy department, age: 40.4 (± 9.2) yr, 19.1 (± 15.1) and 40.3 (± 42.7) months complaints, level of disability: ?, 29% and 56% working	Normative U.S. population grouped by decade (1996)	I. Symptom limited modified Bruce treadmill test II. Symptom limited Åstrand-Ryhming bicycle test III. Symptom limited upper extremity ergometer test	Sample size too small in each decade for firm conclusions, no differences except patients > 50 yr less fit	
Nielens (2001)	55 patients, 54% CLBP, 24% fibromyalgia syndrome, 22% miscellaneous (> 1 yr complaints), 36% male, referred to chronic pain clinic, age: 44 (± 9.3) yr, duration of complaints/level of disability: ?, 20% and 34% working	Normative Belgium population (1994)	Submaximal bicycle test	PWC 65 %/kg lower; 35% (men), 17% (women)	Healthy controls: mean age 5 yr lower

CLBP = chronic low back pain; LBP = low back pain; LBPDI= low back pain disability index; NSLBP = nonspecific low back pain; Os = Oswestry questionnaire; PWC 65% = work capacity at 65% of heart rate.

Studies on fibre types and size are shown in table 4. The following muscle fibres can be identified; type I (slow oxidative), IIA (fast-twitch oxidative glycolytic), IIX (fast-twitch glycolytic) and IIC (intermediate). The results are inconclusive. Several studies suggest the existence of an atrophy of the type II fibres in CLBP, but others state that the type IIX fibres are generally smaller even in healthy back muscles and decrease with growing age and so the observed atrophy may simply resemble an atrophy of this type IIX fibre. Several studies show that the longer the duration of LBP the more fibre-type transformation seems to occur (increased IIX:IIA ratio and decreased I:IIX ratio). But one study with healthy controls matched for age and anthropometry does not show this transformation, although no further distinction in fibre type II is made (Crossman et al. 2004). No study has yet been able to convincingly show significant differences in the ratio of the size of the type I:IIX fibre between patients and matched controls.

Cardiovascular deconditioning – intervention studies.

Cardiovascular training in healthy people refers to exercise with large muscle groups with a duration of 20 to 60 min of continuous exercise performed at an intensity of 55/65% to 90% of maximum heart rate, three to five times a week (ACSM 1998; McArdle et al. 1996). Lower-intensity activity should be conducted over a longer period of time (30 min or more), and individuals training at higher levels of intensity should train for at least 20 min.

The review undertaken by the Paris Task Force on Back Pain reports no study in which only cardiovascular training has been evaluated (Abenhaim et al. 2000). In all the studies shown in table 5, cardiovascular training makes up a significant part of a multi-focal treatment. Studies in which intensive functional restoration programmes are evaluated were not included in our review because the cardiovascular training is only a small component of a large treatment package. Due to lack of information about the exact content (intensity, frequency, etc.) of the cardiovascular training it is often impossible to evaluate whether the training can be beneficial. Furthermore, no information is available whether the cardiovascular capacity improves as a result of the training except in the study of Vollenbroek et al. (2004). But in this study the care as usual patients also report less disability without improving their VO_2max. The general conclusion is that there is no evidence that cardiovascular training as given in the selected studies has an additional effect and the reported benefits may arise from the general fitness training and/or cognitive-behavioral therapy. Only one study with a more intense specific cardiovascular and fitness training programme of nearly 5 h/week during 8 weeks shows some effect on daily activities and work activities (Bendix et al. 2000).
Because no study has been identified in which the effectiveness of pure cardiovascular endurance training is evaluated, no conclusion can be drawn whether this specific treatment is worth the effort when compared to other active or passive treatments.

Table 3: Study characteristics of observational studies concerning muscle weakness (wasting of muscles)

Author	Patient population	Controls	Research method	Compared to controls	Notes
Mayer (1989)	46 patients 3 months after spinal surgery for degenerative disease, 76% male, age: 39 (± 14) yr and 41 (± 16) yr, duration of complaints: ?	19 persons without LBP or spinal surgery, 58% male	CSA and density of erector spinae, psoas muscle (single cut CT-scan lower edge of L3 lamina)	Decrease in muscle density in erector spina and psoas, only trend towards decreased CSA	No normalisation of CSA, not matched for activity level, height and weight
Cooper (1992)	44 CLBP patients (> 18 months), 70 % lumbar surgery, 59% male, age: 44 (28-54) yr and 44 (20-53) yr, 7 (2-20) yr and 8 (2-26) yr complaints	43 LBP patients, 53% male, age: 36 (23-54) yr and 38 (22-50) yr, 0.8 (0.3-1.4) yr and 0.5 (0.2-1.3) yr complaints	CSA of L4 vertebra, paraspinal muscles (erector spinae plus multifidus), psoas, the paraspinal and psoas/L4 CSA ratios	Reduction of muscle dimensions in patients with CLBP (for both sexes)	Many CLBP patients had lumbar surgery, were older and not matched for activity level
Alaranta (1993)	39 CLBP patients (41 % lumbar disc surgery > 1 yr before examination), 41% male, age: 43.8 (± 6.0) yr and 44.3 (± 5.3) yr, duration of complaints: ?	No controls	Amount of fat in lumbar extensor muscles (CT-scan at 3 lowest lumbar intervertebral levels)	Disability associated with more muscle fat content at L5-S1 level in men but not women	No control group, no information about activity level
Hultman (1993)	20 male CLBP patients (> 3 years, no surgery, > 3 months sick leave < 1 yr), age: 45-55 yr, duration of complaints: ?	I. 18 healthy men II. 35 men with intermittent LBP	A. CSA of erector spinae muscle (CT-scan at L3 level) B. Radiological density	A. No difference B. More fat infiltration in CLBP-group	No normalisation of CSA, not matched for activity level
Parkkola (1993)	48 CLBP patients (I), 10 patients with more serious CLBP (II), 53% and 40% male, age: 30-47 yr, duration of complaints: ?, level of disability (MI): 45 and 45 (I), 46 and 53 (2), > 3 months sick leave for LBP < 1 yr	60 healthy persons matched for gender, age, employment and profession	CSA, amount of fat of erector spinae and multifidus, psoas major (MR-imaging at L4-L5 level)	CSA's smaller in I and II, II smaller CSA than I and controls (no statistical data presented), higher amount of fat for I and II (12%) versus controls (9%)	No normalisation of CSA, not matched for activity level, height and weight
Sihvonen (1993)	7 FBS patients (2-5 yr post-surgery) unable to resume work, 29% male, age/duration of complaints: ?	12 succesful spinal surgery (2-5 yr later) 57% male	Density of back muscle (unclear what muscles were included) at operated (L4-L5) area (CT-image)	Lower density in FBS patients	Not matched for activity level, height and weight
Gibbons (1997a)	85 male LBP-patients of varying frequency < 1 yr, recruited; age: 35-63 yr, duration of complaints: ?	43 men without LBP < 1 yr	I. CSA of erector spinae (including the multifidus) and total paraspinal muscles; MR-imaging at L3-L4 level II. T2-weighted and proton weighted density of paraspinal muscles	Patients reporting LBP ≥ once a month: smaller CSA and greater signal intensities (meaning higher content of water or fat), CSA and density were no predictors for future LBP	No normalisation of CSA, not matched for activity level, heigt and weight
Kader (2000)	78 mechanical LBP patients (> 3 months), ?% male, age: 17-72 yr, duration of complaints: ?	None	Visual assessment of atrophy of multifidus muscle (MR-imaging)	80% macroscopic degeneration (89% bilaterally)	No control group, unknown activity level
Danneels (2000)	32 CMLBP patients (> 1 yr) with or without disc protrusion, no surgery, no sports or fitness training < 3 months, 53% male, age: 25-55 yr, 9.16 (± 7.41) yr duration	23 healthy persons matched for age, height, weight, activity level, 56% male, age: 25-55 yr	CSA and amount of fat deposits of total paraspinal muscles, multifidus, psoas (CT-images at upper endplate L3, upper, lower endplate L4) related to CSA of vertebral body	Lower CSA of multifidus at lower end plate L4	
Barker (2004)	48 CULBP patients (> 12 weeks, no surgery), 54% male, age: 44.3 (± 11.8) yr, 15.6 (± 5.8) weeks complaints, level of disability: (Os): 38.4 (± 14.7)	Asymptomatic side of each patient	CSA of multifidus, psoas muscle at clinically indicated spinal segment level and one level above and below, comparison between symptomatic and asymptomatic sides	Reduced CSA of multifidus and psoas at symptomatic side at the clinically indicated segment level (22%, 12%); in levels above (16%, 17%), below (3%, 9%)	No calculation of fat deposits

CLBP = chronic low back pain; CMLBP = chronic mechanical low back pain; CULBP = chronic unilateral low back pain; FBS = failed back surgery; LBP = low back pain; MI = Million index; Os = Oswestry questionnaire; CSA = cross sectional area.

Table 4: Study characteristics of observational studies concerning loss of muscular endurance (changes of fibres types)

Author	Patient population	Controls	Muscle biopsy	Results	Notes
Mattila (1986)	41 LDHS patients, 54% male, age: 41.1 (± 8.1) yr and 43.5 (± 8.0) yr, duration of complaints: ?, sick leave: 67.6 (± 61.5) days and 50.6 (± 50.1) days	12 cadavers without LBP (mean age 40.1 yr, 75% male)	Multifidus at L4-L5 and L5-S1 intervertebral levels during surgery	Selective type II-fibres atrophy, change in internal structure of type I-fibres, type II-fibres in controls also atrophic	No further distinction in type II-fibres
Zhu (1989)	22 LDHS patients, 59% male, age: 37.4 (24-57) yr, 51 (2-360) months complaints	Normative values from cadavers	Erector spinae: side of disc protrusion 1 cm lateral to top of spinous processus during surgery	Angulated, selective atrophy of type II-fibres, higher type IIX/IIA ratio, type IIX-fibres smaller in longer duration group	Controls: no further distinction in type II-fibres, no correction for age, gender and body size
Rantanen (1993)	18 LDHS patients 5 yr post-surgery, 56% male, age: 42.7 (25-53) yr, duration of complaints: ?, categorized by WHOOH in positive (N=10) and negative (N=8) outcome group	Same 18 patients (biopsies at the time of LDHS)	Semi-open biopsy 5 years post-surgery at level of the spinous process of the affected disc; Multifidus at level of intervertebral discs L4-L5, L5-S1 during surgery	No changes in fibre, atrophy and hypertrophy, increased diameter of type I-fibres. Positive group: increased diameter of type II-fibres, also slight decrease abnormalities in type I-fibres but increase in negative group	No further distinction in type II-fibres, no comparison of the preoperative biopsies with normal population
Weber (1997)	I. 43 patients with posterior spondylosyndesis, II. 32 patients with removal of internal fixation, 43 % male, age: 41.2 (14-74) yr, duration of complaints: ?	No controls	Multifidus muscle bilaterally during surgery	Increased proportion of type I-fibres with increasing age, type II atrophy, changes in internal structure type I-fibres. More type II atrophy in patients with higher age and more pain	No comparison of preoperative biopsies with normal population
Mannion (1997)	21 spinal surgery patients, 57% male, age: 22.1 (± 7.8) yr and 35.9 (± 8.5) yr, < 1 yr (24%), 1-3 yr (38%), > 3 yr (38%) complaints	21 healthy persons matched for age, gender, body mass, age: 22.9 (± 5.1) yr and 33.3 (± 11.0) yr	Patients: superficial multifidus muscle, L3-L4 bilaterally during surgery; controls: percutaneous biopsy belly lateral tract left erector spinae at L3 vertebra	Lower proportion of type I, higher proportion of type IIX-fibres in patients, no difference in fibre diameter	Different side of biopsy for patients and controls, not matched for activity level
Mannion (2000)	59 CLBP patients, 49% male, age: 42.5 (± 11.2) yr and 45.1 (± 10.7) yr, 170 (± 114) months and 144 (± 133) months duration, level of disability (RDQ): 6.1 (± 4.8) and 10.1 (± 4.5)	None	I. Percutaneous belly of lateral tract of left erector spinae at L3-L4 vertebra; II. CSA of erector spinae, quadratus lumborum, psoas by MR-imaging at the centre of L3-L4 and L4-L5 discs	Correcting for age, gender and body size: negative correlation between duration of LBP and proportion of type I-fibres, highly positive correlation with proportion of type IIX- fibres, no correlation with mean size of muscles	No biopsy and CSA of the multifidus, no normalisation of CSA
Crossman (2004)	35 male NSCLP patients, age: 41 (± 11) yr, 6.79 (± 4.73) yr complaints, level of disability (VK): low (N=29), high (N=6)	32 healthy male (no LBP < 1 yr), matched for age and anthropometry	Not described properly: percutaneous, belly of the left erector spinae at T10 and L3?	No differences in fibre types in percent number, mean fibre diameter and relative area of muscle sample occupied by type I-fibres	No biopsy of the multifidus, no further distinction in type II-fibres, mainly low disabled patients

CLBP = chronic low back pain; CSA = cross sectional area; LDHS = lumbar disc herniation surgery; NSCLBP= nonspecific chronic low back pain; RDQ= Roland disability questionnaire; VK = Von Korff chronic pain grade; WHOOH = WHO occupational handicap.

Table 5: Study characteristics of RCTs concerning cardiovascular training

Author	Patient population	Treatment	Outcomes	Results	Notes
Bendix (2000)[a]	138 NSLBP patients with precarious work situation, referred by physicians or insurance companies, exclusion: receiving a disability pension; age: I 40 (interquartile range 35-48) yr, II 41 (35-47) yr, duration of complaints: ?, sick leave days < 3 yr: I 200 (interquartile range 52-432) days , II 220 (43-440) days	I. Aerobic and strengthening exercises, ergonomics, back school, recreation activities, pain management, coping, relaxation, visualization, behavioral support, 39 hr/week, 3 weeks, 3 FU-treatment days in 2 months, N=64 II. Aerobic and strengthening exercises (submaximal contractions, 30 repetitions), 3 1½ hr sessions/week, 8 weeks, N=74	Work capability, sick leave, health care contacts, back and leg pain, self-reported activities of daily living, overall assessment	1 yr (N=99, who completed treatment): I better overall assessment; II better work capability; activities of daily living improved in both groups, no change in back and leg pain	High dropout rate (39/138), no intention to treat analysis, blinding of observer (80%), no control group
Frost (1995; 1998)[b]	81 SCLBP patients, referred by orthopaedic surgeons to PT department, exclusion: SPLBP, musculoskeletal disabilities hampering participation, major surgery (< 1 yr), PT (< 3 months), sporting activities (< 6 months), pregnancy; age: I 34.2 (± 9.4) yr, II 38.5 (± 9.3) yr, duration of complaints; I. 26.3 (± 2.4) months, II. 18.7 (± 15.4), level of disability (Os): I 23.6 (± 9.7), II 23.6 (± 12.3)	I. 4 home-exercises (twice daily) and back-school (2 sessions), fitness program: 15 progressive exercises for large muscle groups, light aerobic exercise, CBT-behavioral components, 8 1 hr sessions in 4 weeks, N=36 II. 4 home-exercises (twice daily), back-school (2 sessions), N=35	VAS-pain, Os, PSEQ PLCQ, GHQ, shuttle walking test	Positive change in VAS-pain, Os (6 vs. 1.9), PSEQ, walking distance in I, and GHQ in both groups; 6 months: I improvement Os (8 vs. – 0.7); 24 months: I improvement in Os (7.7 vs. 2.4)	71 of 81 patients in analysis, intention-to-treat only for the 6 and 24 months FU-data, no real control group
Mannion (1999)[a]	148 continuous or recurrent LBP patients (> 3 months), absenteeism from work or seeking medical attention, recruited by local media, exclusion: constant or persistent severe pain, SLBP, pregnancy, spinal surgery, co-morbidity hampering physical training; age: I 46.3 (± 10.1) yr, II 43.7 (± 10.1) yr, III 45.2 (± 9.7) yr, duration of complaints: I 10.0 (± 9.0) yr, II 13.0 (± 10.0) yr, III 9.7 (± 9.1 yr), level of disability (RDQ): I 7.9 (± 4.0), II 8.0 (± 5.1), III 7.7 (± 4.7)	I. Isometric strengthening (Therabands), general strengthening (training devices), ergonomics, home exercises, optional passive therapy, 2 ½ hr individual sessions/week, 3 months, N=49. II. Active rehabilitation (see Kankaapää, table 6): group of 2-3 persons, 2 1 hr session/week, 3 months, N=49 III. Low-impact aerobics: group of maximal 12 persons, 10-20 min. warm-up, 30 min. specific exercises for trunk and leg muscles, gradually increasing, 2 1 hr session/week, 3 months, N=50	VAS-pain, RDQ, FABQ, CSQ, MSPQ, MZDS, BBQ	All groups: reduction in VAS, RDQ (overall 1.3), FABQ-physical activity, less praying/hoping, catastrophizing, no group differences; 6 months: all groups reductions in VAS, RDQ (1.6), FABQ-physical activity, CSQ, except I no reduction in RDQ and FABQ	No control group
Moffett (1999)[a]	187 MLBP patients, referred by general practices, exclusion: see Frost (this table); age: I 41.1 (± 9.2) yr, II 42.6 (± 8.6) yr, duration of complaints:?, level of disability (RDQ): I 6.65 (± 4.02), II 5.56 (± 3.94)	I. Warm up, stretching, 15 progressive exercises large muscle groups, light aerobic exercise, CBT-components, 8 1 hr sessions in 4 weeks, group of maximal 10 patients, N=89 II. Care as usual by general practitioner, N=98	VAS-pain, ABPS, RDQ, EuroQol, Health index (EQ-5D), (FABQ), use of health care services	I decrease in ABPS; 6 months: I decrease in RDQ (2.99 vs. 1.64); 12 months: I decrease in VAS, RDQ (3.19 vs. 1.77), less days of work loss, no difference in cost saving	
Vollenbroek (2004)[a]	163 NSCLBP patients, referred by rehabilitation specialist, exclusion: SLBP, back surgery < 3 months, co-morbidity hampering physical training; age: I 38.5 (± 9.8) yr, II 39.5 (± 9.9) yr, duration of complaints: I 72 months, II 48 months, level of disability (RDQ): I 13.1 (± 4.4), II 12.7 (± 4.5)	I. Each week 3 hours conditional training, 0.5 hours swimming, 1.5 hours OT, 4 hours PT, total 7 weeks, optional psychologist, individual OT and dietician, group of 8 patients, N=79 II. Care as usual during 6 months, after this period they were offered treatment I, N=84	RDQ, EuroQol, isometric leg strength, Åstrand submaximal test (VO_2max), TSK, SCL-90	No differences between groups post-treatment and 6 months; both groups (30 – 50%), improved on RDQ, EuroQol at 6 months; I more change in VO_2max and TSK	142 in intention-to-treat analysis, moderate to severely disabled patients

[a] = studies not included in the Cochrane review on exercise therapy; [b] = studies included in the Cochrane review on exercise therapy; APBS = Aberdeen back pain scale; BBQ = back beliefs questionnaire; CBT = cognitive behavioral therapy; CSQ = coping strategy questionnaire; FABQ = fear avoidance beliefs questionnaire; FU= follow-up; GHQ = general health questionnaire; LBP = low back pain; MLBP = mechanical low back pain; MSPQ = modified somatic perception questionnaire; MZDS = modified Zung-depression scale; NCSLBP = nonspecific (chronic) low back pain; Os = Oswestry questionnaire; OT= occupational therapy; PT = physiotherapy; PLCQ= pain locus of control questionnaire; PSEQ= pain self-efficacy questionnaire; RDQ = Roland disability questionnaire; SCLBP = somatic chronic low back pain; SPLBP = specific low back pain; TENS = transcutaneous electrical nerve stimulation; TSK = Tampa scale for kinesiophobia; SCL-90 = symptom checklist-90; VAS = visual analogue scale.

Low back muscle deconditioning – intervention studies.

The relevant studies are shown in table 6. The Cochrane systematic review on strengthening exercises (van Tulder et al. 2000b), which includes many but not all of the studies tabulated, concludes that there is no evidence to support the use of back muscle strengthening exercises in comparison to other active treatments. According to the theoretical principles of exercise physiology the strengthening exercises should be of sufficient intensity, frequency and duration to achieve improvement of strength and/or endurance (ACSM 1998; Bronfort et al. 1996; Haskell 1994; Hilde and Bo 1998; McArdle et al. 1996; McArdle et al. 2001; van Wingerden 1997). For example, the American College of Sports Medicine (1998) states that muscle strengthening exercises should be done two to three times a week, with at least one set of 8 to 12 repetitions inducing volitional fatigue and the resistance should be progressed over time. Brontfort et al. (1996) calculated that at least 16 h of training is necessary to increase muscle strength. However, in two of the included studies in the Cochrane systematic review the patients probably have had very limited training (Bronfort et al. 1996; Hansen et al. 1993). In another study, not only one but both exercises programmes should be classified as general body strengthening exercises of sufficient intensity and both show reduction of work absenteeism (Ljunggren et al. 1997). Only one study, definitely giving a high intensity programme, especially training the low back extensor muscles shows significant reduction of pain and disability when compared to less intensive treatments (Manniche et al. 1991).

In addition to the studies that had already been included in the Cochrane review on exercise therapy, we included six additional studies (Danneels 2001; Danneels et al. 2001a; Danneels et al. 2001b; Helmhout et al. 2004; Kankaanpaa et al. 1999; Mannion et al. 1999; Petersen et al. 2002; Torstensen et al. 1998). One study, examining a high intensity training program comparable to Manniche shows additional reduction of pain and disability when compared to less intensive treatment (Danneels 2001; Danneels et al. 2001a; Danneels et al. 2001b). A study, using the same exercise protocol as Manniche, although performing only half the amount of sessions, shows no difference between this high intensity programme and McKenzie treatment (Petersen et al. 2002). One study, also of sufficient intensity for specifically training the trunk muscles (back and abdominal muscles) but also using cognitive-behavioral treatment components, shows reduction of pain and disability compared to passive treatment (Kankaanpaa et al. 1999), although this same treatment does not show better results compared to low impact aerobics or modern active physiotherapy (Mannion et al. 1999). Furthermore, another study, using isolated lumbar extensor training of high or low intensity but with a frequency of only 14 sessions in 12 weeks, and so not fulfillling the abovementioned criteria of exercise physiology, clearly shows improvement in back complaints and disability in both groups but not between group differences (Helmhout et al. 2004). Torstensen et al. (1998), using high frequency but low intensity exercises of back, abdomen, legs and arms did not find an additional effect when they compared these exercises to conventional physiotherapy. In conclusion, there is moderate evidence that general muscle strengthening (of the body or trunk) is equally effective as other active treatments. There is moderate evidence that these strengthening exercises are more effective than passive treatments or regular walking. Specific low back muscle strengthening exercises of sufficient intensity and frequency to fulfill the exercise physiology criteria show moderate evidence that they are more effective

Table 6: Study characteristics of RCTs concerning muscle strength or endurance enhancing exercises

	Patient population	Treatment	Outcome	Results	Notes
Brontfort (1996)[b]	174 NSLBP patients (≥ 6 weeks), recruited by advertisement, exclusion: SLBP, organic diseases with referred pain, severe osteopenia, back surgery, co-morbidity hampering physical training, pregnancy, NSAIDs, SMT or PT < 3 months; age: I 41.3 (± 10.5) yr. II 40.3 (± 8.9) yr, III 41.4 (± 9.3) yr, duration of complaints: I 3.0 yr, II 2.0 yr, III 2.3 yr, level of disability (RDQ): 8.2 (± 4.2)* *Recalculated to the normal scoring of 0-24	I. SMT and TSE, N=71; II. SMT and SE, N=51; III. NSAIDs, SMT and TSE, N=52. SMT; 10 sessions of 10-15 min first 5 weeks. NSAID; Naproxen sodium twice a day 500 mg, 5 weeks. TSE; dynamic strengthening of back and leg extensors according to Manniche (see this table), abdominal muscles, first 5 weeks 10 sessions in combination with SMT or NSAID, next 6 weeks 10 1 hr sessions, 20 repetitions, total 20 sessions in 11 weeks. SE: stretching for spine and lower extremities, first 5 weeks 10 sessions in combination with SMT or NSAID, next 6 weeks 10 sessions, twice stretching for 1 min, total 20 sessions in 11 weeks	Low back pain (11-box scale), RDQ, COOP	After 5 and 11 weeks no group differences, clinical important improvement in all groups; pain (50%), RDQ (50%), COOP (17%); 1 yr: no group differences, clinical important improvement in all groups; pain (40%), no data available for RDQ and COOP	RCT not designed to evaluate the individual treatment components, the intensity of TSE is 1/5 of intensity in Manniche, total number of sessions is lower (20 vs. 30 sessions), no control group
Danneels (2001; 2001a; 2001b)[a]	88 CLBP patients referred by a physiatrist, exclusion: spondylolysis/spondylolisthesis, lumbar surgery, lumbar scoliosis > 10º, co-morbidity hampering physical training, fitness training for low back muscles < 3 months; age: I 42.13 (± 12.63) yr, II 43.65 (± 11.35) yr, III 44.84 (± 13.17) yr, duration of complaints: ?, level of disability (Os): I 14.81 (± 11.18), II 11.71 (± 6.13), III 15.05 (± 9.40)	I. Stabilization training (multifidus muscle in co-contraction with transversus abdominis muscle), N= 23. II. Stabilization training plus 3 dynamic (2 sec. concentric, 2 sec. eccentric movement) BSE, N=23. III. Stabilization training plus 3 dynamic-static (2 sec. concentric, 5 sec. static contraction, 2 sec. eccentric ovement) BSE, N=25. All treatments: starting with 10 min. diathermy to low back followed by exercises and massage. II and III: training at 70% of 1-RM, each exercise 3 series of 15-18 repetitions, 3 1 hr sessions/week, 10 weeks	VAS-pain, Os, CSA, amount of fat deposits of total paraspinal muscles, multifidus and psoas (CT-images at upper endplate of L3, upper and lower endplate L4)	All groups: reduction of VAS, Os (9.52, 7.04, 9.96), II and III hypertrophy multifidus, III hypertrophy paraspinal muscles, no correlation between change in-Os and hyperttrophy. 1 yr: II and III reduction of Os. In comparison to baseline: all groups reduction of Os (6.47, 6.96, 10.05)	71 of 88 patients for analysis, no intention to treat analysis, relative low rate of disability, no control group
Hansen (1993)[b]	180 LBP patients (≥ 4 weeks or ≥ 2 episodes per month < 1 yr) recruited by company newspaper, exclusion: SLBP, spondylolisthesis, spinal fusion, neuromuscular disease of the trunk, co-morbidity hampering physical training, pregnancy; age: I 40.5, (21-60) yr, II 39.4 (21-64) yr, III 41.9 (23-59) yr, duration of complaints: ?, level of disability: no one permanently disabled (work): median number of disability days < 1 yr: 4.2 days	I. BSE according to Manniche: 100 repetitions for each exercise, 2 1 hr sessions/week, 4 weeks, N=60. II. PT: manual traction, hot packs, massage, stretching, stabilization exercises and flexibility, coordination, slowly progressive TSE, ergonomics, 2 1 hr sessions/week, 4 weeks, N=59. III. Semihot packs and light traction (10% of body weight), 2 1 hr sessions/week, 4 weeks, N=61	Pain (10-point-scale), overall treatment effect (10-point-scale)	No differences in pain between groups, I - II better overall treatment effect 6 and 12 months: no differences in pain between groups, I - II better overall treatment effect	Total sessions strengthening exercises only 8 sessions in 4 weeks, half of the patients had CLBP, no control group
Helmhout (2004)[a]	81 male NSCLBP Dutch army workers, 86% recruited by advertisement 14% referred by physician, exclusion: spinal surgery < 2 years, severe back pain preventing maximal isometric strength effort, SLBP, spondylolisthesis; age: I 41 (± 10) yr, II 40 (± 9) yr, duration of complaints: < 1 yr; I 12%, II 25 %; 1 - 5 yr: I 39 %, II 35 %; > 5 yr; I 47 %, II 40 %, level of disability (RDQ): I 7.1 (± 4.8), II 7.9 (± 4.8)	I. Progressive resistance training of isolated lumbar extensor muscles, from flexion to extension (2 sec) and back (4 sec), 2 5-10 min sessions/week in first 2 weeks, 1 session/week in next 10 weeks. Initial load 35% of 1-RM, 15-20 repetitions in first 2 weeks; 50 -70 % of 1-RM, 10 -15 repetitions in last 10 weeks, N=41. II. Same as I. only non-progressive with training load ≤ 20% of 1-RM, 15 repetitions in first 2 weeks, 20 repetitions in last 10 weeks, N=40. All treatment groups supervised by one physiotherapist (the same for both groups)	Back complaints RDQ, Os, TSK, SF-36, isometric back extension strength	No differences between groups except I higher mean isometric strength, both showed improvement on all outcome measures up to 6 months FU	No intention to treat analysis, relative low rate of disability with only occasional work absenteeism, no control group
Kankaanpää (1993)[a]	59 NSCLBP patients able to work with occasional absences, referred to occupational health centre, exclusion: SLBP, spondylarthrosis, back surgery, severe scoliosis; age: I 40.7 (± 8.6) yr and 38.9 (± 8.2) yr, II 38.0 (± 6.9) yr and 40.6 (± 8.1) yr, duration of complaints since first LBP episode: I 10.9 (± 8.5) yr and 5.8 (± 6.8) yr. II 6.8 (± 6.1) yr and 7.4 (± 6.7) yr, level of disability (PDI): I 12.3 (± 10.1) and 15.2 (± 10.5), II 10.1 (± 7.6) and 8.5 (± 10.3)	I. Physical exercises (lumbar flexion, extension, lateral flexion and rotation) for muscle function and coordination, low loads first 4 weeks, gradually increase of load and performing within painless range of motion, stretching and relaxation exercises, CBT-support, ergonomic advice and home exercises, 2 1 ½ hr sessions/week, 12 weeks, group of 4-5 patients, N=30 II. Massage, thermal therapy once /week, 4 weeks, N=29	VAS-pain, PDI	I reduction of VAS-pain and PDI (2.4 vs. -1.4) 6 and 12 months: I reduction of VAS-pain and PDI (7.5 vs.. 1.9)	Relative low rate of disability, no control group

(Continued)

Table 6: (Continued).

Ljunggren (1997)[b]	153 LBP patients after completed PT, occupational active, referred by general practitioners, exclusion: SLBP, spondylolysis, spondylolisthesis, inflammatory rheumatic diseases, other ambiguous pathologic diagnoses; age: I 40.2 (± 9.5) yr, II 39.0 (±10.4) yr., duration of complaints: ?	Both treatment groups: 8 sessions/yr supervised by physiotherapist and progressive home exercises 15-30 min, 3 series of 10 repetitions, 3 times/week during 1 yr; I. Conventional PT exercises (9): sit-ups, push ups, TSE, shoulder and chest, stretching, N=64. II. TerapiMaster exercises (9): mobilisation, traction, TSE, shoulder and chest muscles, N=62	Absenteeism from work	I 17.2 days reduction, II 15.4, no group difference 1 yr: I 9.9 days reductions, II 9.3, no group difference	Both treatments use strengthening exercises with same amount of sessions, intensity and frequency, no control group
Manniche (1989; 1991)[b]	105 CLBP patients (≥ 6 months or ALBP 3 times ≤ ½ yr), referred by general practitioners, exclusion: SLBP, spondylolysis, painful hip arthrosis, co-morbidity hampering physical training; age: 45 (10/90 percentile 32-57) yr, duration of complaints: 15 (10/90 percentile 5-34) yr, level of disability (disability section LBPRS): I 10.2 (5.2 -19.4), II 11.4 (5.1-19.8), III 10.3 (7.6-19.6)	I. Hot compresses, massage, isometric exercises (back extension, abdominal contraction, sit-ups, curl-ups), 10 repetitions, 8 times 1 hour sessions in 1 month, N=32 II. Hot pack, 3 dynamic BSE (trunk lifting, leg lifting, pull to neck), 20 repetitions, stretching of muscle, 1 month 3 45 min sessions/week, next 2 months 2 sessions/week, total 30 sessions in 3 months (intensity 1/5 of III), N=31. III. Same as I, starting with 50 repetitions after 2 weeks 100 repetitions each exercise, 1 month 3 1 ½ hr sessions/week, next 2 months 2 sessions/week, total 30 sessions in 3 months, N=27. All groups: 4-5 patients, 2 therapists. II and III: therapists helped patients during last part of movement without eliciting pain and gradual increase of load	LBPRS	III better (14.7) vs. I (2.0) and II (5.7) 3 months: III better (15.0) vs. I (5.5) and II (7.0)	After 3 months all patients could continue (once a week) or start intensive BSE, no control group
Mannion (1999)[a]	For description see table 5				
Petersen (2002)[a]	270 LBP patients (≥ 8 weeks), referred to rheumatology clinic (85% CLBP), exclusion: SLBP, spondylolysis, spondylolisthesis, no previous TSE or McKenzie, psychosocial problems requiring multidisciplinary treatment; age: I 34.5 (23-52) yr, II 35 (24-52) yr, duration of complaints: I 8 (2 - 96) months, II 14 (3 - 114) months, level of disability (LBPRS, %): I 36.7 (19.1 - 60.1), II 39.3 (17.5 - 60)	I. McKenzie; after individual assessment self-mobilizing repeated movements, sustained positions in specific directions, manual overpressure and/or mobilisation by therapist, max 15 sessions in 8 weeks, N=132 II. 4 dynamic TSE, starting with 50 repetitions, increasing to 100 repetitions, 10-15 minutes of warm up, 10 min SE after TSE, 2 60- 90 min sessions/week, max 15 sessions in 8 weeks, group of 6 patients, 1physiotherapist, N=128 All groups were asked to keep on practicing 2 months post-treatment (no control whether this was done)	Disability (LBPRS) Back and leg pain (11-box)	Post-treatment, 2 and 8 months FU: no group difference	Improper exclusion of 10 patients, max 15 sessions of TSE, I almost twice as much sessions as II, high drop out rate (> 30%), no control group
Torstensen (1998)[a]	208 CLBP patients, sick-listed 8-52 weeks, recruited from social security offices, exclusion: SLBP, back surgery, spondylolisthesis, hip arthrosis, co-morbidity hampering physical training; age: I 42.1 (± 11.2) yr, II 43.0 (± 12.0) yr, III 39.9 (± 11.4) yr, duration of complaints: ?, level of disability (Os): I 51.2 (± 10.7), II 49.9 (± 10.5), III 50.0 (± 11.9)	I. MET (high frequency, low intensity): trunk flexion, extension, rotation of back and extremities, abdominal exercises to mobilize and/or stabilize spine, 7-9 exercises 2-3 sets of 20-30 repetitions, 3 1 hr sessions/week, 12 weeks, group of max. 5 patients, N=71. II. PT: heat, cold, massage, stretching, traction, electrotherapy and exercise (no extensive exercises allowed), 3 1 hr sessions/week, 12 weeks, N=67. III. Home exercise: 6 contacts by telephone to check adherence to treatment, 3 1 hr walking 3 sessions/week, 12 weeks, N=70	VAS-pain back, buttock and leg, Os, return to work, total costs (data security office)	I and II reduction of VAS-pain and Os (5.9, and 2.5 vs. –2.7) 1 yr: I and II reduction of VAS-pain leg and Os (7.8 and 6.4 vs. –0.6), no group difference in return to work, I ($ 122.531) and II ($ 254.200) more cost saving	No real control group

[a] = studies not included in the Cochrane review on strengthening exercises; [b] = studies included in the Cochrane review on strengthening exercises; 1-RM = One-repetition maximum; ALBP = acute low back pain; BSE = back and leg extensor strengthening exercises; CBT = cognitive behavioral therapy; CLBP = chronic low back pain; CSA = cross sectional area; COOP = Dartmouth primary care cooperative information project; FU = follow-up; LBP = low back pain; LBPRS = low back pain rating scale; MET = medical exercise therapy; NSAID = nonsteroidal anti-inflammatory drug; NSCLBP = nonspecific chronic low back pain; NSLBP = nonspecific low back pain; Os = Oswestry questionnaire; PDI = pain disability index; PT = physiotherapy; RDQ = Roland disability questionnaire; SE = stretching exercises; SF-36 = short form health survey; SLBP = specific low back pain; SMT= spinal manipulative therapy; TSE = trunk (back and abdominal) and leg strengthening exercises; TSK = Tampa scale for kinesiophobia; VAS = visual analogue scale.

compared to less intensive exercises. But this has only been proven for CLBP patients with low to moderate levels of disability and a comparison with a waiting list control group is missing. When these specific low back muscle strengthening exercises are applied with a lower total amount of sessions (8 to 15 instead of 30 sessions) and/or intensity of workload (15 to 20 instead of 45 to 100 repetitions) there is moderate evidence that they are equally effective as other active treatments.

Two studies, one uncontrolled and one controlled study, show some evidence that intensive strengthening training of the back muscles decreases fatty infiltration (Mooney et al. 1997), and increases the muscle bulk of the paraspinal muscles, including the multifidus muscle (Danneels 2001; Danneels et al. 2001a; Danneels et al. 2001b), but no significant correlation with the decrease of activity limitation can be found. Two studies, also one uncontrolled and one controlled study, show conflicting evidence whether the ratio of type I:II fibres changes, and the association between this change and the decrease of activity limitation is also missing (Käser et al. 2001; Rissanen et al. 1995).

Discussion

The most notable finding of this review is the lack of any strong evidence supporting the existence of physical deconditioning symptoms regarding cardiovascular capacity and paraspinal muscles in CLBP patients. There is only limited evidence that atrophy of the multifidus muscle plays a role in the complex of chronic low back pain.
Likewise this review does not show strong evidence supporting the use of intensive physical training programmes instead of other active therapies in the treatment for CLBP and its resulting disability. In comparison to other active treatments, only the back extensor muscle strengthening exercises show moderate proof for being more effective. The most important key to success seems to be the intensity and duration/frequency of the treatment as several less intense treatments and/or treatments of shorter duration show no additional effectiveness compared to other active treatments (Bronfort et al. 1996; Hansen et al. 1993; Helmhout et al. 2004; Manniche et al. 1988; Manniche et al. 1991; Petersen et al. 2002). The exercise used in the study of Manniche et al. (1988; 1991) resembles the training of the multifidus muscle that is described by Danneels et al. (2001; 2001a; 2001b). The only difference between the treatments is that Danneels used three series of 15 to 18 repetitions with 70% of the 1-Repetition Maximum (1-RM) and Manniche used two series of 50 repetitions with gradual reduction of assistance of the therapist. The effectiveness of the dynamic-static exercise protocol for back extensor muscles as described by Danneels should be studied in a population of patients with a much higher level of disability and in a trial of higher methodological quality.
The third major finding is the lack of high quality research focused on physical deconditioning in CLBP. Specifically, only cross-sectional studies have been used to investigate the link between deconditioning, back pain and disability. Future research should use prospective designs to study this relationship.

This review differs from a standard systematic review because it examines the evidence relating to an alleged association, rather than the effect of a specific intervention. Moreover most of the relevant factors are poorly defined. For example even CLBP is subject to debate, and there is little agreement on the definition of training programmes or the measurement of factors such as muscle strength and disability. Consequently the review itself has to be considered critically.

Reviews like this are subject to potential publication bias, especially the risk of selective publication of positive studies. However, as the publications do not support any of the hypotheses strongly and sometimes not at all, we doubt that publication bias is an important factor.
There are many methodological difficulties in investigating the hypotheses concerning the effect of training programmes. Most studies on the effectiveness of training programmes investigate a variety of treatments simultaneously, and it is almost impossible to separate out the particular features of interest. Future studies investigating treatments for physical deconditioning should use specific interventions on their own, and should not use multiple interventions (Hilde and Bo 1998).

Moreover, the treatments that are subject to the studies reviewed here sometimes are poorly to moderately described and therefore difficult to evaluate. Searching for relevant studies evaluating specific treatments is made difficult through the lack of agreed terminology. It is especially important that future trials use exercises that are grounded on sound theoretical principles concerning exercise physiology and dose response. Protas (1996) already advised that not only the intensity but also the duration, frequency and progression of the aerobic training should be documented in future studies so that one can judge whether the training was of good quality and quantity. For example, to increase cardiovascular capacity, one should train at least three times a week during 20 to 60 minutes for 10 weeks or more (ACSM 1998). To gain strength, a training program of at least two to three sessions a week, during at least 16 h in total seems to be necessary (Bronfort et al. 1996).

The comparability of studies regarding the effectiveness of training and the generalization to daily practice of their results is very difficult because the level of disability at the start of treatment differs considerably and patients with a high degree of disability are less represented. Often different outcome scales for low back pain disability are used that cannot be compared to each other and some studies do not even use a disability scale. Furthermore, the recruitment of patients differs. Patients selectively recruited by advertisement in local newspapers do not share the characteristics of those referred to a rehabilitation centre or pain clinic.

Future trials should describe the population of patients extensively, and especially restrict the selection to patients with chronic nonspecific low back pain that are normally referred for treatment, so excluding as few patients as possible. Furthermore, researchers should at least use the same low back pain disability scale (Deyo et al. 1998).

Despite all these problems, it is striking how little evidence there is to support the hypothesis that patients with CLBP are physically deconditioned, e.g. have a lower level of cardiovascular capacity and/or loss of paraspinal muscle strength and endurance. Maybe neuromuscular changes such as loss of propriocepsis (Brumagne et al. 2000), lower motor control of the stabilizing abdominal and back muscles (Hides et al. 1996; Hides et al. 1994; Hodges and Richardson 1996; Hodges and Richardson 1998; Hodges and Richardson 1999) or lower postural control (Luoto et al. 1999) are important contributing factors in the maintenance of CLBP, but that was not the scope of our review. Nonetheless, the relative lack of evidence for the physical deconditioning theory and despite the growing evidence for the important role of psychosocial factors in nonspecific CLBP disability, many patients and quite a few health care professionals adhere to this more biomedical model of pain instead of the biopsychosocial model (Abenhaim et al. 2000; Buchbinder et al. 2001; Daykin and Richardson 2004; Jette and Jette 1996; Protas 1996; Quittan 2002).
For them the developing physical deconditioning seems to be an attractive biomedical explanation for the persistence of the patients' back pain problems and all the implications that follow. It suggests that a specific physical oriented treatment might alleviate the situation without for example, having to address psychosocial and personal aspects that might also contribute to the level of low back pain disability. On the other hand this review shows that there is scientific evidence for the use of reconditioning treatment when compared to doing nothing or passive treatments, but it also shows that this treatment is no more effective than other active treatments. This is in concordance with other systematic reviews on exercise and cognitive-behavioral therapy that conclude that all these treatments are more effective than doing nothing, but equally effective when compared to other active treatments (Liddle et al. 2004; Ostelo et al. 2005; van Tulder et al. 2000a). Probably the reactivation is the common and important factor in the reduction of CLBP disability. Several supposed mechanisms for this positive effect of reactivation are the reduction of fear avoidance (Mannion et al. 2001), the increase of self-efficacy (Campello et al. 1996), reduction of depressive symptoms (Crews and Landers 1987; Thirlaway and Benton 1992), socialization and reduction of pain anticipation and pain perception (Abenhaim et al. 2000; Ferrell et al. 1997; Gurevich et al. 1994; Mannion et al. 1999; Rainville et al. 2004).

The findings of this review emphasize the need for more research, and preferably longitudinal studies on the physical deconditioning process. In our opinion it is especially important and necessary to keep on searching for better models to try to understand the complex of CLBP and disability. We still should stimulate and motivate our patients to keep active but based on the available evidence we suggest that we should be more careful in the way we explain the role of deconditioning in CLBP and disability to them. Given the increasing recognition that many factors unrelated to physical fitness are associated with the persistence of CLBP and activity limitation, such as social and medical attitudes, socio-economic factors, patient beliefs and emotional status, patients may benefit more from education in which the interplay between biological, social and psychological factors are more important than biological (such as physical deconditioning) alone. For future research it is even more interesting to find out how each treatment exerts its effect. It seems that changes in cognitions such as decrease of catastrophizing, increase of internal control of pain and decrease of fear of injury/movement play a crucial role (Jensen et al. 2001; Klaber Moffett et al. 2004; Spinhoven et al. 2004; Verbunt et al. 2003; Vlaeyen et al. 1995).

Acknowledgements

This study is supported by Zorgonderzoek Nederland/Medische Wetenschappen (ZonMw) grant number 014-32-007.

References

- AAC Committee. NH, New Zealand acute low back pain guide. Wellington: New Zealand, 1997.
- Abenhaim L, Rossignol M, Valat JP, Nordin M, Avouac B, Blotman F, Charlot J, Dreiser RL, Legrand E, Rozenberg S, Vautravers P. The role of activity in the therapeutic management of back pain. Report of the International Paris Task Force on Back Pain. Spine 2000;25:1S-33S.
- ACSM. ACSM position stand on the recommended quantity and quality of exercise for developing and maintaining cardiorespiratory and muscular fitness, and flexibility in adults. Med Sci Sports Exerc 1998;30:975-91.
- Alaranta H, Tallroth K, Soukka A, Heliövaara M. Fat content of lumbar extensor muscles and low back pain disability: a radiographic and clinical comparison. J Spinal Disord 1993;6:137-40.
- Barker KL, Shamley DR, Jackson D. Changes in the cross-sectional area of multifidus and psoas in patients with unilateral back pain: the relationship to pain and disability. Spine 2004;29:E515-9.
- Battie MC, Bigos SJ, Fisher LD, Hansson TH, Nachemson AL, Spengler DM, Wortley MD, Zeh J. A prospective study of the role of cardiovascular risk factors and fitness in industrial back pain complaints. Spine 1989;14:141-7.
- Bendix T, Bendix A, Labriola M, Haestrup C, Ebbehoj N. Functional restoration versus outpatient physical training in chronic low back pain: a randomized comparative study. Spine 2000;25:2494-500.
- Bronfort G, Goldsmith CH, Nelson CF, Boline PD, Anderson AV. Trunk exercise combined with spinal manipulative or NSAID therapy for chronic low back pain: a randomized, observer-blinded clinical trial. J Manipulative Physiol Ther 1996;19:570-82.
- Brumagne S, Cordo P, Lysens R, Verschueren S, Swinnen S. The role of paraspinal muscle spindles in lumbosacral position sense in individuals with and without low back pain. Spine 2000;25:989-94.
- Buchbinder R, Jolley D, Wyatt M. 2001 Volvo Award Winner in Clinical Studies: Effects of a media campaign on back pain beliefs and its potential influence on management of low back pain in general practice. Spine 2001;26:2535-42.
- Campello M, Nordin M, Weiser S. Physical exercise and low back pain. Scand J Med Sci Sports 1996;6:63-72.
- Cooper RG, Clair Forbes WS, Jayson MIV. Radiographic demonstration of paraspinal muscle wasting in patients with chronic low back pain. Br J Rheum 1992;31:389-94.
- Crews DJ, Landers DM. A meta-analytic review of aerobic fitness and reactivity to psychosocial stressors. Med Sci Sports Exerc 1987;19:S114-S120.
- Crombez G, Vervaet L, Baeyens F, Lysens R, Eelen P. Do pain expectancies cause pain in chronic low back patients? A clinical investigation. Behav Res Ther 1996;34:919-25.
- Crossman KBM, Mahon MM, Watson PJPM, Oldham JAPRGN, Cooper RGMDF. Chronic low back pain-associated paraspinal muscle dysfunction is not the result of a constitutionally determined «adverse» fiber-type composition. Spine 2004;29:628-34.
- Danneels LA. Evaluation and rehabilitation of functional spinal stability. Ghent: Ghent University, 2001.
- Danneels LA, Cools AM, Vanderstraeten GG, Cambier DC, Witvrouw EE, Bourgois J, de Cuyper HJ. The effects of three diffrent training modalities on the cross-sectional area of the paravertebral muscles. Scan J Med Sci Sports 2001a;11:335-41.
- Danneels LA, Vanderstraeten GG, Cambier DC, Witvrouw EE, de Cuyper HJ. Ct imaging of trunk muscles in chronic low back pain patients and healthy control subject. Eur Spine J 2000;9:266-72.
- Danneels LA, Vanderstraeten GG, Cambier DC, Witvrouw EE, Bourgois J, Dankaerts W, de Cuyper HJ. Effects of three different training modalities on the cross sectional area of the lumbar multifidus muscle in patients with chronic low back pain. Br J Sports Med 2001b;35:186-91.
- Davis VP, Fillingim RB, Doleys DM, Davis MP. Assessment of aerobic power in chronic pain patients before and after a multidisciplinary treatment program. Arch Phys Med Rehabil 1992;73:726-9.

References

- Daykin AR, Richardson B. Physiotherapists' pain beliefs and their influence on the management of patients with chronic low back pain. Spine 2004;29:783-95.
- Deyo RA, Battie M, Beurskens AJ, Bombardier C, Croft P, Koes B, Malmivaara A, Roland M, Von Korff M, Waddell G. Outcome measures for low back pain research. A proposal for standardized use. Spine 1998;23:2003-13.
- Ferrell BA, Josephson KR, Pollan AM, Loy S, Ferrell BR. A randomized trial of walking versus physical methods for chronic pain management. Aging 1997;9:99-105.
- Flor H, Fydrich T, Turk DC. Efficacy of multidisciplinary pain treatment centres: a meta-analytic review. Pain 1992;49:221-30.
- Fordyce WE. Behavioral methods for chronic pain and illness. St. Louis: Mosby, 1976.
- Frost H, Klaber-Moffett JA, Moser JS, Fairbank JC. Randomised controlled trial for evaluation of fitness programme for patients with chronic low back pain. BMJ 1995;310:151-4.
- Frost H, Lamb SE, Klaber-Moffett JA, Fairbank JC, Moser JS. A fitness programme for patients with chronic low back pain: 2 year follow-up of a randomised controlled trial. Pain 1998;75:273-9.
- Frymoyer JW. Back pain and sciatica. N Engl J Med 1988;318:291-300.
- Gibbons LE, Videman T, Battie MC. Isokinetic and psychophysical lifting strength, static back muscle endurance, and magnetic resonance imaging of the paraspinal muscles as predictors of low back pain in men. Scan J Rehab Med 1997a;29:187-91.
- Gibbons LE, Videman T, Crites Battié M. Determinants of isokinetic and psychophysical lifting strength and static back muscle endurance: A study of male monozygotic twins. Spine 1997b;23:2412-21.
- Gurevich M, Kohn PM, Davis C. Exercise-induced analgesia and the role of reactivity in pain sensitivity. J Sports Sci 1994;12:549-59.
- Hansen FR, Bendix T, Skov P, Jensen CV, Kristensen JH, Krohn L, Schioeler H. Intensive, dynamic back-muscle exercises, conventional physiotherapy, or placebo-control treatment of low-back pain. A randomized, observer-blind trial. Spine 1993;18:98-108.
- Haskell WL. Dose-response issues from a biological perspective. In: C Bouchard, RJ Shephard, T Stephens. Physical activity, fitness, and health: international proceedings and consensus statement. Champaign: Human Kinetics, 1994, pp. 1030-9.
- Helmhout PH, Harts CC, Staal JB, Candel MJ, de Bie RA. Comparison of a high-intensity and a low-intensity lumbar extensor training program as minimal intervention treatment in low back pain: a randomized trial. Eur Spine J 2004;13:537-47.
- Hides JA, Richardson CA, Jull GA. Multifidus muscle recovery is not automatic after resolution of acute, first-episode low back pain. Spine 1996;21:2763-9.
- Hides JA, Stokes MJ, Saide M, Jull GA, Cooper DH. Evidence of lumbar multifidus muscle wasting ipsilateral to symptoms in patients with acute/subacute low back pain. Spine 1994;19:165-72.
- Hilde G, Bo K. Effect of exercise in the treatment of chronic low back pain: a systematic review, emphasising type and dose of exercise. Phys Ther Rev 1998;3:107-17.
- Hodges PW, Richardson CA. Inefficient muscular stabilization of the lumbar spine associated with low back pain: A motor control evaluation of transversus abdominis. Spine 1996;21:2640-50.
- Hodges PW, Richardson CA. Delayed postural contraction of the transversus abdominis in low back pain associated with movement of the lower limb. J Spinal Disord 1998;11:990-1001.
- Hodges PW, Richardson CA. Altered trunk muscle recruitment in people with low back pain with upper limb movement at different speeds. Arch Phys Med Rehabil 1999;80:1005-12.

References

- Hultman G, Nordin M, Saraste H, Ohlsen H. Body composition, endurance, strength, cross-sectional area, and density of MM erector spinae in men with and without low back pain. J Spinal Disord 1993;6:114-23.
- Hurri H. The Swedish back school in chronic low back pain. Part I. Benefits. Scan J Rehab Med 1989;21:33-40.
- Hurri H, Mellin G, Korhonen O, Harjula R, Harkapaa K, Luoma J. Aerobic capacity among chronic low-back-pain patients. J Spinal Disord 1991;4:34-8.
- International Association for the Study of Pain SOT. Classification of chronic pain, description of chronic pain syndromes and definitions of pain terms. Pain 1986;Suppl. 3:1-225.
- Jensen MP, Turner JA, Romano JM. Changes in beliefs, catastrophizing, and coping are associated with improvement in multidisciplinary pain treatment. J Consult Clin Psychol 2001;69:655-62.
- Jette DU, Jette AM. Physical therapy and health outcomes in patients with spinal impairments. Phys Ther 1996;76:730-41.
- Kader DF, Wardlaw D, Smith FW. Correlation between the MRI changes in the lumbar multifidus muscles and leg pain. Clin Radiol 2000;55:145-9.
- Kankaanpaa M, Laaksonen D, Taimela S, Kokko SM, Airaksinen O, Hanninen O. Age, sex, and body mass index as determinants of back and hip extensor fatigue in the isometric Sörensen back endurance test. Arch Phys Med Rehabil 1998;79:1069-75.
- Kankaanpaa M, Taimela S, Airaksinen O, Hanninen O. The efficacy of active rehabilitation in chronic low back pain. Effect on pain intensity, self-experienced disability, and lumbar fatigability. Spine 1999;24:1034-42.
- Käser L, Mannion AF, Rhyner A, Weber E, Dvorak J, Muntener M. Active therapy for chronic low back pain. Part 2. Effects on paraspinal muscle cross-sectional area, fiber type size, and distribution. Spine 2001;26:909-19.
- Keller A, Gunderson R, Reikeraas O, Brox JI. Reliability of computed tomography measurements of paraspinal muscle cross-sectional area and density in patients with chronic low back pain. Spine 2003;28:1455-60.
- Keller A, Johansen JG, Hellesnes J, Brox JI. Predictors of isokinetic back muscle strength in patients with low back pain. Spine 1999;24:275-80.
- Kellett KM, Kellett DA, Nordholm LA. Effects of an exercise program on sick leave due to back pain. Phys Ther 1991;71:283-91.
- Klaber-Moffett JA, Carr J, Howarth E. High fear-avoiders of physical activity benefit from an exercise program for patients with back pain. Spine 2004;29:1167-72.
- Lariviere C, Arsenault B, Gravel D, Gagnon D, Loisel P, Vadeboncoeur R. Electromyographic assessment of back muscle weakness and muscle composition: reliability and validity issues. Arch Phys Med Rehabil 2002;83:1206-14.
- Liddle SD, Baxter GD, Gracey JH. Exercise and chronic low back pain: what works? Pain 2004;107:176-90.
- Ljunggren AE, Weber H, Kogstad O, Thom E, Kirkesola G. Effect of exercise on sick leave due to low back pain. A randomized, comparative, long-term study. Spine 1997;22:1610-7.
- Luoto S, Taimela S, Hurri H, Alaranta H. Mechanisms explaining the association between low back trouble and deficits in information processing. A controlled study with follow-up. Spine 1999;24:255-61.
- Manniche C, Hesselsoe G, Bentzen L, Christensen I, Lundberg E. Clinical trial of intensive muscle training for chronic low back pain. Lancet 1988;2:1473-6.
- Manniche C, Lundberg E, Christensen I, Bentzen L, Hesselsoe G. Intensive dynamic back exercises for chronic low back pain: a clinical trial. Pain 1991;47:53-63.
- Mannion AF, Käser L, Weber E, Rhyner A, Dvorak J, Müntener M. Influence of age and duration of symptoms on fiber distribution and size of the back muscles in chronic low back pain patients. Eur Spine J 2000;9:273-81.

References

- Mannion AF, Muntener M, Taimela S, Dvorak J. A randomized clinical trial of three active therapies for chronic low back pain. Spine 1999;24:2435-48.
- Mannion AF, Müntener M, Taimela S, Dvorak J. Comparison of three active therapies for chronic low back pain: results of a randomized clinical trial with one-year follow-up. Rheumatology 2001;40:772-8.
- Mannion AF, Weber BR, Dvorak J. Fiber type characteristics of the lumbar paraspinal muscles in normal healthy subjects and in patients with low back pain. J Orthop Res 1997;15;881-7.
- Mattila M, Hurme M, Alaranta H, Paljärvi L, Kalimo H, Falcks B, Lehto M, Einalo S, Jarvinen M. The multifidus muscle in patients with lumbar disc herniation. A histochemical and morphometric analysis of intraoperative biopsies. Spine 1986;11:732-38.
- Mayer TG, Gatchel RJ. Functional restoration for spinal disorders: the sports medicine approach. Philadelphia: Lea & Febiger,1988.
- Mayer TG, Polatin PB, Gatchel RJ. Functional restoration and other rehabilitation approaches to chronic musculoskeletal pain disability syndromes. Crit Rev Physical Rehab Med 1998;10:209-21.
- Mayer TG, Vanharanta H, Gatchel RJ, Mooney V, Barnes D, Judge L, Smith S, Terry A. Comparison of CT scan muscle measurements and isokinetic trunk strength in postoperative patients. Spine 1989;14:33-6.
- McArdle W, Catch FI, Katch VL. Training for anaerobic and aerobic power. In: W McArdle, FI Catch, VL Katch. Exercise physiology. Baltimore: Williams & Wilkins, 1996, pp. 393-415.
- McArdle WD, Katch FI, Katch VL. Exercise Physiology: energy, nutrition, and human performance. In: W McArdle, FI Catch, VL Katch. Exercise physiology. Baltimore: Lippincott Wiliams & Wilkins, 2001, pp. 222-48.
- Meerding WJ, Bonneux L, Polder JJ, Koopmanschap MA, van der Maas PJ. Demographic and epidemiological determinants of healthcare costs in Netherlands: cost of illness study. BMJ 1998;317:111-5.
- Moffett JK, Torgerson D, Bell-Syer S, Jackson D, Llewlyn-Phillips H, Farrin A, Barber J. Randomised controlled trial of exercise for low back pain: clinical outcomes, costs, and preferences. BMJ 1999;319:279-83.
- Mooney V, Gulick J, Perlman M, Levy D, Pozos R, Leggett SH, Resnick D. Relationships between myoelectric activity, strength, and MRI of lumbar extensor muscles in back pain patients and normal subjects. J Spinal Disord 1997;10:348-56.
- Morley S, Eccleston C, Williams A. Systematic review and meta-analysis of randomized controlled trials of cognitive behaviour therapy and behaviour therapy for chronic pain in adults, excluding headache. Pain 1999;80:1-13.
- Newton M, Thow M, Somerville D, Henderson I, Waddell G. Trunk strength testing with isokinetic lifting-machines. Part 2: Experimental evaluation of the Cybec II back testing system in normal subjects and patients with chronic low back pain. Spine 1993;18:812-24.
- Nielens H, Plaghki L. Evaluation of physical adaptation to exercise of chronic pain patients by step-test procedure. Pain Clin 1991;4:21-5.
- Nielens H, Plaghki L. Perception of pain and exertion during exercise on a cycle ergometer in chronic pain patients. Clin J Pain1994;10:204-9.
- Nielens H, Plaghki L. Cardiorespiratory fitness, physical activity level, and chronic pain: are men more affected than women? Clin J Pain 2001;17:129-37.
- Ostelo RW, Tulder M, Vlaeyen J, Linton S, Morley S, Assendelft W. Behavioural treatment for chronic low-back pain. Cochrane Database Syst Rev 2005:Cd002014.

References

- O'Sullivan PB, Phyty GD, Twomey LT, Allison GT. Evaluation of specific stabilizing exercise in the treatment of chronic low back pain with radiologic diagnosis of spondylolysis or spondylolisthesis. Spine 1997;22:2959-67.
- Panjabi MM. The stabilizing system of the spine. Part I. Function, dysfunction, adaptation, and enhancement. J Spinal Disord 1992a;5:383-9.
- Panjabi MM. The stabilizing system of the spine. Part II. Neutral zone and instability hypothesis. J Spinal Disord 1992b;5:390-6.
- Parkkola R, Rytokoski U, Kormano M. Magnetic resonance imaging of the discs and trunk muscles in patients with chronic low back pain and healthy control subjects. Spine 1993;18:830-6.
- Petersen T, Kryger P, Ekdahl C, Olsen S, Jacobsen S. The effect of McKenzie therapy as compared with that of intensive strengthening training for the treatment of patients with subacute or chronic low back pain: A randomized controlled trial. Spine 2002;27:1702-9.
- Pincus T, Burton AK, Vogel S, Field AP. A systematic review of psychological factors as predictors of chronicity/disability in prospective cohorts of low back pain. Spine 2002;27:E109-20.
- Protas EJ. Aerobic exercise in the rehabilitation of individuals with chronic low back pain: a review. Crit Rev Physical Rehab Med 1996;8:283-95.
- Quittan M. Management of back pain. Disabil Rehabil 2002;24:423-34.
- Rainville J, Hartigan C, Jouve C, Martinez E. The influence of intense exercise-based physical therapy program on back pain anticipated before and induced by physical activities. Spine J 2004;4:176-83.
- Rantanen J, Rissanen A, Kalimo H. The lumbar multifidus muscle five years after surgery for a lumbar intervertebral disc herniation. Spine 1993;18:568-74.
- Richardson CA, Jull GA, Hodges PW, Hides JA. Therapeutic exercise for spinal segmental stabilization in low back pain. Scientific basis and clinical approach. London: Churchill Livingstone, 1999.
- Richardson CA, Snijders CJ, Hides JA, Damen L, Pas MS, Storm J. The relation between the transversus abdominis muscles, sacroiliac joint mechanics, and low back pain. Spine 2002;27:399-405.
- Rissanen A, Kalimo H, Alaranta H. Effect of intensive training on the isokinetic strength and structure of lumbar muscles in patients with chronic low back pain. Spine 1995;20:333-40.
- Schmidt AJM. Cognitive factors in the performance level of chronic low back pain patients. J Psychosom Res 1985;29:183-9.
- Sihvonen T, Herno A, Paljärvi L, Airaksinen O, Partanen J, Tapaninaho A. Local denervation atrophy of paraspinal muscles in postoperative failed back syndrom. Spine 1993;18:575-81.
- Simmonds MJ, Olson SL, Jones S, Hussein T, Lee CE, Novy D, Radwan H. Psychometric characteristics and clinical usefulness of physical performance tests in patients with low back pain. Spine 1998;23:2412-21.
- Spinhoven P, Ter Kuile M, Kole-Snijders AM, Hutten-Mansfeld M, Den Ouden DJ, Vlaeyen JW. Catastrophizing and internal pain control as mediators of outcome in the multidisciplinary treatment of chronic low back pain. Eur J Pain 2004;8:211-9.
- Thirlaway K, Benton D. Participation in physical activity and cardiovascular fitness have different effects on mental health and mood. J Psychosom Res 1992;36:657-65.
- Torstensen TA, Ljunggren AE, Meen HD, Odland E, Mowinckel P, Geijerstam S. Efficiency and costs of medical exercise therapy, conventional physiotherapy, and self-exercise in patients with chronic low back pain. A pragmatic, randomized, singleblinded, controlled trial with 1-year follow-up. Spine 1998;23:2616-24.
- van der Velde G, Mierau D. The effect of exercise on percentile rank aerobic capacity, pain, and self-rated disability in patients with chronic low-back pain: a retrospective chart review. Arch Phys Med Rehabil 2000;81:1457-63.

References

- van Tulder M, Malmivaara A, Esmail R, Koes B. Exercise therapy for low back pain. A systematic review within the framework of the Cochrane Collaboration Back Review Group. Spine 2000a;25:2784-96.
- van Tulder MW, Koes BW, Bouter LM. A cost-of-illness study of back pain in The Netherlands. Pain 1995;62:233-40.
- van Tulder MW, Malmivaara A, Esmail R, Koes BW. Exercise therapy for low back pain. Cochrane Database Syst Rev 2000b:Cd000335.
- van Wingerden BAM. Trainingsprincipes. In: BAM van Wingerden. Bindweefsel in de revalidatie. Schaan: Scipro Verlag, 1997, pp. 291-324.
- Verbunt JA, Seelen HA, Vlaeyen JWS, van der Heijden GJ, Heuts PH, Pons K, Knottnerus JA. Disuse and deconditioning in chronic low back pain: concepts and hypothesis on contributing mechanisms. Eur J Pain 2003;7:9-21.
- Vlaeyen JWS, Kole-Snijders AM, Rotteveel AM, Ruesink R, Heuts PHTG. The role of fear of movement/(re)injury in pain disability. J Occup Rehabil 1995;5:235-52.
- Vlaeyen JWS, Linton SJ. Fear-avoidance and its consequences in chronic musculoskeletal pain: a state of the art. Pain 2000;85:317-32.
- Vollenbroek-Hutten MM, Hermens HJ, Wever D, Gorter M, Rinket J, IJzerman MJ. Differences in outcome of a multidisciplinary treatment between subgroups of chronic low back pain patients defined using two multiaxial assessment instruments: the multidimensional pain inventory and lumbar dynamometry. Clin Rehabil 2004;18:566-79.
- Waddell G. A new clinical model of low back pain and disability. The back pain revolution. London: Churchill Livingstone, 1998.
- Watson PJ. Non-psychological determinants of physical performance in musculoskeletal pain. In: M Mitchell. Pain 1999- an updated review: Refresher course syllabus 9th World Congress on Pain. Seattle: IASP Press, 1999, pp. 153-8.
- Weber B, Grob D, Dvorak J. Posterior surgical approach to the lumbar spine and its effect on the multifidus muscle. Spine 1997;22:1765-72.
- Wittink H, Hoskins Michel T, Wagner A, Sukiennik A, Rogers W. Deconditioning in patients with chronic low back pain: fact or fiction? Spine 2000a;25:2221-8.
- Wittink H, Michel TH, Kulich R, Wagner A, Sukiennik A, Maciewicz R, Rogers W. Aerobic fitness testing in patients with chronic low back pain: which test is best? Spine 2000b;25:1704-10.
- Zhu XZ, Parnianpour M, Nordin M. Histochemistry and morphology of erector spinae muscle in lumbar disc herniation. Spine 1989;14:391-7.

Chapter 3

Do patients with chronic low back pain have a lower level of aerobic fitness than healthy controls? Are pain, disability, fear of injury, working status or level of leisure time activity associated with the difference in aerobic fitness level?

Rob J.E.M. Smeets, Harriët Wittink, Alita Hidding, J. André Knottnerus

Published as:
Smeets RJEM, Wittink H, Hidding A, Knottnerus JA. Do patients with chronic low back pain have a lower level of aerobic fitness than healthy controls? Are pain, disability, fear of injury, working status or level of leisure time activity associated with the difference in aerobic fitness level? Spine 2006;31:90-97.

Abstract

Study design. Prospective case series with historical controls (normative data).
Objectives. To compare the aerobic fitness level of patients with chronic low back pain (CLBP) with healthy controls matched for gender, age and level of sport activity and to evaluate the association of the difference in aerobic fitness level with pain intensity, duration and degree of disability, fear of injury, and level of activity during work including household, and leisure time.
Summary and background Data. Controversy exists whether patients with CLBP have a lower level of aerobic fitness and whether this level may partly depend on the patients' activity level.
Methods. A total of 108 CLBP patients completed questionnaires regarding pain, disability, fear of injury, and activity level and performed a modified Åstrand submaximal cycling test. The maximum oxygen consumption (VO_2max) was calculated and compared with normative data. Multiple linear regression analysis was performed with the difference of the observed and expected level of aerobic fitness as dependent variable.
Results. VO_2max could be calculated in 78% of the patients. Both men and women with CLBP had significant lower VO_2max than the healthy referents (10 mL/kg LBM.min^{-1} and 5.6 mL/kg LBM.min^{-1} respectively, $P < 0.001$), and this difference was significantly greater in men ($P = 0.03$). Multiple regression analysis showed that the level of aerobic fitness was not associated with the presumed variables. The patients who stopped the test prematurely were older ($P = 0.02$) and more disabled ($P = 0.01$).
Conclusion. CLBP patients, especially men seem to have a reduced aerobic fitness level compared with the normative population. No explanatory factor for that loss could be identified.

Introduction

Current rehabilitation for patients with chronic low back pain (CLBP) is often based on the assumption that these patients have low levels of aerobic fitness and are deconditioned (Abenhaim et al. 2000; Jette and Jette 1996; Mayer and Gatchel 1988). The deconditioning syndrome was postulated in the mid 1980s as a factor contributing to the intolerance to physical activities and subsequent further loss of function and disability in patients with CLBP. More recently, physical disuse has been presented as one of the factors that perpetuate chronic pain in the fear avoidance model (Verbunt et al. 2003a; Verbunt et al. 2003b; Vlaeyen et al. 1995b). Because of catastrophizing, the patient experiencing pain might become fearful, start to diminish activities and then become more and more deconditioned, disabled and depressed.

Despite the long existence of the deconditioning theory, it is still not clear whether symptoms of physical deconditioning, especially a reduced level of aerobic fitness, develop or even exist in patients with CLBP. A number of studies have investigated aerobic fitness in patients with CLBP with conflicting results (Nielens 2003; Simmonds 2002; Verbunt et al. 2003a; Wittink 1999; Wittink et al. 2000a). However, authors have been consistent in suggesting that differences in levels of physical activity may contribute to their discrepant findings. For instance, Nielens et al. (2001) and Wittink et al. (2000a) found male-female differences in VO_2max and suggested that these differences are attributable to different levels of physical activity, especially regarding work, household, sport, and leisure time. Data to confirm their hypothesis are not available to date. Although it is known that VO_2max has a moderate correlation with sport activity (Jacobs et al. 1993), no study has been published in which patients and healthy controls are matched for their level of sport activities before comparing their aerobic fitness levels. As it seems plausible that sport and physical activity matter to the level of aerobic fitness, it is important that studies in patients with CLBP incorporate information about the preceding as well as the current level of physical activities (Nielens 2003; Verbunt et al. 2003a; Wittink et al. 2000a; Wittink et al. 2000b).

According to the deconditioning theory and fear avoidance model of pain, other relevant factors related to the level of aerobic fitness might be pain intensity, the duration and degree of disability, and fear of injury. Controversy exists as some cross-sectional studies showed a correlation between these factors and the level of aerobic fitness and others did not (Hurri et al. 1991; McQuade et al. 1988; van der Velde and Mierau 2000; Verbunt et al. 2003b; Wittink et al. 2000a).

We hypothesized that 1) patients with CLBP have a reduced level of aerobic fitness after adjustment for age, gender and level of sport activity as compared with healthy controls; and 2) if there is a difference in the level of aerobic fitness between the CLBP patients and their healthy controls, this difference is associated with pain intensity, duration and degree of disability, fear of injury, and especially with the level of activities at work including household, and leisure time.

Methods

Participants and setting

A total of 115 patients of the eligible 150 patients who were referred by their general practitioner or medical specialist during the period of April 2002 until September 2003 for treatment to an outpatient unit of three rehabilitation centres in the province of Noord-Brabant in The Netherlands agreed to participate in a randomized controlled trial to evaluate the effectiveness of different treatments for CLBP disability. The inclusion criteria were: 1) first referral to a rehabilitation centre to reduce disability due to nonspecific low back pain existing at least 3 months, 2) age between 18 and 65 years, and 3) ability to walk at least 100 m. Exclusion criteria were lumbar disc herniation with neurological symptoms, inflammatory or neoplastic disease, fracture of the spine, spondylolisthesis or spondylolysis, lumbar spondylodesis, major psychiatric, cardiac and/or pulmonary conditions, severe addiction to drugs, narcotics or alcohol, and pregnancy. A total of 35 patients did not participate in the randomized controlled trial. The reasons for not participating were as follows: not willing to participate in this research (n = 4), not meeting criteria (n = 11), preference for a particular treatment (n = 10), logistic problems to attend treatment (n = 8), waiting time too long (n = 2).
For this study on the level of aerobic fitness, 7 of the 115 patients who agreed to participate in the randomized controlled trial were excluded because they were taking medication that influenced heart rate (e.g. ß-blockers). A total of 108 patients remained. All data were obtained during a prerandomization testing session. Patients completed questionnaires and performed a submaximal bicycle test and several performance tasks. In this study, the results of several questionnaires and the submaximal test will be presented. All patients gave written informed consent to participate. The medical ethics committee of the Rehabilitation Foundation Limburg and Institute for Rehabilitation Research at Hoensbroek in The Netherlands approved the research protocol.

Testing procedure

Modified submaximal Åstrand bicycle test

To predict the maximum oxygen consumption (VO_2max in mL/kg lean body mass/min. [mL/kg LBM.min^{-1}]), a modified submaximal Åstrand bicycle test was performed. In this test, developed for CLBP patients by the Department of Rehabilitation of the University Hospital Groningen, The Netherlands, the workload is gradually increased. Sufficient test-retest reliability (r = 0.87) and validity (r = 0.84, when compared with maximal VO_2-uptake measured with a maximal exercise test) have been established (Hodselmans et al. 2001; Schotanus 1999). Patients' LBM was measured according to the protocol of Durnin and Womersley (1974) by using a skinfold caliper (Servier Nederland BV, Leiden, The Netherlands). The participants performed the test on a calibrated cycle ergometer (Tunturi E3, Turku, Finland; Jaeger ER 800, Bitz, Germany; Lode Concord 1.0, Groningen, The Netherlands). The heart rate was monitored by a monitor placed on the patients' chest (Polar Favor, Kempele, Finland). The patients started cycling with a workload of 0.5 W/kg LBM at a constant rate of 60 rpm. After 2 minutes the workload was increased to 1.5 W/kg LBM. If the heart rate was still below 120 beats/min the workload was increased to 2.0 W/kg LBM and, when

necessary to 2.5 W/kg LBM. When the heart rate exceeded 120 beats/min, the patient cycled 6 minutes with a fixed workload to reach a steady state phase, meaning that the heart rate did not vary more than ± 5 beats/min during the last minute of exercise. The average heart rate during the last minute was calculated. The VO_2 max was estimated by using the Åstrand's nomogram based on the linear association between heart rate and increase in oxygen uptake (Åstrand and Rohdahl 1986).

If the heart rate during the last minute varied more than ± 5 beats/min, no VO_2max could be estimated. The test was stopped if the patient did not reach a heart rate of at least 120 beats/min, the heart rate exceeded the predefined maximum rate ([220 – age] x 0.85), the blood pressure reached the level of 220/115 mm Hg, or if the patient showed signs of serious cardiovascular or pulmonary difficulties.

Normative data

Normative data on the aerobic fitness level of healthy Dutch people from all over the country were gathered from 1991 until 2003 using the submaximal Åstrand bicycle test (Vos 2004). This database includes 18,082 healthy people 17 to 70 years of age. All healthy people were already classified into groups specified by age, gender and the level of sport activity during the preceding half year. People undertaking 1 to 2 hours of sport weekly for at least half a year were categorized as 'recreational', and those sporting less than that or not at all were categorized as 'untrained'.
In the normative dataset, mean values and standard deviations (SD) of VO_2 max in mL/kg LBM.min^{-1} are available for groups stratified by gender, age and level of sport activity.

To match patients and controls for sport activity during the preceding half year, the patients completed the Baecke Physical Activity Questionnaire (BPAQ) (Baecke et al. 1982). Based on the answers on the questions concerning sport activity the patients were stratified using the same criteria as used for the normative sample.

Instruments, reliability, and validity

Demographic data and clinical characteristics

For each patient the age, gender, duration of complaints and disability, extent of radiation of pain to leg, and the history of back surgery and/or trauma were recorded.

Pain

A 100-mm-long visual analogue scale with on the left side 'no pain' and on the right side 'unbearable pain' was used to measure the pain intensity. Relevance, validity, and reliability have been sufficiently tested for patients with low back pain (Carlsson 1983; Revill et al. 1976; Sriwatanakul et al. 1983).

Perceived disability
The Dutch version of the Roland Disability Questionnaire (RDQ) measures perceived low back pain disability. The questionnaire consists of 24 items with yes or no answers and total score ranging from 0 to 24. The higher the score, the more disabled a patient is. The RDQ is sufficiently valid and reliable in CLBP (Beurskens et al. 1996; Beurskens et al. 1995; Gommans et al. 1997; Roland and Morris 1983; Stratford et al. 1994; Stratford et al. 1998). In addition, the Dutch version of the Quebec Back Pain Disability Scale (QBPDS) was used (Kopec et al. 1995; Kopec et al. 1996). This questionnaire consists of 20 items with a total score ranging from 0 to 100, with 0 meaning not being disabled and 100 being maximal disabled. The validity and reliability for the Dutch version in CLBP are good (Schoppink et al. 1996).

Fear of injury
The Dutch version of the Tampa Scale for Kinesiophobia (TSK), which measures fear of injury and movement and consists of 17 items with a total score ranging from 17 to 68, was used. A higher score on the TSK indicates more fear. The questionnaire is considered reliable and valid in CLBP (Goubert et al. 2000; Vlaeyen et al. 1995a).

Level of activity at work and during leisure time
The BPAQ is used to quantify the amount of physical activity during the preceding year (Baecke et al. 1982; Jacobs et al. 1993). This questionnaire consists of 19 items addressing the three main types of physical activity; work, sport and leisure time. For each type of activity, an index is calculated. For each question, the patient is asked to score on a 5-point Likert-scale ranging from 'never' to "always' or 'very often'. In addition, the patient has to report his main occupation, which is then categorized into light, moderate, or heavy work according to the level of energy expenditure. The work index is calculated by adding the work-intensity score to the score of seven questions regarding work-related sitting, standing, walking, heavy lifting, tiredness, sweating, comparison with others of same age and dividing this total score by eight. The leisure time index is calculated by summing the scores of four questions about watching television, walking and cycling and dividing this score by four. The reliability (Baecke et al. 1982; Jacobs et al. 1993; Pereira et al. 1997; Philippaerts and Lefevre 1998), and validity in healthy populations appear to be good (Philippaerts et al. 1999) and sufficient for patients with at least 1 month of low back pain (Jacob et al. 2001).

Since household activities might be responsible for maintaining a higher level of aerobic fitness, we categorized household duties exceeding 5 hours a week as work. Because the number of working hours is not assessed in the BPAQ, patients were additionally asked how many hours a week they worked and performed household activities. An alternative work index was calculated by multiplying the work-intensity score with the amount of hours of work or household activities per week. It was decided that all work and household activities performed during the 3 months previous to the testing contributed to the alternative work index; for example, when a person was sick listed for 4 weeks he was assigned the full score for 8 weeks and no score for 4 weeks on the alternative work index. This calculation is in accordance with the way the sport index is calculated in the BPAQ.

Statistical analysis
For all variables, the mean and SD were calculated. Because only the mean and SD of VO_2max of healthy controls were available per group, matching on an individual basis was not possible. For each patient the following formula was used: $Z_{patient} = (VO_2max\ observed_{patient} - VO_2max\ expected_{group})/SD_{group}$.

Next, all $Z_{patient}$ scores were summarized (Z_{total}) and SD_{Ztotal} and Standard Error$_{Ztotal}$ (SE_{Ztotal}) were calculated. To test the null hypothesis that patients have the same level of aerobic fitness as their healthy controls ($Z_{total} = 0$), a one sample Student's *t*-test with a two-sided alpha of 0.05 was performed. Based on an estimated SD of 13.35 (50% higher than the mean SD of the healthy controls), 75 patients are needed to give a power of 90% to detect a real difference of 5 mL/kg LBM.min^{-1}. Comparisons between two groups (men and women, those who completed the test and those who stopped the test because of pain/fatigue) were performed by using the Student's *t*-test for unpaired observations with a normal distribution of the data. In case of non-normal distribution of the data the Mann-Whitney U test for unpaired observations was used. For normative data, χ^2 tests were used.

Multiple linear regression analysis (stepwise regression with backward elimination) was performed to define the contribution of independent variables to the difference in observed versus expected level of aerobic fitness in CLBP patients. The $Z_{patient}$ score was the dependent variable. Visual analogue scale-pain, RDQ, duration of disability, TSK, BPAQ-work index and BPAQ-leisure index were the six independent variables. In addition alternative regression analysis was performed using the QBPDS instead of the RDQ and using the alternative work index instead of the BPAQ-work index. Standardized beta coefficients and significance were tested under the null hypothesis that the coefficient did not differ from zero. For performing the multiple linear regression analysis with 6 independent variables, the number of the variables times ten (60 patients) are needed, as recommended for multiple regression analysis by Dawson-Saunders and Trap (1998). All statistical analyses were performed with SPSS software, version 11.5.

Results

Demographic data and clinical characteristics
Data for the whole sample and specified for men and women are presented in table 1. Except for weight and length ($P < 0.001$), there were no statistically significant differences between men and women. According to the scores on the RDQ, most patients were moderate to severely disabled, and more than half of them was not working at all. This is in accordance with the Dutch state of affairs in which usually moderately to severely disabled patients are treated in outpatient rehabilitation centres for their CLBP disability (VRIN/VRA 2000).

Table 1: Demographic data and clinical characteristics

Variables	Total Sample (n=108)	Men (n=63) (58%)	Women (n=45) (42%)
Age (yr)	41 ± 10 (20-61)	42 ± 9 (20-56)	39 ± 11 (20-61)
Weight (kg)	81 ± 15 (49-133)	87 ± 14 (57-133)	72 ± 11 (49-101)*
Height (cm)	177 ± 9 (157-200)	182 ± 6 (170-200)	169 ± 7 (157-184)*
Duration of LBP (mo)	62 ± 76 (3-396)	67 ± 80 (3-396)	54 ± 71 (6-384)
Radiation of pain (%)			
no radiation	10.0	11.0	9.0
above knee	34.0	33.0	35.0
below knee	56.0	56.0	56.0
Duration of functional limitations (mo)	35 ± 39 (3-240)	38 ± 42 (3-240)	32 ± 35 (3-180)
Previous back surgery (%)	19.0	18.0	22.0
Trauma preceding LBP (%)	19.0	18.0	20.0
Work status (%)			
full time	31.0	29.0	33.0
modified hours	9.0	10.0	9.0
modified work	7.0	8.0	5.0
full sick leave	23.0	22.0	24.0
disability payment	21.0	25.0	16.0
no job / retired	9.0	6.0	13.0
Sport activity (%)			
untrained	75.0	81.0	67.0
recreational	25.0	19.0	33.0
Disability (RDQ)	14.2 ± 3.9 (3-21)	14.0 ± 4.3 (3-21)	14.4 ± 3.3 (7-20)

Values are mean ± SD (range) or percentage. LBP = low back pain; RDQ = Roland Disability Questionnaire. * P < 0.001.

Table 2: Comparison between participants and nonparticipants

	Participants (n=115)	Nonparticipants (n=35)	Significance
Age (yr)	41 ± 10 (20-62)	39 ± 11 (19-59)	0.36
Gender (% male)	57.0	60.0	0.72
RDQ	14.1 ± 3.5 (5-21)	12.8 ± 4.6 (6-20)	0.16

Values are mean ± SD (range) or percentage. RDQ = Roland Disability Questionnaire.

For the patients who did not participate, data concerning the age, gender and level of disability at the moment they were referred for participation in the trial were available. Comparison between those who did not (n = 35) and those who did participate (n = 115) showed no significant difference regarding age, gender and level of disability (table 2).

Level of aerobic fitness

For 84 patients (78%) the VO_2max could be calculated, but 13 patients (12%) stopped the test prematurely due to pain or fatigue, and in 11 patients (10%) the VO_2 max could not be calculated due to medical (exceeding predefined heart rate, reaching predefined blood pressure or paleness of patient, n = 5) or technical reasons (no steady state phase reached, incomplete skinfold measures, n = 6).

The results of the 84 patients for whom the VO_2max could be calculated are presented in table 3. The CLBP patients had a significantly lower level of aerobic fitness compared with their matched healthy controls, with a mean lower VO_2max of 10.0 mL/kg LBM.min^{-1} (20%) in men and 5.6 (11%) in women. Men with CLBP had a significantly greater difference in observed versus expected level of aerobic fitness than women with CLBP. Furthermore, 72 patients (86%) had a lower level of aerobic fitness than was predicted on the basis of their gender, age and level of sport activity. Only 12 patients (14%) had an equal or higher level of aerobic fitness compared with the healthy control group.

Determinants of the difference between observed and expected level of aerobic fitness

Multiple linear regression analysis as well as the alternative multiple regression analyses showed that none of the hypothesized determinants (pain, level and duration of disability, fear of injury and activity level during work, household and leisure time) was significantly associated with the difference between the observed and expected level of aerobic fitness of the patients with CLBP (table 4).

Difference between completed and prematurely stopped submaximal test

The patients who stopped the test prematurely due to pain or fatigue were significantly older and more disabled (higher RDQ and QBPDS scores) but were not significantly different regarding their level of activity during work and leisure time (table 5). Also gender, fear of injury, pain, level of radiating pain, or duration of symptoms did not significantly differ between the groups, although it should be taken into account that the number of patients that stopped the submaximal test prematurely was small.

Table 3: Results of modified submaximal Åstrand bicycle test

	Observed VO_2max (ml/kg LBM.min^{-1})	Expected VO_2max (mL/kg LBM.min^{-1})
CLBP patients		
Total (n=84)	42.1 ± 8.1 (18.6-61.2)	50.2 ± 5.0 (41.8-60.5)
Men (n=46)	40.0 ± 7.9 (18.6-58.4)	50.0 ± 4.4 (43.7-59.7)
Women (n=38)	44.8 ± 7.7 (31.1-61.2)	50.4 ± 5.6 (41.8-60.5)
	Mean Z-score (95% CI)	Significance
CLBP patients vs. controls		
Total (n=84)	-0.87 (-1.06 to -0.69)	< 0.001
Men (n=46)	-1.06 (-1.30 to -0.82)	< 0.001
Women (n=38)	-0.64 (-0.93 to -0.35)	< 0.001
Within CLBP patients		
Difference between men and women	-0.42 (-0.78 to -0.05)	0.03

Values are mean ± SD (range) for VO_2max data. Expected VO_2max is based on normative data, with each patient being matched for age, gender, and sport activity.
LBM = lean body mass; CI = confidence interval; CLBP = chronic low back pain.

Table 4: Multiple linear regression analysis for Z as dependent variable and pain, duration and level of disability, activities during work and leisure time, and fear of injury as independent variables

Independent variable	Standardized ß	F ratio	Significance
		0.31	0.93
Pain	0.06		0.70
Duration of disability	-0.10		0.39
Activity at work (BPAQ)	0.03		0.82
Activity during leisure time	0.01		0.91
Perceived disability (RDQ)	-0.02		0.90
Fear of injury	-0.07		0.62
		0.32	0.92
Pain	0.06		0.67
Duration of disability	-0.10		0.41
Activity at work and household (alternative work score)	0.04		0.76
Activity during leisure time	0.02		0.90
Perceived disability (RDQ)	-0.01		0.94
Fear of injury	-0.07		0.54

BPAQ = Baecke Physical Activity Questionnaire; RDQ = Roland Disability Questionnaire; QBPDS = Quebec Back Pain Disability Scale.
Regression analyses with QBPDS instead of RDQ showed similar results.

Table 5: Comparison of patients with completed versus prematurely stopped submaximal test

Variables	Completed test (n=84)	Prematurely stopped test because of pain/fatigue (n=13)	Significance
Gender (% male)	55.0	69.0	0.33
Age (yr)	39 ± 10 (20-57)	46 ± 7 (31-58)	0.02*
Pain (VAS, in mm)	51 ± 24 (1-95)	64 ± 24 (16-99)	0.11
Duration of LBP (mo)	62 ± 80 (3-396)	84 ± 69 (6-200)	0.12
Radiation of pain (%)			0.69
no radiation	8.0	23.0	
above knee	37.0	15.0	
below knee	55.0	62.0	
Duration of disability (mo)	35 ± 40 (3-240)	45 ± 44 (3-120)	0.65
Back surgery (%)	19.0	31.0	0.33
Trauma preceding LBP (%)	20.0	8.0	0.28
RDQ	13.6 ± 4.0 (3-21)	16.3 ± 2.8 (9-19)	0.01*
QBPDS	44.4 ± 15.0 (2-77)	54.5 ± 14.4 (28-71)	0.02*
TSK	38.4 ± 6.9 (26-57)	40.8 ± 5.7 (33-51)	0.17
BPAQ-work index	2.15 ± 1.16 (0.75-4.38)	2.47 ± 1.07 (0.75-3.75)	0.36
Alternative work index	37.1 ± 49.1 (0-210)	26.2 ± 24.9 (0-84)	0.95
BPAQ-leisure index	2.90 ± 0.63 (1.5-4.25)	2.58 ± 0.65 (1.5-3.5)	0.14

Values are means ± SD (range) or percentage. VAS = visual analog scale; LBP = low back pain; RDQ = Roland Disability Questionnaire; QBPDS = Quebec Back Pain Disability Scale; TSK = Tampa Scale for Kinesiophobia; BPAQ = Baecke Physical Activity Questionnaire.
*P < 0.05.

Discussion

This study showed a significantly lower level of VO_2max in patients with CLBP compared with healthy controls matched for age, gender, and level of sport activity during the preceding half year. The overall mean difference was 8 mL/kg LBM.min^{-1}, equivalent to 2.3 METS (Metabolic EquivalenT). This is a clinically relevant lower aerobic fitness level considering that most household activities have an energy cost of two METS more than standing or working at a desk. The finding that 86% of all CLBP patients had a lower observed than expected level of VO_2max additionally emphasizes that a lower level of aerobic fitness is present in many patients with CLBP.

No study could be found in which patients with CLBP were compared with controls matched for their level of sport activity. This study is the first to show significantly lower levels of aerobic fitness in CLBP when comparing them to matched healthy controls. The earlier studies differed greatly in their testing method, reported data in different units, included dissimilar populations and reported by gender or for the whole population. When no appropriate controls were available, often less relevant normative data were used for comparison. Schmidt (1985), for example, only used total testing time and heart rate as measure of aerobic fitness level. Several other studies included patients with recurrent or previous back pain and only a small proportion had CLBP (Battie et al. 1989; Kellett et al. 1991). One study did not present results separately for gender (Davis et al. 1992). Several studies used normative data that were not valid (Davis et al. 1992; Harkapaa et al. 1989; Hurri et al. 1991). The studies that used valid controls showed some resemblance with our data (Keller et al. 2001; Nielens and Plaghki 1991; Nielens and Plaghki 1994; Nielens and Plaghki 2001; van der Velde and Mierau 2000). In the studies of Nielens and Plaghki (1991; 1994; 2001), using a submaximal bicycle or step test, men with CLBP had a lower level and women a normal level of aerobic fitness although only 50% to 60% of the sample had CLBP, mostly chronic lumbo-radicular pain. Wittink et al. (2000a; 2000b) found that men with CLBP while performing a maximal treadmill test had a aerobic fitness level equal to sedentary healthy men and women with CLBP had a level equal to active healthy women. In the studies of van der Velde and Mierreau (2000) and Keller et al. (2001) using a submaximal step test and bicycle test, respectively, patients with CLBP had a significantly lower level of aerobic fitness.

This study has some potential weaknesses. By using the nomogram to calculate the VO_2max, one may tend to under- or overestimate VO_2max by 15% in normal subjects (Åstrand and Rohdahl 1986). The use of a maximal test with direct calorimetry would give more accurate VO_2max values, although the validity and reliability of the maximal testing can be questioned because maximal testing is strongly influenced by motivation, fear and pain (Nielens 2003; Nielens et al. 2002). In congruence with this statement, analysis of the data of the patients who underwent maximal treadmill testing in the study of Wittink et al. (2000b), showed that the average CLBP patient did not reach the criteria of maximal performance (Nielens 2003). Based on these results and because the normative data were already collected by using the Åstrand submaximal bicycle test, it seemed reasonable to use a submaximal bicycle test for the patients.

Although the reliability of the submaximal Åstrand bicycle test is reported to be very good in CLBP patients (Keller et al. 2001), serious problems were expected when using this test in the present study population. In the submaximal Åstrand test, the workload is intensively increased during the first 1 to 2 minutes until the heart rate exceeds 120 beats/min. The research assistant determines the height of the workload on the basis of the increase of the heart rate and not according to a predefined increase of workload, making it difficult to choose the right workload. In daily practice we noticed that many CLBP patients could not finish the Åstrand test because the initial workload was too high. Since reducing the workload is not allowed once the Åstrand test is started, we expected that we could not calculate the VO_2max for many patients, which would reduce the power of our study.
The modified Åstrand submaximal test we used was particularly developed to test CLBP patients presented for rehabilitation and proved to be reliable and valid (Hodselmans et al. 2001; Schotanus 1999). Since the patients are tested by the modified test and the healthy controls by the Åstrand submaximal test, the comparability of both tests is an important issue. In both tests the patient has to reach a steady state phase in which the heart rate does not vary more than ± 5 beats/min. The calculation of the mean heart rate during the last minute of this steady state phase and the extrapolation of the corresponding VO_2max is exactly the same for both tests. The only difference is that in the modified test the workload, at which the patient has to cycle during 6 minutes, is reached in 2 to 6 minutes instead of the 1 to 2 minutes used in the Åstrand test. Although we did not perform a comparability study of both tests in CLBP patients, from an exercise physiology point of view it is not expected that the slower increase of workload leads to a higher or lower mean heart rate during the steady state phase. Based on this, we think that the results of the calculated VO_2max of both tests are comparable.
A reasonable percentage (80%) of patients with a moderate to high degree of disability was able to perform the submaximal testing. Older patients with a higher level of disability had more difficulty in performing the modified submaximal test.
A potential selection bias between patients who participated and those who did not participate in this study is not likely because no significant differences regarding age, gender and level of low back pain disability at the moment of referral were present.

By collecting data of 84 patients, we ensured sufficient power to test our two hypotheses. Due to the small sample size of the subgroup of 13 patients that stopped the test prematurely due to pain or fatigue, no firm conclusions can be made about the significance of the variables we hypothesized to play a role, although the level of disability and age were significantly higher in the group that stopped the test prematurely. For these patients an even less robust increase of workload should probably be used.

The fact that men showed an even lower level of observed versus expected aerobic fitness than women is intriguing. Nielens and Plaghki (1991; 1994; 2001) postulated that sociocultural factors could account for these findings as women were thought to be generally more active once they were at home as they were still engaged in childcare and various household duties. In the present study, the level of work and household activities was specifically checked for by using the BPAQ complemented by questions regarding hours of work, sick leave, and household during the last 3 months. The level of these activities did not significantly differ between men and women. While performing multiple regression analyses, the level of physical activity during work, household and leisure time was not associated to the difference of the level of aerobic fitness at all, so the hypothesis of Nielens and Plaghki (1991; 1994; 2001) could not be confirmed. It is also intriguing that, in this cross-sectional study, we did not find any association between the variables pain intensity, level and duration of disability and fear of injury after adjusting for age, gender and level of sport activity. It should be kept in mind, however, that this is a cross-sectional study and the results should be treated with caution, especially regarding the cause-and-effect relationship between the postulated factors and the difference in the level of aerobic fitness.

Possible explanations for the reduced level of aerobic fitness might be that the intensity (lower speed, less power etc.), duration and frequency (decline in activities) of the activities performed by patients were much lower than before the low back pain started. In the questionnaires, we gathered no information regarding the quality and decline of the activities. Otherwise it is known that, although patients state that they are moderately or severely disabled and less active, they still perform activities on a rather normal level (Protas 1996; Verbunt et al. 2001). In this study, we showed that patients who stopped the submaximal test prematurely reported higher levels of disability but were not significantly different regarding the level of physical activity, although the number of those who stopped the test was small. It is known that the VO_2max has a moderate correlation with the sport index of the BPAQ in healthy persons but only a minor correlation with the work and leisure activity score (Jacobs et al. 1993). It might thus be possible that the questionnaires we used could not detect differences and changes in the level of activities that correlate with the lower level of aerobic fitness in CLBP. In future research, monitoring activities in daily life, preferably combined with methods to measure the total energy consumption, might provide more answers on how active patients really are (Philippaerts et al. 1999; Philippaerts et al. 2001; Verbunt et al. 2001; Westerterp 1999). Still, the potential decline of activities cannot be measured by these methods and we have to rely on self-report methods.

Another explanation might be that patients with CLBP already have a lower level of aerobic fitness level before developing CLBP, as it is known that the VO_2max is explained for 40% by genetic factors (McArdle et al. 2001). Prospective research, however, has not identified a lower aerobic fitness level or being physically inactive as a risk factor for developing CLBP (Battie et al. 1989; Picavet and Schuit 2003). In conclusion, no satisfactory explanation of the findings in this study and no proof for a part of the fear avoidance model of pain could be found. In order to get more insight in the development and impact of loss of the level of aerobic fitness, longitudinal studies should be performed in patients with acute low back.

Conclusion

This study provides evidence for an association between a lower level of aerobic fitness and chronic low back pain but does not support the associations with the 'usual suspects', namely fear of injury, pain, low level of activities, or duration and severity of disability.

Acknowledgements

The authors thank Frederieke Berendsen and all research assistants of the rehabilitation centres Blixembosch, Leijpark and Breda for their contribution in the assessment of patients, Ria Laros for providing administrative support, Hubert Schouten and Arnold Kester for their statistical advice.

References

• Abenhaim L, Rossignol M, Valat JP, Nordin M, Avouac B, Blotman F, Charlot J, Dreiser RL, Legrand E, Rozenberg S, Vautravers P. The role of activity in the therapeutic management of back pain. Report of the International Paris Task Force on Back Pain. Spine 2000;25:1S-33S.
• Åstrand P, Rohdahl K. Textbook of work physiology, physiological bases of exercise. New York: McGraw-Hill, 1986, pp. 354-91.
• Baecke JA, Burema J, Frijters JE. A short questionnaire for the measurement of habitual physical activity in epidemiological studies. Am J Clin Nutr 1982;36:936-42.
• Battie MC, Bigos SJ, Fisher LD, Hansson TH, Nachemson AL, Spengler DM, Wortley MD, Zeh J. A prospective study of the role of cardiovascular risk factors and fitness in industrial back pain complaints. Spine 1989;14:141-7.
• Beurskens AJ, de Vet HC, Koke AJ. Responsiveness of functional status in low back pain: a comparison of different instruments. Pain 1996;65:71-6.
• Beurskens AJ, de Vet HC, Koke AJ, van der Heijden GJ, Knipschild PG. Measuring the functional status of patients with low back pain. Assessment of the quality of four disease-specific questionnaires. Spine 1995;20:1017-28.
• Carlsson AM. Assessment of chronic pain. I. Aspects of the reliability and validity of the visual analog scale. Pain 1983;16:87-101.
• Davis VP, Fillingim RB, Doleys DM, Davis MP. Assessment of aerobic power in chronic pain patients before and after a multi disciplinary treatment program. Arch Phys Med Rehabil 1992;73:726-9.
• Dawson-Saunders B, Trapp RG. Basic and clinical biostatistics. Norwalk: Appleton & Lange, 1998, pp. 207-29.
• Durnin JVGA, Womersley J. Body fat assessed from total body density and its estimation from skinfold thickness: measurements on 481 men and women aged 16 to 72 years. Br J Nutr 1974;32:77-97.
• Gommans IHB, Koes BW, van Tulder MW. Validiteit en responsiviteit Nederlandstalige Roland Disability Questionnaire. Vragenlijst naar functionele status bij patiënten met lage rugpijn. Ned T Fysiother 1997;107:28-33.
• Goubert L, Crombez G, Vlaeyen JWS, van Damme S, van den Broeck A, van Houdenhove B. De Tampa schaal voor kinesiofobie: psychometrische karakteristieken en normering. Gedrag & Gezondheid 2000;28:54-62.
• Harkapaa K, Jarvikoski A, Mellin G, Hurri H. A controlled study on the outcome of inpatient and outpatient treatment of low back pain. Part I. Pain, disability, compliance, and reported treatment benefits three months after treatment. Scan J Rehab Med 1989;21:81-9.
• Hodselmans AP, Jaegers SM, Goeken LN. Short-term outcomes of a back school program for chronic low back pain. Arch Phys Med Rehabil 2001;82:1099-105.
• Hurri H, Mellin G, Korhonen O, Harjula R, Harkapaa K, Luoma J. Aerobic capacity among chronic low-back-pain patients. J Spinal Disord 1991;4:34-8.
• Jacob T, Baras M, Zeev A, Epstein L. Low back pain: reliability of a set of pain measurement tools. Arch Phys Med Rehabil 2001;82:735-42.
• Jacobs DR, Ainsworth BE, Hartman TJ, Leon AS. A simultaneous evaluation of 10 commonly used physical activity questionnaires. Med Sci Sports Exerc 1993;25:81-91.
• Jette DU, Jette AM. Physical therapy and health outcomes in patients with spinal impairments. Phys Ther 1996;76:730-41.
• Keller A, Hellesnes J, Brox JI. Reliability of the isokinetic trunk extensor test, Biering-Sorensen test, and Åstrand bicycle test: assessment of intraclass correlation coefficient and critical difference in patients with chronic low back pain and healthy individuals. Spine 2001;26:771-7.
• Kellett KM, Kellett DA, Nordholm LA. Effects of an exercise program on sick leave due to back pain. Phys Ther 1991;71:283-91.

References

- Kopec JA, Esdaile JM, Abrahamowicz M, Abenhaim L, Wood-Dauphinee S, Lamping DL, Williams JI. The Quebec Back Pain Disability Scale. Measurement properties. Spine 1995;20:341-52.
- Kopec JA, Esdaile JM, Abrahamowicz M, Abenhaim L, Wood-Dauphinee S, Lamping DL, Williams JI. The Quebec Back Pain Disability Scale: conceptualization and development. J Clin Epidemiol 1996;49:151-61.
- Mayer TG, Gatchel RJ. Functional restoration for spinal disorders: the sports medicine approach. Philadelphia: Lea & Febiger, 1988.
- McArdle WD, Katch FI, Katch VL. Exercise Physiology: energy, nutrition, and human performance. In: W McArdle, FI Catch, VL Katch. Exercise physiology. Baltimore: Lippincott Wiliams & Wilkins, 2001, pp. 222-48.
- McQuade KJ, Turner JA, Buchner DM. Physical fitness and chronic low back pain. An analysis of the relationships among fitness, functional limitations, and depression. Clin Orthop 1988:198-204.
- Nielens H. Physical capacity evaluation and reconditioning of patients with chronic pain. Bruxelles: Université Catholique de Louvain, Faculté de médicine, 2003.
- Nielens H, Cornet D, Rigot F. Maximal versus submaximal testing protocols to evaluate fitness of patients with chronic low back pain; abstract. Med Sci Sports Exerc 2002;34:5.
- Nielens H, Plaghki L. Evaluation of physical adaptation to exercise of chronic pain patients by step-test procedure. Pain Clin 1991;4:21-5.
- Nielens H, Plaghki L. Perception of pain and exertion during exercise on a cycle ergometer in chronic pain patients. Clin J Pain 1994;10:204-9.
- Nielens H, Plaghki L. Cardiorespiratory fitness, physical activity level, and chronic pain: are men more affected than women? Clin J Pain 2001;17:129-37.
- Pereira MA, FitzGerald SJ, Gregg EW, Joswiak ML, Ryan WJ, Suminski RR, Utter AC, Zmuda JM. A collection of physical activity questionnaires for health-related research. Med Sci Sports Exerc 1997;29:S3-205.
- Philippaerts RM, Lefevre J. Reliability and validity of three physical activity questionnaires in Flemish males. Am J Epidemiol 1998;147:983-90.
- Philippaerts RM, Westerterp KR, Lefevre J. Doubly labelled water validation of three physical activity questionnaires. Int J Sports Med 1999;20:284-9.
- Philippaerts RM, Westerterp KR, Lefevre J. Comparison of two questionnaires with a tri-axial accelerometer to assess physical activity patterns. Int J Sports Med 2001;22:34-9.
- Picavet HS, Schuit AJ. Physical inactivity: a risk factor for low back pain in the general population? J Epidemiol Community Health 2003;57:517-8.
- Protas EJ. Aerobic exercise in the rehabilitation of individuals with chronic low back pain: a review. Crit Rev Physical Rehab Med 1996;8:283-95.
- Revill SI, Robinson JO, Hogg MIJ. The reliability of a linear analog for evaluating pain. Anaesthesia 1976;31:1191-8.
- Roland M, Morris R. A study of the natural history of back pain. Part I: development of a reliable and sensitive measure of disability in low-back pain. Spine 1983;8:141-4.
- Schmidt AJM. Cognitive factors in the performance level of chronic low back pain patients. J Psychosom Res 1985;29:183-9.
- Schoppink LE, van Tulder MW, Koes BW, Beurskens SA, de Bie RA. Reliability and validity of the Dutch adaptation of the Quebec Back Pain Disability Scale. Phys Ther 1996;76:268-75.

References

- Schotanus J. Het meten van belastbaarheid bij patiënten met chronische lage rugklachten. Groningen: Rijksuniversiteit Groningen,The Netherlands, 1999.
- Simmonds MJ. Exercise and activity for individuals with nonspecific back pain. In: JN Myers, WG Herbert, R Humphrey. ACSM's resources for clinical exercise physiology: musculoskeletal, neuromuscular, neoplastic, immunologic, and hematologic conditions. Philadelphia: Lippincot Williams and Wilkins, 2002, pp. 125-38.
- Sriwatanakul K, Kelvie W, Lasagna L, Calimlim JF, Weis OF, Mehta G. Studies with different types of visual analog scales for measurement of pain. Clin Pharmacol Ther 1983;34:234-9.
- Stratford PW, Binkley J, Solomon P, Gill C, Finch E. Assessing change over time in patients with low back pain. Phys Ther 1994;74:528-33.
- Stratford PW, Binkley JM, Riddle DL, Guyatt GH. Sensitivity to change of the Roland-Morris Back Pain Questionnaire: part 1. Phys Ther 1998;78:1186-96.
- van der Velde G, Mierau D. The effect of exercise on percentile rank aerobic capacity, pain, and self-rated disability in patients with chronic low-back pain: a retrospective chart review. Arch Phys Med Rehabil 2000;81:1457-63.
- Verbunt JA, Seelen HA, Vlaeyen JWS, van der Heijden GJ, Heuts PH, Pons K, Knottnerus JA. Disuse and deconditioning in chronic low back pain: concepts and hypothesis on contributing mechanisms. Eur J Pain 2003a;7:9-21.
- Verbunt JA, Seelen HA, Vlaeyen JWS, van der Heijden GJ, Knottnerus JA. Fear of injury and physical deconditioning in patients with chronic low back pain. Arch Phys Med Rehabil 2003b;84:1217-32.
- Verbunt JA, Westerterp KR, van der Heijden GJ, Seelen HA, Vlaeyen JW, Knottnerus JA. Physical activity in daily life in patients with chronic low back pain. Arch Phys Med Rehabil 2001;82:726-30.
- Vlaeyen JWS, Kole-Snijders AM, Boeren RG, van Eek H. Fear of movement/(re)injury in chronic low back pain and its relation to behavioral performance. Pain 1995a;62:363-72.
- Vlaeyen JWS, Kole-Snijders AM, Rotteveel AM, Ruesink R, Heuts PHTG. The role of fear of movement/(re)injury in pain disability. J Occup Rehabil 1995b;5:235-52.
- Vos JA. Ergometrie en trainingsbegeleiding. Amersfoort: Nederlands Paramedisch Instituut, 2004.
- VRIN/VRA. Productdefiniëring Revalidatiezorg: registratiepilot revalidatie-instellingen januari 1999 - juli 2000. Utrecht: VRIN, 2000.
- Westerterp KR. Assessment of physical activity level in relation to obesity: current evidence and research issues. Med Sci Sports Exerc 1999;31:S522-5.
- Wittink H. Physical fitness, function and physical therapy in patients with pain: Clinical measures of aerobic fitness and performance in patients with low back pain. In: M Mitchell. Pain 1999- an updated review. Refresher course syllabus 9th world congress on pain. Seattle: IASP press, 1999, pp. 137-44.
- Wittink H, Hoskins Michel T, Wagner A, Sukiennik A, Rogers W. Deconditioning in patients with chronic low back pain: fact or fiction? Spine 2000a;25:2221-8.
- Wittink H, Michel TH, Kulich R, Wagner A, Sukiennik A, Maciewicz R, Rogers W. Aerobic fitness testing in patients with chronic low back pain: which test is best? Spine 2000b;25:1704-10.

Chapter 4

Active rehabilitation for chronic low back pain: Cognitive-behavioral, physical, or both? First direct post-treatment results from a randomized controlled trial [ISRCTN22714229].

Rob JEM Smeets, Johan WS Vlaeyen, Alita Hidding, Arnold DM Kester, Geert JMG van der Heijden, Tineke CM van Geel, J André Knottnerus

Published as:
Smeets RJEM, Vlaeyen JWS, Hidding A, Kester ADM, van der Heijden GJMG, van Geel ACM, Knottnerus JA. Active rehabilitation for chronic low back pain: Cognitive-behavioral, physical, or both? First direct post-treatment results form a randomized controlled trial [ISRCTN22714229]. BMC Musculoskeletal Disorders 2006;7:5.

Abstract

Background
The treatment of nonspecific chronic low back pain is often based on three different models regarding the development and maintenance of pain and especially functional limitations: the deconditioning model, the cognitive-behavioral model and the biopsychosocial model.
There is evidence that rehabilitation of patients with chronic low back pain is more effective than no treatment, but information is lacking about the differential effectiveness of different kinds of rehabilitation. A direct comparison of a physical, a cognitive-behavioral treatment and a combination of both has never been carried out so far.
Methods
The effectiveness of active physical, cognitive-behavioral and combined treatment for chronic nonspecific low back pain compared with a waiting list control group was determined by performing a randomized controlled trial in three rehabilitation centres.
Two hundred and twenty three patients were randomized, using concealed block randomization to one of the following treatments, which they attended three times a week for 10 weeks: Active Physical Treatment (APT), Cognitive-Behavioral Treatment (CBT), Combined Treatment of APT and CBT (CT), or Waiting List (WL). The outcome variables were self-reported functional limitations, patient's main complaints, pain, mood, self-rated treatment effectiveness, treatment satisfaction and physical performance including walking, standing up, reaching forward, stair climbing and lifting. Assessments were carried out by blinded research assistants at baseline and immediately post-treatment. The data were analyzed using the intention-to-treat principle.
Results
For 212 patients, data were available for analysis. After treatment, significant reductions were observed in functional limitations, patient's main complaints and pain intensity for all three active treatments compared to the WL. Also, the self-rated treatment effectiveness and satisfaction appeared to be higher in the three active treatments. Several physical performance tasks improved in APT and CT but not in CBT. No clinically relevant differences were found between the CT and APT, or between CT and CBT.
Conclusions
All three active treatments were effective in comparison to no treatment, but no clinically relevant differences between the combined and the single component treatments were found.

Background

Chronic nonspecific low back pain (CLBP) and the resulting functional limitations have become an epidemic health and socioeconomic problem (Meerding et al. 1998; van Tulder et al. 1995). Many models and therapies have been postulated and applied in order to reduce the burden placed on the individuals with CLBP and society (Abenhaim et al. 2000; Guzman et al. 2002; Morley et al. 1999; Ostelo et al. 2005; van Tulder et al. 2000; van Tulder et al. 1997). In the literature, three frequently used models regarding the development and maintenance of CLBP functional limitations are described:

1) The physical deconditioning model assuming that loss of muscle strength and endurance including aerobic capacity is responsible for reduced activity levels and hence functional limitations (Mayer et al. 1998; Verbunt et al. 2003).
2) The cognitive-behavioral model postulating that functional limitations result from maladaptive beliefs and avoidance behaviors that are maintained by learning processes (Fordyce 1976; Turk et al. 1983; Turk and Okifuji 2002; Vlaeyen et al. 1995).
3) The biopsychosocial model assuming that loss of functional abilities results from both the deconditioning and the cognitive-behavioral model (Waddell 1998).

There is growing evidence that strengthening exercises combined with aerobic exercises as well as cognitive-behavioral treatment (CBT) are worth the effort when compared to no treatment or waiting list control. But there is insufficient evidence for the effectiveness of strengthening and aerobic exercises versus other active therapies (Hayden et al. 2005a; Liddle et al. 2004; van Tulder et al. 2000), however just recently, Hayden et al. (2005b) showed that exercise therapy consisting of individually designed programs, including stretching and strengthening, improves pain and function. Furthermore, a recent review showed that there is moderate evidence for the strengthening of deep low back muscles (Smeets et al. 2006). Controversy exists regarding the effectiveness of CBT when compared to alternative active treatments (Morley et al. 1999; Ostelo et al. 2005; van Tulder et al. 2000). Multidisciplinary treatment of at least 100 hours, combining exercise therapy, functional restoration and CBT appeared promising in comparison to other non-multidisciplinary treatments, whereas multidisciplinary rehabilitation programs of less than 30 hours failed to prove improvements on several relevant outcome measures. It should be taken into account although, that there is no consensus about the content, intensity and frequency of the different training sessions and the results are based on a relative low number of studies (Guzman et al. 2002).

However, taking a closer look at the studies included in the meta-analyses, it appears that many therapies are not solely based on one of the three models mentioned above. For example in the studies regarding the effectiveness of CBT, exercise therapy was used in order to increase a patient's level of activity while applying the operant learning principles. As a result, it is not clear whether the improvement is reached by the CBT itself, the exercise therapy or the combination

of both. Furthermore, many exercise therapies were not of sufficient intensity, frequency and duration to fulfil the physiologic training principles and therefore should not be classified as real reconditioning or strengthening therapies (ACSM 1998; Haskell 1994; Hilde and Bo 1998; Liddle et al. 2004; McArdle et al. 1996). On the basis of the current available research, it appears that the evidence for the effectiveness of model-based treatments is still scarce and many programs used several different treatment techniques often based on several models without knowing what treatment elements or combinations are really necessary to reach positive treatment results.
Therefore we designed treatments that are exclusively based on the deconditioning model, the cognitive-behavioral model and the biopsychosocial model.

The aim of the current study was to compare the effectiveness of a physical treatment (APT), a cognitive-behavioral treatment (CBT) and a combination of both (CT) by means of a randomized controlled trial. It is hypothesized that all active treatments are more effective in reducing functional limitations compared to waiting list controls (WL). Furthermore, based on the biopsychosocial model it is assumed that the patients with CLBP are physically deconditioned or have to relearn healthy behaviors. Even both problems might be present. So it might be possible that patients only receiving physical treatment will enhance their aerobic capacity, muscle strength and endurance, but there might also be patients who do not resume their normal daily activities because of for example maladaptive beliefs or avoidance behavior.These problems might even hamper successful physical training. Otherwise people receiving cognitive-behavioral treatment might be willing to increase their activity level but physical deconditioning might prevent this. By combining both treatments it seems plausible that more people might decrease their level of functional limitations and therefore it is hypothesized that the combination of APT and CBT shows a larger difference than APT or CBT alone.This paper reports on the immediate post-treatment effects. One year follow-up results for the three active therapies will be presented later.

Methods

Study population

Between April 2002 and December 2004, patients for the first time referred by general practitioners and medical specialists to three outpatient rehabilitation centres in The Netherlands were invited by their consulting rehabilitation physician to participate. Inclusion criteria were: age between 18 and 65 years, nonspecific low back pain (CLBP) with or without radiation to leg for more than 3 months resulting in functional limitations (Roland Disability Questionnaire score > 3) (Roland and Morris 1983), ability to walk at least 100 meters without interruption. Exclusion criteria were: vertebral fracture, spinal inflammatory disease, spinal infections or malignancy, current nerve root pathology, spondylolysis or spondylolisthesis, lumbar spondylodesis, medical co-morbidity making intensive exercising impossible (e.g. cardiovascular or metabolic disease), ongoing diagnostic procedures or treatment for their CLBP at the time of referral or a clear treatment preference. Patients were requested to stop other treatments for their low back complaints,

except pain medication. The Symptom Checklist (SCL-90) (Arrindell and Ettema 1986) and the Dutch Personality Questionnaire (NPV) (Luteijn et al. 1985) were used to check for psychopathology that would hamper individual or group processes (Vlaeyen et al. 1996). Further exclusion criteria were: not proficient in Dutch, pregnancy and substance abuse that could interfere with the rehabilitation treatment.

To control for expectation bias, patients were told that the study was being performed to compare three currently used treatments for CLBP, of which the exact efficacy had not yet been established and in case they would be randomized to the WL, they would receive a treatment consisting of similar treatment components as in the trial.

The Medical Ethics Committee of the Rehabilitation Foundation Limburg and the Institute for Rehabilitation Research, Hoensbroek, The Netherlands approved the study protocol.

Randomization

To ensure balance with regard to the number of patients receiving a specific treatment, for each rehabilitation centre clusters of four consecutive patients were randomized using permuted blocks of size eight. For each rehabilitation centre a randomization list was generated by computer under supervision of an independent statistician. Before recruitment of patients the main researcher prepared sealed opaque envelopes for each rehabilitation centre and numbered them sequentially. Furthermore, in each rehabilitation centre the randomization list was handed over to the employee who was responsible for planning the treatments. After the start of recruitment of patients, this employee was the only one who had access to the randomization list. Once the research group had recruited four patients in a participating centre, this employee was informed and asked to make arrangements for the first assessment, followed by the start of the allocated treatment within one or two weeks. After the first assessment, the blinded research assistant handed over the sealed envelope to the patient. In order to make sure that the research assistant stayed blinded, the patient was asked not to open the envelope before leaving the building and under no circumstance tell the research assistant what treatment he was allocated to. The participating therapists, research assistants and referring physicians were not aware of this randomization procedure.

Interventions

The overall goal of the active treatments was to improve functioning (decrease of functional limitations). Emphasis was put on the responsibility of the patient for making plans to keep on being active after the treatment (generalization). Each treatment lasted 10 weeks and started with the explanation of the rationale of that particular treatment.
A written summary of the rationale was given to the patients. In order to assure sufficient contrast between the three different treatments and to avoid incorporating possible confounding elements, all therapists were instructed not to discuss general aspects concerning back pain origin, anatomy and ergonomics. In the fourth and tenth week, the rehabilitation physician responsible for the whole treatment, together with the patient evaluated the treatment and

checked the generalization plans.
No other interventions than those that were chosen for the APT, CBT or CT took place. In case of acute and severe psychosocial stress or pathology (severe depression, high risk for suicide or personal problems the patient did not wish to discuss during the group treatment), a consultation of a clinical psychologist or social worker was possible. During this consultation the therapist tried to find out what the exact problem was and consecutively, when judged necessary, arranged for professional help outside the rehabilitation centre.
All therapists received an extensive training before the start of the trial. They attended refresher courses; two one-day courses during the first year, and one each year in the next two years of the trial. The clinical psychologists and social workers had at least five years of experience in treating CLBP patients.

Active Physical Treatment (APT)
APT was based on the assumption that a reduced aerobic capacity and muscle deconditioning/disuse, especially of the deep lumbar extensor muscles (multifidus muscle) are present (Danneels et al. 2001). The duration and intensity of APT were chosen according to the physiologic principles of training (ACSM 1998). The APT consisted of aerobic training, and three dynamic-static strengthening exercises.
APT started with half an hour of *aerobic training* on a bicycle; 5-minute warming up, 20 minutes performing at 65% to 80% of the maximum heart rate (HRmax) followed by a 5-minute cooling down. Before training, the VO_2max was calculated based on a slightly modified submaximal Åstrand bicycle test (Åstrand and Rohdahl 1986). The target HR was calculated using the formula of Karvonen for patients with an aerobic fitness level lower than or equal to the non-trained Dutch population: HRtarget = HRrest + 50% to 60% (HRmax – HRrest) (ACSM 1998; Karvonen et al. 1957). For patients with an aerobic fitness level higher than the non-trained Dutch population the target rate was: HRtarget = HRrest + 55% to 65% (HRmax – HRrest).
During the training the patient judged the perceived exertion by using the Borg scale ranging from 6-20 (Borg 1990). When the patient scored above 14, the next HRtarget was lowered, with a score of 14, the middle level of the two calculated HRtargets was chosen, and with a score lower than 14, the upper HRtarget was aimed at (Dishman 1994). After two and four weeks the percentage in the abovementioned formula was increased by 5%. From week three on, the patient also had to sprint three times during one minute to achieve a HR calculated by increasing the percentage in the formula with an additional 10%. After the aerobic training the patient stretched the trunk and leg muscles during five minutes. *Three dynamic-static exercises* were performed at 70% of the 1-Repetition Maximum (1-RM), which allowed 15 to 18 repetitions until muscular fatigue occurred. Each repetition was performed in a standardized and controlled matter allowing two seconds of concentric movement, five seconds of static contraction and two seconds of the eccentric movement. The patient performed three sessions of 15 to 18 repetitions. The exercises started gradually with in the first week one exercise per session, two exercises in the second week and three in the third week. After five sessions of performing an exercise, the patient performed a test to establish the 70% of the 1-RM again. The three exercises consisted of leg extension while sitting on knees and hands, trunk lifting and lifting both legs while lying prone on a

couch. During the exercises assistive weight by a pulley system or extra weight placed on the body of the patient were used depending on the calculated load (70% of 1-RM). The increment of load was based on the performance of the patient and not on the judgment of the therapists. Only when a patient reported change in pain pattern (e.g. radiation to one leg), the rehabilitation physician was asked whether the training had to be adjusted (dynamic instead of dynamic-static, lowering of load or temporary not performing an exercise). After the dynamic-static exercises the patient stretched the trunk and leg muscles again during five minutes. Two physiotherapists guided a maximum of four patients at a time. Each total APT session took 1¾ hours and was given three times a week.

Cognitive-Behavioral Treatment (CBT)
CBT was based on the assumption that how individuals with chronic pain behave is a resultant of learning, both through environmental contingencies as through information processing (Morley et al. 1999). CBT was aimed to help patients to reach their individual daily life goals, to increase their activity level and to modify dysfunctional beliefs.
In this trial, CBT consisted of operant behavioral graded activity training (Fordyce 1976; Sanders 1996) and problem solving training (van den Hout et al. 2003).
During the graded activity (GA), the therapist focused on a time contingent gradual increase or pacing of activities being important and relevant for the patient's personal situation. The patient selected three activities that were of the highest importance but compromised by the pain problem. After the establishment of a baseline, the activity tolerance level was calculated and final treatment goals were set. The patient started performing the selected activities following quotas for each day, starting from 70% to 80% of the baseline with gradually increasing activity levels towards the final treatment goals. The patient was instructed only to perform the agreed amount of activity and not perform less or more, even when he felt capable of doing so. During the training sessions the patient performed one or more of the selected activities, but the most important part of the treatment session was the evaluation of the amount of the activities the patient had performed at home. The patient graphically registered in a personal diary his daily performance. The therapists were instructed to discuss these graphs regularly with the patient, while positively reinforcing any progress towards the pre-set goals. In order to create as much contrast as possible with the APT, no physical training element (e.g. muscle strength or aerobic exercises) was incorporated.
Graded activity consisted of two introductory group meetings followed by 18 individual sessions guided by a skilled physiotherapist or occupational therapist. The frequency of the sessions gradually decreased from three to one session a week. In total 11½ hours of treatment were given. The partner was invited to attend the first session and a session in the fourth week of treatment.
The problem *solving training* (PST) started with three initial sessions in which the rationale of training and the skill of positive problem orientation were discussed. Sessions four to ten focused on problem definition and formulation, generation of alternatives, decision-making, implementation and evaluation. Patients received a course book with additional information, a summary of each session and homework assignments. The training of the skills and application were the main focus of the training, rather than one specific problem area. Patients were free to select their

own personal problem areas. After each session, homework was provided in order to practice the skills in everyday life. A clinical psychologist or social worker, specifically trained to guide this training, provided 10 sessions of 1½ hours to a maximum of four patients at a time.

Combined Treatment (CT)
According to the biopsychosocial approach, CT aimed at restoring functional ability through increased fitness, the reinforcement of health behaviors and the modification of dysfunctional beliefs. CT consisted of APT in combination with PST, both in the same frequency and duration as described before. The patient was told that he first had to gain enough aerobic fitness and muscle strength before increasing his activities. The GA was not started until the third week, and began with the selection of the three relevant activities. By the end of the fourth week the final goals and daily quota were set. In total 19 sessions, with a total duration of 11 hours were given.

Waiting List (WL)
The patients assigned to the WL were requested to wait 10 weeks after which they were offered a regular individual rehabilitation treatment. During the waiting period, patients were not allowed to participate in diagnostic or therapeutic procedures because of their CLBP.

Assessment
Assessments (questionnaires and physical performance tasks) were carried out before treatment and immediately after ten weeks of active treatment, and six and twelve months after completion of the treatment. They were supervised and carried out by blinded research assistants who received a special training and who attended regular refresher courses two or three times a year. The WL patients were only assessed before and after ten weeks of waiting.

Baseline data
During the pre-treatment assessment data were collected on age, gender, level of education, employment status, duration of complaints and functional limitations, previous low back surgery, previous treatment, level of radiation of pain to leg, traumatic onset of low back pain, fear of injury and movement (Tampa Scale for Kinesiophobia; TSK) (Goubert et al. 2000; Vlaeyen et al. 1995) and physical activity (Baecke Physical Activity Questionnaire; BPAQ) (Baecke et al. 1982; Jacob et al. 2001).

Primary outcome measure
The level of low back pain associated functional limitations was measured by the Roland Disability Questionnaire (RDQ) which proved to be a valid and reliable instrument in the evaluation of chronic low back pain treatment (Beurskens et al. 1996; Gommans et al. 1997; Roland and Morris 1983).

Secondary outcome measures: questionnaires
1) The severity of the three patient-specific main complaints by using the approach method from Tugwell et al. (1987; Beurskens et al. 1999). At baseline the patient selected three activities he performed frequently, which he perceived as important in his daily life, and which LBP made difficult for him. The severities of these main complaints were rated on a 100-mm visual analogue scale (VAS). This is a valid and reliable method with sufficient responsiveness (Beurskens et al. 1996; Ostelo et al. 2004).
2) Current pain by using a 100-mm VAS for pain at this moment and the Pain Rating Index (total score) of the McGill Pain Questionnaire, a reliable measure of pain intensity (Melzack and Katz 1992; van der Kloot et al. 1995).
3) Depression measured by the Beck Depression Inventory (Beck et al. 1979), a reliable, valid and widely used questionnaire (Beck et al. 1988).
4) Patient's global assessment of overall result measured by a transitional seven-point ordinal scale (1= vastly worsened, 7 = completely recovered) (Beurskens et al. 1996; Ostelo et al. 2004).
5) Treatment satisfaction measured by using a 100-mm VAS for the overall treatment provided to the patient.

Secondary outcome measures: physical performance tasks
Six performance tests were selected out of several performance task batteries described in detail by Simmonds et al. (1998), Harding et al. (1994) and Mayer et al. (1988; 1990). All seem to have a fairly good validity and reliability in healthy persons and pain patients and most of them also in CLPB patients. The tests include: 1) five-minutes walking (meters), 2) fifty-foot walking (seconds), 3) five times sit to stand, performed twice; average time needed to perform a series of five (seconds), 4) loaded forward reaching by holding a stick with a weight of 2.25 or 4.5 kg in front of the body at shoulder height and extend as far as possible (centimetres), 5) one-minute stair climbing (number of stairs), 6) PILE-test weight lifting from floor to waist; the patient has to lift a box with a weight four times within 20 seconds from floor up to a 75 cm high table. After each round of four lifting cycles the weight was increased in a standardized way. The test was stopped when the patient could not lift the weight four times within 20 seconds, the HR exceeded 85% of the maximal HR (0.85 x [220 – age]), or the maximal amount of the weight that could be lifted safely (0.6 x body weight) was reached or the research assistant considered the lifting unsafe. The total number of fully completed cycles of lifting was registered. The research assistants were specifically instructed not to encourage the patient to increase his effort. The patient was only asked to perform the tasks as quickly as possible or to walk or reach as far or lift as much as possible.

Manipulation check
In order to check whether aerobic capacity and problem solving skills were exclusively manipulated in the APT/CT and CBT/CT, respectively, the following assessments were carried out:
1) Predicted VO_2max in ml per kg lean body mass by using a modified Åstrand submaximal bicycle test (Hodselmans et al. 2001; Schotanus 1999). It is hypothesized that the aerobic training increases the VO_2max.
2) Problem solving skills by using a recently validated short form of the Social Problem Solving Inventory-Revised (SPSI-R) which consists of three scales: rational problem solving (RPS), negative problem orientation (NPO); and impulsive/careless style (IMP) (D'Zurilla et al. 1997). It is hypothesized that the problem solving training decreases the NPO.

Treatment compliance and co-interventions
In order to check whether patients were compliant with the allocated treatment, each therapist kept records on the presence during treatment, the amount of exercise (duration, intensity of exercising by monitoring HR during cycling and amount of repetitions and weight displaced during muscle training) and choice of activities and increase in time of these activities for the GA. Also adverse effects and extra appointments with the rehabilitation physician and therapists were registered. Furthermore, cost diaries were introduced at the first assessment and patients were asked to fill these out during the treatment period. This was an additional way to check for extra appointments at the rehabilitation centre and to check whether patients received additional diagnostic or therapeutic procedures outside the centre during the treatment period.
It was decided that each patient not attending at least 2/3 of all possible treatment sessions for each training element would be classified as having a protocol deviation. Furthermore, during a consensus meeting, two members of the research team, blinded for the allocated treatment examined protocol deviations reported in the cost diaries. Relevant protocol deviations were defined as; visit to chiropractor or physiotherapist performing manipulation, more than one visit to a regular physiotherapist not performing manipulation, visit to general physician or medical specialist other than just for diagnostic questions (e.g. prescription of other pain medication or facet joint injection) or more than two visits to alternative medicine.

Sample size
A difference of 2.5 points change in score of the primary outcome measure (RDQ) between each of the three active treatment groups and the waiting list group was considered to be clinically relevant (Roland and Fairbank 2000). Based on a 2-sided α of 0.05 and a 1-ß of 0.90, with a standard deviation (SD) of the RDQ change of 4, a minimum of 220 patients (55 patients per group) needed to be recruited.

Statistics
To compensate for possibly skewed randomization results, demographic and baseline variables and outcome measures at pre-test were compared between treatment groups. Variables for which differences between groups at baseline were found ($P < 0.1$) were added to the regression equation as a covariate.
To account for possible dependence of the outcomes within the clusters of four patients who were randomized together, a random intercept term for these patient clusters was included in all models, using multilevel analyses (SPSS mixed linear).

Multiple linear regression analyses were executed in order to test the hypothesis that APT, CBT and CT were more effective than WL, and whether CT showed the strongest effect, as to the outcome measures.
The initial regression model included the following independent variables:
pre-treatment measurement of the outcome variable, type of treatment, age, gender, centre of treatment, variables that turned out to be unequally divided between treatment groups despite randomization (covariates), potential prognostic factors and potential effect modifiers such as fear of injury and movement, level of functional limitations, and level of pain intensity. When the interaction variables (treatment x fear of injury/movement, treatment x level of functional limitations, treatment x pain intensity) were non-significant ($P > 0.05$), they were removed from the model. In case an interaction variable turned out to be significant ($P < 0.05$), analyses with regard to the concerning outcome variable were repeated within strata. Next, non-significant co-variables ($P > 0.05$) were eliminated one by one.

Statistical analyses were carried out according to the intention-to-treat principle: all patients, including withdrawals from treatment and patients with poor compliance remained in the treatment group they were randomized to. If data on outcome measures were missing, the baseline value carried forward method was used and a worst case analysis by imputing the tenth percentile score of the outcome measure at post-treatment of the respondents was performed as well.

In order to check for manipulation, pre-post difference in the calculated VO_2max was tested by means of paired *t*-tests within each group receiving sufficient aerobic training (APT and CT) and each group receiving no aerobic training (CBT and WL), and by an independent *t*-test for the difference in VO_2max between these two groups. The same was done for the groups receiving a sufficient number of problem solving skills training sessions (CBT and CT) and those receiving no such training at all (APT and WL) by testing the differences on the SPSI-NPO subscale.
SPSS statistical software, version 12.0 was used for the statistical analyses (SPSS, Inc., Chicago).

Results

Study population

Of the 309 eligible patients, 82 patients were not included. The reasons for not including these patients are shown in figure 1. At the start of treatment, four patients, who were randomized to APT (n = 1), CBT (n = 2) and CT (n = 1), respectively, appeared not to fulfill the selection criteria (see figure 1). Because these medical problems already existed before the treatment started, these patients were excluded from further analysis.

Furthermore three patients, although they had a RDQ score of > 3 at the time of inclusion appeared to have a score of ≤ 3 at pre-treatment assessment. One patient completed the APT, another patient stopped the CT after nine sessions because he had only minor functional limitations and could not combine his job with the intensity of treatment. The third patient only attended one session of CBT and stopped further treatment because her complaints were almost completely resolved. These three patients were all included in the intention-to-treat analysis.

Another three patients (one in APT, one in CBT and one in CT) never showed up for treatment but all attended the follow-up assessment and were also included in the intention-to-treat analysis.

For eleven patients (APT: n = 1, CBT: n = 3, CT: n = 6, WL: n =1) no data of questionnaires at post-treatment were available (5%). Ten of these patients dropped out of treatment and the reasons for not filling out questionnaires are shown in figure 1. The questionnaire of one patient who completed the treatment got lost in the mail.

For the performance tasks the data were complete for 200 of the 223 patients (90%). The data were missing for five patients in APT, ten in CBT, six in CT and two in WL; one patient was unreachable (CT), five had other non-LBP associated medical or psychological problems (one in APT, CBT, WL and two in CT), ten rejected treatment (two in CT, three in APT and five in CBT), five reported logistic problems (one in CT and WL, three in CBT), one had increase of pain (APT) and one had no complaints anymore (CBT). So only a small percentage of data was missing and the reasons for not responding were not related to the content of treatment (e.g. adverse effects). Nevertheless, a worst case analysis was performed by imputing the missing data by the tenth percentile score of the available data.

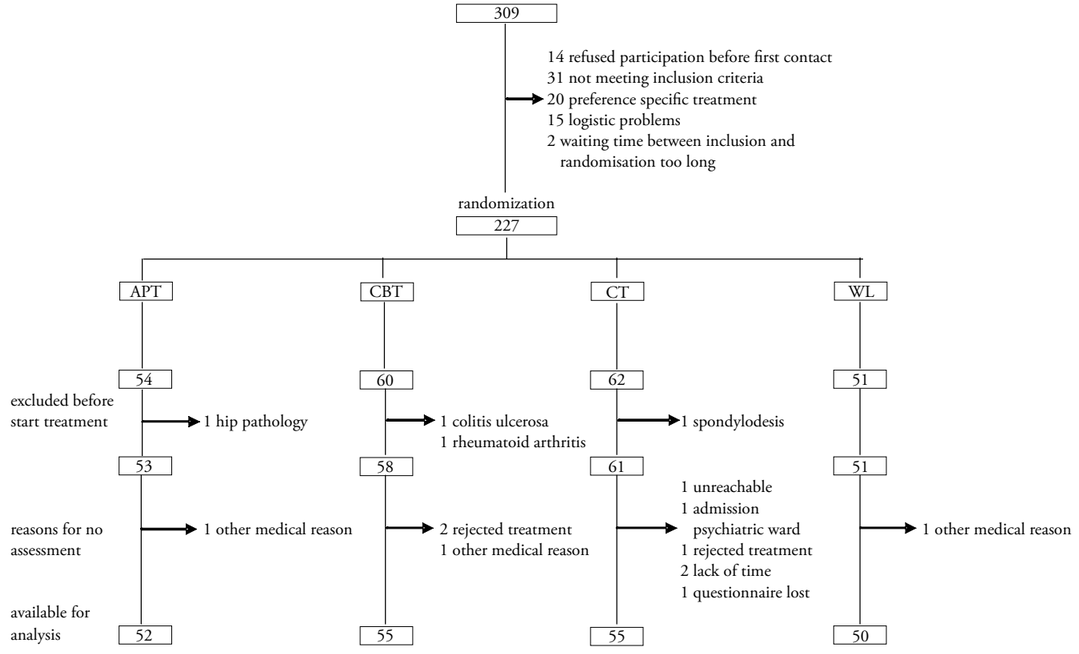

Figure 1
Flow diagram summarizing the formation of the study group and the number of lost to follow-up.
APT= active physical treatment, CBT= cognitive-behavioral treatment, CT= combination treatment, WL= waiting list.

Comparability at baseline

Demographic variables had similar distribution in the treatment groups (see table 1). Regarding disease characteristics only the duration of functional limitations was not similar distributed in the treatment groups, and was entered into the regression analysis as a covariate. Regarding the outcome measures, the WL performed significantly better at baseline on the forward reach test than the three active treatments. The total group had moderate to severe functional limitations and quite a high percentage of the patients was on sick leave or disability pension because of their low back pain and most of them (97%) already had undergone some sort of physical or medical treatment for their CLBP.
There were no statistically significant differences regarding the baseline variables between the responders and non-responders on the questionnaires. The non-responders on the performance tasks had a significantly lower baseline level of sport activity, fast walking and stair climbing.

Treatment compliance and co-interventions

In the APT the mean number of sessions for the total group of randomized patients was 24.5 (maximum of 30) for aerobic and strengthening training. In the CBT the mean number of sessions for the GA was 14.3 (maximum of 20) and 7.7 (maximum of 10) for the PST. The CT patients attended a mean of 21.8 physical training sessions, 11.9 (maximum of 19) GA-sessions and 7.1 PST-sessions.
In APT, 83% of all patients met the criterion of at least 2/3 attendance. Of all CBT patients, 78% and 76% had a sufficient number of sessions of GA and PST, respectively. For the CT patients, 72% had sufficient physical training and 62% sufficient GA and PST. The main reasons for not completing a sufficient number of treatments were rejection of treatment (n=23; 4 in APT, 8 in CBT and 11 in CT) or other non-LBP associated medical or psychological problems (n=11; 3 in APT, 1 in CBT, 7 in CT). For only six patients (three in APT and three in CT), increase of pain was the reason to stop the treatment (see adverse effects). Logistic reasons were only reported by patients attending the CBT or CT (n=12; 6 in each treatment), mainly because the time of treatment changed regularly or the higher frequency of different training elements and the resulting longer total treatment time. One patient (CBT) quitted treatment because she had no complaints anymore.
Analysis of the cost diaries that had been returned by 209 patients (94%; APT, n = 51; CBT, n = 52; CT, n = 56; WL, n = 50) showed that an equal number of patients did not completely adhere to the treatment protocol. There were no statistically significant differences between groups regarding the percentage of patients visiting the general physician, medical specialist, radiologist, physiotherapist or alternative medicine and the mean number of visits during the treatment period (table 2). Relevant protocol deviations as defined earlier were reported in 27 patients (5 in APT, 8 in CBT, 4 in CT and 10 in WL).

Table 1: Baseline variables for the total population and the four therapy groups (total n = 223)

Variables	APT (n=53)	CBT (n=58)	CT (n=61)	WL (n=51)
Age (yr)	42.68 ± 9.06	42.52 ± 9.67	40.67 ± 10.14	40.55 ± 11.17
Gender (% male)	58.5	41.4	62.3	49.0
Education (%)				
low	67.9	62.1	57.4	62.7
middle - high	32.1	37.9	42.6	37.3
Work status (%)				
full time	32.1	36.2	24.6	25.5
partial sick leave /disability pension	18.9	24.1	32.8	17.6
full sick leave / disability pension	43.4	31.0	36.1	43.1
no job / retired	5.7	8.6	6.6	13.7
Duration of LBP (mo)	56.91 ± 75.86	68.33 ± 74.21	56.13 ± 67.50	43.75 ± 70.82
Duration of functional limitations (mo)	28.85 ± 37.43	49.12 ± 61.40	36.90 ± 53.57	23.55 ± 32.76*
Radiation of pain (%)				
no radiation	9.4	17.2	13.1	15.7
above knee	35.8	27.6	42.6	41.2
below knee	54.7	55.2	44.3	43.1
Previous back surgery (%)	17.0	10.3	14.8	19.6
Trauma preceding LBP (%)	17.0	19.0	11.5	27.5
Sport activity (BPAQ-sport)	2.07 ± 0.67	2.14 ± 0.78	2.04 ± 0.65	2.10 ± 0.67
Leisure activity (BPAQ-leisure)	2.96 ± 0.69	2.79 ± 0.74	2.84 ± 0.60	2.78 ± 0.67
TSK	39.02 ± 6.52	38.72 ± 6.88	39.69 ± 7.08	37.75 ± 6.97
VO_2max (ml/kg/min)	27.26 ± 6.76	28.23 ± 6.69	30.06 ± 7.88	29.32 ± 7.07
SPSI-NPO	13.37 ± 7.42	13.90 ± 6.29	13.34 ± 7.71	15.10 ± 7.66
RDQ	14.15 ± 3.70	13.74 ± 3.65	13.51 ± 3.92	13.96 ± 3.88
Main complaints	74.52 ± 14.59	74.71 ± 16.19	72.44 ± 17.03	77.42 ± 11.35
Current pain	51.23 ± 26.55	48.84 ± 23.51	45.98 ± 23.95	51.02 ± 25.40
PRI-T	18.34 ± 11.32	17.86 ± 9.94	18.08 ± 9.04	17.37 ± 8.52
BDI	10.38 ± 7.62	10.45 ± 7.06	9.75 ± 6.68	9.78 ± 7.67
Walking (m)	375.37 ± 68.03	385.29 ± 75.36	372.40 ± 88.03	383.48 ± 101.75
Fast walking (sec)	9.92 ± 2.27	10.28 ± 2.40	10.31 ± 2.51	10.02 ± 2.80
Sit to stand (sec)	20.98 ± 7.33	22.62 ± 9.99	21.86 ± 7.65	21.19 ± 8.98
Loaded forward reach (cm)	52.04 ± 12.27*	50.12 ± 14.97*	55.00 ± 11.24*	59.62 ± 13.55*
Stair climbing (number of stairs)	72.02 ± 18.22	75.10 ± 24.76	71.26 ± 21.95	78.35 ± 26.90
Lifting (stages)	4.31 ± 2.58	4.29 ± 3.22	4.16 ± 3.05	4.06 ± 2.64

Values presented as means and standard deviation or percentage. * Significant differences between therapy groups ($P < 0.05$).
APT = active physical treatment; CBT = cognitive-behavioral treatment; CT = combination treatment; WL = waiting list; LBP = low back pain; BPAQ = Baecke Physical Activity Questionnaire; TSK = Tampa Scale for Kinesiophobia; SPSI-NPO = Social Inventory Problem Solving - Negative Problem Orientation; RDQ = Roland Disability Questionnaire; PRI-T = Pain Raiting Index Total score; BDI = Beck Depression Inventory.

Table 2: Protocol deviations; total percentage and mean number of visits

	APT (n=51)		CBT (n=52)		CT (n=56)		WL (n=50)	
	%	Mean	%	Mean	%	Mean	%	Mean
General physician	23.5	0.37	25.0	0.81	26.8	0.39	24.0	0.44
Medical specialist	7.8	0.12	7.7	0.15	8.9	0.11	8.0	0.08
Radiologist	0	0	0	0	0	0	0	0
Physiotherapy	11.8	0.53	13.5	0.77	5.4	0.39	18.0	0.90
Alternative medicine	0	0	3.8	0.23	5.4	0.21	4.0	0.16

APT = active physical treatment; CBT = cognitive-behavioral treatment; CT = combination treatment; WL = waiting list.

Adverse effects and other co-morbidity

Three patients in the APT and three in the CT stopped the treatment because of increased pain in the lower back or radiating leg pain. One of these patients, attending the APT developed three days after a training session a herniated disc with neurological deficits needing neurosurgical intervention. Furthermore, one patient attending the APT stopped the aerobic training because of knee complaints. One patient attending the APT developed pain complaints in both legs during cycling which appeared to be caused by vascular problems that could be resolved by vascular surgery.

Outcome measures

The observed change on the RDQ, the primary outcome measure without any correction was +0.04 ± 2.90 for the WL, -2.25 ± 4.51 for APT, -2.65 ± 4.66 for CBT and -2.27 ± 4.19 for CT, respectively. In table 3 the results of the multiple linear regression analyses with correction for dependence within patient clusters, gender, age, centre of treatment, duration of functional limitations, and significant prognostic variables are presented. First the mean scores of the WL at post-treatment are shown and next, the mean difference of the three active therapies compared with the score of the WL and the 95% confidence intervals (CI) are presented. The presented differences represent the real difference between the active treatments and WL after correction for any time-effects in the WL. All active treatments showed significant improvement on functional limitations, main complaints and pain intensity measured by using the VAS. The CBT and CT significantly improved on the global improvement measure and the APT almost reached significance. No difference was found on the total score of the Pain Rating Index and only APT showed a significant reduction of depression. Because the outcome measures fast walking and sit to stand were not normally distributed at baseline, the inverse of these measures was used. Compared to the WL, both APT and CT showed significant improvements on walking, sit to stand, and stair climbing. Lifting improved significantly in APT and fast walking in CBT and CT.
In table 4 the results of the comparison of CT versus APT and CBT are presented. No differences were found, except that the BDI was significantly more decreased in APT and walking improved significantly more in CT than CBT.
Since the treatment effect for overall satisfaction was modified by the baseline RDQ (interaction), the satisfaction scores for three different percentiles of the baseline RDQ are presented. It appeared that the level of satisfaction was significantly higher in APT compared to WL when the patient had a lower level of functional limitations at pre-treatment. For the ninetieth percentile score (RDQ = 19) this difference was not significant. CBT and CT showed a

significantly higher level of satisfaction compared to WL, and the higher the baseline RDQ score, the greater this difference became. Only for the ninetieth percentile score, CT showed significantly greater satisfaction than APT. No differences were evident between CT and CBT.

The alternative analyses that replaced the missing data according to the baseline value carried forward method showed very similar results, except that the differences on the lifting task and sit to stand task were no longer significant for the APT- versus WL-group. The worst case analysis did not show great differences either. The mean difference between CT and WL regarding current pain was no longer significant (-5.852 [95% CI: -14.192 to 2.489]), as well as the global improvement (0.495 [95% CI: -0.29 to 1.018]). For the performance tasks a few changes occurred: in comparison to the WL, the sit to stand task was no longer significant in APT and CT, the fast walking task was no longer significant in CBT and CT, and the stair-climbing task just did not reach significance anymore in APT.

The multilevel regression analyses showed that the dependence within patient clusters was usually small, intraclass correlations (ICCs) being never larger than 0.15 with only two exceptions; the regression model on loaded forward reach showed an ICC of 0.17 and stair climbing 0.20.

Table 3: Effects of APT, CBT and CT as compared with WL

Dependent variable	WL mean ± SD	APT mean difference (95% CI)†	CBT mean difference (95% CI)†	CT mean difference (95% CI)†
RDQ††	13.88 ± 4.78	-2.40 (-4.14 to -0.65)**	-3.05 (-4.80 to -1.30)**	-2.56 (-4.27 to -0.85)**
Main complaints††	74.25 ± 14.7	-11.19 (-20.07 to -2.31)*	-16.36 (-25.13 to -7.60)**	-17.84 (-26.54 to -9.14)**
Current pain††	53.35 ± 22.6	-8.68 (-16.87 to -0.48)*	-14.76 (-23.00 to -6.52)**	-8.23 (-16.37 to -0.10)*
PRI-T††	17.28 ± 10.48	0.90 (-2.94 to 4.74)	-2.04 (-5.91 to 1.83)	-0.33 (-4.14 to 3.48)
BDI	9.42 ± 7.81	-2.09 (-3.86 to -0.32)*	-1.65 (-3.42 to 0.12)	0.04 (-1.71 to 1.79)
Global improvement††	3.78 ± 0.91	0.47 (-0.08 to 1.09)	0.90 (0.36 to 1.44)**	0.70 (0.17 to 1.24)*
Satisfaction$				
10th percentile of baseline RDQ (=9)	45.65 ± 25.30	32.38 (14.19 to 50.57)**	18.34 (0.51 to 36.16)*	19.33 (2.01 to 36.65)*
50th percentile of baseline RDQ (=14)	46.67 ± 25.30	19.30 (6.74 to 31.88)**	22.68 (10.10 to 35.27)**	23.57 (11.28 to 35.86)**
90th percentile of baseline RDQ (=19)	47.69 ± 25.30	6.23 (-12.04 to 24.50)	27.03 (8.37 to 45.69)**	27.81 (9.54 to 46.08)**
Walking (m)††	386.60 ± 86.62	27.85 (5.35 to 50.34)*	12.82 (-9.75 to 35.40)	35.65 (13.85 to 57.45)**
Inversion fast walking (1/sec)	0.10 ± 0.02	0.005 (-0.002 to 0.012)	0.009 (0.002 to 0.016)*	0.008 (0.001 to 0.015)*
Inversion sit to stand (1/sec)	0.06 ± 0.02	0.007 (0.001 to 0.013)*	0.003 (-0.003 to 0.010)	0.007 (0.001 to 0.014)*
Loaded forward reach (cm)††	59.07 ± 12.55	1.30 (-4.40 to 7.00)	1.19 (-4.58 to 6.96)	0.98 (-4.54 to 6.51)
Stair climbing (stairs)	80.76 ± 24.80	7.83 (0.67 to 14.98)*	3.58 (-3.53 to 10.69)	9.25 (2.28 to 16.21)*
Lifting (stages)††	4.13 ± 2.56	0.96 (0.06 to 1.85)*	0.48 (-0.41 to 1.37)	0.68 (-0.16 to 1.52)

† Values are the mean difference between this group and the WL. The WL score is the score at post-treatment. The mean differences and confidence intervals (CIs) and corresponding P values were estimated adjusting for age, gender, centre of treatment, baseline score of outcome measure, duration of functional limitations.
†† Additional correction for relevant prognostic co-variates: RDQ - work status and TSK; Main complaints/Lifting - TSK; Current pain/PRI-T - work status; Global improvement/loaded forward reach - current pain; Walking - radiation and TSK.
$ Data presented for different strata for baseline RDQ, mean and SD for WL are estimated by use of regression model.
* P < 0.05; ** P < 0.01.
APT = active physical treatment; CBT = cognitive-behavioral treatment; CT = combination treatment; WL = waiting list; RDQ = Roland Disability Questionnaire; PRI-T = Pain Rating Index Total score; BDI = Beck Depression Inventory.

Table 4: Effects of APT and CBT as compared with CT

Dependent variable	CT mean ± SD	APT mean difference (95% CI)†	CBT mean difference (95% CI)†
RDQ††	11.40 ± 5.25	0.16 (-1.52 to 1.85)	-0.49 (-2.17 to 1.19)
Main complaints††	54.68 ± 21.79	6.65 (-1.96 to 15.26)	1.48 (-7.04 to 9.99)
Current pain††	42.31 ± 25.56	-0.45 (-8.41 to 7.52)	-6.53 (-14.48 to 1.43)
PRI-T††	17.53 ± 10.53	1.23 (-2.51 to 4.97)	-1.71 (-5.45 to 2.03)
BDI	9.07 ± 6.53	-2.13 (-3.84 to -0.42)*	-1.69 (-3.41 to 0.03)
Global improvement††	4.53 ± 1.33	-0.23 (-0.77 to 0.30)	0.20 (-0.33 to 0.73)
Satisfaction$			
10th percentile of baseline RDQ (=9)	64.98 ± 25.30	13.05 (-4.87 to 30.97)	-0.99 (-18.55 to 16.56)
50th percentile of baseline RDQ (=14)	70.24 ± 25.30	-4.26 (-16.51 to 7.98)	-0.89 (-13.16 to 11.28)
90th percentile of baseline RDQ (=19)	75.50 ± 25.30	-21.58 (-39.56 to -3.59)*	-0.78 (-19.11 to 17.55)
Walking (m)††	419.33 ± 66.44	-7.80 (-29.60 to 13.99)	-22.83 (-44.72 to -0.93)*
Inversion fast walking (1/sec)	0.11 ± 0.02	-0.003 (-0.010 to 0.004)	0.001 (-0.006 to 0.008)
Inversion sit to stand (1/sec)	0.06 ± 0.02	-0.001 (-0.007 to 0.005)	-0.004 (-0.010 to 0.002)
Loaded forward reach (cm)††	58.70 ± 12.80	0.32 (-5.19 to 5.19)	0.21 (-5.30 to 5.71)
Stair climbing (stairs)	84.50 ±22.31	-1.42 (-8.37 to 5.53)	-5.67 (-12.58 to 1.25)
Lifting (stages)††	4.87 ± 2.99	0.28 (-0.57 to 1.12)	-0.19 (-1.04 to 0.64)

† Values are the mean difference between this group and the CT. The CT score is the score at post-treatment. The mean differences and confidence intervals (CIs) and corresponding P values were estimated adjusting for age, gender, centre of treatment, baseline score of outcome measure, duration of functional limitations.

†† Additional correction for relevant prognostic co-variates: RDQ - work status and TSK; Main complaints/Lifting - TSK; Current pain/PRI-T - work status; Global improvement/loaded forward reach - current pain; Walking - radiation and TSK.

$ Data presented for different strata for baseline RDQ, mean/SD for CT estimated by use of regression model.

* P < 0.05.

APT = active physical treatment; CBT = cognitive-behavioral treatment; CT = combination treatment. RDQ = Roland Disability Questionnaire; PRI-T = Pain Rating Index Total score; BDI = Beck Depression Inventory.

Manipulation check

The patients who received a sufficient number of aerobic training showed a significant VO_2max improvement of 4.84 mL/kg LBM.min^{-1} versus a non-significant improvement of 0.29 mL/kg LBM.min^{-1} in the CBT and WL-group. The mean difference of 4.55 mL/kg LBM.min^{-1} (95% CI: -2.88 to -6.21 mL/kg LBM.min^{-1}) was highly significant indicating that the aerobic training was highly effective in increasing the VO_2max in patients receiving a sufficient number of aerobic training sessions.
Both patients who received a sufficient number of PST and those who did not receive this training at all, showed a significant change in the SPSI-NPO score (-1.80 and -1.92, respectively) indicating that the negative problem orientation decreased. But the mean difference of -0.1 (95% CI: - 1.18 to 1.43) was not significant, suggesting that the PST had no additional effect in decreasing the SPSI-NPO score.

Discussion

Although our patients had moderate limitations and most of them were already treated previously, all theory-based treatments, as hypothesized, were more effective than WL. The uncorrected data already showed a relevant decrease of the RDQ score for all active treatments. After correction for non-balance regarding baseline variables and adjustment for patient cluster dependence and centre of treatment, CBT and CT even showed a clinically relevant decrease of ≥ 2.5 points on the RDQ and APT just did not reach this clinically relevant level (2.4 points). Furthermore, hardly any adverse effects were reported. Several secondary outcomes such as main complaints, current pain and global improvement further confirmed this conclusion. The more the patients had functional limitations the higher they reported treatment satisfaction after attending CBT or CT. Patients attending APT were more satisfied than WL patients although this difference turned non-significant when pre-treatment functional limitations were high. This indicates that based on the patient's overall satisfaction, administration of APT in patients with moderate to severe functional limitations was less effective.
The performance tasks that seemed to be more physically demanding, improved significantly in treatments with a physical modality.

While comparing CT with APT or CBT, CT showed a significantly higher walking distance than CBT, but it can be debated whether a difference of 22.8 on 419 meters is clinically relevant. APT showed a greater reduction of depression than CT, although this change was not clinically relevant since the mean baseline score of the BDI was already low in all treatment groups. Regarding all other outcome measures, CT was not more effective than APT or CBT, respectively, so the hypothesis that CT has a stronger effect than APT and CBT, could not be supported.

Until now, it was not known whether changes in functional limitations are more effectively reduced by purely psychological or physical interventions (Mannion et al. 2001; van Tulder et al. 1997). To our knowledge the present study is the first trial in CLBP comparing explicitly theory-based treatments to WL and CT to APT and CBT, respectively, trying to address these problems. Both purely physical and psychological treatments as well as CT showed clinically relevant improvement. The overall results of all active treatment groups are comparable to the results presented in the most recent reviews and meta-analysis on different active treatments (Guzman et al. 2002; Hayden et al. 2005a; Hayden et al. 2005b; Ostelo et al. 2005).

The lack of great differential effects of different active treatments has been attributed to the fact that these treatments are not sufficiently theory-driven (Fuhrer 2003). But even in our study, CT showed no additional effect.This might be caused by the relative low compliance rate in CT. Furthermore, the total treatment intensity might have been a crucial factor for obtaining an additional effect. Although CT had a total duration of 78 hours, this is still lower than the 100 hours of therapy in daily intensive CT-programs showing an additional effect when compared to non-multidisciplinary

treatment (Guzman et al. 2002). Furthermore, no functional restoration was applied. On the other hand, we cannot rule out that the mixture of APT and CBT had an oppositional effect. For example, the increase of exercise load in APT was based on training physiology, and the increase of activity in CBT was based on time-contingency, which could have obscured the supplemental effect of both treatments.

By using the WL as reference treatment it cannot be ruled out that the positive effects of all active treatments were caused by other nonspecific factors such as attention, a standardized treatment program, or emphasis on active participation. Otherwise, by using the WL we were able to control for time effects showing that the WL did not improve or deteriorate on all outcome measures. Furthermore, due to ethical regulations, it was not possible to use an attention-control/placebo treatment once a patient had been given an indication for rehabilitation treatment.
By including a total number of 223 patients and having 212 patients available for analyses, sufficient power was assured for the comparison between the three active treatments and the WL regarding the primary outcome measure, functional limitations. Otherwise, the power might have been insufficient to find differences between the active treatments although the point-estimates showed no clear tendency in favor of CT.

Although the patient compliance was not very high, 95% of all randomized patients were assessed. This means that most patients, also those who did not have a sufficient intensity or showed serious protocol deviations, were included in the analyses. In this way the intention-to-treat method was approached as much as possible, ensuring that the real effectiveness of theory-based treatments in comparison to WL was determined. The results appeared to be quite robust since the alternative analyses hardly changed our results.

At randomization, all groups were quite similar on demographics and patient characteristics. Only the duration of functional limitations was definitely not equally distributed, but this variable was additionally controlled for in the statistical analyses.

In this study quite liberal inclusion criteria were used. In other studies, for example patients with psychosocial problems were excluded (Petersen et al. 2002), were treated as inpatients and should not have had previous back surgery, ongoing somatic or psychiatric disease or generalized disc degeneration (Brox et al. 2003). The level of functional limitations of our patients was relative high compared to that reported in other studies (Frost et al. 1995; Mannion et al. 1999; Moffett et al. 1999; Torstensen et al. 1998), but in accordance with the Dutch health care system in which CLBP patients with moderate to severe functional limitations are treated in outpatient rehabilitation centres (VRIN/VRA 2000). This means that the generalizability of the results for clinical practice is very high.

To improve the quality of the interventions, all treatments were highly structured by using detailed treatment protocols, and given by well-trained and skilled therapists. In order to avoid possible confounding within all treatment elements,

therapists were trained only to deliver one specific treatment element in the single and combination treatment as well. Therefore it was not possible to keep the therapists blinded. The patients could not be blinded because of ethical reasons. However, before randomization the patients were told that all active treatments are effective, but that the exact effectiveness is not yet clear. Furthermore, patients with an absolute preference for one treatment were excluded. Concealment of randomization was successfully achieved since no one of the referring physicians was aware of the type of treatment the referred patient would be randomized to. The blinding of the research assistants also seemed to be successfully maintained.

The compliance with treatment protocol by patients and clinicians are rarely assessed or adequately reported in RCTs (Prescott et al. 1999). Therefore additional registration forms for the therapists and diaries for the patients were used. Inspection of these forms indicated that there were no statistically significant differences between all treatment groups regarding protocol deviations. The treatment quality was judged to be sufficient in those patients who received at least 2/3 of all possible treatment sessions. This was further confirmed by the significant increase in VO_2max in the treatments using a physical training modality. The lack of additional improvement in negative problem solving after completion of PST was also reported by van den Hout et al. (2003). Despite this lack of improvement, van den Hout et al. found a decreased level of functional limitations in the problem solving group at 12 month follow-up, meaning that PST probably exerted its effect otherwise.
Furthermore, in only a limited number of patients the reasons for insufficient adherence were exclusively related to the type of treatment. The reported rate of compliance was quite similar to a few other RCTs using comparable treatments, ranging from 68% to 73% (Kole-Snijders et al. 1999; Moffett et al. 1999) and even only 69% in regular multidisciplinary rehabilitation programs (Woby et al. 2004).

Several investigators recommend that besides subjective questionnaires, more objective outcome measures such as performance tasks should be used (Simmonds et al. 1998; Waddell 1998). The improvement on performance tasks mainly occurred in APT and CT. This might be due to an increase of aerobic capacity or endurance strength. Otherwise, several patients did not report difficulties in, for example stair climbing, and for them a ceiling effect for this particular task seemed to apply. Furthermore, the performance tasks were possibly not specific enough to detect a change in the ability to perform patient relevant activities. For instance when a patient wanted to improve his walking distance, he would have increased the distance and not the speed. Because of this, the five-minute walking task probably would not change dramatically. Another explanation can be that the subjective experience of functional limitations and main complaints, as rated by questionnaires changes more easily, while changes in performance tasks take more time to establish (Ljungquist et al. 2003). It is not known what these results mean for clinical practice, since besides our own study, only one RCT using similar performance tasks was identified and the only conclusion was that the number of patients showing improvement on these performance tasks was lower than for the self-rated outcome measures (Ljungquist et al. 2003).

Since all active treatment groups showed similar effectiveness, the question arises how the effects are mediated. Previous studies have shown that irrespective of the treatment modality, improvement was mediated by the reduction of pain catastrophizing and the increase in experienced pain control (Jensen et al. 2001; Spinhoven et al. 2004; Vlaeyen and Morley 2005). Also in our study, patients in all treatment conditions have been exposed to situations that may have challenged their catastrophic beliefs that pain is a serious threat to their health. Results of such a mediation analysis based on our data will be reported later.

In order to improve the effectiveness it is not only necessary to find out how the treatment exerts its effect but also the question "what works for whom?" A way to further explore this might be to look for subgroups of patients by using objective, valid and reliable criteria to enhance the effectiveness of treatment programs (Vollenbroek-Hutten et al. 2004).

Conclusion

The results showed that the three theory-based treatments were more effective than WL. However, CT did not show greater differences than APT and CBT, respectively. These findings show that APT and CBT are as effective in reducing the personal experienced level of functioning as CT. Given the treatment intensity one could prefer the less intensive treatments instead of the more intensive and therefore more expensive CT. However, one can only decide on cost-effectiveness by comparing APT/CBT with CT after a longer period of follow-up with a proper cost-effectiveness analysis taking the total health care and work-related costs into account. This analysis will be carried out and presented once the one year follow-up will be completed. Based on the patients' overall satisfaction, CBT is to be preferred when the patient has moderately to severely functionally limitations.
Further research on theory-based treatments to confirm our findings, to investigate mediation and to develop more effective treatments is warranted.

Acknowledgments

The authors would like to thank the rehabilitation centres Breda, Leypark and Blixembosch and the many people who contributed to the successful completion of this study, particularly the participating patients, the rehabilitation centre contact persons, therapists, research assistants, especially Ria Laros, Frederieke Berendsen and Saskia Beelen, and Hubert Schouten for his statistical advice regarding the study design and sample size calculation. This study is supported by Zorgonderzoek Nederland/Medische Wetenschappen (ZonMw) grant number 014-32-007 and the Rehabilitation Centre Blixembosch.

References

- Abenhaim L, Rossignol M, Valat JP, Nordin M, Avouac B, Blotman F, Charlot J, Dreiser RL, Legrand E, Rozenberg S, Vautravers P. The role of activity in the therapeutic management of back pain. Report of the International Paris Task Force on Back Pain. Spine 2000;25:1S-33S.
- ACSM. ACSM position stand on the recommended quantity and quality of exercise for developing and maintaining cardiorespiratory and muscular fitness, and flexibility in adults. Med Sci Sports Exerc 1998;30:975-91.
- Arrindell WA, Ettema JHM. SCL-90. Handleiding bij een multidemensionele psychopathologie indicator. Lisse: Swets & Zeitlinger, 1986.
- Åstrand P, Rohdahl K. Textbook of work physiology, physiological bases of exercise. New York: McGraw-Hill, 1986, pp. 354-91.
- Baecke JA, Burema J, Frijters JE. A short questionnaire for the measurement of habitual physical activity in epidemiological studies. Am J Clin Nutr 1982;36:936-42.
- Beck AT, Rush AJ, Shaw BF, Emmery G. Cognitive therapy of depression. New York: Wiley and sons, 1979.
- Beck AT, Steer RA, Garbin MG. Psychometric properties of the Beck Depression Inventory: Twenty-five years of evaluation. Clin Psychol Rev 1988;8:77-100.
- Beurskens AJ, de Vet HC, Koke AJ. Responsiveness of functional status in low back pain: a comparison of different instruments. Pain 1996;65:71-6.
- Beurskens AJ, de Vet HC, Koke AJ, Lindeman E, van der Heijden GJ, Regtop W, Knipschild PG. A patient-specific approach for measuring functional status in low back pain. J Manipulative Physiol Ther 1999;22:144-8.
- Borg G. Psychophysical scaling with applications in physical work and the perception of exertion. Scand J Work Environ Health 1990;16 Suppl 1:55-8.
- Brox IJ, Sorensen R, Friis A, Nygaard O, Indahl A, Keller A, Ingebrigtsen T, Eriksen HR, Holm I, Koller AK, Riise R, Reikeras O. Randomized clinical trial of lumbar instrumented fusion and cognitive intervention and exercises in patients with chronic low back pain and disc degeneration. Spine 2003;28:1913-21.
- Danneels LA, Cools AM, Vanderstraeten GG, Cambier DC, Witvrouw EE, Bourgois J, de Cuyper HJ. The effects of three different training modalities on the cross-sectional area of the paravertebral muscles. Scan J Med Sci Sports 2001;11:335-41.
- Dishman RK. Prescribing exercise intensity for healthy adults using perceived exertion. Med Sci Sports Med 1994;26:1087-94.
- D'Zurilla TJ, Nezu AM, Maydeu-Olivares A. Manual for the Social Problem Solving Inventory-Revised. North Tonawanda, NY: Multi-Health Systems, 1997.
- Fordyce WE. Behavioral methods for chronic pain and illness. St. Louis: Mosby, 1976.
- Frost H, Klaber-Moffett JA, Moser JS, Fairbank JC. Randomised controlled trial for evaluation of fitness programme for patients with chronic low back pain. BMJ 1995;310:151-4.
- Fuhrer MJ. Overview of clinical trials in medical rehabilitation: impetuses, challenges, and needed future directions. Am J Phys Med Rehabil 2003;82:S8-15.
- Gommans IHB, Koes BW, van Tulder MW. Validiteit en responsiviteit Nederlandstalige Roland Disability Questionnaire. Vragenlijst naar functionele status bij patiënten met lage rugpijn. Ned T Fysiother 1997;107:28-33.
- Goubert L, Crombez G, Vlaeyen JWS, van Damme S, van den Broeck A, van Houdenhove B. De Tampa schaal voor kinesiofobie: psychometrische karakteristieken en normering. Gedrag & Gezondheid 2000;28:54-62.
- Guzman J, Esmail R, Karjalainen K, Malmivaara A, Irvin E, Bombardier C. Multidisciplinary bio-psycho-social rehabilitation for chronic low back pain. Cochrane Database Syst Rev 2002:CD000963.

References

- Harding VR, Williams AC, Richardson PH, Nicholas MK, Jackson JL, Richardson IH, Pither CE. The development of a battery of measures for assessing physical functioning of chronic pain patients. Pain 1994;58:367-75.
- Haskell WL. Dose-response issues from a biological perspective. In: C Bouchard, RJ Shephard, T Stephens. Physical activity, fitness, and health: international proceedings and consensus statement. Champaign: Human Kinetics, 1994, pp. 1030-9.
- Hayden JA, van Tulder MW, Malmivaara AV, Koes BW. Meta-analysis: exercise therapy for nonspecific low back pain. Ann Intern Med 2005a;142:765-75.
- Hayden JA, van Tulder MW, Tomlinson G. Systematic review: strategies for using exercise therapy to improve outcomes in chronic low back pain. Ann Intern Med 2005b;142:776-85.
- Hilde G, Bo K. Effect of exercise in the treatment of chronic low back pain: a systematic review, emphasising type and dose of exercise. Phys Ther Rev 1998;3:107-17.
- Hodselmans AP, Jaegers SM, Goeken LN. Short-term outcomes of a back school program for chronic low back pain. Arch Phys Med Rehabil 2001;82:1099-105.
- Jacob T, Baras M, Zeev A, Epstein L. Low back pain: reliability of a set of pain measurement tools. Arch Phys Med Rehabil 2001;82:735-42.
- Jensen MP, Turner JA, Romano JM. Changes in beliefs, catastrophizing, and coping are associated with improvement in multidisciplinary pain treatment. J Consult Clin Psychol 2001;69:655-62.
- Karvonen M, Kentala K, Mustala O. The effects of training hearth rate: a longitudinal study. Ann Med Exp Biol Fenn 1957;35:307-315.
- Kole-Snijders AM, Vlaeyen JW, Goossens ME, Rutten-van Molken MP, Heuts PH, van Breukelen G, van Eek H. Chronic low back pain: what does cognitive coping skills training add to operant behavioral treatment? Results of a randomized clinical trial. J Consult Clin Psychol 1999;67:931-44.
- Liddle SD, Baxter GD, Gracey JH. Exercise and chronic low back pain: what works? Pain 2004;107:176-90.
- Ljungquist T, Nygren A, Jensen I, Harms Ringdahl K. Physical performance tests for people with spinal pain - sensitivity to change. Disabil Rehabil 2003;25:856-66.
- Luteijn G, Starren J, van Dijk H. Handleiding bij de Nederlandse Persoonlijkheids Vragenlijst. Lisse: Swets & Zeitlinger, 1985.
- Mannion AF, Junge A, Taimela S, Muntener M, Lorenzo K, Dvorak J. Active therapy for chronic low back pain: part 3. Factors influencing self-rated disability and its change following therapy. Spine 2001;26:920-9.
- Mannion AF, Muntener M, Taimela S, Dvorak J. A randomized clinical trial of three active therapies for chronic low back pain. Spine 1999;24:2435-48.
- Mayer TG, Barnes D, Kishino ND, Nichols G, Gatchel RJ, Mayer H, Mooney V. Progressive isoinertial lifting evaluation. I. A standardized protocol and normative database. Spine 1988;13:993-7.
- Mayer TG, Gatchel RJ, Barnes D, Mayer H, Mooney V. Progressive isoinertial lifting evaluation. Erratum notice. Spine 1990;15:5.
- Mayer TG, Polatin PB, Gatchel RJ. Functional restoration and other rehabilitation approaches to chronic musculoskeletal pain disability syndromes. Crit Rev Physical Rehab Med 1998;10:209-21.
- McArdle W, Catch FI, Katch VL. Training for anaerobic and aerobic power. In: W McArdle, FI Catch, VL Katch. Exercise physiology. Baltimore: Williams & Wilkins, 1996, pp. 393-415.
- Meerding WJ, Bonneux L, Polder JJ, Koopmanschap MA, van der Maas PJ. Demographic and epidemiological determinants of healthcare costs in Netherlands: cost of illness study. BMJ 1998;317:111-5.

References

- Melzack R, Katz J. The McGill Pain Questionnaire: appraisal and current status. In: DC Turk, R Melzack. Handbook of pain assessment. New York: Guilford Press, 1992, pp. 152-68.
- Moffett JK, Torgerson D, Bell-Syer S, Jackson D, Llewlyn-Phillips H, Farrin A, Barber J. Randomised controlled trial of exercise for low back pain: clinical outcomes, costs, and preferences. BMJ 1999;319:279-83.
- Morley S, Eccleston C, Williams A. Systematic review and meta-analysis of randomized controlled trials of cognitive behaviour therapy and behaviour therapy for chronic pain in adults, excluding headache. Pain 1999;80:1-13.
- Ostelo RW, de Vet HC, Knol DL, van den Brandt PA. 24-item Roland-Morris Disability Questionnaire was preferred out of six functional status questionnaires for post-lumbar disc surgery. J Clin Epidemiol 2004;57:268-76.
- Ostelo RW, Tulder M, Vlaeyen J, Linton S, Morley S, Assendelft W. Behavioural treatment for chronic low-back pain. Cochrane Database Syst Rev 2005:Cd002014.
- Petersen T, Kryger P, Ekdahl C, Olsen S, Jacobsen S. The effect of McKenzie therapy as compared with that of intensive strengthening training for the treatment of patients with subacute or chronic low back pain: A randomized controlled trial. Spine 2002;27:1702-9.
- Prescott RJ, Counsell CE, Gillespie WJ, Grant AM, Russell IT, Kiauka S, Colthart IR, Ross S, Shepherd SM, Russell D. Factors that limit the quality, number and progress of randomised controlled trials. Health Technol Assess 1999;3:1-143.
- Roland M, Fairbank J. The Roland-Morris Disability Questionnaire and the Oswestry Disability Questionnaire. Spine 2000;25:3115-24.
- Roland M, Morris R. A study of the natural history of back pain. Part I: development of a reliable and sensitive measure of disability in low-back pain. Spine 1983;8:141-4.
- Sanders SH. Operant conditioning with chronic pain: back to basics. In: RJ Gatchel, DC Turk. Psychological approaches to pain management. A practitioner's handbook. New York: The Guilford press, 1996, pp. 112-30.
- Schotanus J. Het meten van belastbaarheid bij patiënten met chronische lage rugklachten. Groningen: Rijksuniversiteit Groningen, The Netherlands, 1999.
- Simmonds MJ, Olson SL, Jones S, Hussein T, Lee CE, Novy D, Radwan H. Psychometric characteristics and clinical usefulness of physical performance tests in patients with low back pain. Spine 1998;23:2412-21.
- Smeets RJ, Wade D, Hidding A, van Leeuwen PJ, Vlaeyen JW, Knottnerus JA. The association of physical deconditioning and chronic low back pain: A hypothesis-oriented systematic review. Disabil Rehabil 2006;28:673-93.
- Spinhoven P, Ter Kuile M, Kole-Snijders AM, Hutten Mansfeld M, Den Ouden DJ, Vlaeyen JW. Catastrophizing and internal pain control as mediators of outcome in the multidisciplinary treatment of chronic low back pain. Eur J Pain 2004;8:211-9.
- Torstensen TA, Ljunggren AE, Meen HD, Odland E, Mowinckel P, Geijerstam S. Efficiency and costs of medical exercise therapy, conventional physiotherapy, and self-exercise in patients with chronic low back pain. A pragmatic, randomized, single-blinded, controlled trial with 1-year follow-up. Spine 1998;23:2616-24.
- Tugwell P, Bombardier C, Buchanan WW, Goldsmith CH, Grace E, Hanna B. The MACTAR Patient Preference Disability Questionnaire - an individualized functional priority approach for assessing improvement in physical disability in clinical trials in rheumatoid arthritis. J Rheumatol 1987;14:446-51.
- Turk DC, Meichenbaum D, Genest M. Pain and behavioral medicine: A cognitive behavioral perspective. New York: Guilford Press, 1983.
- Turk DC, Okifuji A. Psychological factors in chronic pain: evolution and revolution. J Consult Clin Psychol 2002;70:678-90.

References

- van den Hout JH, Vlaeyen JW, Heuts PH, Zijlema JH, Wijnen JA. Secondary prevention of work-related disability in nonspecific low back pain: does problem-solving therapy help? A randomized clinical trial. Clin J Pain 2003;19:87-96.
- van der Kloot WA, Oostendorp RA, van der Meij J, van den Heuvel J. De Nederlandse versie van 'McGill pain questionnaire': een betrouwbare pijnvragenlijst. Ned Tijdschr Geneeskd 1995;139:669-73.
- van Tulder M, Malmivaara A, Esmail R, Koes B. Exercise therapy for low back pain. A systematic review within the framework of the Cochrane Collaboration Back Review Group. Spine 2000;25:2784-96.
- van Tulder MW, Koes BW, Bouter LM. A cost-of-illness study of back pain in The Netherlands. Pain 1995;62:233-40.
- van Tulder MW, Koes BW, Bouter LM. Conservative treatment of acute and chronic nonspecific low back pain. A systematic review of randomized controlled trials of the most common interventions. Spine 1997;22:2128-56.
- Verbunt JA, Seelen HA, Vlaeyen JWS, van der Heijden GJ, Heuts PH, Pons K, Knottnerus JA. Disuse and deconditioning in chronic low back pain: concepts and hypothesis on contributing mechanisms. Eur J Pain 2003;7:9-21.
- Vlaeyen JWS, Kole-Snijders AM, Boeren RG, van Eek H. Fear of movement/(re)injury in chronic low back pain and its relation to behavioral performance. Pain 1995;62:363-72.
- Vlaeyen JWS, Morley S. Cognitive-behavioral treatments for chronic pain: what works for whom? Clin J Pain 2005;21:1-8.
- Vlaeyen JWS, Teeken-Gruben NJ, Goossens ME, Rutten-van Molken MP, Pelt RA, van Eek H, Heuts PH. Cognitive-educational treatment of fibromyalgia: a randomized clinical trial. I. Clinical effects. J Rheumatol 1996;23:1237-45.
- Vollenbroek-Hutten MM, Hermens HJ, Wever D, Gorter M, Rinket J, IJzerman MJ. Differences in outcome of a multidisciplinary treatment between subgroups of chronic low back pain patients defined using two multiaxial assessment instruments: the multidimensional pain inventory and lumbar dynamometry. Clin Rehabil 2004;18:566-79.
- VRIN/VRA. Productdefiniëring Revalidatiezorg: registratiepilot revalidatie-instellingen januari 1999 - juli 2000. Utrecht: VRIN, 2000.
- Waddell G. A new clinical model of low back pain and disability. The back pain revolution. London: Churchill Livingstone, 1998.
- Woby SR, Watson PJ, Roach NK, Urmston M. Are changes in fear-avoidance beliefs, catastrophizing, and appraisals of control, predictive of changes in chronic low back pain and disability? Eur J Pain 2004;8:201-10.

Chapter 5

Physical, cognitive-behavioral or combination treatment for chronic low back pain? The one year post-treatment results of a randomized, controlled trial.

Rob JEM Smeets, Johan WS Vlaeyen, Alita Hidding, Arnold DM Kester, Geert JMG van der Heijden, J André Knottnerus

Accepted for publication in Pain, pending minor revisions.

Abstract

Several models for the maintenance of chronic low back pain have been postulated but the effectiveness of explicitly model-based treatments has never been directly compared. To examine whether a combination of a physical and a cognitive-behavioral treatment program is more effective than either alone in the long-term, a cluster randomized controlled trial was conducted. In total 172 patients, 18 to 65 years of age, with chronic disabling nonspecific low back pain referred for rehabilitation treatment, were randomized in clusters of four consecutive patients to 10 weeks of aerobic training and muscle strengthening of back extensors (active physical treatment; APT), 10 weeks of gradual assumption of patient relevant activities based on operant behavioral principles and problem solving training (cognitive-behavioral treatment; CBT), or APT combined with CBT (combination treatment; CT).
The primary outcome was the Roland Disability Questionnaire adjusted for centre of treatment, cluster, and baseline scores. Secondary outcomes were patients' main complaints, pain intensity, self-perceived improvement, depression and six physical performance tasks.
During the one year follow-up, an increasing small to moderate, but non-significant difference for Roland Disability scores in favor of the single treatments compared to CT was found. The additional self-perceived improvement for CBT and APT compared to CT reached significance one year post-treatment: 0.65 (95 % CI 0.10 to 1.21) and 0.61 (95 % CI 0.05 to 1.16), respectively.
No significant differences were found regarding the other secondary outcomes. We conclude that the combination treatment consisting of both physical and cognitive-behavioral treatment is not a more useful treatment option.

Introduction

Chronic nonspecific low back pain (CLBP) and its resulting disability have become a huge and epidemic health and socioeconomic problem (Meerding et al. 1998; van Tulder et al. 1995). To reduce the burden of this health problem, three frequently used models regarding the development and maintenance of CLBP associated disability, and evolving treatments have been postulated and applied.

1) The physical deconditioning model, assuming that loss of muscle strength and endurance including aerobic capacity is responsible for reduced activity levels and hence disability (Mayer et al. 1998; Verbunt et al. 2003).
2) The cognitive-behavioral model, postulating that disability results from maladaptive beliefs and avoidance behaviors that are maintained by learning processes (Fordyce 1976; Turk et al. 1983; Turk and Okifuji 2002; Vlaeyen et al. 1995).
3) The biopsychosocial model, assuming that disability results from both physical deconditioning and maladaptive beliefs and behaviors (Waddell 1998).

Strengthening combined with aerobic exercises, cognitive-behavioral treatment (CBT), and multidisciplinary treatment are superior to no treatment (Abenhaim et al. 2000; Guzman et al. 2002; Morley et al. 1999; Ostelo et al. 2005; van Tulder et al. 1997; van Tulder et al. 2000). Nevertheless, strong evidence is lacking that one of these treatments is more effective than the other (Guzman et al. 2002; Hayden et al. 2005a; Hayden et al. 2005b; Liddle et al. 2004; Morley et al. 1999; Ostelo et al. 2005; Smeets et al. 2006d; van Tulder et al. 2000). Examining the content of treatments, suggests that many treatments were not solely based on one of these three models. For example, in studies regarding the effectiveness of CBT, additional exercise therapy was used for increasing a patient's level of activity. As a result, it is not clear whether the improvement is reached by CBT, exercise therapy or the combination of both. Furthermore, many exercise therapies were of insufficient intensity, frequency and duration to fulfil physiologic training principles and should not have been classified as reconditioning or strengthening treatments (ACSM 1998; Haskell 1994; Hilde and Bo 1998; Liddle et al. 2004; McArdle et al. 1996). The evidence for the effectiveness of model-based treatments is still scarce and we do not know what treatment elements are really necessary to reach positive treatment results.

We therefore designed three treatments, each exclusively based on one of the abovementioned models. Since the combination treatment addresses physical deconditioning as well as maladaptive beliefs and avoidance behavior, we investigate in a randomized controlled trial the hypothesis that this combination treatment is more effective in reducing disability than the treatments solely based on the deconditioning or the cognitive-behavioral model.
The immediate post-treatment results showed that all three treatments were more effective than a waiting list group, but the combination treatment was not more effective than the single treatments (Smeets et al. 2006b). This paper describes the effects of the three active treatments 6 and 12 months post-treatment.

Methods

Patients

Between April 2002 and December 2004, patients for the first time referred by general practitioners or medical specialists to three Dutch outpatient rehabilitation centres were invited by their rehabilitation physician to participate. A research assistant checked the inclusion criteria: age between 18 and 65 years, nonspecific low back pain for more than 3 months resulting in disability (Roland Disability Questionnaire score [RDQ] > 3) (Roland and Morris 1983), and ability to walk at least 100 meters. The exclusion criteria were: vertebral fracture, spinal inflammatory disease or infection, malignancy, current nerve root pathology, spondylolysis, spondylolisthesis, lumbar spondylodesis, medical co-morbidity making exercising impossible (e.g. cardiovascular disease), clear treatment preference, not proficient in Dutch, pregnancy, and substance abuse interfering with treatment. The Symptom Checklist (SCL-90) (Arrindell and Ettema 1986) and the Dutch Personality Questionnaire (NPV) (Luteijn et al. 1985) were used to check for psychopathology hampering individual or group processes (Vlaeyen et al. 1996). Patients meeting the inclusion criteria were informed about the purpose and procedures of the study and were enrolled after giving informed consent.
The Medical Ethics Committee of the Rehabilitation Foundation Limburg and the Institute for Rehabilitation Research in Hoensbroek, the Netherlands approved the study protocol. The trial was assigned an international trial identification number (www.controlled-trials.com; ISRCTN22714229), and the results are reported according the Consolidated Standards of Reporting Trials (CONSORT)(Moher et al. 2001).

Randomization

Each cluster of four consecutive patients was assigned to one of the three active treatments or a waiting list, using block randomization. Randomization was stratified for rehabilitation centre, used blocks of size eight and was supervised by an independent statistician. Opaque, sequentially numbered, sealed envelopes were prepared for each rehabilitation centre before enrollment started. The envelope contained a sheet of paper indicating one of the four treatments. After the patient completed the baseline measurement, the research assistant handed over the envelope.

Masking

The research assistants collecting data were blinded to treatment allocation. All patients were repeatedly asked not to reveal information about their treatment allocation. Patients and therapists were not blinded to treatment allocation.

Interventions

All active treatments started with a group meeting with a maximum of four patients, during which the treatment rationale was explained. Emphasis was put on the responsibility of patients for making plans how to stay active after treatment. Each treatment lasted ten weeks. Patients were allowed to continue medication prescribed at baseline, but other co-interventions were discouraged. All therapists received an extensive training before the start of the trial and attended annual refresher courses.

Active Physical Treatment (APT)
In a group of maximum four, patients were invited to perform 30 minutes of aerobic training on a bicycle (65% to 80% heart rate maximum) and 75 minutes of strength and endurance training of their lower back and upper leg muscles (three series of 15 to 18 repetitions in a dynamic-static manner with a training intensity of 70% of the 1-Repetition Maximum, which was reassessed every fifth session), three times a week during ten weeks. Two physiotherapists supervised the training.

Cognitive-Behavioral Treatment (CBT)
CBT consisted of operant behavioral graded activity training (GA) (Fordyce 1976; Sanders 1996) and problem solving training (PST) (van den Hout et al. 2003). The GA started with three group sessions followed by a maximum of 17 individual sessions of 30 minutes. During the GA, a skilled physiotherapist or occupational therapist focused on a time contingent gradual increase or pacing of three patient relevant activities. The PST started with three sessions in which the rationale and the skill of positive problem orientation were discussed. Sessions four to ten focused on problem definition and formulation, generation of alternatives, decision making, implementation and evaluation. Patients received a course book with a summary of each session and homework assignments. A clinical psychologist or social worker, specifically trained to guide this intervention, provided ten sessions of 1½ hours to a maximum of four patients at a time.

Combination Treatment (CT)
CT consisted of APT and PST, both offered in the same frequency and duration as described before. Patients were told that they first had to gain sufficient aerobic fitness and strength before increasing their activities. GA started in the third week with a total of 19 sessions.

Waiting List (WL)
Patients were requested to wait ten weeks after which they were offered regular individual rehabilitation treatment. During the waiting period, patients were not allowed to participate in diagnostic or therapeutic procedures for their CLBP. For these patients data were collected before and immediately after the ten weeks of no treatment only.

Data collection

During the pre-treatment assessment, data were collected on age, gender, level of education, employment status, duration of complaints and disability, previous treatment, level of radiating pain, traumatic onset, pain-related fear (Tampa Scale for Kinesiophobia) (Goubert et al. 2000; Vlaeyen et al. 1995), VO_2max (modified Åstrand submaximal bicycle test) (Smeets et al. 2006e), negative problem orientation (Social Problem Solving Inventory-Revised) (D'Zurilla et al. 1997), and physical activity (Baecke Physical Activity Questionnaire) (Baecke et al. 1982; Jacob et al. 2001). Immediately after the explanation of the treatment rationale, treatment credibility and expectancy were assessed (Credibility/Expectancy Questionnaire) (Devilly and Borkovec 2000).

Outcome measures
Outcome measures were recorded at baseline, immediately post-treatment and 6 and 12 months post-treatment. The primary outcome measure was the RDQ, a reliable, valid, and responsive outcome measure (Beurskens et al. 1996; Gommans et al. 1997; Roland and Morris 1983).
Secondary outcomes included the severity of three individual main complaints on 100-mm visual analogue scales (Beurskens et al. 1999), current back pain on a 100-mm visual analogue scale and the Pain Rating Index of the McGill Pain Questionnaire (Melzack and Katz 1992; van der Kloot et al. 1995), self-perceived improvement of disability (7-point Likert scale, 1 = vastly worsened, 7 = completely recovered) (Beurskens et al. 1996), and depression according to the Beck Depression Inventory (Beck et al. 1988).
Six performance tests were used (Harding et al. 1994; Mayer et al. 1988; Mayer et al. 1990; Simmonds et al. 1998): 1) five-minute walking (meters), 2) fifty-foot walking (seconds), 3) five times sit to stand (seconds), 4) forward reaching by holding a stick with a weight of 4.5 kg at shoulder height (centimeters), 5) one-minute stair climbing (number of stairs), 6) progressive isoinertial lifting evaluation; the patient lifts a box with a weight four times within 20 seconds from floor up to a 75 cm high table. After each cycle of four lifting movements, the weight is increased in a standardized way (completed cycles).

Content of treatment
Treatment content was validated by check lists filled out by therapists after each treatment session. Patient attending at least 2/3 of all assigned treatment sessions for each training element, were classified as having sufficient treatment compliance. Co-interventions were documented by cost diaries filled out by patients during the treatment period (Goossens et al. 2000).

Statistical analyses
A difference of 2.5 points in RDQ score between treatment groups was considered to be clinically relevant (Roland and Fairbank 2000). Based on a 2-sided α of 0.05 and a 1-β of 0.90, with a standard deviation (SD) of the RDQ change of 4, 55 patients per group were needed.
Distribution of baseline characteristics were calculated to determine the prognostic similarity of groups.
Outcome measures were analyzed by using longitudinal analysis of covariance. The follow-up measurement was the dependent variable and the baseline value of the particular outcome was added as covariate. The coefficients of the longitudinal analysis of covariance were estimated using a mixed linear model with a random intercept for individuals to allow for dependence within patients (Twisk 2003), and with a random intercept for patient clusters to account for possible dependence of the outcome within the clusters of four patients who were randomized together. Adjustments were made for age, gender, centre of treatment, and baseline variables that were unequally divided between treatment groups despite randomization (covariates). Furthermore, the interaction effect of treatment with pain-related fear, level of disability and pain intensity was checked. Finally, potential prognostic factors were forced in the model and

eliminated one by one when non-significant ($P > 0.05$).
All statistical analyses were performed according to the intention-to-treat principle, using SPSS statistical software, version 12.0 (SPSS, Inc., Chicago, Illinois).
Additionally, for the primary outcome (RDQ) alternative analyses were performed by replacing missing values by the previous available value (last value carried forward) and the 10th percentile score of the total group (worst case scenario), respectively.

Results

Recruitment and follow-up
Of the 309 eligible patients, 82 patients were excluded. Reasons for exclusion are shown in figure 1. Additionally, four patients were excluded before start of treatment because of another medical diagnosis preventing participation (one in APT, two in CBT, one in CT). A total of 223 patients were randomly assigned to either APT (n = 53), CBT (n = 58), CT (n = 61), or WL (n = 51). Since no 6 and 12 months post-treatment data for WL were collected, only the results of the patients allocated to an active treatment will be further discussed (n = 172).
Follow-up rates for the questionnaires remained high, even at 12 months (156 of 172 patients). The reasons for not responding (n = 16), were not clearly related to the type of treatment: stopped treatment prematurely due to low baseline level of disability (1 in CBT, 2 in CT), no more complaints post-treatment but other medical problem (1 in CT), no more complaints 6 months post-treatment (3 in CBT), dissatisfaction (2 in both APT and CBT, 1 in CT), unreachable (2 in CT), other medical problem (1 in CT), and logistic reasons (1 in CT). The non-responders were significantly younger (mean, 34.3 versus 42.7 years) and showed lower sufficient treatment compliance (25% versus 74%). The number of patients with missing data of performance tasks decreased to 56 (of 172), 12 months post-treatment (15 in APT, 17 in CBT, 24 in CT). The reasons for not performing tasks were also not clearly associated with the type of treatment. The non-responders were significantly younger (mean, 37.9 versus 43.9 years), had lower sufficient treatment compliance (39% versus 84%), and lower 5-minute walking baseline score (mean, 357.9 versus 387.5 meters).

Comparability at baseline
Baseline status of patients did not differ significantly between treatment groups (table 1).

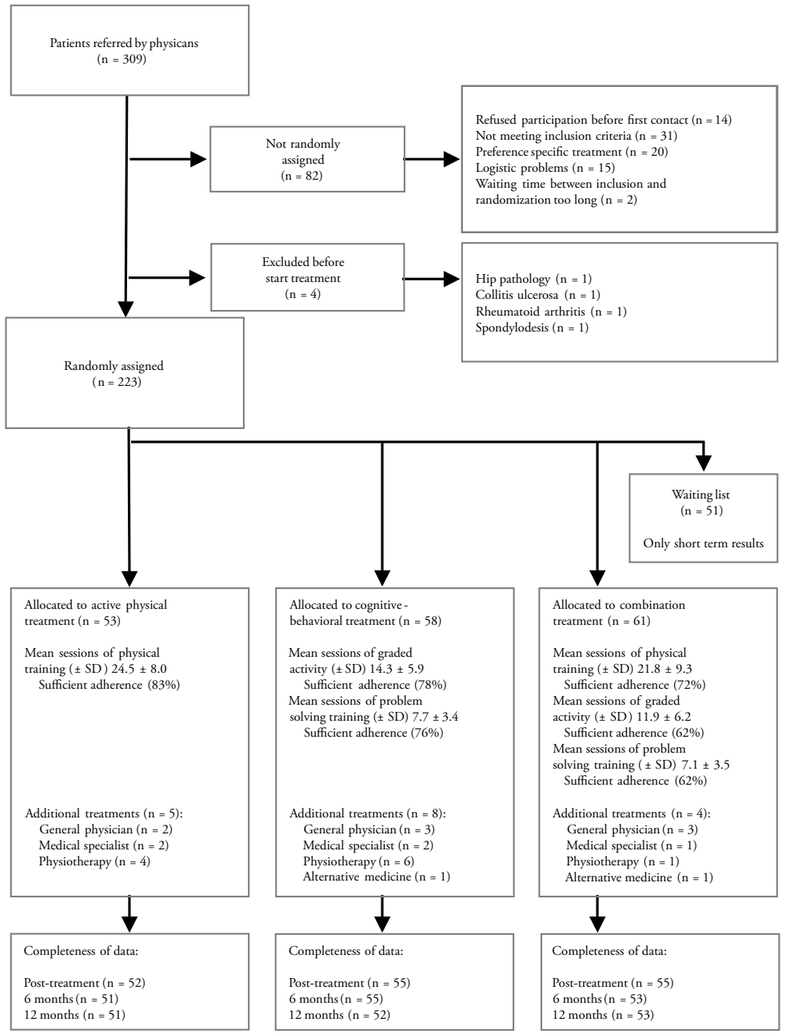

Figure 1
Flow chart describing the progress of patients through the trial.

Content of treatment and side effects
Compared with CT, the percentage of patients with sufficient treatment compliance was 11% and 15% higher in APT and CBT, respectively (figure 1). The reasons for insufficient compliance were: dissatisfaction (4 in APT, 8 in CBT, 11 in CT), other medical problem (3 in APT, 1 in CBT, 7 in CT), logistic problem (6 in both CBT and CT), and no complaints anymore (1 in CBT). Three patients (all CT) stopped treatment because of increased back pain.
One patient (APT) developed nerve root pathology needing neurosurgical intervention, one patient (APT) stopped aerobic training because of knee complaints, and another one (APT) developed vascular leg pain during cycling, that could be resolved by vascular surgery.

Analysis of the cost diaries of 159 patients (92%; 51 in APT, 52 in CBT, 56 in CT), showed that 17 patients (5 in APT, 8 in CBT, 4 in CT) reported additional treatments during treatment (figure 1).

Effectiveness of the interventions
Twelve months post-treatment, the observed change on the RDQ without any correction, was -3.20 ± 4.81 for APT, -3.15 ± 4.33 for CBT, and -1.74 ± 4.55 for CT, respectively. In tables 2 and 3, the results of the longitudinal random coefficient analysis with correction for dependence within patient clusters, gender, age, centre of treatment and significant prognostic variables are presented. The results represent the difference in adjusted improvement over time within each treatment group, and between the CT and the single treatments.
During follow-up, the improvement regarding disability, self-perceived improvement and main complaints persisted or even increased for APT and CBT, and diminished slightly to modestly for CT. All treatment groups showed small significant but clinically irrelevant persistent improvements on depression and all performance tasks except lifting. Fast walking and loaded forward reach did not improve significantly 12 months post-treatment in CT. Although all treatments showed a modest decrease of pain immediately post-treatment, this decrease almost disappeared 12 months post-treatment, and even showed an increase in CT. Follow-up measurements consistently did not show that CT was more effective than APT or CBT. At 6 months, the level of disability, self-perceived improvement and pain favored the single treatments. CBT showed higher reduction of main complaints and significantly higher self-perceived improvement than CT (0.76 [95% CI 0.21 to 1.31]). These differences between the single treatments and CT became even more apparent 12 months post-treatment, reaching significance for self-perceived improvement for CBT (0.65 [95% CI 0.10 to 1.21]) and APT (0.61 [95% CI 0.05 to 1.16]). The difference on the RDQ almost reached significance for CBT. No differences were found for depression and performance tasks.

Table 1: Baseline variables for the total population and the three therapy groups (total n=172)

Variables	Total (n=172)	APT (n=53)	CBT (n=58)	CT (n=61)
Age (yr)	41.91 ± 9.65	42.68 ± 9.06	42.52 ± 9.67	40.67 ± 10.14
Gender (% male)	54.1	58.5	41.4	62.3
Education (%)				
low	62.2	67.9	62.1	57.4
middle - high	37.8	32.1	37.9	42.6
Work status (%)				
full time	30.8	32.1	36.2	24.6
partial sick leave / disability pension	25.6	18.9	24.1	32.8
full sick leave / disability pension	36.6	43.4	31.0	36.1
no job / retired	7.0	5.7	8.6	6.6
Duration of LBP (mo)	60.48 ± 72.22	56.91 ± 75.86	68.33 ± 74.21	56.13 ± 67.50
Duration of functional limitations (mo)	38.54 ± 52.51	28.85 ± 37.43	49.12 ± 61.40	36.90 ± 53.57
Radiation of pain (%)				
no radiation	13.4	9.4	17.2	13.1
above knee	35.5	35.8	27.6	42.6
below knee	51.2	54.7	55.2	44.3
Previous physical/medical treatment (%)	96.5	94.3	96.6	98.4
Previous back surgery (%)	14.0	17.0	10.3	14.8
Trauma preceding LBP (%)	15.7	17.0	19.0	11.5
BPAQ-sport index score (1-5)	2.08 ± 0.70	2.07 ± 0.67	2.14 ± 0.78	2.04 ± 0.65
BPAQ leisure time index score (1-5)	2.86 ± 0.68	2.96 ± 0.69	2.79 ± 0.74	2.84 ± 0.60
TSK (17-68)	39.16 ± 6.81	39.02 ± 6.52	38.72 ± 6.88	39.69 ± 7.08
VO_2max (ml/kg/min)	28.61 ± 7.23	27.26 ± 6.76	28.23 ± 6.69	30.06 ± 7.88
SPSI-NPO (0-40)	13.54 ± 7.12	13.37 ± 7.42	13.90 ± 6.29	13.34 ± 7.71
Treatment credibility (3-27)	19.78 ± 3.61	20.28 ± 3.27	19.19 ± 3.65	19.96 ± 3.82
Treatment expectancy (3-27)	16.35 ± 4.41	16.89 ± 4.03	15.36 ± 4.43	16.84 ± 4.60
RDQ (0-24)	13.78 ± 3.75	14.15 ± 3.70	13.74 ± 3.65	13.51 ± 3.92
Main complaints (0-100)	73.84 ± 15.97	74.52 ± 14.59	74.71 ± 16.19	72.44 ± 17.03
Current pain (0-100)	48.56 ± 24.58	51.23 ± 26.55	48.84 ± 23.51	45.98 ± 23.95
PRI-T (0-63)	18.09 ± 10.03	18.34 ± 11.32	17.86 ± 9.94	18.08 ± 9.04
BDI (0-63)	10.18 ± 7.07	10.38 ± 7.62	10.45 ± 7.06	9.75 ± 6.68
Walking (m)	377.61 ± 77.80	375.37 ± 68.03	385.29 ± 75.36	372.40 ± 88.03
Fast walking (sec)	10.18 ± 2.40	9.92 ± 2.27	10.28 ± 2.40	10.31 ± 2.51
Sit to stand (sec)	21.85 ± 8.41	20.98 ± 7.33	22.67 ± 9.99	21.86 ± 7.65
Loaded forward reach (cm)	52.44 ± 13.00	52.04 ± 12.27	50.12 ± 14.97	55.00 ± 11.24
Stair climbing (number of stairs)	72.79 ± 21.84	72.02 ± 18.22	75.10 ± 24.76	71.26 ± 21.95
Lifting (stages)	4.25 ± 2.96	4.31 ± 2.58	4.29 ± 3.22	4.16 ± 3.05

Values presented are means ± SD or percentage.
APT = active physical treatment; CBT = cognitive-behavioral treatment; CT = combination treatment; LBP = low back pain; BPAQ = Baecke Physical Activity Questionnaire;
TSK = Tampa Scale for Kinesiophobia; SPSI-NPO = Social Problem Solving Inventory - Negative Problem Orientation; RDQ = Roland Disability Questionnaire;
PRI-T = Pain Raiting Index Total score; BDI = Beck Depression Inventory.

Table 2: Mean improvement from baseline, and differences of mean improvement between APT-CT and CBT-CT, respectively, in primary and secundary outcome measures assessed by questionnaires

Outcome measure	Participants n	Mean improvement APT (95% CI)*	Mean improvement CBT (95% CI)*	Mean improvement CT (95% CI)*	Between APT-CT difference (95% CI)*	Between CBT-CT differerence (95% CI)*
RDQ†						
Post-treatment	162	2.42 (1.14 to 3.69)	3.04 (1.79 to 4.29)	2.47 (1.25 to 3.86)	-0.05 (-1.71 to 1.62)	0.58 (-1.08 to 2.24)
6 months follow-up	159	3.15 (1.88 to 4.43)	3.65 (2.40 to 4.90)	2.54 (1.31 to 3.76)	0.62 (-1.06 to 2.30)	1.11 (-0.56 to 2.79)
12 months follow-up	156	3.28 (2.00 to 4.58)	3.74 (2.48 to 5.01)	2.12 (0.89 to 3.36)	1.16 (-0.52 to 2.84)	1.62 (-0.06 to 3.31)
Main complaints†						
Post-treatment	161	11.11 (4.11 to 18.10)	15.74 (8.98 to 22.51)	17.48 (10.78 to 24.18)	-6.38 (-16.00 to 3.25)	-1.74 (-11.23 to 7.75)
6 months follow-up	159	10.58 (3.60 to 17.57)	17.67 (10.90 to 24.44)	12.85 (6.08 to 19.62)	-2.26 (-11.94 to 7.41)	4.82 (-4.74 to 14.38)
12 months follow-up	154	12.27 (5.28 to 19.26)	20.19 (13.32 to 27.05)	11.95 (5.15 to 18.74)	0.33 (-9.36 to 10.02)	8.24 (-1.40 to 17.88)
Self-perceived improvement† ‖						
Post-treatment	162	4.33 (3.91 to 4.74)	4.71 (4.30 to 5.11)	4.53 (4.13 to 4.93)	-0.20 (-0.75 to 0.35)	0.18 (-0.37 to 0.73)
6 months follow-up	159	4.38 (3.97 to 4.80)	4.76 (4.36 to 5.17)	4.00 (3.60 to 4.40)	0.38 (-0.18 to 0.94)	0.76 (0.21 to 1.31)§
12 months follow-up	156	4.50 (4.08 to 4.91)	4.54 (4.13 to 4.95)	3.89 (3.49 to 4.29)	0.61 (0.05 to 1.16)§	0.65 (0.10 to 1.21)§
Current pain†						
Post-treatment	162	4.72 (-1.87 to 11.32)	10.25 (3.81 to 16.69)	4.90 (-1.52 to 11.32)	-0.18 (-9.34 to 8.99)	5.35 (-3.73 to 14.42)
6 months follow-up	158	2.81 (-3.82 to 9.44)	4.08 (-2.41 to 10.56)	-2.17 (-8.70 to 4.35)	4.98 (-4.29 to 14.25)	6.25 (-2.94 to 15.44)
12 months follow-up	156	2.31 (-4.32 to 8.94)	3.15 (-3.41 to 9.71)	-5.73 (-12.25 to 0.79)	8.04 (-1.23 to 17.31)	8.88 (-0.36 to 18.13)
PRI-T†						
Post-treatment	162	0.50 (-2.25 to 3.25)	3.52 (0.82 to 6.22)	1.45 (-1.19 to 4.10)	-0.95 (-4.59 to 2.68)	2.06 (-1.55 to 5.68)
6 months follow-up	159	3.13 (0.36 to 5.89)	2.21 (-0.50 to 4.91)	1.15 (-1.53 to 3.84)	1.97 (-1.71 to 5.65)	1.05 (-2.60 to 4.71)
12 months follow-up	156	1.70 (-1.06 to 4.46)	1.84 (-0.91 to 4.60)	-0.94 (-3.63 to 1.74)	2.64 (-1.04 to 6.32)	2.79 (-0.91 to 6.48)
BDI†						
Post-treatment	161	2.86 (1.42 to 4.29)	2.31 (0.91 to 3.72)	0.69 (-0.71 to 2.09)	2.17 (0.18 to 4.17)§	1.62 (-0.36 to 3.61)
6 months follow-up	158	2.63 (1.18 to 4.07)	2.41 (1.00 to 3.81)	2.14 (0.71 to 3.57)	0.49 (-1.54 to 2.51)	0.26 (-1.74 to 2.27)
12 months follow-up	156	3.23 (1.78 to 4.67)	2.08 (0.65 to 3.52)	2.17 (0.75 to 3.60)	1.05 (-0.97 to 3.07)	-0.09 (-2.11 to 1.93)

*Adjusted for the baseline value of the outcome measure, age, gender and treatment centre based on a longitudinal random coefficient analysis with an extra random intercept for clusters of four patients being randomized together.
† Additional correction for relevant prognostic covariates: RDQ - work status, duration disability and TSK; Main complaints - TSK;
Self-perceived improvement - work status, current pain and duration disability; Current pain- RDQ, TSK and BDI; PRI-T - work status, RDQ and duration disability; BDI - RDQ;
‖ Mean score on scale ranging from 1 to 7.
APT = active physical treatment; CBT = cognitive-behavioral treatment; CT = combination treatment; RDQ = Roland Disability Questionnaire;
PRI-T = Pain Rating Index Total score; BDI = Beck Depression Inventory.
§P < 0.05

Table 3: Mean improvement from baseline, and differences of mean improvement between APT-CT and CBT-CT, respectively in performance tasks

Outcome measure	Participants n	Mean improvement APT (95% CI)*	Mean improvement CBT (95% CI)*	Mean improvement CT (95% CI)*	Between APT-CT difference (95% CI)*	Between CBT-CT differerence (95% CI)*
Walking (m)†						
Post-treatment	149	30.10 (14.26 to 45.93)	16.85 (0.88 to 32.82)	35.14 (20.18 to 50.10)	-5.00 (-26.75 to 16.66)	-18.30 (-40.19 to 3.60)
6 months follow-up	122	37.90 (21.49 to 54.30)	27.83 (11.41 to 44.25)	37.63 (21.16 to 54.09)	0.27 (-22.90 to 23.44)	-9.80 (-33.12 to 13.53)
12 months follow-up	116	35.76 (18.96 to 52.56)	25.09 (9.58 to 43.00)	35.76 (19.35 to 52.16)	0.00 (-23.40 to 23.40)	-10.67 (-34.07 to 12.74)
Inversion fast walking (1/sec)†‡						
Post-treatment	150	0.003 (-0.002 to 0.009)	0.007 (0.002 to 0.012)	0.005 (0.000 to 0.009)	-0.001 (-0.008 to 0.006)	0.003 (-0.005 to 0.010)
6 months follow-up	123	0.005 (0.000 to 0.010)	0.007 (0.002 to 0.012)	0.004 (-0.001 to 0.010)	0.001 (-0.007 to 0.008)	0.003 (-0.005 to 0.010)
12 months follow-up	116	0.008 (0.002 to 0.013)	0.006 (0.001 to 0.012)	0.004 (-0.001 to 0.010)	0.004 (-0.004 to 0.011)	0.002 (-0.005 to 0.010)
Inversion sit to stand (1/sec)†‡						
Post-treatment	150	0.014 (0.009 to 0.019)	0.010 (0.006 to 0.015)	0.013 (0.008 to 0.017)	0.002 (-0.005 to 0.008)	-0.002 (-0.009 to 0.004)
6 months follow-up	123	0.017 (0.012 to 0.021)	0.013 (0.008 to 0.018)	0.013 (0.009 to 0.018)	0.003 (-0.004 to 0.010)	-0.001 (-0.008 to 0.006)
12 months follow-up	115	0.018 (0.013 to 0.023)	0.016 (0.011 to 0.021)	0.013 (0.008 to 0.018)	0.005 (-0.002 to 0.012)	0.003 (-0.004 to 0.010)
Loaded forward reach (cm)†						
Post-treatment	149	4.32 (0.83 to 7.82)	4.98 (1.46 to 8.51)	4.63 (1.35 to 7.91)	-0.31 (-5.07 to 4.45)	0.35 (-4.46 to -5.16)
6 months follow-up	119	3.70 (0.01 to 7.39)	2.27 (-1.41 to 5.95)	0.39 (-3.34 to 4.12)	3.31 (-1.90 to 8.53)	1.88 (-3.36 to 7.12)
12 months follow-up	115	5.42 (1.64 to 9.19)	3.80 (0.10 to 7.51)	2.35 (-1.38 to 6.07)	3.07 (-2.21 to 8.36)	1.46 (-3.81 to 6.72)
Stair climbing (stairs)						
Post-treatment	150	11.02 (5.56 to 16.48)	5.85 (0.48 to 11.22)	11.65 (6.54 to 16.77)	-0.64 (-8.07 to 6.80)	-5.81 (-13.19 to 1.57)
6 months follow-up	121	10.63 (5.06 to 16.21)	6.54 (1.06 to 12.03)	10.19 (4.76 to 15.62)	0.44 (-7.30 to 8.18)	-3.65 (-11.35 to 4.05)
12 months follow-up	115	11.38 (5.72 to 17.04)	8.21 (2.68 to 13.74)	7.83 (2.40 to 13.26)	3.54 (-4.26 to 11.34)	0.37 (-7.36 to 8.10)
Lifting (cycles)						
Post-treatment	147	0.89 (0.23 to 1.55)	0.36 (-0.29 to 1.00)	0.51 (-0.08 to 1.09)	0.38 (-0.50 to 1.26)	-0.15 (-1.02 to 0.71)
6 months follow-up	116	0.59 (-0.11 to 1.29)	0.33 (-0.35 to 1.00)	0.16 (-0.49 to 0.81)	0.43 (-0.52 to 1.38)	0.17 (-0.77 to 1.10)
12 months follow-up	112	0.29 (-0.42 to 0.10)	-0.44 (-1.11 to 0.24)	-0.03 (-0.70 to 0.63)	0.32 (-0.65 to 1.29)	-0.40 (-1.34 to 0.55)

*Adjusted for the baseline value of the outcome measure, age, gender and treatment centre based on a longitudinal random coefficient analysis with an extra random intercept for clusters of four patients being randomized together.
† Additional correction for relevant prognostic covariates: Walking/inversion fast walking/inversion sit to stand - RDQ and education level; Loaded forward reach - TSK.
‡ Due to not normal distribution at baseline, the inverse of these measures was used; a higher inverse score indicates a faster performance.
APT = active physical treatment; CBT = cognitive-behavioral treatment; CT = combination treatment; RDQ = Roland Disability Questionnaire.

Alternative analysis

The last value carried forward analysis showed similar results. Although the amount of missing data due to attrition is limited and appears to be mostly at random, a worst case scenario was performed showing that the reduction of the RDQ reached significance for APT versus CT (1.84 [95% CI 0.04 to 3.64]), but not for CBT versus CT (1.23 [95% CI -0.55 to 3.01]), 12 months post-treatment.

Discussion

Although our patients were moderately to severely disabled and most of them had been treated previously, all treatments showed improvement of disability and most secondary outcomes over time. However, the hypothesis that the combination treatment is more effective than the single treatments could not be confirmed. At 12 months, APT and CBT compared with CT, showed a higher, but both statistically and clinically insignificant reduction of RDQ: 1.16 [95% CI -0.52 to 2.84], and 1.62 [95% CI -0.06 to 3.31], respectively. However, several researchers argue that a 1 to 1.5 point additional reduction of the RDQ, is worth the effort (Moffett et al. 1999; UK BEAM Team 2004).
Since we did not find a tendency of CT being better, it is not likely that a higher powered study would change our conclusion that CT is not more effective than the single treatments. The two alternative analyses also confirmed our findings. Furthermore, consecutive post-treatment measurements showed an increasing small to moderate difference in pain in favor of the single treatments, and the improvement of main complaints favored CBT. Self-perceived improvement reached statistical significance at 12 months for CBT and APT. No relevant differences were found regarding depression and performance tasks. Given that CT implies a higher burden for patients, our study does not support CT as a more useful treatment option.

To our knowledge this is the first trial in CLBP comparing explicitly model-based treatments. The overall results of the treatments are comparable to those presented in recent reviews and meta-analyses (Guzman et al. 2002; Hayden et al. 2005a; Hayden et al. 2005b; Ostelo et al. 2005). The interventions were delivered according the study protocol with only a few protocol deviations. CT consisted of the APT and CBT program combined, and was carried out by the same, well-trained therapists as in the single treatments. Treatment groups did not differ on baseline characteristics, treatment credibility and expectancy. That CT is not more effective than the single treatments is puzzling, and suggests opposing rather than synergetic effects. First, CT required a higher effort of patients which seems to be reflected by the lower compliance rate. On the other hand, the compliance rate was quite similar to multidisciplinary treatments used in RCTs (Kole-Snijders et al. 1999; Moffett et al. 1999), and daily practice (Woby et al. 2004). Second, the different rationales of the CT elements might have been counteractive. The increase of exercise load in APT was based on training physiology, and in CBT a time-contingent increase of activities was used. This could have confused patients, resulting in insufficient generalization of the learned principles during follow-up. Otherwise, the total treatment duration of 78 hours in CT might not have been sufficient, since it is still lower that the 100 hours of

multidisciplinary treatment with functional restoration, showing higher effectiveness than non-multidisciplinary treatment (Guzman et al. 2002). Finally, the APT was more effective during follow-up than originally expected (Hayden et al. 2005a; Liddle et al. 2004). APT might have resulted in both attitude and behavior changes similar to those achieved by CBT, making the addition of CBT to CT redundant. Recently, we found that APT, although not using cognitive-behavioral techniques, resulted in a similar decrease of pain catastrophizing which also mediated outcome, as did CBT and CT (Smeets et al. 2006c). This is in line with the suggestion that not improved physical fitness, but cognitive processes are responsible for the effectiveness of exercise training (Helmhout et al. 2004; Mannion et al. 1999).
The total group was moderately to severely disabled and a high percentage was on sick leave or disability pension because of low back pain, making the results highly applicable to patients normally being referred for rehabilitation, at least in The Netherlands (VRIN/VRA 2000).

The study also has limitations; the high number of non-responders on the performance tasks for whom we did not impute missing values. Otherwise, the reasons for not responding were equally distributed across treatment groups. Although the use of objective performance tasks is recommended (Simmonds et al. 1998; Waddell 1998), this high percentage of non-responders as in other studies (Kole-Snijders et al. 1999; Ljungquist et al. 2003), and the rather low capability of most performance tasks to detect clinically relevant changes in CLBP samples (Smeets et al. 2006a), urges us to develop more patient-specific performance tasks. Furthermore, blinding of patients was not possible, but since no differences regarding treatment credibility and expectancy between treatment groups were found, we consider bias unlikely.

We conclude, that the single treatments are at least equally effective as the combination treatment one year post-treatment. Future research might focus on the question whether subgroups of patients can be identified, who benefit most from one specific treatment. Although we did not find any interaction between the level of disability, pain and pain-related fear and the different treatments, future trials could include treatments that target more systematically on putative mechanisms of change, such as pain catastrophizing. For example, exposure-in-vivo is shown to be a strong intervention in reducing pain catastrophizing in highly fearful patients (Vlaeyen et al. 2002).
Furthermore, in daily rehabilitation practice, the treatment is fine-tuned by selecting treatment modalities based on patient's characteristics, problems and needs. A next step could be comparing our highly structured treatments to these individually designed treatments.

Acknowledgements

The authors would like to thank the rehabilitation centres Breda, Leypark and Blixembosch and the many people who contributed to the successful completion of this study, particularly the participating patients, the rehabilitation centre contact persons, therapists, research assistants, especially Ria Laros, Frederieke Berendsen and Saskia Beelen.

References

• Abenhaim L, Rossignol M, Valat JP, Nordin M, Avouac B, Blotman F, Charlot J, Dreiser RL, Legrand E, Rozenberg S, Vautravers P. The role of activity in the therapeutic management of back pain. Report of the International Paris Task Force on Back Pain. Spine 2000;25:1S-33S.
• ACSM. ACSM position stand on the recommended quantity and quality of exercise for developing and maintaining cardiorespiratory and muscular fitness, and flexibility in adults. Med Sci Sports Exerc 1998;30:975-91.
• Arrindell WA, Ettema JHM. SCL-90. Handleiding bij een multidemensionele psychopathologie indicator. Lisse: Swets & Zeitlinger, 1986.
• Baecke JA, Burema J, Frijters JE. A short questionnaire for the measurement of habitual physical activity in epidemiological studies. Am J Clin Nutr 1982;36:936-42.
• Beck AT, Steer RA, Garbin MG. Psychometric properties of the Beck Depression Inventory: Twenty-five years of evaluation. Clin Psychol Rev 1988;8:77-100.
• Beurskens AJ, de Vet HC, Koke AJ. Responsiveness of functional status in low back pain: a comparison of different instruments. Pain 1996;65:71-6.
• Beurskens AJ, de Vet HC, Koke AJ, Lindeman E, van der Heijden GJ, Regtop W, Knipschild PG. A patient-specific approach for measuring functional status in low back pain. J Manipulative Physiol Ther 1999;22:144-8.
• Devilly GJ, Borkovec TD. Psychometric properties of the credibility/expectancy questionnaire. J Behav Ther Exp Psychiatry 2000;31:71-86.
• D'Zurilla TJ, Nezu AM, Maydeu-Olivares A. Manual for the Social Problem Solving Inventory-Revised. North Tonawanda, NY: Multi-Health Systems, 1997.
• Fordyce WE. Behavioral methods for chronic pain and illness. St. Louis: Mosby, 1976.
• Gommans IHB, Koes BW, van Tulder MW. Validiteit en responsiviteit Nederlandstalige Roland Disability Questionnaire. Vragenlijst naar functionele status bij patiënten met lage rugpijn. Ned T Fysiother 1997;107:28-33.
• Goossens ME, Rutten-van Molken MP, Vlaeyen JW, van der Linden SM. The cost diary: a method to measure direct and indirect costs in cost-effectiveness research. J Clin Epidemiol 2000;53:688-95.
• Goubert L, Crombez G, Vlaeyen JWS, Van Damme S, Van den Broeck A, van Houdenhove B. De Tampa schaal voor kinesiofobie: psychometrische karakteristieken en normering. Gedrag & Gezondheid 2000;28:54-62.
• Guzman J, Esmail R, Karjalainen K, Malmivaara A, Irvin E, Bombardier C. Multidisciplinary bio-psycho-social rehabilitation for chronic low back pain. Cochrane Database Syst Rev 2002:CD000963.
• Harding VR, Williams AC, Richardson PH, Nicholas MK, Jackson JL, Richardson IH, Pither CE. The development of a battery of measures for assessing physical functioning of chronic pain patients. Pain 1994;58:367-75.
• Haskell WL. Dose-response issues from a biological perspective. In: C Bouchard, RJ Shephard, T Stephens. Physical activity, fitness, and health: international proceedings and consensus statement Champaign: Human Kinetics, 1994, pp. 1030-9.
• Hayden JA, van Tulder MW, Malmivaara AV, Koes BW. Meta-analysis: exercise therapy for nonspecific low back pain. Ann Intern Med 2005a;142:765-75.
• Hayden JA, van Tulder MW, Tomlinson G. Systematic review: strategies for using exercise therapy to improve outcomes in chronic low back pain. Ann Intern Med 2005b;142:776-85.
• Helmhout PH, Harts CC, Staal JB, Candel MJ, de Bie RA. Comparison of a high-intensity and a low-intensity lumbar extensor training program as minimal intervention treatment in low back pain: a randomized trial. Eur Spine J 2004;13:537-47.

References

- Hilde G, Bo K. Effect of exercise in the treatment of chronic low back pain: a systematic review, emphasising type and dose of exercise. Phys Ther Rev 1998;3:107-17.
- Jacob T, Baras M, Zeev A, Epstein L. Low back pain: reliability of a set of pain measurement tools. Arch Phys Med Rehabil 2001;82:735-42.
- Kole-Snijders AM, Vlaeyen JW, Goossens ME, Rutten-van Molken MP, Heuts PH, van Breukelen G, van Eek H. Chronic low back pain: what does cognitive coping skills training add to operant behavioral treatment? Results of a randomized clinical trial. J Consult Clin Psychol 1999;67:931-44.
- Liddle SD, Baxter GD, Gracey JH. Exercise and chronic low back pain: what works? Pain 2004;107:176-90.
- Ljungquist T, Nygren A, Jensen I, Harms Ringdahl K. Physical performance tests for people with spinal pain--sensitivity to change. Disabil Rehabil 2003;25:856-66.
- Luteijn G, Starren J, van Dijk H. Handleiding bij de Nederlandse Persoonlijkheids Vragenlijst. Lisse: Swets & Zeitlinger, 1985.
- Mannion AF, Muntener M, Taimela S, Dvorak J. A randomized clinical trial of three active therapies for chronic low back pain. Spine 1999;24:2435-48.
- Mayer TG, Barnes D, Kishino ND, Nichols G, Gatchel RJ, Mayer H, Mooney V. Progressive isoinertial lifting evaluation. I. A standardized protocol and normative database. Spine 1988;13:993-7.
- Mayer TG, Gatchel RJ, Barnes D, Mayer H, Mooney V. Progressive isoinertial lifting evaluation. Erratum notice. Spine 1990;15:5.
- Mayer TG, Polatin PB, Gatchel RJ. Functional restoration and other rehabilitation approaches to chronic musculoskeletal pain disability syndromes. Crit Rev Physical Rehab Med 1998;10:209-21.
- McArdle W, Catch FI, Katch VL. Training for anaerobic and aerobic power. In: W McArdle, FI Catch, VL Katch. Exercise physiology. Baltimore: Williams & Wilkins, 1996, pp. 393-415.
- Meerding WJ, Bonneux L, Polder JJ, Koopmanschap MA, van der Maas PJ. Demographic and epidemiological determinants of healthcare costs in Netherlands: cost of illness study. BMJ 1998;317:111-5.
- Melzack R, Katz J. The McGill Pain Questionnaire: appraisal and current status. In: DC Turk, R Melzack. Handbook of pain assessment. New York: Guilford Press, 1992, pp. 152-68.
- Moffett JK, Torgerson D, Bell-Syer S, Jackson D, Llewlyn-Phillips H, Farrin A, Barber J. Randomised controlled trial of exercise for low back pain: clinical outcomes, costs, and preferences. BMJ 1999;319:279-83.
- Moher D, Schulz KF, Altman DG. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomised trials. Lancet 2001;357:1191-4.
- Morley S, Eccleston C, Williams A. Systematic review and meta-analysis of randomized controlled trials of cognitive behaviour therapy and behaviour therapy for chronic pain in adults, excluding headache. Pain 1999;80:1-13.
- Ostelo RW, Tulder M, Vlaeyen J, Linton S, Morley S, Assendelft W. Behavioural treatment for chronic low-back pain. Cochrane Database Syst Rev 2005:Cd002014.
- Roland M, Fairbank J. The Roland-Morris Disability Questionnaire and the Oswestry Disability Questionnaire. Spine 2000;25:3115-24.
- Roland M, Morris R. A study of the natural history of back pain. Part I: development of a reliable and sensitive measure of disability in low-back pain. Spine 1983;8:141-4.
- Sanders SH. Operant conditioning with chronic pain: back to basics. In: RJ Gatchel, DC Turk. Psychological approaches to pain management. A practitioner's handbook. New York: The Guilford press, 1996, pp. 112-30.

References

- Simmonds MJ, Olson SL, Jones S, Hussein T, Lee CE, Novy D, Radwan H. Psychometric characteristics and clinical usefulness of physical performance tests in patients with low back pain. Spine 1998;23:2412-21.
- Smeets RJ, Hijdra HJ, Kester AD, Hitters WG, Knottnerus JA. The usability of six physical performance tasks in a rehabilitation population with chronic low back pain. Clin Rehabil 2006a;20:987-96
- Smeets RJ, Vlaeyen JW, Hidding A, Kester AD, van der Heijden GJ, van Geel AC, Knottnerus JA. Active rehabilitation for chronic low back pain: Cognitive-behavioral, physical, or both? First direct post-treatment results from a randomized controlled trial [ISRCTN22714229]. BMC Musculoskelet Disord 2006b;7:5.
- Smeets RJ, Vlaeyen JW, Kester AD, Knottnerus JA. Reduction of pain catastrophizing mediates the outcome of both physical and cognitive-behavioral treatment in chronic low back pain. J Pain 2006c;7:261-71.
- Smeets RJ, Wade D, Hidding A, van Leeuwen PJ, Vlaeyen JW, Knottnerus JA. The association of physical deconditioning and chronic low back pain: A hypothesis-oriented systematic review. Disabil Rehabil 2006d;28:673-93.
- Smeets RJ, Wittink H, Hidding A, Knottnerus JA. Do patients with chronic low back pain have a lower level of aerobic fitness than healthy controls? Are pain, disability, fear of injury, working status, or level of leisure time activity associated with the difference in aerobic fitness level? Spine 2006e;31:90-7.
- UK BEAM Team. United Kingdom back pain exercise and manipulation (UK BEAM) randomised trial: effectiveness of physical treatments for back pain in primary care. BMJ 2004;329:1377-81.
- Turk DC, Meichenbaum D, Genest M. Pain and behavioral medicine: A cognitive behavioral perspective. New York: Guil ford Press, 1983.
- Turk DC, Okifuji A. Psychological factors in chronic pain: evolution and revolution. J Consult Clin Psychol 2002;70:678-90.
- Twisk JWR. Applied longitudinal data analysis for epidemiology. A practical guide. Cambridge: Cambridge University Press, 2003.
- van den Hout JH, Vlaeyen JW, Heuts PH, Zijlema JH, Wijnen JA. Secondary prevention of work-related disability in nonspecific low back pain: does problem-solving therapy help? A randomized clinical trial. Clin J Pain 2003;19:87-96.
- van der Kloot WA, Oostendorp RA, van der Meij J, van den Heuvel J. De Nederlandse versie van 'McGill pain questionnaire': een betrouwbare pijnvragenlijst. Ned Tijdschr Geneeskd 1995;139:669-73.
- van Tulder MW, Koes BW, Bouter LM. A cost-of-illness study of back pain in The Netherlands. Pain 1995;62:233-40.
- van Tulder MW, Koes BW, Bouter LM. Conservative treatment of acute and chronic nonspecific low back pain. A systematic review of randomized controlled trials of the most common interventions. Spine 1997;22:2128-56.
- van Tulder MW, Malmivaara A, Esmail R, Koes BW. Exercise therapy for low back pain. Cochrane Database Syst Rev 2000:Cd000335.
- Verbunt JA, Seelen HA, Vlaeyen JWS, van der Heijden GJ, Heuts PH, Pons K, Knottnerus JA. Disuse and deconditioning in chronic low back pain: concepts and hypothesis on contributing mechanisms. Eur J Pain 2003;7:9-21.
- Vlaeyen JWS, de Jong J, Geilen M, Heuts PHTG, van Breukelen G. The treatment of fear of movement/(re)injury in chronic low back pain: further evidence on the effectiveness of exposure in vivo. Clin J Pain 2002;18:251-61.
- Vlaeyen JWS, Kole-Snijders AM, Boeren RG, van Eek H. Fear of movement/(re)injury in chronic low back pain and its relation to behavioral performance. Pain 1995;62:363-72.
- Vlaeyen JWS, Teeken-Gruben NJ, Goossens ME, Rutten-van Molken MP, Pelt RA, van Eek H, Heuts PH. Cognitive-educational treatment of fibromyalgia: a randomized clinical trial. I. Clinical effects. J Rheumatol 1996;23:1237-45.

References

- VRIN/VRA. Productdefiniëring Revalidatiezorg: registratiepilot revalidatie-instellingen januari 1999 - juli 2000. Utrecht: VRIN, 2000.
- Waddell G. A new clinical model of low back pain and disability. The back pain revolution. London: Churchill Livingstone, 1998.
- Woby SR, Watson PJ, Roach NK, Urmston M. Are changes in fear-avoidance beliefs, catastrophizing, and appraisals of control, predictive of changes in chronic low back pain and disability? Eur J Pain 2004;8:201-10.

Chapter 6

Reduction of pain catastrophizing mediates the outcome of both physical and cognitive-behavioral treatment in chronic low back pain.

Rob J.E.M. Smeets, Johan W.S. Vlaeyen, Arnold D.M. Kester, J. André Knottnerus

Published as:
Smeets RJEM, Vlaeyen JWS, Kester ADM, Knottnerus JA. Reduction of pain catastrophizing mediates the outcome of both physical and cognitive-behavioral treatment in chronic low back pain. Journal of Pain 2006;7:261-271.

Abstract

The aim of this study was to examine whether treatments based on different theories change pain catastrophizing and internal control of pain, and whether changes in these factors mediate treatment outcome. Participants were 211 patients with nonspecific chronic low back pain (CLBP) participating in a RCT, attending active physical treatment (APT, n = 52), cognitive-behavioral treatment (CBT, n = 55), treatment combining the APT and CBT (CT, n = 55), or waiting list (WL, n = 49). Pain catastrophizing decreased in all three active treatment groups, and not in the WL. There was no difference in the change in internal control across all four groups. In all the active treatment groups, patients improved regarding perceived disability, main complaints and current pain at post-treatment and no changes were observed in the WL-group. Depression only changed significantly in the APT-group. Change in pain catastrophizing mediated the reduction of disability, main complaints and pain intensity. In the APT condition, pain catastrophizing also mediated the reduction of depression. Not only cognitive-behavioral treatments but also a physical treatment produced changes in pain catastrophizing that seemed to mediate the outcome of the treatment significantly.
The implications and limitations of these results are discussed.

Perspective

This article shows that treatment elements that do not deliberately target cognitive factors can reduce pain catastrophizing. Reduction in pain catastrophizing seemed to mediate the improvement of functioning in chronic low back pain patients. The results might contribute to the development of more effective interventions.

Introduction

In the treatment of chronic low back pain (CLBP), different treatments have been developed and studied. There is substantial evidence that exercise therapy, cognitive-behavioral therapy, and multidisciplinary treatment are more effective than doing nothing or a waiting list control treatment. But there is no strong evidence that one of the abovementioned active treatments is more effective than the other (Guzman et al. 2002; Hayden et al. 2005a; Morley et al. 1999; Ostelo et al. 2005). Furthermore, the treatment effects are not impressive as most effect sizes are not greater than 0.5, and a substantial proportion of patients do not seem to benefit at all. To improve the effectiveness/efficacy of treatment, we need insight into what mechanisms are responsible for the desired outcomes and what kind of treatment is the best for a particular patient.
Recently, Vlaeyen and Morley (2005) proposed to approach these questions by using the moderator-mediator distinction as described by Baron and Kenny (1986). Moderators provide the answer to the question “in what circumstances does the treatment work?”, whereas mediators concern the question “how does the treatment work?”. Some examples of moderators are dose of treatment and client expectancy, whereas examples of mediators are changing beliefs and behavior contingencies.
Pain catastrophizing and internal control seem to be important factors in the development of chronic pain and disability (Jensen et al. 2001a; Sullivan et al. 2001; Turk and Okifuji 2002) and also have been shown to mediate the outcome of multidisciplinary treatment (Burns et al. 2003a; Burns et al. 1998; Jensen et al. 2001b; Jensen et al. 1994; Keefe et al. 1990). Most studies consisted of uncontrolled designs, but Spinhoven et al. (2004) were the first to show that pain catastrophizing and internal control mediated the outcome of cognitive-behavioral treatment as compared with a waiting list control group. Unfortunately, the outcome measures were composite scores, and a common measure for disability, the main outcome of rehabilitation treatment, was not available. So far, no study examined the mediating role of pain catastrophizing and internal control in a pure physical treatment for CLBP.
Studying the potential mediating factors of a physical, cognitive-behavioral, and multidisciplinary treatment might increase our knowledge regarding the working mechanisms of such treatments and further improve the development of more effective treatments.

Recently, we performed a randomized controlled trial (RCT) in which 3 active treatments based on 3 frequently used theories regarding the development and treatment of CLBP (the deconditioning theory, the cognitive-behavioral theory, and the biopsychosocial theory) were compared with a waiting list control treatment (WL). The treatments were an active physical therapy (APT) to improve aerobic fitness level and low back muscles strength/endurance, a cognitive-behavioral therapy with operant treatment principles and problem solving techniques to improve coping with daily problems, stress and pain (CBT), and multidisciplinary therapy combining the APT and CBT (CT).

The aim of the current analysis was to test whether pain catastrophizing and internal control mediated the outcome in terms of pain, depression, and functional disability in patients with CLBP who received APT, CBT, or CT.

Materials and methods

This study is part of a larger RCT (ISRCTN22714229) of which the methods are extensively described elsewhere (Smeets et al. 2006). The main aim of this RCT was to test the hypotheses that the 3 active treatments are more effective than the WL, and that the combination of both single treatment elements is more effective than the single treatments. A short description of participants, treatments, measures, and study design is presented with special emphasis on the putative mediating factors.

Participants

Patients with CLBP were referred by general practitioners and medical specialists to 3 Dutch outpatient rehabilitation centres. The main inclusion criteria for participation in the RCT were disability caused by nonspecific LBP of more than 3 months, age between 18 to 65 years, and ability to walk at least 100 meters without interruption. Exclusion criteria were vertebral fracture, spinal inflammatory disease, spinal infections or malignancy, current nerve root pathology, spondylolysis or spondylolisthesis, lumbar spondylodesis, medical comorbidity making intensive exercising impossible (e.g. cardiovascular or metabolic disease), ongoing diagnostic procedures or treatment for their CLBP at the time of referral, psychopathology that would hamper individual or group processes, not proficient in Dutch language, pregnancy, and substance abuse that could interfere with the rehabilitation treatment. Patients were requested to stop other therapies for their back complaints, except pain medication.
Of the 309 patients who were referred to the study, 82 (26%) did not participate. The main reasons were not meeting the criteria (n = 31), clear preference for a particular treatment (n = 20), and logistic problems (n = 15). Another 4 patients were excluded immediately after randomization because they did not meet the criteria (other medical diagnosis). Of the 223 patients who started the treatment (APT; n = 53, CBT, n = 58, CT; n = 61, WL; n = 51), 11 patients did not complete any questionnaire immediately after the end of treatment (5%), and 1 patient did not complete the Pain Cognition List (PCL) questionnaire immediately after treatment. Complete results were available for 211 patients (APT; n = 52; CBT, n = 55; CT, n = 55; WL, n = 49). A summary of the baseline data of the total study population is presented in table 1. Comparison of the baseline variables between those who did and those who did not complete questionnaires immediately after treatment did not show significant differences.

Study design

After referral by a consultant in rehabilitation medicine, the patient was invited for a meeting with the research assistant, during which further information about the trial was provided, and inclusion criteria were checked. If eligible for participation, the patient was asked to give his/her informed consent for participation. One to several weeks later the patient was invited to participate in the first assessment (pre-treatment). After this assessment participants were allocated at random to 1 of the 3 possible interventions or the WL condition. Randomization took place per group of 4 patients at a time in each rehabilitation centre. After 10 weeks of therapy the immediately post-treatment assessment took place.

Table 1: Baseline variables for all patients in the four therapy groups (Total n=211)

Variables	APT (n=52)	CBT (n=55)	CT (n=55)	WL (n=49)
Age (yr)	43.00 ± 8.84	42.02 ± 9.47	41.58 ± 10.07	40.63 ± 11.29
Gender (% male)	59.6	40.0	61.8	49.0
Education (%)				
low	67.3	60.0	60.0	63.3
middle - high	32.7	40.0	40.0	36.7
Work status (%)				
full time	32.7	32.7	25.5	24.5
partial sick leave / disability pension	19.2	25.5	34.5	18.4
full sick leave / disability pension	42.3	32.7	32.7	42.9
no job / retired	5.8	7.3	7.3	14.3
Duration of LBP (mo)	57.87 ± 76.28	69.87 ±75.86	56.18 ± 70.62	44.65 ± 72.07
Duration of disability (mo)	29.27 ± 37.67*	50.05 ± 62.94*	35.65 ± 55.71*	24.20 ± 33.26*
Radiation of pain (%)				
no radiation	9.6	16.4	14.5	16.3
above knee	36.5	27.3	41.8	42.9
below knee	53.8	56.4	43.6	40.8
Previous back surgery (%)	17.3	10.9	14.5	16.3
Trauma preceding LBP (%)	17.3	20.0	10,9	28.6
PCL-catastrophizing	40.44 ± 13.94	41.09 ± 11.92	38.82 ± 11.85	38.27 ± 11.92
PCL-internal control	16.77 ± 3.58	15.67 ± 4.09	16.53 ± 3.21	15.91 ± 4.02
RDQ	14.15 ± 3.74	13.87 ± 3.55	13.67 ± 3.66	13.82 ± 3.86
PSC	74.25 ± 14.60	74.84 ± 16.24	71.95 ± 17.14	77.18 ± 11.08
Current pain	51.06 ± 26.78	49.07 ± 23.92	45.62 ± 23.94	50.73 ± 25.86
BDI	10.52 ± 7.62	10.67 ± 7.11	9.35 ± 6.28	9.85 ± 7.81

Values presented as means and standard deviation or percentage.
APT = active physical therapy; CBT = cognitive-behavioral therapy; CT = combined therapy; WL = waiting list;
LBP = low back pain; PCL = Pain Cognition List; RDQ = Roland Disability Questionnaire;
PSC = Patient-Specific Complaints; BDI = Beck Depression Inventory.
* Significant differences between therapy groups (P < 0.05).

The medical ethics committee of the Rehabilitation Foundation Limburg and Institute for Rehabilitation Research at Hoensbroek in The Netherlands approved the study protocol.

Interventions

All interventions started with a group meeting of a maximum of 4 patients during which the rationale of the particular therapy was explained. Emphasis was put on the responsibility of the patient for making plans to keep on being active after the treatment. Each treatment lasted 10 weeks and started with the explanation of the rationale of that particular treatment. A written summary of the rationale was given to the patients.

All patients were asked to attend as much as possible sessions offered during the treatment.To assure sufficient contrast between the 3 different treatments and to avoid incorporating possible confounding elements, all therapists were instructed not to discuss general aspects concerning back pain origin, anatomy, and ergonomics. No other interventions than those that were chosen for the APT, CBT, or CT took place. In case of acute and severe psychosocial stress or pathology (severe depression, high risk for suicide, or personal problems the patient did not wish to discuss during the group treatment), a consultation of a clinical psychologist or social worker was possible. During this consultation the therapist tried to find out what the exact problem was and consecutively, when judged necessary, arranged for professional help outside the rehabilitation centre. All therapists received an extensive training before the start of the trial. They attended refresher courses; two 1-day courses during the first year and 1 each year in the next 2 years of the trial. For each intervention detailed treatment manuals were used by all therapists. The clinical psychologists and social workers had at least 5 years of experience in treating CLBP patients.

APT

In a group of 4 maximum, patients were invited to perform 30 minutes of aerobic training on a bicycle (65% to 80% heart rate maximum) and 75 minutes of strength and endurance training of their lower back and upper leg muscles (3 series of 15 to 18 repetitions in a dynamic-static manner), 3 times a week during 10 weeks. The exercises consisted of leg extension while sitting on knees and hands, trunk lifting and lifting both legs while lying prone on a couch. During the exercises assistive weight by a pulley system or extra weight placed on the body of the patient was used, depending on the calculated load (70% of 1-Repetition Maximum, which allowed 15 to 18 repetitions until muscular fatigue occurred). No kind of cognitive-behavioral intervention was given, and patients did not receive homework assignments. The training was supervised by 2 physiotherapists.

CBT

From a theoretical point of view, Aldrich et al. (2000) hypothesized that patients with chronic pain tend to persevere in their attempt to solve an unsolvable problem, namely pain, despite the experience of repeated failure. This perseverance may keep them hypothetically stuck in a vicious circle of chronic pain. To intervene in this process it might be more important to help patients to identify and cope with the consequences of pain in everyday life. Problem solving

training (PST) can help the patients to redefine their problem(s) and focus more on other individually relevant daily life goals that can be achieved by using graded activity (GA) techniques. Therefore CBT consisted of operant behavioral GA training (Fordyce 1976; Sanders 1996) and PST (Nezu 1986; Nezu and Perri 1989; van den Hout et al. 2003). The GA started with 3 group sessions followed by a maximum of 17 individual sessions of 30 minutes. During the GA, the therapist focused on a time-contingent gradual increase or pacing of 3 activities that were considered important and relevant for the patient's personal situation. After a baseline, activity tolerance level was calculated and final treatment goals were set, and the patient started performing the selected activities following quotas for each day, starting from 70% to 80% of the baseline with gradually increasing activity levels toward the final treatment goals. The patient was instructed only to perform the agreed amount of activity and not to perform less or more, even when he/she felt capable of doing so. The daily performance was graphically registered in a personal diary, and therapists were instructed to discuss these graphs regularly with the patient, while positively reinforcing any progress toward the preset goals. Contrary to GA therapy that has been studied in other trials and in order to create as much contrast as possible with the APT, no physical therapy elements (e.g. muscle strengthening or aerobic exercises) were incorporated.

The PST started with 3 initial sessions in which the rationale of training and the skill of positive problem orientation were discussed. Sessions 4 to 10 focused on problem definition and formulation, generation of alternatives, decision-making, implementation and evaluation. Patients received a course book with additional information, a summary of each session, and homework assignments. The training of the skills and application were the main focus of the therapy, rather than one specific problem area. Patients were free to select their own personal problem areas. After each session, homework assignments were given to practice the skills in everyday life situations. A clinical psychologist or social worker, specifically trained to guide this intervention, provided 10 sessions of 1½ hours to a maximum of 4 patients at a time.

CT

CT consisted of APT in combination with the PST, both offered in the same frequency and duration as described before. The patient was told that he/she first had to gain enough aerobic fitness and strength before increasing his/her activities. The GA was not started until the third week, and started with the selection of the 3 patient-relevant activities. By the end of the fourth week the final goals and daily quota were set. In total 19 sessions, with a total duration of 11 hours were given.

WL

The patients assigned to the WL were requested to wait 10 weeks, after which they were offered a regular individual rehabilitation treatment. During the waiting period, patients were not allowed to participate in diagnostic or therapeutic procedures because of their CLBP.

Assessment
Assessments (questionnaires) were carried out before treatment and immediately after 10 weeks of active treatment or waiting by research assistants who were not aware of the patient's assigned treatment.

Outcome measures
Disability
The Roland Disability Questionnaire (RDQ) was used. This questionnaire has been shown to be a valid and reliable instrument in the evaluation of CLBP treatment (Beurskens et al. 1996; Gommans et al. 1997; Roland and Morris 1983; Stratford et al. 1994; Stratford et al. 1998). In this sample the internal consistency was sufficient (Cronbach alpha = 0.73).

Patient-specific complaints
The patient-specific main complaints were identified by using the patient-specific approach method from Tugwell et al. (1987; Beurskens et al. 1999). The patient has to indicate the 3 most limited functional activities and to rate the difficulty in performing these activities during the previous week on a 100-mm visual analogue scale (VAS). This method appeared to be valid, reliable and of sufficient responsiveness (Beurskens et al. 1996; Ostelo et al. 2004).

Pain
A 100-mm VAS with 'no pain' on the left side and 'unbearable pain' on the right side was used to measure the current pain intensity. Relevance, validity and reliability have been sufficiently tested for patients with LBP (Carlsson 1983; Revill et al. 1976; Sriwatanakul et al. 1983).

Depression
The level of depression was measured by the Beck Depression Inventory (BDI) (Beck et al. 1979; Zitman et al. 1989), a reliable, valid and widely used questionnaire in CLBP research (Beck et al. 1988). The internal consistency in this population was excellent (Cronbach alpha = 0.86).

Mediating factors
Pain catastrophizing and internal control of pain
The subscales pain catastrophizing and internal control of pain of the Pain Cognition List (PCL) were used to measure pain catastrophizing and internal control of pain. The reliability and stability of this questionnaire has been proven to be sufficient in CLBP patients (Vlaeyen et al. 1990; Vlaeyen et al. 2003). The validity of the subscales was supported by the meaningful pattern of correlations with other relevant constructs (Vlaeyen et al. 2003). For example, the subscale of pain catastrophizing showed correlations of approximately 0.70 with the catastrophizing subscale of the Coping Strategies Questionnaire (CSQ) (Rosenstiel and Keefe 1983). Furthermore, pain catastrophizing correlated with negative emotions, such as depression (0.66 with BDI)(Beck et al. 1979) and fear (0.51 with Tampa Scale for Kinesiophobia [TSK]) (Vlaeyen et al. 1995b). The subscale internal control of pain correlated positively with other measures of internal locus of control measures, 0.37 with subscale perceived control of the CSQ and 0.32 with Multidimensional Health Locus of Control Scale (MHLC) (Wallston et al. 1978).
The PCL is a 39-item self-report questionnaire with a 5-point Likert scale answering categories ranging from "completely disagree" to "completely agree". The pain catastrophizing subscale consists of 16 items and the total score ranges from 16 to 80. Some items are "My thoughts are always concentrating on the pain", "I feel like an unlucky person", "I think that fate has struck me". The higher the score, the more the person is catastrophizing. Internal consistency appeared to be excellent in our sample (Cronbach alpha = 0.89). The internal control of pain subscale consists of 5 items. Typical items are: "I know a way to decrease my pain a little", "I think that I can influence my pain positively". Two questions need to be scored inversely (e.g. "I think that I can't do anything against my pain"), and next a total score of these 5 items is calculated. The total score ranges from 5 to 25, with a higher score indicating increased control of pain. The internal consistency also appeared to be sufficient (Cronbach alpha = 0.68).

Compliance and treatment integrity
To check whether patients were compliant with the allocated treatment and the therapists did follow the treatment manual instructions, each therapist kept precise records on the presence during treatment, the amount of exercise (duration, intensity of exercising by monitoring heart rate during cycling and amount of repetitions and weight displaced during muscle training), and choice of activities and increase in time of these activities for the GA training. At the end of treatment each therapist made copies of the graphs in which the patient recorded the amount of activities he/she performed during the treatment and at home.
In accordance with other studies and guidelines, the presence at the treatment had to be at least 2/3 of the maximal amount of administered sessions in order to conclude that the treatment was of sufficient intensity (ACSM 1998; van den Hout et al. 2003).
During the trial an independent researcher made several observations of all active treatments (on-site, audio and video recordings). The independent researcher analyzed whether the therapists did not engage in delivering treatments outside the manual, for example decatastrophizing during the APT.

Statistical analysis

All statistical analyses were carried out according to the intention-to-treat principle. All patients, including withdrawals from treatment and patients with poor compliance remained in the group to which they were randomized.

To check whether treatment differentially influenced pain catastrophizing and internal pain control, linear regression analyses were performed with the score of the putative mediating factor at the post-treatment assessment as the dependent variable. The variables age, gender, treatment centre, the baseline score of the putative mediating factor, and the baseline value of the outcome of interest always stayed in the regression analysis as covariates. Variables for which, despite randomization, differences between treatment groups at baseline were found ($P < 0.10$) were added to the regression equation as a covariate.

To account for possible dependence of the outcomes within the groups of 4 patients who were randomized together, a random intercept term for these groups was included in all models by using multilevel analyses (SPSS mixed linear; SPSS Inc, Chicago, Ill).

Dummy variables for the 3 active treatments were made and the coefficients for these dummies were the estimated treatment effects with respect to WL (reference group). Differences between CT and APT and CBT, respectively, were estimated similarly. Effect sizes are estimated as Hedges's g (Hedges and Olkin 1985).

Mediation can be investigated by the 3-step method described by Baron and Kenny (1986).

Mediation is suggested when the change in the putative mediating factors is significantly related to treatment as the independent variable, outcome is significantly related to treatment as the independent variable and finally, the relationship between outcome and treatment decreases (or goes to zero) when the change in the mediating factor is entered into the equation.

Three successive multilevel regression analyses were conducted. The first analysis was performed to test whether pain catastrophizing and internal control of pain significantly changed while comparing the 3 active treatment groups to the WL (reference treatment). The treatment was the independent variable, and pain catastrophizing or internal control immediately post-treatment was the dependent variable, thereby controlling for age, gender, baseline value of pain catastrophizing or internal control of pain, treatment centre, and variables that turned out to be unequally divided between treatment groups at baseline. When the change in pain catastrophizing or internal control was not significantly different between the WL and one of the 3 active treatments ($P < 0.05$), the procedure was terminated at this point because this variable could no longer be regarded as a potential mediating factor for that particular active treatment.

The second multiple regression analysis was performed by using the outcome measure immediately post-treatment as dependent variable, and treatment as independent variable thereby controlling for age, gender, baseline value of outcome measure, baseline score for the mediator, treatment centre and variables that turned out to be unequally divided between treatment group at baseline. The WL was the reference treatment.

The final regression model was analyzed with the putative mediating factor added to the independent variables of the

second model and the outcome measure immediately post-treatment as the dependent variable. Mediation was established when the regression coefficient of the effect of treatment on the outcome measure with the mediating factor taken into the analysis, was substantially lower than the regression coefficient of treatment in the second regression analysis. The change in regression coefficients for treatment between models 2 and 3 was tested for significance by using Sobel's *t*-test (Aroian 1944; McKinnon et al. 2002; Sobel 1982).

Results

Randomization check and group dependence

The baseline characteristics are shown in table 1. Demographic variables showed similar distribution for all treatment groups ($P > 0.10$), except for gender, which was controlled for in the analyses. Regarding disease characteristics, only the duration of disability showed a significant difference between treatment groups, and was entered into the regression analyses as a covariate.

The multilevel regression analyses showed that the dependence within randomization groups was usually small, intraclass correlations (ICCs) being never larger than 0.10 with only one exception; the second regression model on the patient specific complaints showed an ICC of 0.15.

Compliance and integrity

When applying the criterion of at least 2/3 attendance of all possible sessions, 84.6% of all patients attending the APT had a training of sufficient intensity. Of all CBT-patients, 81.8% and 80% had a sufficient number of sessions of GA and PST, respectively. For the CT patients, 80% had sufficient physical training and 69.1% sufficient GA and PST.

Further analysis of the records showed that the APT therapists seemed to have adhered very well to the treatment manual. They kept accurate records of the adjustments of the training and explained the reasons for adjustment carefully. Adjustments were also in accordance with the treatment manual. Analysis of the records of the GA therapists and graphs filled out by the patients showed that the selection of 3 activities, the activity tolerance level and the gradual increase or pacing of these activities was overall satisfactory. The same applies to the PST therapists.

The independent researcher analyzed the observations (on-site, video and audio tapes) and concluded that the APT therapists did not use additional decatastrophizing methods, and the PST and GA therapists did not incorporate physiologic training principles.

Treatment effects on mediating and outcome measures

Mediating factors

In table 2, the mean PCL scores in the WL at post-treatment and the mean difference of this score for each active treatment in comparison to the WL are presented.

Pain catastrophizing significantly decreased in all 3 active therapies when compared with the WL. Because internal control of pain did not change significantly in all active treatments and therefore could not be a mediating factor, no further analyses for this factor were carried out.

Table 2: Effects of APT, CBT and CT as compared with WL on putative mediating factors and outcome measures and effect sizes

Dependent variable	WL mean ± SD	APT mean difference (95% CI)†	APT effect size (95% CI)	CBT mean difference (95% CI)†	CBT effect size (95% CI)	CT mean difference (95% CI)	CT effect size (95% CI)
PCL-pain catastrophizing	38.71± 12.94	-5.18 (-8.73 to -1.62)**	0.43 (0.13 to 0.72)	-4.23 (-7.79 to -0.67)*	0.35 (0.06 to 0.64)	-4.23 (-7.73 to -0.73)*	0.35 (0.06 to 0.63)
PCL-internal control of pain	16.16 ± 3.65	0.37 (-1.20 to 1.94)	0.10 (-0.31 to 0.51)	0.69 (-0.86 to 2.23)	0.18 (-0.23 to 0.53)	0.70 (-0.84 to 2.24)	0.18 (-0.22 to 0.59)
RDQ	13.91 ± 4.82	-2.38 (-4.19 to -0.56)*	0.44 (0.10 to 0.77)	-3.09 (-4.89 to -1.28)**	0.56 (0.23 to 0.89)	-2.50 (-4.29 to -0.72)**	0.46 (0.13 to 0.78)
Current pain	53.35 ± 22.6	-9.14 (-17.74 to -0.55)*	0.36 (0.02 to 0.70)	-15.64 (-24.23 to -7.06)**	0.61 (0.28 to 0.95)	-8.60 (-17.08 to -0.11)*	0.34 (0.01 to 0.67)
PSC	74.59 ± 14.59	-10.31 (-19.38 to -1.23)*	0.49 (0.06 to 0.92)	-15.55 (-24.51 to -6.60)**	0.74 (0.31 to 1.17)	-16.97 (-25.85 to -8.08) **	0.81 (0.39 to 1.23)
BDI	9.63 ± 7.89	-2.09 (-3.86 to -0.32)*#	0.32 (0.05 to 0.59)	-1.65 (-3.42 to 0.12)	0.25 (-0.02 to 0.52)	0.04 (-1.73 to 1.79)	-0.01 (-0.27 to 0.26)

Mean differences, confidence interval (CIs), and corresponding P values were estimated while adjusting for age, gender, treatment centre, baseline score of outcome measure, duration of disability, and dependence within randomization groups.
APT = active physical therapy; CBT = cognitive-behavioral therapy; CT = combined therapy; WL = waiting list; SD = standard deviation; PCL = Pain Cognition List; RDQ = Roland Disability Questionnaire; PSC = Patient-Specific Complaints; BDI = Beck Depression Inventory.
† Values are the mean difference between treatment and WL, WL-score is the score at post-treatment.
* P < 0.05; ** P < 0.01; # APT > CT, P < 0.05.

Outcome measures

The changes of outcome measures in each active treatment in comparison with the post-treatment value of the WL, as well as effect-sizes, are shown in table 2. Significant differences between CT and APT or CBT, respectively, are indicated. Although the active treatments were quite different, all 3 treatments were apparently equally effective regarding the reduction of disability, patient-specific complaints, and current pain. APT was the only treatment that was effective in reducing the depression score, even when compared with the CT. Because CBT and CT did not significantly change the BDI score in comparison with the WL, no further analyses regarding mediation on BDI for these 2 treatments were performed. No other significant differences between CT and the other 2 active treatments were found.

Change in pain catastrophizing as a mediating factor of treatment outcome

Table 3 displays the regression coefficients and their standard errors as well as Sobel's *t*-test and level of significance for the change of the coefficient after correction for pain catastrophizing for each active therapy when compared with the WL group. Correcting for pain catastrophizing significantly changed the regression coefficients in APT, CBT, and CT for the RDQ, patient-specific complaints, and current pain, suggesting that pain catastrophizing mediated the decrease in disability, major complaints and pain. Because APT was the only treatment that showed a significant improvement regarding depression when compared with WL, the mediating role of pain catastrophizing on depression could only be investigated and was also established for this treatment.

Table 3: Regression coefficients and standard errors (SEs) of three successive multiple regression analyses on the mediation effect of pain catastrophizing on disability, pain, patient-specific complaints and depression

	Regression coefficient 1	SE	Regression coefficient 2	SE	Regression coefficient 3	SE	Sobel's *t*	P value
RDQ								
APT vs WL	-5.252	1.745	-2.415	0.905	-1.284	0.820	-2.754	0.006
CBT vs WL	-4.214	1.749	-3.140	0.905	-2.217	0.814	-2.269	0.023
CT vs WL	-4.158	1.718	-2.524	0.889	-1.636	0.801	-2.278	0.022
Current pain (VAS)								
APT vs WL	-5.192	1.791	-9.282	4.287	-4.703	4.042	-2.526	0.012
CBT vs WL	-4.159	1.794	-15.826	4.302	-12.165	4.027	-2.103	0.035
CT vs WL	-4.043	1.766	-8.685	4.231	-5.124	3.958	-2.081	0.037
PSC*								
APT vs WL	-5.104	1.779	-10.389	4.536	-6.723	4.313	-2.472	0.013
CBT vs WL	-4.177	1.782	-15.647	4.489	-12.665	4.246	-2.102	0.036
CT vs WL	-4.110	1.759	-17.010	4.436	-14.192	4.200	-2.096	0.036
BDI*†								
APT vs WL	-4.803	1.711	-2.010	0.901	-0.605	0.751	-2.698	0.007

NOTE. Model 1, pain catastrophizing as dependent variable and treatment as independent variable; Model 2, outcome as dependent variable and treatment as independent variable; Model 3, outcome as dependent variable and treatment as well as pain catastrophizing as independent variables.
APT = active physical therapy; CBT = cognitive-behavioral therapy; CT = combined therapy; WL = waiting list; RDQ = Roland Disability Questionnaire; PSC = Patient-Specific Complaints; BDI = Beck Depression Inventory.
* One patient did not complete PSC at post-treatment, 2 patients did not complete BDI at pre-treatment.
† Data for CBT vs WL and CT vs WL not presented because there was no significant difference on BDI-score.

Discussion

The aim of this study was to test whether pain catastrophizing and internal control of pain mediated the outcome of a physical treatment, a cognitive-behavioral treatment and a combined treatment in CLBP. While being compared with the waiting list, all 3 active treatments significantly decreased the level of pain catastrophizing but unexpectedly not the level of internal control of pain. The active treatments also showed a significant decrease of disability, current pain and patient-specific complaints. The depression score only significantly decreased in the physical treatment (APT). The reduction of pain catastrophizing significantly mediated disability, patient-specific complaints and current pain when the patients received either one of the active treatments. In the APT, pain catastrophizing also mediated the level of depression, but it should be noted that the mean baseline score of the BDI was already relatively low (mean score 10.1 ± 7.18), and by using the cutoff score of ≥ 21 as recommended by Geisser et al. (2000), only 7.6% of all patients could be classified as being depressed at baseline. Therefore, the clinical relevance of the improvement of the BDI score might be questioned.

In general, the overall effect-sizes for the outcome measures, are comparable to the results presented in recent reviews and meta-analysis on different active treatments (Guzman et al. 2002; Hayden et al. 2005a; Hayden et al. 2005b; Morley et al. 1999; Ostelo et al. 2005).

The results regarding the effect of CBT and CT on pain catastrophizing are comparable with the results of a previous RCT in CLBP (Spinhoven et al. 2004) and an uncontrolled study in fibromyalgia (Nielson and Jensen 2004). The mean effect-sizes of pain catastrophizing are also comparable to the mean effect size of negative coping, in a meta-analysis on CBT for chronic pain (Morley et al.1999).
In another study, comparing a cognitive-behavioral treatment, an operant behavioral treatment and a waiting list control group, the use of coping strategies (pain catastrophizing decreased and use of hoping and praying increased) was an important factor in the final treatment results (Turner and Clancy 1988). Interestingly, this change was not only found in the cognitive-behavioral treatment, which was explicitly aimed at increasing the use of these coping skills, but also in the operant behavioral treatment. In contrast, Ostelo et al. (2003) found that a behavioral-GA program and care as usual in patients with LBP of more than 6 weeks past first time lumbar disk surgery did not significantly change pain catastrophizing in both groups. In an uncontrolled study of 54 patients of the originally 83 patients who started a CBT program, Woby et al. (2004) showed that changes in increased perception of pain control were related to the reduction of disability, and not changes in pain catastrophizing.

This is the first study showing that the effects of a physical treatment in which no cognitive-behavioral treatment elements were incorporated, seemed to be mediated by the decrease in pain catastrophizing. One explanation might be that patients were exposed to a number of rigorous activities, thereby challenging the idea that pain is a sign of

impending threat. Similar findings were reported by Mannion et al. (2001), who found that an active therapy program appeared capable of modifying fear-avoidance beliefs, possibly as a result of the positive experience of completing the prescribed exercises without undue harm.
It was suggested earlier that changes in pain impact and pain catastrophizing occurred following improved motor functioning during physical training (Vlaeyen et al. 1995a), which is in line with the more general assumption that cognitions can best be altered by treatment techniques producing changes in motor behavior (Bandura 1977). This change in motor functioning might have resulted in increased confidence in one's own body and consequently created greater possibilities regarding performing activities. Otherwise, it is possible that increased physical capacity (muscle strength, muscle endurance and aerobic capacity) might have had a positive effect on performing activities, which eventually resulted in a reduced level of pain catastrophizing.

Apparently the interventions which are supposed to alter maladaptive thoughts and improve coping with pain and pain disability, did not show a greater decrease of pain catastrophizing than the physical treatment. It is possible that when combined with the operant behavioral GA treatment, the specific cognitive training element (PST) was not able to also change pain catastrophizing. A recent study showed that the same PST added to an operant GA treatment did have an additional effect in decreasing sick leave and disability in patients with subacute sick leave due to LBP, although no differential effects were found for pain catastrophizing which decreased in both interventions (van den Hout et al. 2003). One possibility for the lack of differential effects in our study might be the compliance rate. A substantially high number of patients did not attend sufficient sessions of PST and GA (20% in CBT and 31% in CT). On the other hand, in APT and CT, 15% and 20%, respectively, of all patients also did not attend sufficient physical training sessions. Although this difference in attendance rate between APT and CBT/CT, respectively, is not very substantial, we cannot rule out that this is a possible reason why APT is equally effective in decreasing pain catastrophizing.
Another explanation might be insufficient treatment integrity. Although the independent observer did not encounter the use of physiologic training principles in the GA and PST and all therapists had at least 5 years of experience in treating CLBP patients and attended training and supervision sessions throughout the study, it should be taken into account that the observations were performed by only one independent observer, and the reliability of his observations has not been tested. Therefore, it cannot be totally ruled out that the therapists did use the most proper treatments principles. Otherwise, compared with the study of van den Hout et al. (2003), our CBT and CT showed at least an equal and even a tendency to a larger decrease of pain catastrophizing, indicating at least a comparable effectiveness.
A further possibility might be that the cognitive changes described in the abovementioned studies on CBT were not caused by the CBT specifically. In these multidisciplinary treatments, also physical and occupational therapies were incorporated which could have caused cognitive changes. Several other studies suggest that improvements in physical capacities (muscle strength, cardiovascular capacity and so on) indeed account for favorable outcomes (Burns et al. 1998; Hazard et al. 1989; Mayer et al. 1995; Mellin et al. 1990).

Furthermore, the present study results might be caused at least partially by nonspecific factors of the treatment such as a clear treatment rationale, a highly structured and standardized treatment program, and emphasis on active participation and responsibility for generalization (Kole-Snijders et al. 1999; Spinhoven et al. 2004).
In conclusion, this study shows that treatment elements that do not deliberately target cognitive factors but instead concentrate on letting the patients experience that performing physical exercises or normally hampered activities in a controlled environment is still possible are responsible for a substantial cognitive change.
The lack of change in internal control of pain was not expected. This might be attributed to the fact that the cognitive treatment consisting of the PST and the operant behavioral GA training was not specifically aimed at changing the internal control of pain. In the study of Spinhoven et al. (2004), which did show an increase of control of pain, the cognitive treatment specifically aimed at increasing pain control and self-efficacy beliefs. Another explanation might be the internal control of pain subscale of the PCL itself. Currently there is no information about the responsiveness of this scale. In addition, the baseline score seemed already moderate to high in all treatment groups (mean score ranged from 15.6 to 16.6). However, in the absence of normative values for the healthy population, it is difficult to judge whether there existed a ceiling effect for this measure.

A weakness of this study and most previous studies alike is that the results are based on self-report only, which may be influenced by a number of self-serving biases. On the other hand, by using the 3-step process for assessing mediation as described by Baron and Kenny (1986) and statistically testing the results with Sobel's *t*-test, the mediating effect of pain catastrophizing seems quite solid. To examine whether we are not merely measuring correlation and that the change in mediator really precedes the change in outcome, Burns et al. (2003a; 2003b) suggested using cross-lagged correlation. In their uncontrolled study on chronic pain patients attending a multidisciplinary treatment they collected data on pain catastrophizing, helplessness, depression and the outcome (multidimensional pain inventory) at before, during, and after treatment. An early change in cognitions appeared to predict late treatment change of pain and disability but not the other way around.
Although we performed a controlled study, we were not able to collect data at mid-treatment and a cross-lagged correlation was not possible. Therefore, the direction of causality among the variables remains ambiguous. Taking the results of Burns et al. (2003a; 2003b) into consideration, it seems that the direction of mediation is from pain catastrophizing to outcome measures and not the other way around. More research, especially on temporal precedence of factors that mediate the effects of our rehabilitation treatments for CLBP, is warranted. Once we have detected the most potent mediators, we can start to develop treatments that show the best results in changing these mediators. It might contribute to the development of interventions that have increased cost-effectiveness and can be geared onto the mechanisms that produce the effects at which we aim (Vlaeyen and Morley 2005).

Acknowledgements

We want to thank Bart Poels of the TRANSFORM-research group who was responsible for the observations of treatments and the additional analyses of treatment integrity.
This study is supported by Zorgonderzoek Nederland/Medische Wetenschappen (ZonMw) grant number 014-32-007.

- ACSM. ACSM position stand on the recommended quantity and quality of exercise for developing and maintaining cardiorespiratory and muscular fitness, and flexibility in adults. Med Sci Sports Exerc 1998;30:975-91.
- Aldrich S, Eccleston C, Crombez G. Worrying about chronic pain: vigilance to threat and misdirected problem solving. Behav Res Ther 2000;38:457-70.
- Aroian LA. The probability function of the product of two normally distributed variables. Annals of Mathematical Statistics 1944;18:265-71.
- Bandura A. Self-efficacy: towards a unifying theory of behaviour change. Psychol Rev 1977;84:191-215.
- Baron RM, Kenny DA. The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. J Pers Soc Psychol 1986;51:1173-82.
- Beck AT, Rush AJ, Shaw BF, Emmery G. Cognitive therapy of depression. New York: Wiley and sons, 1979.
- Beck AT, Steer RA, Garbin MG. Psychometric properties of the Beck Depression Inventory: Twenty-five years of evaluation. Clin Psychol Rev 1988;8:77-100.
- Beurskens AJ, de Vet HC, Koke AJ. Responsiveness of functional status in low back pain: a comparison of different instruments. Pain 1996;65:71-6.
- Beurskens AJ, de Vet HC, Koke AJ, Lindeman E, van der Heijden GJ, Regtop W, Knipschild PG. A patient-specific approach for measuring functional status in low back pain. J Manipulative Physiol Ther 1999;22:144-8.
- Burns JW, Glenn B, Bruehl S, Harden RN, Lofland K. Cognitive factors influence outcome following multidisciplinary chronic pain treatment: a replication and extension of a cross-lagged panel analysis. Behav Res Ther 2003a;41:1163-82.
- Burns JW, Johnson BJ, Mahoney N, Devine J, Pawl R. Cognitive and physical capacity process variables predict long-term outcome after treatment of chronic pain. J Consult Clin Psychol 1998;66:434-9.
- Burns JW, Kubilus A, Bruehl S, Harden RN, Lofland K. Do changes in cognitive factors influence outcome following multidisciplinary treatment for chronic pain? A cross-lagged panel analysis. J Consult Clin Psychol 2003b;71:81-91.
- Carlsson AM. Assessment of chronic pain. I. Aspects of the reliability and validity of the visual analog scale. Pain 1983;16:87-101.
- Fordyce WE. Behavioral methods for chronic pain and illness. St. Louis: Mosby, 1976.
- Geisser ME, Roth RS, Theisen ME, Robinson ME, Riley JL 3rd. Negative affect, self-report of depressive symptoms, and clinical depression: relation to the experience of chronic pain. Clin J Pain 2000;16:110-20.
- Gommans IHB, Koes BW, van Tulder MW. Validiteit en responsiviteit Nederlandstalige Roland Disability Questionnaire. Vragenlijst naar functionele status bij patiënten met lage rugpijn. Ned T Fysiother 1997;107:28-33.
- Guzman J, Esmail R, Karjalainen K, Malmivaara A, Irvin E, Bombardier C. Multidisciplinary bio-psycho-social rehabilitation for chronic low back pain. Cochrane Database Syst Rev 2002:CD000963.
- Hayden JA, van Tulder MW, Malmivaara AV, Koes BW. Meta-analysis: exercise therapy for nonspecific low back pain. Ann Intern Med 2005a;142:765-75.
- Hayden JA, van Tulder MW, Tomlinson G. Systematic review: strategies for using exercise therapy to improve outcomes in chronic low back pain. Ann Intern Med 2005b;142:776-85.
- Hazard RG, Fenwick JW, Kalisch SM, Redmond J, Reeves V, Reid S, Frymoyer JW. Functional restoration with behavioral support. A one-year prospective study of patients with chronic low-back pain. Spine 1989;14:157-61.
- Hedges LV, Olkin I. Statistical methods for meta-analysis. Orlando: Academic Press, 1985.
- Jensen IB, Bergstrom G, Ljungquist T, Bodin L, Nygren AL. A randomized controlled component analysis of a behavioral medicine rehabilitation program for chronic spinal pain: are the effects dependent on gender? Pain 2001a;91:65-78.

- Jensen MP, Turner JA, Romano JM. Changes in beliefs, catastrophizing, and coping are associated with improvement in multidisciplinary pain treatment. J Consult Clin Psychol 2001b;69:655-62.
- Jensen MP, Turner JA, Romano JM, Lawler BK. Relationship of pain-specific beliefs to chronic pain adjustment. Pain 1994;57:301-9.
- Keefe FJ, Caldwell DS, Williams DA, Gil KM, Mitchell D, Robertson C, Martinez S, Nunley J, Beckham JC, Crisson JE, Helms M. Pain coping skills training in the management of osteoarthritic knee pain: A comparative study. Behav Ther 1990;21:49-62.
- Kole-Snijders AM, Vlaeyen JW, Goossens ME, Rutten-van Molken MP, Heuts PH, van Breukelen G, van Eek H. Chronic low back pain: what does cognitive coping skills training add to operant behavioral treatment? Results of a randomized clinical trial. J Consult Clin Psychol 1999;67:931-44.
- Mannion AF, Müntener M, Taimela S, Dvorak J. Comparison of three active therapies for chronic low back pain: results of a randomized clinical trial with one-year follow-up. Rheumatology 2001;40:772-8.
- Mayer TG, Polatin P, Smith B, Smith C, Gatchel R, Herring SA, Hall H, Donelson RG, Dickey J, English W. Spine rehabilitation. Secondary and tertiary nonoperative care. Spine 1995;20:2060-6.
- McKinnon DP, Lockwood C, Hoffman JM, West SG, Sheets V. A comparison of methods to test mediation and other intervening variable effects. Psychol Methods 2002;7:81-104.
- Mellin G, Harkapaa K, Hurri H, Jarvikoski A. A controlled study on the outcome of inpatient and outpatient treatment of low back pain. Part IV. Long-term effects on physical measurements. Scand J Rehabil Med 1990;22:189-94.
- Morley S, Eccleston C, Williams A. Systematic review and meta-analysis of randomized controlled trials of cognitive behaviour therapy and behaviour therapy for chronic pain in adults, excluding headache. Pain 1999;80:1-13.
- Nezu AM. Efficacy of a social problem-solving therapy approach for unipolar depression. J Consult Clin Psychol 1986;54:196-202.
- Nezu AM, Perri MG. Social problem-solving therapy for unipolar depression: an initial dismantling investigation. J Consult Clin Psychol 1989;57:408-13.
- Nielson WR, Jensen MP. Relationship between changes in coping and treatment outcome in patients with Fibromyalgia Syndrome. Pain 2004;109:233-41.
- Ostelo RW, de Vet HC, Berfelo MW, Kerckhoffs MR, Vlaeyen JW, Wolters PM, van den Brandt PA. Effectiveness of behavioral graded activity after first-time lumbar disc surgery: short term results of a randomized controlled trial. Eur Spine J 2003;12:637-44.
- Ostelo RW, de Vet HC, Knol DL, van den Brandt PA. 24-item Roland-Morris Disability Questionnaire was preferred out of six functional status questionnaires for post-lumbar disc surgery. J Clin Epidemiol 2004;57:268-76.
- Ostelo RW, Tulder M, Vlaeyen J, Linton S, Morley S, Assendelft W. Behavioural treatment for chronic low-back pain. Cochrane Database Syst Rev 2005:Cd002014.
- Revill SI, Robinson JO, Hogg MIJ. The reliability of a linear analog for evaluating pain. Anaesthesia 1976;31:1191-8.
- Roland M, Morris R. A study of the natural history of back pain. Part I: development of a reliable and sensitive measure of disability in low-back pain. Spine 1983;8:141-4.
- Rosenstiel AK, Keefe FJ. The use of coping strategies in chronic low back pain patients: relationship to patient characteristics and current adjustment. Pain 1983;17:33-44.
- Sanders SH. Operant conditioning with chronic pain: back to basics. In: RJ Gatchel, DC Turk. Psychological approaches to pain management. A practitioner's handbook. New York: The Guilford press, 1996, pp. 112-30.

References

- Smeets RJ, Vlaeyen JW, Hidding A, Kester AD, van der Heijden GJ, van Geel AC, Knottnerus JA. Active rehabilitation for chronic low back pain: Cognitive-behavioral, physical, or both? First direct post-treatment results from a randomized controlled trial [ISRCTN22714229]. BMC Musculoskelet Disord 2006;7:5.
- Sobel ME. Asymptomatic confidence intervals for indirect effects in structural equations models. In: W Leinhart,. Sociological methodology. San Fransisco: Jossy-Bass, 1982, pp. 290-312.
- Spinhoven P, Ter Kuile M, Kole-Snijders AM, Hutten-Mansfeld M, Den Ouden DJ, Vlaeyen JW. Catastrophizing and internal pain control as mediators of outcome in the multidisciplinary treatment of chronic low back pain. Eur J Pain 2004;8:211-9.
- Sriwatanakul K, Kelvie W, Lasagna L, Calimlim JF, Weis OF, Mehta G. Studies with different types of visual analog scales for measurement of pain. Clin Pharmacol Ther 1983;34:234-9.
- Stratford PW, Binkley J, Solomon P, Gill C, Finch E. Assessing change over time in patients with low back pain. Phys Ther 1994;74:528-33.
- Stratford PW, Binkley JM, Riddle DL, Guyatt GH. Sensitivity to change of the Roland-Morris Back Pain Questionnaire: part 1. Phys Ther 1998;78:1186-96.
- Sullivan MJ, Rodgers WM, Kirsch I. Catastrophizing, depression and expectancies for pain and emotional distress. Pain 2001;91:147-54.
- Tugwell P, Bombardier C, Buchanan WW, Goldsmith CH, Grace E, Hanna B. The MACTAR Patient Preference Disability Questionnaire--an individualized functional priority approach for assessing improvement in physical disability in clinical trials in rheumatoid arthritis. J Rheumatol 1987;14:446-51.
- Turk DC, Okifuji A. Psychological factors in chronic pain: evolution and revolution. J Consult Clin Psychol 2002;70:678-90.
- Turner JA, Clancy S. Comparison of operant behavioral and cognitive-behavioral group treatment for chronic low back pain. J Consult Clin Psychol 1988;56:261-6.
- van den Hout JH, Vlaeyen JW, Heuts PH, Zijlema JH, Wijnen JA. Secondary prevention of work-related disability in nonspecific low back pain: does problem-solving therapy help? A randomized clinical trial. Clin J Pain 2003;19:87-96.
- Vlaeyen JWS, Geurts SM, Kole-Snijders AMJ, Schuerman JA, Groenman NH, van Eek H. What do chronic pain patients think of their pain? Towards a pain cognition questionnaire. Br J Clin Psychol 1990;29:383-94.
- Vlaeyen JWS, Haazen IW, Schuerman JA, Kole-Snijders AM, van Eek H. Behavioural rehabilitation of chronic low back pain: comparison of an operant treatment, an operant-cognitive treatment and an operant-respondent treatment. Br J Clin Psychol 1995a;34:95-118.
- Vlaeyen JWS, Kole-Snijders AM, Rotteveel AM, Ruesink R, Heuts PHTG. The role of fear of movement/(re)injury in pain disability. J Occup Rehabil 1995b;5:235-52.
- Vlaeyen JWS, Morley S. Cognitive-behavioral treatments for chronic pain: what works for whom? Clin J Pain 2005;21:1-8.
- Vlaeyen JWS, van Breukelen G, Nooyen-Haazen IWCJ, Stomp-van den Berg SGM, Kole-Snijders AMJ. De Pijn Cognitie Lijst. Maastricht: Pijn Kennis Centrum academisch ziekenhuis Maastricht, 2003.
- Wallston KA, Wallston BS, DeVellis R. Development of the Multidimensional Health Locus of Control (MHLC) Scales. Health Educ Monogr 1978;6:160-70.
- Woby SR, Watson PJ, Roach NK, Urmston M. Are changes in fear-avoidance beliefs, catastrophizing, and appraisals of control, predictive of changes in chronic low back pain and disability? Eur J Pain 2004;8:201-10.
- Zitman FG, Griez EJL, Hooijer C. Standaardisering depressievragenlijsten. Tijdschr Psychiatr 1989;31:114-35.

Chapter 7

Treatment expectancy and credibility are associated with the outcome of both physical and cognitive-behavioral treatment in chronic low back pain.

Rob JEM Smeets, Saskia Beelen, Mariëlle E.J.B. Goossens, Erik G.W. Schouten, J André Knottnerus, Johan WS Vlaeyen

Submitted.

Abstract

Patients' initial beliefs about the success of a given pain treatment are shown to affect final treatment outcome. The Credibility/Expectancy Questionnaire (CEQ) has recently been developed as measure of treatment credibility and expectancy. The objectives of this study were, 1) to investigate the factor structure of the CEQ in a sample of chronic low back pain (CLBP) patients by means of a confirmatory factor analysis (CFA), 2) to examine the association between treatment credibility and expectancy and patient characteristics, and 3) to assess whether treatment expectancy and credibility are associated with the outcome of rehabilitation treatment. CLBP patients (n = 167) were randomized to either Active Physical Therapy (APT; n = 51), Cognitive-Behavioral Therapy (CBT: n = 57), or a Combination Therapy (CT: n = 59), and completed the CEQ after a careful explanation of the treatment rationale. CFA supported the two-factor structure (credibility/expectancy) of the CEQ. Lower credibility was associated with higher pain-related fear and lower internal control of pain, and lower expectancy with higher levels of pain-related fear and no radiating pain. Multiple linear regression analyses revealed that after controlling for age, gender, treatment centre, pain-intensity at baseline, duration of disability, and irrespective of the treatment offered, expectancy was significantly associated with disability and satisfaction. Credibility was significantly associated with patient-specific complaints and satisfaction. For global perceived effect, treatment expectancy was predictive in APT only, and treatment credibility was a significant predictor in CT only.
Although the associations found were low to modest, these results underscore the importance of expectancy and credibility for the outcome of different active interventions for CLBP and might contribute to the development of more effective treatments.

Introduction

Active treatments in chronic low back pain (CLBP) such as exercise therapy, cognitive-behavioral therapy and multidisciplinary therapy are being used to reduce the burden of pain disability patients, and society as a whole, have to bear. Although there is substantial evidence supporting the effectiveness of these treatments, there remains room for improvement (Guzman et al. 2002; Hayden et al. 2005; Morley et al. 1999; Ostelo et al. 2005a; Smeets et al. 2006b). One way to approach this problem is to identify modifiable predictors of treatment outcome. Treatment credibility and expectancy have been postulated to be such modifiable variables (Vlaeyen and Morley 2005). Treatment credibility refers to how believable, convincing and logical the treatment seems to the patient, whereas treatment expectancy refers to improvements the patient believes will be achieved for him/her personally. Recently, Goossens et al. (2005) showed that treatment expectancy, assessed with the questionnaire developed by Borkovec and Nau (1972) and measured before the start of cognitive-behavioral treatment of chronic pain sufferers, significantly predicted outcome. Similar issues were addressed in studies regarding acute, subacute and chronic back pain patients (Kalauokalani et al. 2001; Kole-Snijders et al. 1999; Ostelo et al. 2005b; Skargren and Oberg 1998). However, psychometric data of the expectancy questionnaires used in these previous studies are lacking.
More recently, Devilly and Borkovec (2000) developed the Credibility/Expectancy Questionnaire (CEQ), a modification of the original Borkovec and Nau measure which measures both treatment credibility and expectancy. The CEQ showed high internal consistency and good test-retest reliability. Although other adaptations of the Borkovec and Nau questionnaire to measure treatment expectancy as well as credibility have been used in studies of cognitive-behavioral treatments for rheumatoid arthritis and osteoarthritis (Bradley et al. 1987; Keefe et al. 1990), these measures as well as the CEQ have never been used in CLBP. Several researchers have suggested that treatment expectancy is likely to be a stronger predictor of treatment outcome than treatment credibility (Borkovec and Costello 1993; Devilly and Borkovec 2000; Goossens et al. 2005). However, this assumption is solely based on the study of Devilly and Borkovec (2000) in patients with anxiety and post-traumatic stress disorders being treated with cognitive-behavioral treatment. Treatment expectancy appeared to be related to more outcome measures than treatment credibility, but a sound theoretical framework for this hypothesis is not presented. Devilly and Borkovec suspected that the credibility scale taps more into the more logical thought processes of clients, whereas the expectancy scale ("how much do you expect to improve?") is functionally more related to affective processes similar to those involved in hope or faith.

Furthermore, many studies only assessed treatment expectancy immediately before the start of treatment meaning that patients might have difficulties in judging what the resulting outcome might be. As far as we know, no study has been carried out in which treatment expectancy and credibility were assessed simultaneously and immediately after the treatment rationale was explained to patients with CLBP entering rehabilitation.

In case treatment credibility and expectancy appear to be associated with treatment outcome, ways of strengthening treatment credibility and expectancy might also increase the effectiveness of rehabilitative interventions. Several modifiable psychological variables such as low level of internal control of pain and high levels of pain catastrophizing and pain-related fear have been suggested to be related to a lower level of treatment expectancy (Goossens et al. 2005; Metcalfe and Klaber Moffett 2005; Turk and Rudy 1990). However, similar research on treatment credibility is lacking.

Since we adapted the CEQ for use in CLBP treatment and translated it into Dutch, the first aim of this study is to test the factor structure of the CEQ, using confirmatory factor analysis (CFA). The second aim is to explore which demographic, sickness and psychological variables are associated with treatment expectancy and credibility. Third, we want to examine whether treatment expectancy and credibility are associated with outcome in terms of self-reported disability, satisfaction and pain in patients with CLBP who received different types of rehabilitation treatment. Since relevant literature on musculoskeletal disorders is rare, we wanted to explore whether treatment expectancy, as in cognitive-behavioral interventions for problem areas other than pain, shows a higher association with treatment outcome than credibility.

Materials and methods

This study is part of a larger RCT [ISRCTN22714229] of which the methods and first results are extensively described elsewhere (Smeets et al. 2006a). The main aim of this RCT was to test the hypothesis that an active physical therapy (APT), cognitive-behavioral therapy (CBT) and a combined therapy (CT) are more effective in reducing disability and several other outcome measures in nonspecific CLBP patients, than a waiting list (WL). Furthermore, it was hypothesized that the combination of both single treatment elements is more effective than the single treatments alone. A description of participants, interventions, measures, and study design is presented with special emphasis on treatment credibility and expectancy.

Participants

Patients for the first time referred by general practitioners or medical specialists to three outpatient rehabilitation centres in The Netherlands were invited by their consulting rehabilitation physician to participate in the trial. They were told that the trial was being performed to compare three currently used interventions for CLBP, of which the exact efficacy had not yet been established. Patients assigned to the WL were requested to wait 10 weeks after which they were offered a regular individual rehabilitation treatment. The main inclusion criteria for participation in the RCT were: nonspecific LBP of more than 3 months resulting in disability (Roland Disability Questionnaire score > 3) (Roland and Morris 1983), age between 18 to 65 years, and ability to walk at least 100 meters. Patients were requested to stop other treatments for their back complaints, except pain medication. Randomization per group of four patients at a time to one of the interventions took place in each of three participating rehabilitation centres. The medical ethics committee of

the Rehabilitation Foundation Limburg and Institute for Rehabilitation Research at Hoensbroek in The Netherlands approved the study protocol.
Of the 309 patients who were referred to the study, 86 (28%) did not participate because of not meeting the criteria (n = 35), clear preference for a particular treatment (n = 20), and logistic problems (n = 15). The remaining 223 patients were randomized and started treatment (APT; n = 53, CBT, n = 58, CT; n = 61, WL; n = 51). Of the 172 patients receiving an active intervention, 167 patients (APT, n = 51; CBT, n = 57; CT, n = 59) completed the CEQ. Three patients never showed up for treatment (in each intervention one) and of two patients (one in APT and one in CT) the questionnaire got lost. One of these patients fully completed the treatment, and the other one terminated treatment prematurely because of logistic problems. The patients who were randomized to the WL did not participate in the current study.
Nine patients did not complete post-treatment questionnaires (APT, n = 1; CBT, n = 3; CT, n = 5). There were no statistically significant differences regarding baseline variables, baseline outcome and credibility/expectancy scores between the responders and non-responders.

Interventions

All interventions started with a group meeting with a maximum of four patients during which the rationale of the allocated therapy was carefully explained. Emphasis was put on the responsibility of the patient for making plans for being active after the treatment. Each intervention lasted 10 weeks. A written summary of the rationale was given to the patients.

Active Physical Therapy (APT)
In a group of maximum four, patients were invited to perform 30 minutes of aerobic training on a bicycle (65% to 80% heart rate maximum) and 75 minutes of strength and endurance training of their lower back and upper leg muscles (3 series of 15 to 18 repetitions in a dynamic-static manner), three times a week. Two physiotherapists supervised the training.

Cognitive-Behavioral Therapy (CBT)
CBT consisted of operant behavioral graded activity training (Fordyce 1976; Sanders 1996) and problem solving training (Nezu 1986; Nezu and Perri 1989; van den Hout et al. 2003). The graded activity started with three group sessions, followed by a maximum of 17 individual sessions of 30 minutes. During the graded activity, the therapist focused on a time contingent gradual increase or pacing of three activities that were considered important and relevant for the patient's personal situation. The problem solving training focused on the skills of positive problem orientation, problem definition and formulation, generation of alternatives, decision-making, implementation and evaluation. Patients received a course book with a summary of each session and homework assignments. A clinical psychologist or social worker, provided 10 sessions of 1½ hours to a maximum of four patients at a time.

Combined Therapy (CT)
CT consisted of APT in combination with the problem solving training, both offered in the same frequency and duration as described before. The patient was told that he/she first had to gain sufficient aerobic fitness and strength before increasing his/her activities. The graded activity was not started until the third week with a total of 19 sessions.

Assessment
Assessments were carried out in the week before the start of treatment and immediately after 10 weeks of active treatment by research assistants who were not aware of the patient's assigned intervention.

Baseline data
During pre-treatment assessment, data were collected on age, gender, level of education, employment status, duration of complaints and disability, previous low back surgery, level of radiating pain to leg, traumatic onset of low back pain, depression (Beck Depression Inventory; BDI) (Beck et al. 1979), pain-related fear (Tampa Scale for Kinesiophobia; TSK) (Vlaeyen et al. 1995), pain catastrophizing and internal control of pain (subscales of the Pain Cognition List; PCL) (Vlaeyen et al. 1990; Vlaeyen et al. 2003).

Treatment credibility and expectancy
The CEQ developed by Devilly and Borkovec (2000), was translated into Dutch and adapted for the use in CLBP treatment. The original questionnaire showed high internal consistency and good test-retest reliability. CEQ predicted treatment outcome as expected, and discriminative validity for rationales possessing unequal theoretical basis was promising. Because any treatment credibility/expectancy scale should minimally include the main purpose of the intervention, the patient is asked to judge whether he/she thinks and/or feels that the treatment will reduce his/her low back pain associated functional limitations.
The questionnaire consists of six items (table 2). Introductory instructions tell the patient that beliefs about how well the therapy might help contains both thoughts and feelings about the therapy and that these may be the same or different.
In the original CEQ of Devilly and Borkovec, the first three items belong to the credibility factor (three 'think' questions) and the other three to the expectancy factor (one 'think' and two 'feel' questions). The CEQ utilizes two rating scales, one from 1 (not at all) to 9 (very much) and another from 0% (not at all) to 100% (very much), and therefore the percentage ratings have been subjected to linear transformation with a minimum of 1 and a maximum of 9, and a sum score has been formed for each factor ranging from 3 to 27.
In our study, participants completed the CEQ immediately after attending the first treatment session in which the treatment rationale was explained. To maintain confidentiality and to encourage honest responding, the patients were told that their responses would not be shared with the therapists. Therapists collecting the CEQ were unaware of the scoring method of the CEQ and were specifically instructed to refrain from inspecting the scoring sheets.

Outcome measures
Disability
The Roland Disability Questionnaire (RDQ) was used. This questionnaire has been shown to be a valid and reliable instrument in the evaluation of CLBP treatment (Beurskens et al. 1996; Gommans et al. 1997; Roland and Morris 1983; Stratford et al. 1994; Stratford et al. 1998). RDQ was administered twice, before and after treatment.

Patient-Specific Complaints (PSC)
The patient-specific main complaints were identified by using the patient-specific approach method developed by Tugwell et al. (1987; Beurskens et al. 1999). Patients indicate the three most limited functional activities and rate the difficulty in performing these activities during the previous week on a 100-mm visual analogue scale (VAS), with on the left side 'no difficulty' and on the right side 'impossible'. The average of the three VAS-scores is calculated. This method appeared to be valid, reliable and of sufficient responsiveness (Beurskens et al. 1996; Ostelo et al. 2004). PSC was administered twice, before and after treatment.

Pain
A 100-mm VAS with on the left side 'no pain' and on the right side 'unbearable pain' was used to measure the current pain intensity. Relevance, validity and reliability have been sufficiently tested for patients with LBP (Carlsson 1983; Revill et al. 1976; Sriwatanakul et al. 1983). Pain-VAS was administered twice, before and after treatment.

Global Perceived Effect (GPE)
Patient's global assessment of overall result regarding their disability was measured by a transitional seven-point Numerical Rating Scale (NRS) with 1 = 'vastly worsened' to 7 = 'completely recovered' (Beurskens et al. 1996; Ostelo et al. 2004). The GPE was measured post-treatment.

Satisfaction
Treatment satisfaction was measured by using a 100-mm VAS for the question "How satisfied are you with the total treatment?" with on the left side 'not satisfied at all' and on the right side 'completely satisfied'. Treatment satisfaction was measured post-treatment.

Statistical analysis
Pearson product moment correlations between treatment credibility and expectancy were calculated and the stability of the previously reported factor solution of the CEQ was tested by means of confirmatory factor analysis (CFA) using LISREL, version 8.30 (Jöreskog and Sörbom 1999). The goodness-of-fit for the factor solution was evaluated by calculating several indices: (a) the Comparative Fit Index (CFI); (b) the Goodness of Fit Index (GFI); (c) the Standardized Root Mean Square Residual (SRMR); and (d) the Root Mean Square Error of Approximation (RMSEA).

To evaluate the goodness-of-fit we used the combination rules recommended by Hu and Bentler (1999). For studies with a total number of patients of less than 250, a good fit should be concluded if both the CFI-value is ≥ 0.95 and the SRMR-value is ≤ 0.09.

To examine which baseline variables are associated with treatment credibility and expectancy, Pearson product moment correlations (r) for continuous and dichotomous variables and correlation ratio Eta's (η) for categorical variables and the credibility and expectancy scores, respectively, were calculated. These variables were selected based on the work of several other researchers (Goossens et al. 2005; Metcalfe and Klaber Moffett 2005; Turk and Rudy 1990), and consisted of age, gender, type of intervention, duration and level of disability, duration of complaints, pain intensity, level of radiating pain to leg, traumatic onset of pain, pain catastrophizing, internal control of pain, depression, pain-related fear, work status and level of education. Next, only those variables significantly correlating with the dependent variables (CEQ credibility and expectancy scores), were entered as independent variables in multiple linear regression analyses with a stepwise backward elimination. Non-significant variables (P > 0.05) were removed from the model one by one. At each step, collinearity check (Variance Inflation Factor (VIF) > 3) was done.

To evaluate the influential role of credibility and expectancy on outcome, stepwise forward multiple linear regression analyses were performed by using post-treatment scores of the outcome of interest as dependent variables.
Demographic variables (age, gender), the baseline score of the outcome measure (not available for GPE and satisfaction), type of intervention (two dummy variables), centre of treatment (two dummy variables), pain intensity at baseline and duration of disability (unequally distributed across intervention groups) were entered as independent variables in the first block.
Although no significant differences regarding outcome were found between CT and APT and CT and CBT, respectively (Smeets et al. 2006a), effect modification was assessed by adding interaction variables (intervention x credibility, and intervention x expectancy) to the model. In case an interaction variable was significant (P ≤ 0.05), analyses were performed for each intervention separately, otherwise (P > 0.05), the interaction variables were removed from the model.

Next, the expectancy score was entered as a second block. Subsidiary analyses explored the relative importance of credibility. In these analyses the credibility score was entered as a third block (exploratory analysis 1). The order of entry was then reversed by entering the credibility score as the second block and the expectancy score as the third block (exploratory analysis 2). This sequence enabled us to examine the relative importance of the expectancy and credibility measures: the value of the R^2-change for these two variables when entered as the second block indicates their relative magnitude as contributing factors.

All statistical analyses were carried out according to the intention-to-treat principle. All patients, including withdrawals from treatment and patients with poor compliance remained in the group they were randomized to.

Results

Baseline characteristics
Table 1 displays the baseline characteristics of the study sample as well as the (change) scores of the outcome measures for each intervention. There were no significant differences between the three interventions, except for the duration of disability. This variable was entered as a covariate in the regression analyses testing the influence of credibility/expectancy on treatment outcome.

Credibility, expectancy and outcome scores
Credibility scores ranged from 9 to 27 (mean score, 19.78 ± 3.61) and the expectancy scores ranged from 3.8 to 26 (mean score, 16.35 ± 4.41). Both scales approximated a normal distribution (kurtosis: -0.52 and -0.23, respectively; skewness: -0.10 and -0.34, respectively). The level of treatment expectancy and credibility did not significantly differ between the three interventions.

Confirmatory factor analysis of CEQ
Despite the fact that the two factors had a moderate to high level of correlation (r = 0.56, and r = 0.62 after correction for attenuation between the two factors), the goodness-of-fit of the 2-factor model revealed an acceptable fit to the data (table 2). However, the modification indices suggested that the goodness-of-fit could be improved by adding an error covariance between the recoded item 4 and recoded item 6. With this error covariance the fit turned out to be good (table 2). This finding is in line with the high Pearson correlation of 0.83 between these two items which is attributable to the fact that they differ only with respect to 'think' and 'feel'. Both questions load on the expectancy scale. In accordance with Devilly and Borkovec (2000), credibility was found to be derived from the first three 'think' questions and expectancy was derived from the fourth 'think' and the two, 'feel' questions. The internal consistency of the credibility and expectancy scale (Cronbach's alpha = 0.82 and 0.84, respectively) also appeared to be high.

Pre-treatment associations with credibility and expectancy
Although correlations were low, several patient characteristics were significantly associated with treatment credibility: female patients (r = 0.16; P < 0.05), patients who reported less pain-related fear (TSK: r = -0.33; P < 0.01), who were less catastrophizing (PCL-catastrophizing: r = - 0.16; P < 0.05), and who experienced more internal control of pain (PCL-internal control: r = 0.26; P < 0.01) reported higher level of treatment credibility. Furthermore, work status appeared to be significantly correlated with treatment credibility (η = 0.229; P < 0.05); those having partial sick leave reported the highest credibility score (mean score, 20.92 ± 3.03), followed by those working full time (mean score, 20.17 ± 3.53), and those without a job (mean score, 19.07 ± 4.41). The patients on full time sick leave or disability pension scored lowest (mean score, 18.89 ± 3.63). The multiple linear regression analysis with stepwise backward

elimination revealed that pain-related fear (Standardized β = - 0.267; P = 0.001) and internal control of pain (Standardized β = 0.165; P = 0.036), significantly explained 11.9% of the variance in treatment credibility. The other variables failed to remain in the final model.
For treatment expectancy even fewer variables were significantly associated: patients with a lower level of pain-related fear (TSK: r = - 0.30; P < 0.01), who reported more internal control of pain (PCL-internal control: r = 0.23; P < 0.01), and who experienced less depression (BDI: r = - 0.18; P < 0.05), reported higher treatment expectancy. Those who had no radiating pain to the leg had lower expectancy scores (mean score, 14.10 ± 4.25) than those with radiation below knee (mean score, 16.62 ± 4.44) or above knee (mean score, 16.83 ± 4.23; η = 0.206; P < 0.05). The multiple linear regression analysis revealed that pain-related fear (Standardized β = - 0.302; P < 0.001) and radiating pain to the leg (no radiation versus radiation above knee: Standardized β = - 0.304; P = 0.007; no radiation versus radiation below knee: Standardized β = - 0.258; P = 0.022) significantly explained 11.7% of the variance in treatment expectancy.
All variation inflation factors in both multiple regression analyses were within acceptable range (VIFs < 3), suggesting no problem with collinearity.

The association between treatment credibility/expectancy and outcome

A significant interaction between treatment expectancy and type of intervention was present for global perceived effect (P = 0.021), but not for the other outcome measures. Therefore, the stepwise forward multiple linear regression analyses for global perceived effect were performed for each intervention separately, and for the other outcome measures for all interventions together.

Table 1: Baseline variables and (change) scores of outcome measures (mean ± SD) for the total sample and and the three interventions (total n=167)

Variables	Total (n=167)	APT (n=51)	CBT (n=57)	CT (n=59)
Age (yr)	41.78 ± 9.72	42.37 ± 9.10	42.56 ± 9.75	40.53 ± 10.22
Gender (% male)	54.5	58.8	42.1	62.7
Education (%)				
low	61.7	67.7	61.4	57.6
middle - high	38.3	33.3	38.6	42.4
Work status (%)				
full time	32.3	33.3	40.3	23.7
partial sick leave / disability pension	22.8	17.7	19.3	30.5
full sick leave / disability pension	36.5	39.2	31.6	39.0
no job / retired	8.4	9.8	8.8	6.8
Duration of LBP (mo)	58.41 ± 68.24	58.41 ± 76.93	65.74 ± 72.17	51.32 ± 55.60
Duration of disability (mo)	37.13 ± 46.94	29.35 ± 38.02*	49.77 ± 61.74*	31.64 ± 33.73*
Radiation of pain (%)				
no radiation	13.8	9.8	17.6	13.5
above knee	35.9	37.3	26.3	44.1
below knee	50.3	52.9	56.1	42.4
Trauma preceding LBP (%)	15.6	15.7	19.3	11.9
BDI	10.19 ± 7.10	10.39 ± 7.77	10.26 ± 6.98	9.93 ± 6.71
TSK	39.16 ± 6.85	38.96 ± 6.58	38.74 ± 6.94	39.73 ± 7.06
PCL-catastrophizing	39.97 ± 12.68	40.25 ± 14.20	40.11 ± 11.78	39.59 ± 12.34
PCL-internal control	16.26 ± 3.65	16.73 ± 3.60	15.65 ± 4.00	16.44 ± 3.30
RDQ	13.83 ± 3.69	14.22 ± 3.70	13.65 ± 3.61	13.66 ± 3.79
PSC	74.40 ± 15.64	75.46 ± 13.89	74.56 ± 16.30	73.33 ± 16.59
Current pain	48.66 ± 24.32	51.14 ± 26.01	48.49 ± 23.56	46.69 ± 23.75
Treatment credibility	19.78 ± 3.61	20.28 ± 3.27	19.19 ± 3.65	19.92 ± 3.82
Treatment expectancy	16.35 ± 4.41	16.88 ± 4.03	15.36 ± 4.43	16.84 ± 4.60
RDQ-change	2.51 ± 4.42	2.40 ± 4.54	2.70 ± 4.69	2.41 ± 4.11
Pain intensity-change	7.47 ± 22.69	6.58 ± 22.73	11.50 ± 23.01	4.28 ± 22.15
PSC-change	16.15 ± 21.13	12.29 ± 22.34	17.40 ± 22.17	18.42 ± 18.72
GPE	4.49 ± 1.35	4.24 ± 1.48	4.65 ± 1.26	4.57 ± 1.30
Satisfaction	68.14 ± 25.53	64.92 ± 32.38	69.43 ± 22.85	69.89 ± 20.51

Values presented as means and standard deviation or percentage. * Significant differences between interventions (P < 0.05).
APT = active physical therapy; CBT = cognitive-behavioral therapy; CT = combination therapy; low back pain; BDI = Beck Depression Inventory; TSK = Tampa Scale for Kinesiophobia; PCL = Pain Cognition List; RDQ = Roland Disability Questionnaire; PCS = Patient-Specific Complaints, GPE = General Perceived Effect.

Table 2: Factor loadings and goodness-of-fit indices for two factor solutions as obtained by means of confirmatory factor analysis of the CEQ without and with the error covariance between item 4R and 6R (n=167)

	Without error covariance between item 4R and 6R		With error covariance between item 4R and 6R	
Item	Credibility	Expectancy	Credibility	Expectancy
1. How logical does the therapy offered to you seem	0.68		0.70	
2. How successfully do you think this treatment will be in reducing your limitations due to back pain	0.94		0.92	
3. How confident would you be in recommending this treatment to a friend with the same problems	0.71		0.73	
4R. By the end of the therapy period, how much improvement in your limiatiations due to back pain do you think will occur (recoded from 0% to 100% to 1 to 9)		0.90		0.69
5. At this point, how much do you really feel that the therapy will help to reduce your limitations due to back pain		0.64		0.81
6R. By the end of the therapy period, how much improvement in your limiatiations due to back pain do you feel will occur (recoded from 0% to 100% to 1 to 9)		0.92		0.72
Standardized error covariance between item 4R and 6R			0.33	
Correlation, corrected for attenuation between the factors	0.62		0.76	
Fit indices				
GFI	0.93		0.96	
CFI	0.94		0.98	
SRMR	0.08		0.03	
RMSEA	0.15		0.11	

CEQ = Credibility/Expectancy Questionnaire; GFI = Goodness of Fit Index; CFI = Comparative Fit Index; SRMR = Standardized Root Mean Square Residual; RMSEA = Root Mean Square Error of Approximation.

Table 3 shows the results for self-reported disability (RDQ score) at post-treatment as the dependent variable and the expectancy score being entered as the second block. In the first column the standardized Beta's for the final model with all three blocks entered are reported. The remaining columns report the adjusted R^2 and R^2-change and the significance of F-change for each successively entered block. The expectancy score was significantly associated with the RDQ score at post-treatment, explaining an additional 2.3% of the total variance (P = 0.02; one-point increase in expectancy score was associated with 0.2 decrease in RDQ score) after age, gender, centre of treatment, type of intervention, pain intensity at baseline, duration of disability and baseline score of RDQ had already been accounted for. In the first subsidiary explorative analysis, the introduction of the credibility score (block 3) did not add explanatory power after the effect of expectancy had been accounted for. While reversing the order of entry (second explorative analysis), the introduction of the credibility score (block 2) did not contribute significantly to the RDQ score at post-treatment (P = 0.059). Introduction of expectancy (block 3) did not significantly change the explanatory power of the model. So only treatment expectancy was significantly associated with the post-treatment RDQ score when entered as the second block, but when credibility was controlled for, expectancy did not additionally contribute significantly to the explanatory power. Furthermore, the value of the R^2-change for these two variables when entered as the second block showed that expectancy explained more of the total variance than credibility (R^2-change of 0.023 and 0.015, respectively), although this difference was rather small.

For patient-specific complaints (PSC), almost completely reversed results were obtained (table 4). Treatment credibility score was significantly associated with the PSC score at post-treatment, explaining an additional 2.3% of the total variance (P = 0.035; one-point increase in credibility score was associated with 1.03 decrease in PSC score). Based on the magnitude of the R^2-change (0.023 compared to 0.019, respectively), credibility explained only slightly more of the total variance than expectancy.

Both expectancy and credibility were significantly associated with treatment satisfaction. Expectancy, even being entered as the third block, significantly added further explanatory variance to the regression equation (table 5). This, beside the larger R^2-change (0.107 compared to 0.058), indicates that expectancy was more strongly associated with treatment satisfaction than credibility. One-point increase of expectancy and credibility score was associated with 1.98 and 1.83 increase of treatment satisfaction, respectively.

Neither expectancy (Standardized β = - 0.103, R^2-change = 0.009, P = 0.146) nor credibility (standardized β = 0.009, R^2-change = 0.002, P = 0.452) added explanatory power to the multiple regression equation for pain intensity. These results held irrespective of the measures being entered as the second or third block in the analysis.

In the analysis for global perceived effect, in APT (table 6a), expectancy (P = 0.002) and not credibility (P = 0.589) was significantly associated, explaining an additional 17.6% of the total variance. One-point increase in expectancy was associated with 0.16 increase of global perceived effect (scoring range from 1 to 7). Expectancy even showed an additional predictive power when entered as the third block, so after credibility had been adjusted for. Furthermore, the R^2-change of 0.176 for expectancy was clearly higher than the 0.058 for credibility.
For CBT, no significant association of expectancy (Standardized β = - 0.097, R^2-change =0.022, P = 0.276), nor credibility (Standardized β = - 0.103, R^2-change =0.022, P = 0.273) was found.
In CT (table 6b), credibility was significantly associated and explained an additional 8.7% of the total variance. One-point increase in credibility was associated with 0.12 increase of global perceived effect. The R^2-change of 0.087 for credibility was higher than the 0.041 for expectancy.

Table 3: Summary of multiple regression for RDQ and expectancy and credibility (n=158)

Model steps	Variables entered	Standardized Beta	Adj. R^2	R^2-change	F-change	Sig. F-change
Block 1						
Demographics	baseline score RDQ	0.561	0.335	0.376	9.904	< 0.001
	age	0.068				
	gender	0.003				
Centre of treatment	centre Tilburg	0.039				
	centre Breda	-0.006				
Intervention	APT	0.001				
	CBT	-0.067				
Disease specifics	duration disability	0.067				
	pain intensity	0.065				
Block 2	expectancy	-0.125	0.358	0.023	5.551	0.020
Exploratory analyses						
1 Block 3	credibility	-0.058	0.355	0.002	0.448	0.486
2 Block 2	credibility	-0.058	0.350	0.015	3.715	0.056
Block 3	expectancy	-0.125	0.356	0.009	2.135	0.146

The reported standardized Beta's are for the final model (3 blocks together). The remaining colums report the adjusted R^2 and R^2-change for successive blocks and the significance of the predictive value of the model. In the exploratory analyses 2, credibility was entered in block 2 and expectancy in block 3.
Centre of treatment: 2 dummies, Blixembosch is the reference centre. Intervention: 2 dummies, CT is reference intervention. APT = active physical therapy; CBT = cognitive-behavioral therapy; CT = combination therapy. RDQ = Roland Disability Questionnaire.

Table 4: Summary of multiple regression for PSC and expectancy and credibility (n=157)

Model steps	Variables entered	Standardized Beta	Adj. R^2	R^2-change	F-change	Sig. F-change
Block 1						
Demographics	baseline score PSC	0.406	0.200	0.247	5.346	< 0.001
	age	0.121				
	gender	0.039				
Centre of treatment	centre Tilburg	0.013				
	centre Breda	0.035				
Intervention	APT	0.138				
	CBT	0.029				
Disease specifics	duration disability	-0.072				
	pain intensity	-0.090				
Block 2	expectancy	-0.081	0.215	0.019	3.831	0.053
Exploratory analyses						
1 Block 3	credibility	-0.113	0.218	0.007	1.467	0.228
2 Block 2	credibility	-0.113	0.219	0.023	4.508	0.035
Block 3	expectancy	-0.081	0.218	0.004	0.791	0.375

The reported standardized Beta's are for the final model (3 blocks together). The remaining colums report the adjusted R^2 and R^2-change for successive blocks and the significance of the predictive value of the model. In the exploratory analyses 2, credibility was entered in block 2 and expectancy in block 3.
Centre of treatment: 2 dummies, Blixembosch is the reference centre. Intervention: 2 dummies, CT is reference intervention. APT = active physical therapy; CBT = cognitive-behavioral therapy; CT = combination therapy. PSC = Patient-Specific Complaints.

Table 5: Summary of multiple regression for satisfaction and expectancy and credibility (n=153)

Model steps	Variables entered	Standardized Beta	Adj. R^2	R^2-change	F-change	Sig. F-change
Block 1						
Demographics	age	0.133	0.004	0.056	1.071	0.387
	gender	0.066				
Centre of treatment	centre Tilburg	-0.151				
	centre Breda	0.070				
Intervention	APT	-0.093				
	CBT	0.045				
Disease specifics	duration disability	0.006				
	pain intensity	-0.097				
Block 2	expectancy	0.295	0.111	0.107	18.342	< 0.001
Exploratory analyses						
1 Block 3	credibility	0.083	0.109	0.004	0.711	0.401
2 Block 2	credibility	0.083	0.058	0.058	9.288	0.003
Block 3	expectancy	0.295	0.109	0.054	9.244	0.003

The reported standardized Beta's are for the final model (3 blocks together). The remaining colums report the adjusted $R^2$and R^2-change for successive blocks and the significance of the predictive value of the model. In the exploratory analyses 2, credibility was entered in block 2 and expectancy in block 3. Centre of treatment: 2 dummies, Blixembosch is the reference centre. Intervention: 2 dummies, CT is reference intervention. APT = active physical therapy; CBT = cognitive-behavioral therapy; CT = combination therapy.

Table 6a: Summary of multiple regression for GPE and expectancy and credibility in APT (n=50)

Model steps	Variables entered	Standardized Beta	Adj. R^2	R^2-change	F-change	Sig. F-change
Block 1						
Demographics	age	-0.042	0.027	0.146	1.223	0.314
	gender	-0.117				
Centre of treatment	centre Tilburg	-0.165				
	centre Breda	-0.078				
Disease specifics	duration disability	0.076				
	pain intensity	-0.283				
Block 2	expectancy	0.531	0.209	0.176	10.897	0.002
Exploratory analyses						
1 Block 3	credibility	-0.210	0.221	0.026	1.642	0.207
2 Block 2	credibility	-0.210	0.010	0.006	0.297	0.589
Block 3	expectancy	0.531	0.221	0.196	12.330	0.001

The reported standardized Beta's are for the final model (3 blocks together). The remaining colums report the adjusted $R^2$and R^2-change for successive blocks and the significance of the predictive value of the model. In the exploratory analyses 2, credibility was entered in block 2 and expectancy in block 3. Centre of treatment: 2 dummies, Blixembosch is the reference centre. Intervention: 2 dummies, CT is reference intervention. APT = active physical therapy; CBT = cognitive-behavioral therapy; CT = combination therapy; GPE = Global Perceived Effect.

Table 6b: Summary of multiple regression for GPE and expectancy and credibility in CT (n=54)

Model steps	Variables entered	Standardized Beta	Adj. R^2	R^2-change	F-change	Sig. F-change
Block 1						
Demographics	age	-0.012	-0.060	0.060	0.496	0.808
	gender	-0.185				
Centre of treatment	centre Tilburg	-0.119				
	centre Breda	-0.062				
Disease specifics	duration disability	0.055				
	pain intensity	-0.161				
Block 2	expectancy	0.099	-0.036	0.041	2.117	0.152
Exploratory analyses						
1 Block 3	credibility	0.282	0.002	0.052	2.764	0.103
2 Block 2	credibility	0.282	0.017	0.087	4.684	0.036
Block 3	expectancy	0.099	0.002	0.006	0.344	0.560

The reported standardized Beta's are for the final model (3 blocks together). The remaining colums report the adjusted R^2 and R^2-change for successive blocks and the significance of the predictive value of the model. In the exploratory analyses 2, credibility was entered in block 2 and expectancy in block 3. Centre of treatment: 2 dummies, Blixembosch is the reference centre. Intervention: 2 dummies, CT is reference intervention. APT = active physical therapy; CBT = cognitive-behavioral therapy; CT = combination therapy. GPE = Global Perceived Effect.

Discussion

The 2-factor structure of credibility and expectancy of the Dutch version of the CEQ which was adapted for CLBP rehabilitation could be confirmed. Both scales also showed a high internal consistency. Although no normative values are available yet, the mean credibility score seemed to be sufficiently high (73% of maximum score), and the expectancy score only moderate (60.5% of maximum score). One possible explanation might be that these patients had already undergone at least some kind of unsuccessful treatment and therefore were more reserved regarding the interventions offered, especially regarding the effectiveness of the intervention for themselves. Otherwise, 20 patients stated that they had a clear preference for a particular intervention and therefore refused to participate in the trial. As a consequence some kind of positive selection might have been present, causing higher mean scores on the CEQ. This potential overestimation of the treatment credibility and expectancy might have caused restriction of range resulting in weaker associations.
Higher pain-related fear was associated with both lower treatment expectancy and credibility, and higher internal control with higher treatment credibility. The hypothesis that having radiating pain to the leg which is normally interpreted as having a more difficult to cure problem, and therefore associated with less expectancy (Goossens et al. 2005; Metcalfe and Klaber Moffett 2005; Turk and Rudy 1990), was not supported. On the contrary, an inverse association was found. The level of pain and disability, two other expressions of the severity of the low back pain problem were unexpectedly not associated with the level of expectancy. We have no satisfying explanation for these results and future research should further elucidate this issue. Our results regarding the associated variables for expectancy are in line with the results of Goossens et al. (2005), who identified pain-related fear as the single significant contributor accounting for 6.5% of the variance of pre-treatment expectancy.

The results indicate that both expectancy and credibility are associated with outcome in rehabilitation treatment for CLBP. Expectancy accounted for a significant explained variance in disability (2.3%) and satisfaction (10.7%). For patients attending APT, expectancy explained 17.6% of the total variance in global perceived effect. Credibility accounted for a significant explained variance in patient-specific complaints (2.3%) and satisfaction (5.8%), and for patients attending CT, also for global perceived effect (8.7%). These percentages of explained variance are low to moderate with the exception for treatment expectancy in global perceived effect in APT. Future research may highlight which percentage of variance predicted by credibility/expectancy is also clinically relevant, however in our opinion every further reduction of disability, even a small one is worth the effort. Furthermore, percentages of 8.7% and higher, are in our opinion certainly not trivial, especially because this level of variance was calculated after several other variables (age, gender, pain intensity, duration of disability, treatment centre and type of intervention) were accounted for.

Both expectancy and credibility did not predict pain intensity. This is probably attributable to the fact that in the CEQ, patients were specifically asked to judge whether they think and/or feel that the treatment will reduce functional limitations due to back pain and not pain intensity.

The differential influence of expectancy on global perceived effect is interesting, but cannot be explained by earlier research, so the following explanations are only speculative. Since most patients seemed to understand what aerobic training and strengthening exercises are like, patients might have been better able to judge the potential effectiveness of the physical treatment for them personally. This in contrast to the cognitive-behavioral training elements, with which they were unfamiliar. The finding that expectancy was only differentially predictive for global perceived effect, might be the result of item overlap or resemblance (GPE: "how much did your low back pain associated functional limitations change compared with the situation before treatment?"; CEQ :"how much they think/feel that treatment will reduce back pain associated functional limitations?").

Devilly and Borkovec (1993; 2000) speculated that although a patient still thinks that a new treatment is credible, this may differ from what the patient really feels is the case. They hypothesized that expectancy seems to be more affectively driven as opposed to credibility which is more cognitively based, and may therefore have a higher association with the final outcome of treatment, at least in psychotherapy (Devilly and Borkovec 2000). However, based on the difference of the magnitude of R^2-change for expectancy and credibility, we did not find clear indications for their hypothesis.

Notable is that the CEQ was completed immediately after the treatment rationale was explained, making the patient more able to judge treatment credibility and expectancy. The percentages of explained variance of expectancy for several outcome measures are similar to studies in which the patients rated expectancy after the explanation of the rationale of cognitive-behavioral treatment for CLBP and fibromyalgia (Goossens et al. 2005), post lumbar discectomy pain (Ostelo et al. 2005b), and social phobia (Safren et al. 1997). Our results are also in line with intervention studies for patients with mixed duration of neck and/or low back pain (Skargren and Oberg 1998), and CLBP (Kalauokalani et al. 2001) in which the treatment expectancy was rated before the treatment rationale had been explained.

Also some potential bias regarding this study can be identified. We did not collect the CEQ for five patients and the post-treatment questionnaires for nine patients. Since the total number of missing data was low and the reasons for the missing outcome data seldom were related to the type of intervention, we decided not to impute missing data. Nevertheless, we are quite confident that our results are sufficiently reliable and the total number of patients of whom the dataset was complete, guaranteed sufficient power to perform the described analyses.

Our results were obtained from patients who agreed to participate in a trial and therefore unconditional generalization to daily rehabilitation practice is not possible. Otherwise, the baseline characteristics (e.g. level of disability, disease characteristics and previous treatments) of the study population are similar to the population that is normally being treated in outpatient rehabilitation clinics (VRIN/VRA 2000).

Conclusions

Despite the relative high correlation between the two subscales of the CEQ, our results provide evidence for the existence, and differential, predictive validity of two separate constructs, namely treatment credibility and expectancy in CLBP treatment. Although the predictive power was low to modest, these results underscore the importance of treatment expectancy and credibility for the outcome of different active interventions for CLBP. Future research might find a better way to further discern both constructs, but until then, we recommend the use of the CEQ at least for research purposes. Furthermore, future studies might examine the effects of (pre)-treatments that focus more specifically on increasing treatment credibility and expectancy. The explanation of a clear treatment rationale at the start of treatment can be a valuable starting point, but motivational interviewing and counseling addressing internal control of pain and pain-related fear before the start of treatment might be important in order to increase the effectiveness of treatment in CLBP (Jensen 1996; Jensen et al. 2003; Kerns et al. 2006). We hope that this study will contribute to the use and further development of assessments, as well as interventions that can help a greater amount of patients and further increase the cost-effectiveness of treatment.

References

- Beck AT, Rush AJ, Shaw BF, Emmery G. Cognitive therapy of depression. New York: Wiley and sons, 1979.
- Beurskens AJ, de Vet HC, Koke AJ. Responsiveness of functional status in low back pain: a comparison of different instruments. Pain 1996;65:71-6.
- Beurskens AJ, de Vet HC, Koke AJ, Lindeman E, van der Heijden GJ, Regtop W, Knipschild PG. A patient-specific approach for measuring functional status in low back pain. J Manipulative Physiol Ther 1999;22:144-8.
- Borkovec TD, Costello E. Efficacy of applied relaxation and cognitive-behavioral therapy in the treatment of generalized anxiety disorder. J Consult Clin Psychol 1993;61:611-9.
- Borkovec TD, Nau SD. Credibility of analogue therapy rationales. J Behav Ther Exp Psychiatry 1972;3:257-60.
- Bradley LA, Young LD, Anderson KO, Turner RA, Agudelo CA, McDaniel LK, Pisko EJ, Semble EL, Morgan TM. Effects of psychological therapy on pain behavior of rheumatoid arthritis patients. Treatment outcome and six-month followup. Arthritis Rheum 1987;30:1105-14.
- Carlsson AM. Assessment of chronic pain. I. Aspects of the reliability and validity of the visual analog scale. Pain 1983;16:87-101.
- Devilly GJ, Borkovec TD. Psychometric properties of the credibility/expectancy questionnaire. J BehavTher Exp Psychiatry 2000;31:71-86.
- Fordyce WE. Behavioral methods for chronic pain and illness. St. Louis: Mosby, 1976.
- Gommans IHB, Koes BW, van Tulder MW. Validiteit en responsiviteit Nederlandstalige Roland Disability Questionnaire. Vragen lijst naar functionele status bij patiënten met lage rugpijn. Ned T Fysiother 1997;107:28-33.
- Goossens ME, Vlaeyen JW, Hidding A, Kole Snijders A, Evers SM. Treatment expectancy affects the outcome of cognitive-behavoral interventions in chronic pain. Clin J Pain 2005;21:18-26.
- Guzman J, Esmail R, Karjalainen K, Malmivaara A, Irvin E, Bombardier C. Multidisciplinary bio-psycho-social rehabilitation for chronic low back pain. Cochrane Database Syst Rev 2002:CD000963.
- Hayden JA, van Tulder MW, Malmivaara AV, Koes BW. Meta-analysis: exercise therapy for nonspecific low back pain. Ann Intern Med 2005;142:765-75.
- Hu Lt, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. Structural Equation Modeling 1999;6:1-55.
- Jensen MP. Enhancing motivation to change in pain treatment. In: RJ Gatchel, DC Turk. Psychological approaches to pain management. A practitioner's handbook New York, NY: The Guilford Press, 1996, pp. 78-111.
- Jensen MP, Nielson WR, Kerns RD. Toward the development of a motivational model of pain self-management. J Pain 2003;4:477-92.
- Jöreskog KG, Sörbom D. LISREL 8.30. Scientific International Inc., 1999.
- Kalauokalani D, Cherkin DC, Sherman KJ, Koepsell TD, Deyo RA. Lessons from a trial of acupuncture and massage for low back pain. Spine 2001;26:1418-24.
- Keefe FJ, Caldwell DS, Williams DA, Gil KM. Pain coping skills training in the management of osteoarthritic knee pain: A comparative study. Behavior Therapy 1990;21:49-62.
- Kerns RD, Jensen MP, Nielson WR. Motivational issues in pain self-management. In: H Flor, E Kalso, JO Dostrovsky. Proceedings of the 11th World Congress on Pain. Seattle: IASP Press, 2006, pp. 555-65.
- Kole-Snijders AM, Vlaeyen JW, Goossens ME, Rutten-van Molken MP, Heuts PH, van Breukelen G, van Eek H. Chronic low-back pain: what does cognitive coping skills training add to operant behavioral treatment? Results of a randomized clinical trial. J Consult Clin Psychol 1999;67:931-44.

References

- Metcalfe CJ, Klaber Moffett JA. Do patients' expectations of physiotherapy affect treatment outcome? Part 1: Baseline data. Int J Ther Rehabil 2005;12:55-62.
- Morley S, Eccleston C, Williams A. Systematic review and meta-analysis of randomized controlled trials of cognitive behaviour therapy and behaviour therapy for chronic pain in adults, excluding headache. Pain 1999;80:1-13.
- Nezu AM. Efficacy of a social problem-solving therapy approach for unipolar depression. J Consult Clin Psychol 1986;54:196-202.
- Nezu AM, Perri MG. Social problem-solving therapy for unipolar depression: an initial dismantling investigation. J Consult Clin Psychol 1989;57:408-13.
- Ostelo RW, de Vet HC, Knol DL, van den Brandt PA. 24-item Roland-Morris Disability Questionnaire was preferred out of six functional status questionnaires for post-lumbar disc surgery. J Clin Epidemiol 2004;57:268-76.
- Ostelo RW, Tulder M, Vlaeyen J, Linton S, Morley S, Assendelft W. Behavioural treatment for chronic low-back pain. Cochrane Database Syst Rev 2005a:Cd002014.
- Ostelo RW, Vlaeyen JW, van den Brandt PA, de Vet HC. Residual complaints following lumbar disc surgery: prognostic indicators of outcome. Pain 2005b;114:177-85.
- Revill SI, Robinson JO, Hogg MIJ. The reliability of a linear analog for evaluating pain. Anaesthesia 1976;31:1191-8.
- Roland M, Morris R. A study of the natural history of back pain. Part I: development of a reliable and sensitive measure of disability in low-back pain. Spine 1983;8:141-4.
- Safren SA, Heimberg RG, Juster HR. Clients' expectancies and their relationship to pretreatment symptomatology and outcome of cognitive-behavioral group treatment for social phobia. J Consult Clin Psychol 1997;65:694-8.
- Sanders SH. Operant conditioning with chronic pain: back to basics. In: RJ Gatchel, DC Turk. Psychological approaches to pain management. A practitioner's handbook. New York: The Guilford press, 1996, pp. 112-30.
- Skargren EI, Oberg BE. Predictive factors for 1-year outcome of low-back and neck pain in patients treated in primary care: comparison between the treatment strategies chiropractic and physiotherapy. Pain 1998;77:201-7.
- Smeets RJ, Vlaeyen JW, Hidding A, Kester AD, van der Heijden GJ, van Geel AC, Knottnerus JA. Active rehabilitation for chronic low back pain: Cognitive-behavioral, physical, or both? First direct post-treatment results from a randomized controlled trial [ISRCTN22714229]. BMC Musculoskelet Disord 2006a;7:5.
- Smeets RJ, Wade D, Hidding A, van Leeuwen PJ, Vlaeyen JW, Knottnerus JA. The association of physical deconditioning and chronic low back pain: A hypothesis-oriented systematic review. Disabil Rehabil 2006b;28:673-93.
- Sriwatanakul K, Kelvie W, Lasagna L, Calimlim JF, Weis OF, Mehta G. Studies with different types of visual analog scales for measurement of pain. Clin Pharmacol Ther 1983;34:234-9.
- Stratford PW, Binkley J, Solomon P, Gill C, Finch E. Assessing change over time in patients with low back pain. Phys Ther 1994;74:528-33.
- Stratford PW, Binkley JM, Riddle DL, Guyatt GH. Sensitivity to change of the Roland-Morris Back Pain Questionnaire: part 1. Phys Ther 1998;78:1186-96.
- Tugwell P, Bombardier C, Buchanan WW, Goldsmith CH, Grace E, Hanna B. The MACTAR Patient Preference Disability Questionnaire--an individualized functional priority approach for assessing improvement in physical disability in clinical trials in rheumatoid arthritis. J Rheumatol 1987;14:446-51.
- Turk DC, Rudy TE. Neglected factors in chronic pain treatment outcome studies--referral patterns, failure to enter treatment, and attrition. Pain 1990;43:7-25.

References

- van den Hout JH, Vlaeyen JW, Heuts PH, Zijlema JH, Wijnen JA. Secondary prevention of work-related disability in nonspecific low back pain: does problem-solving therapy help? A randomized clinical trial. Clin J Pain 2003;19:87-96.
- Vlaeyen JWS, Geurts SM, Kole-Snijders AMJ, Schuerman JA, Groenman NH, van Eek H. What do chronic pain patients think of their pain? Towards a pain cognition questionnaire. Br J Clin Psychol 1990;29:383-94.
- Vlaeyen JWS, Kole-Snijders AM, Boeren RG, van Eek H. Fear of movement/(re)injury in chronic low back pain and its relation to behavioral performance. Pain 1995;62:363-72.
- Vlaeyen JWS, Morley S. Cognitive-behavioral treatments for chronic pain: what works for whom? Clin J Pain 2005;21:1-8.
- Vlaeyen JWS, van Breukelen G, Nooyen-Haazen IWCJ, Stomp-van den Berg SGM, Kole-Snijders AMJ. De Pijn Cognitie Lijst. Maastricht: Pijn Kennis Centrum academisch ziekenhuis Maastricht, 2003.
- VRIN/VRA. Productdefiniëring Revalidatiezorg: registratiepilot revalidatie-instellingen januari 1999 - juli 2000. Utrecht: VRIN, 2000.

Chapter 8

The usability of six physical performance tasks in a rehabilitation population with chronic low back pain.

Rob J.E.M. Smeets, Helma J.M. Hijdra, Arnold D.M. Kester, Minou W.G.C. Hitters, J. André Knottnerus

Published as:
Smeets RJEM, Hijdra HJM, Kester ADM, Hitters MWGC, Knottnerus JA. The usability of six physical performance tasks in a rehabilitation population with chronic low back pain. Clinical Rehabilitation, 2006;20:987-996.

Abstract

Objective: To examine the influence of task experience on the difference between test and retest, and to assess test-retest reliability and limits of agreement of six performance tasks in chronic low back pain patients. These measures will be used to define the clinical usability.
Design: Test-retest of six performance tasks in a group of patients with no experience and a group of patients after previous experience with these tasks.
Setting: Three rehabilitation centres.
Subjects: 53 patients with nonspecific chronic low back pain.
Main Measures: Five-minute walking, 50-foot fast walking, sit to stand, loaded forward reach, one-minute stair climbing, Progressive Isoinertial Lifting Evaluation (PILE). To assess the influence of task experience, differences between test and retest between both groups were tested using Mann-Whitney test. For both groups together, intraclass correlation coefficients (ICCs) and the limits of agreement using Bland and Altman plots were calculated.
Results: 30 patients with no task experience and 23 patients already having undertaken the tasks on at least two occasions participated. Both groups showed similar differences between test and retest. The test-retest reliability for the total study population was good to very high: ICC varied from 0.74 to 0.99. For the total study population, the limits of agreement expressed as percentage of the mean score of each task was low to moderate for five-minute walking and one-minute stair climbing (21% and 20%, respectively), moderate for 50-foot fast walking, sit to stand and forward reach (33%, 29% and 36% respectively) and high for the PILE (48%).
Conclusions: Task experience did not significantly influence test-retest differences. All tasks showed sufficient test-retest reliability. Based on the natural variability of the tasks, the five-minute walking and stair climbing task, and to a lesser degree the 50-foot walking, sit to stand and loaded forward reach, seem clinically useful. There are major concerns about the usability of the PILE.

Clinical messages

- Differences between test and retest were not affected by previous task experience.
- All tasks showed good to excellent test-retest reliability.
- Based on the limits of agreement, five-minute walking and one-minute stair climbing are clinical useful, but PILE is not.
- Increase of pain before retesting did not influence test-retest results.

Introduction

The major aim of rehabilitation treatment of patients with chronic low back pain (CLBP) is to improve their level of activities and participation. In daily practice and scientific research, mostly self-reported questionnaires have been used to assess daily functioning. This yields information that may not necessarily reflect the real capability of patient's performance. In order to improve objectivity, measures of body function, such as mobility and muscle strength, have been used, although the correlation between these measures and the level of activity in daily life is very weak (Parks et al. 2003; Simmonds et al. 1998; Waddell 1998; Wittink 2005). Furthermore, there is substantial concern about the reliability and validity of these measures (Gibbons et al. 1997; Kankaanpaa et al. 1998; Keller et al. 2001; Newton et al. 1993; Simmonds et al. 1998). Many have urged the use of objective and direct measures of the functional capacity (Dworkin et al. 2005; Harding et al. 1994; WHO 2001; Simmonds et al. 1998; Watson 1999; Wittink 2005). Several performance tests have been developed such as obstacle courses and Functional Capacity Evaluation. However, psychometric data on obstacle courses are lacking (Mayer and Gatchel 1988; Simmonds et al. 1998). Recently published psychometric data of a Functional Capacity Evaluation indicated that some of the tasks are reliable, but others are not (Brouwer et al. 2003; Reneman et al. 2002). Unfortunately, the whole sequence of tasks in a Functional Capacity Evaluation is time consuming and expensive, as is the training of the test observer.

Other researchers have described tasks resembling daily activities that might be hampered by low back pain, but that are less time consuming, easier to administer and do not use sophisticated and expensive equipment (Harding et al. 1994; Ljungquist et al. 1999; Magnussen et al. 2004; Mayer et al. 1988a; Mayer et al. 1990; Simmonds et al. 1998; Strand et al. 2002). In 2001, at the start of a randomized controlled trial (RCT) in which three different treatment strategies for patients with nonspecific CLBP are compared, six such easy-to-administer and back pain-relevant performance tasks were selected (Smeets et al. 2006). Due to unfamiliarity with the tasks, patients who perform the tasks for the first time might perform worse than patients being tested after previous experience of these tasks. When these inexperienced patients are offered a retest they may show more improvement on the tasks than the more experienced ones, but such data are not available.
For some tasks test-retest reliability data are available, although these were assessed in patients with mixed musculoskeletal problems (Harding et al. 1994) or healthy subjects (Mayer et al. 1988a; Mayer et al. 1988b; Mayer et al. 1990). For other tasks the test-retest reliability was assessed in CLBP patients with a relative low level of disability (Simmonds et al. 1998). Since it was expected that the group of patients being referred to our trial would be moderately to severely disabled, there were concerns about the generalization. Furthermore, because the instructions for the patient and test observer for several tasks were unclear, we modified the protocols, making the generalization even more difficult.

The first aim of this study was to assess whether patients being tested for the first time showed more improvement between the test and retest for each task in comparison to patients being tested for the third or fourth time. Because of the high resemblance of the tasks with daily activities such as walking, stair climbing and lifting, we hypothesized that there would be no difference between the groups regarding the improvement between test and retest.
Next, for the total study population we wanted to assess the test-retest reliability and to define the limits of agreement for quantifying the natural variation for each performance task. Finally, it is known that the level of back pain in chronic patients might vary and even increase by performing tasks. We therefore wanted to examine whether the change in back pain after the first testing influenced the retest.

Methods

Study design

Six performance tasks were selected out of performance task batteries described in detail by others (Harding et al. 1994; Lee et al. 2001; Mayer et al. 1988a; Mayer et al. 1988b; Mayer et al. 1990; Novy et al. 2002; Novy et al. 1999; Simmonds et al. 1998). The task protocols were slightly adjusted with regard to the instructions given to the patients. The test observers were carefully instructed not to respond to the patients' behavior, to always stick to the testing protocol and to only use the verbal instruction as indicated in the protocol. This prevented solicitous behavior as far as possible. All test observers received special training and attended refresher courses two or three times a year to ensure reliable testing results. For all patients a retest was done five to nine days after the first test, at the same time of day and by the same test observer. In order to examine the influence of task experience we selected two groups of patients based on their experience with the performance tasks. One group (group A) consisted of patients who had no previous experience with the tasks and another group (group B) consisted of patients who had already undertaken the tasks on two or three occasions.
In order to assess the test-retest reliability and limits of agreement, both groups of patients (group A and B) were taken together.

Participants

The test-retest study included 53 patients (30 patients in group A and 23 patients in group B) with more than three months existing disability due to nonspecific low back pain (Roland Morris Disability Questionnaire score > 3) (Roland and Morris 1983), who also participated in our trial. They were referred by general practitioner or medical specialist for rehabilitation treatment in one of the three participating rehabilitation centres. Exclusion criteria were: lumbar spondylolisthesis, spondylodesis, fracture, severe psychopathology preventing group treatment, medical disorders preventing physical exercise, severe substance abuse, pregnancy and illiteracy. All patients participating in our trial were assessed using the performance tasks before treatment, immediately post-treatment and also at 6 and 12 months post-treatment. The details of our trial have been described elsewhere (Smeets et al. 2006). After the patients agreed to

participate in our trial, they were also invited to participate in the test-retest study and informed consent was obtained (group A). Patients who were invited for their regular 6 or 12 month follow-up assessment after treatment offered in our trial, were also invited to participate in the test-retest study and were sent additional written information. Once they appeared for their regular assessment they were asked to give their informed consent (group B). Patients who were assessed immediately post-treatment were not included in this test-retest study because we thought it was likely that their level of performance was not stable yet.
Baseline characteristics at the start of participation in this test-retest study of the whole sample and per group are presented in table 1. The Medical Ethics Committee of the Rehabilitation Foundation Limburg and the Institute for Rehabilitation Research, Hoensbroek, The Netherlands approved the study protocol.

Assessment of physical performance

Five-minute walking
Participants were asked to walk as fast and far as possible, no running, for a five-minute period. The use of walking aids was prohibited. The circuit was 30 meters long and eight-shaped. Participants were permitted to take a rest on a chair. The distance covered was recorded in meters.

Fifty-foot walking
Participants had to walk as fast as possible, no running, until they got back to the starting point. The use of walking aids was prohibited. The circuit was 50 feet long and eight-shaped. The time needed was recorded in seconds.

Sit to stand
Participants were asked to perform a sit to stand movement from a chair without arms five times as fast as possible. The task was performed twice and the average time in seconds needed to perform the task was calculated.

Loaded forward reach
Participants were instructed to hold a stick with a weight of 4.5 kg at shoulder height and width in both hands. Participants had to reach forward as far as possible without lifting the heels off the floor. The distance reached was recorded in centimeters.

One-minute stair climbing
Participants were asked to walk a stair up and down for one minute. The circuit was five stairs high and eight-shaped. The number of stairs climbed was recorded.

Progressive Isoinertial Lifting Evaluation (PILE)
Participants had to lift a box with a weight four times within 20 seconds from floor up to a 75 cm high table. The starting weight for women was 3.6 kg and 5.85 kg for men (weight of box included). After each completed lifting cycle the weight for women was increased by 2.25 kg and for men by 4.5 kg. Starting weights and incremental weights were different for men and women to compensate for the greater strength of men and thus to equalize the sensitivity of the test. The test stopped when the participant could not lift the box four times within 20 seconds, the participant decided to stop because of tiredness or pain, the heart rate exceeded 85% of the maximal heart rate, maximal amount of the weight that could safely be lifted was reached (60% of body weight), or the test observer considered the lifting unsafe. The amount of fully completed lifting stages was recorded instead of the total weight lifted to obtain a more normally distributed test outcome.

Change in pain
Before the retest each patient judged whether his or her low back pain had changed (7-point ordinal scale; 1 = very much decreased, 7 = very much increased) compared with his/her level of back pain before the first test. Patients reporting 'very much' or 'much decrease' of pain were categorized as 'decreased pain', those reporting 'little decrease', 'no change' and 'little increase' as 'unchanged' and those reporting 'much' or 'very much increase' as 'increased pain'.

Statistics
To study the influence of task experience, differences between the test and retest were calculated and compared between the two groups by using the Mann-Whitney test. A P-level of 0.05 was considered as significant. Because the time-scored tasks (50-foot fast walking, sit to stand) were not normally distributed at the first test, the inverse of these measures was used for further analysis. To estimate the degree of agreement in the total study population, the intraclass correlation coefficient (ICC (1,1)) was used (Rankin and Stokes 1998; Shrout and Fleiss 1979). This model is based on one-way analysis of variance, in which all variation between occasions is regarded as measurement error. An ICC above 0.75 was considered as very good to excellent reliability (Fleiss 1986). For the total study population, plots of the difference between test and retest against the mean of test and retest for each task and the limits of agreement were constructed according to Bland and Altman's recommendations (1986). The limits of agreement were calculated by 1.96 x SD of the differences between test and retest. This defines the natural variation for quantifying stability over time of the performance tasks. Stability over time, in the absence of treatment, is influenced by within-patient variation and/or random error. The limits of agreement provide an interval within which 95% of all differences between two measurements of the same person will lie. The interpretation of the size of the limits of agreement must depend upon the clinical circumstances and it is not possible to use statistics to define acceptable agreement (Altman 1991). Data analyses were performed using SPSS, version 12.0.

Table 1: Baseline characteristics for total sample and per group at moment of testing; Group A had no experience with the tasks, group B had previous experience on two or three occasions

Characteristics	Total (n=53)	Group A (n=30)	Group B (n=23)
Age (yr)	43.19 ± 9.27	42.50 ± 11.01	44.09 ± 6.50
Gender (% male)	25 (47.2%)	16 (53.3%)	9 (39.1%)
Education (%)			
low	27 (50.9%)	13 (43.3%)	14 (60.9%)
middle	25 (47.2%)	16 (53.3%)	9 (39.1%)
high	1 (1.9)	1 (3.3%)	0 (0.0%)
Work status (%)			
full time	17 (32.1%)	9 (30.0%)	8 (34.8%)
partial sick leave/ disability pension	14 (26.4%)	8 (26.7%)	6 (26.1%)
full sick leave/ disability pension	20 (37.7%)	12 (40.0%)	8 (34.8%)
no job/retired	2 (3.8%)	1 (3.3%)	1 (4.3%)
Duration of complaints (mo)	53.36 ± 67.70	46.00 ± 66.98*	62.96 ± 68.90*
Duration of disability (mo)	31.77 ± 45.12	29.10 ± 53.20*	35.26 ± 32.53*
Radiation of pain (%)			
no radiation	10 (18.9%)	5 (16.7%)	5 (21.7%)
above knee	16 (30.2%)	11 (36.7%)	5 (21.7%)
below knee	27 (50.9%)	14 (46.7%)	13 (56.5%)
RDQ	13.17 ± 4.21	14.63 ± 2.93*	11.26 ± 4.88*
Current pain (VAS 0-100)	44.45 ± 23.45	43.8 ± 24.50	45.30 ± 22.51

Values presented as means and standard deviation or percentage. * Significant differences between groups (P < 0.05). RDQ = Roland Disability Questionnaire.

Results

Group B had a significantly longer duration of complaints (P = 0.023) and lower level of disability (P = 0.006) (table 1). Since no significant differences between both groups were found at the start of our trial (data not presented), these differences were due to the fact that for this test-retest study, group B was tested 6 or 12 months later than group A. Group B also had a significantly lower level of disability; mean RDQ score 3.37 points lower (95% CI: 1.21 to 5.54). At the start of our trial, group B had exactly the same mean RDQ score as group A, indicating that the patients of group B decreased their level of disability during and/or after trial treatment and before participation in this test-retest study. The mean scores of all tasks during test and retest and the difference between test and retest per group of patients are shown in table 2. In order to facilitate interpretation of the results, the scores of 50-foot fast walking and sit to stand task are also presented in seconds. At retest, group A performed slightly better on all but one task: the PILE. Group B showed similar results, except at retest the performance on the sit to stand and one-minute stair climbing was a bit worse. There were no significant differences between the two groups, confirming our hypothesis that patients without task experience did not perform better at retest than the more experienced ones.

The overall test-retest results and the level of ICCs for the total study population are presented in table 3. The ICC for loaded forward reach did not quite reach 0.75, but was very good to excellent for all other tasks.

The Bland Altman plots including the mean of difference and the limits of agreement are presented in figure 1. The mean difference for five-minute walking was -2.43 meter (SD 42.2), for the inverse of 50-foot walking and sit to stand -0.0027 seconds^{-1} (SD 0.017) and 0.003 seconds^{-1} (SD 0.008), respectively, -2.75 cm (SD 9.5) for loaded reach, -1.66 stairs (SD 7.5) for one-minute stair climbing, and 0.26 stages (SD 1.0) for the PILE. The limits of agreement for five-minute walking were ± 82.7 meters, for 50-foot fast walking and sit to stand (recalculated from inverse for the mean score of sample) ± 3.9 seconds and ± 7.6 seconds, respectively, ± 18.7 cm for loaded forward reach, ± 14.7 stairs for one-minute stair climbing, and ± 2 stages for the PILE. The limits of agreement expressed as percentage of the mean score of each task (limits of agreement%) appeared to be low to moderate for five-minute walking and one-minute stair climbing, 21% and 20%, respectively, but only moderate for 50-foot fast walking, sit to stand and forward reach (33%, 29% and 36%, respectively). The limits of agreement of 2 stages for the PILE (48%) means that within-person variance and/or random errors have led to a large instability in measurement results. Approximately 95% of all differences within persons will lie between ± 2 stages as the mean score is only 4 stages.

Of the total study population 50 patients reported on the change of pain before the retest. Ten were categorized as having 'increased pain' and 40 were classified 'unchanged'. No significant differences regarding the test-retest results between both groups were found, although patients reporting increase in pain showed a tendency of less performance on retest. ICCs and limits of agreement for the group of patients reporting to be unchanged were quite similar to the results of the whole sample (data not presented). Due to the low number of patients reporting increased pain (n = 10), calculation of ICCs and limits of agreement for this group of patients was not meaningful.

Discussion

As hypothesized, task experience did not influence the difference between test and retest results of tasks resembling daily life activities. The ICCs represent a high level of reliability. Despite strict testing protocol, same testing time at test and retest and well trained test observers, the limits of agreement are less satisfying.

The ICCs found are comparable to the results reported earlier for the five-minute walking (0.87), 50-foot fast walking (0.80), and sit to stand (0.89). Only the ICC for the loaded forward reach of 0.74 is lower than the 0.91 reported by Simmonds et al. (1998). Our sample is more and longer disabled than that of Simmonds et al. (1998), with a RDQ score of 13.1 versus 8.3 and duration of complaints of 53.4 months versus 12.4 months. The Pearson correlation for two-minute stair climbing reported by Harding et al. (1994) is 0.94, which although not completely comparable, resembles the ICC of 0.96 in our sample. Recently, Lygren et al. (2005) reported an ICC of 0.91 for the PILE for a group of 31 patients with CLBP, which is totally in agreement with our results.

Even when ICC values are excellent, this does not imply that the tasks are clinically useful. The problem with the use of ICC is that it measures the strength of the relation between two variables, but not the agreement between them and that it is strongly influenced by the variation between patients (Bland and Altman 1986). The coefficient depends on that quantity in the study sample; when the patients perform very differently on a task and causing much variation in test results, the ICC will tend to be higher. Therefore it is advised to also calculate the limits of agreement, which in this study indeed show less favorable results from the ICCs.

This is the first study that calculated the limits of agreement for several easy-to-administer and easy-to-perform tasks for patients with chronic nonspecific low back pain. Limits of agreement can also be used as a cut-off score for change in an intervention study or in daily practice. In an individual patient the change due to treatment should exceed these limits of agreement before one can state that a treatment has been clinically effective (Brouwer et al. 2004). The limits of agreement% for the five-minute walking and one-minute stair climbing are in our opinion sufficient to advise clinical use although it should be taken into account that a patient still has to improve considerably. For example the walking distance should improve at least 83 meters after treatment in order to conclude that the change is larger than could

be due to natural variation of the task. The limits of agreement% of the 50-foot fast walking, sit to stand and loaded forward reach varied from 29% to 38% and are quite high, but in our opinion are still useful for clinical practice. It should, however, be considered that the patient has to improve substantially to conclude that he really made an improvement beyond chance. Lygren et al. (2005) were to first to present the limits of agreement for the PILE. Since they did not use the number of lifting stages but the final amount of weight lifted, we recalculated their data to the number of lifting stages. The limits of agreement appeared to be 2.23 stages (52% of the mean score), which is almost exactly the same as in our sample. They reported no conclusions at all about the clinical usefulness of this task. As already stated there is no cut-off point available for the limits of agreement% (Altman 1991), but in our opinion the 52% in the study of Lygren et al. (2005) and 48% in our study, is too high for the PILE to be of any clinical use. This is also in agreement with Keller et al. (2001) who calculated the limits of agreement for the Åstrand bicycle test, isokinetic trunk extensor test and Biering-Sørensen test in CLBP patients and judged a test with a limits of agreement of ≤ 33% as reliable, and ≥ 42% as unreliable.

We checked whether the increase of pain after the first test resulted in more variance and therefore different levels of ICC and limits of agreement. No significant differences were found between test and retest for the group with increased pain and the group with unchanged pain. However, the number of patients reporting an increase of pain was too low to guarantee sufficient precision to calculate ICCs and limits of agreement for this particular group. Nevertheless, the ICCs and limits of agreement of the group with exclusion of the 10 patients who reported increase of pain did not differ from the results of the total study population. This is in agreement with earlier research that reported only a low correlation between the level of performance and pain intensity (Simmonds et al. 1998). By including the patients who reported an increase of pain, we approached the normal situation that a patient's level of pain might vary even in the absence of treatment. Furthermore, this increases the possibility for the generalization of the results to our total trial population and clinical practice.

A sample of at least 50 patients is needed to calculate the limits of agreement (Altman 1991; Hopkins 2000). By including a total number of 53 patients, the power of our study is sufficient. So based on the abovementioned reasons, we are confident that our results are representative and reliable for the population that was included for the trial, being patients with a moderate to high level of disability due to chronic nonspecific low back pain.

Weaknesses of this study should be mentioned. The variation of patients' outcome on the performance tasks was quite large, and for several patients, floor or ceiling effects were present. For example, 10 patients even could not complete one stage of the PILE at test and retest. Furthermore, some patients did not report difficulties in stair climbing. Others reported that not the walking speed (which is measured with the five-minute walking task), but the reduced walking distance in daily life was their main problem. This indicates that the selected tasks were not always specific enough for individual patients and in our opinion future research should be more focused on tasks measuring activities that are specifically hampered for a particular patient.

Table 2: Mean scores of test, retest and mean difference of test versus retest per group

Task	Group A Test mean (SD)	Group A Retest mean (SD)	Group A Mean difference (SD)	Group B Test mean (SD)	Group B Retest mean (SD)	Group B Mean difference (SD)
Five-minute walking (m)	388.12 (74.3)	389.96 (94.0)	-1.84 (50.8)	384.60 (91.1)	387.80 (95.1)	-3.20 (28.5)
Inverse fifty-foot fast walking (1/sec)	0.101 (0.02)	0.103 (0.02)	-0.002 (0.02)	0.094 (0.03)	0.098 (0.03)	-0.004 (0.02)
Untransformed (sec)	10.45 (2.8)	10.33 (2.9)	0.12 (2.8)	11.56 (3.8)	11.25 (5.1)	0.31 (2.6)
Inverse five times sit to stand (1/sec)	0.048 (0.02)	0.053 (0.02)	-0.005 (0.01)	0.059 (0.02)	0.060 (0.02)	-0.001 (0.01)
Untransformed (sec)	23.88 (10.2)	22.43 (12.8)	1.45 (7.7)	19.46 (9.38)	19.75 (9.4)	-0.29 (3.0)
Loaded forward reach (cm)	50.79 (12.9)	53.51 (12.5)	-2.72 (10.3)	50.39 (15.7)	53.18 (14.5)	-2.79 (8.6)
One-minute stair climbing (steps)	73.90 (28.8)	76.87 (28.6)	-2.97 (8.8)	74.65 (24.8)	74.61 (25.3)	0.04 (5.1)
Pile test (lifting stages)	4.43 (2.8)	4.11 (2.5)	0.32 (1.2)	4.27 (2.8)	4.09 (2.8)	0.18 (0.8)

Table 3: Performance tasks results during test and retest and ICCs for the total study population

Task	Test mean (SD)	Retest mean (SD)	*n*	ICC (1,1) (95 % CI)
Five-minute walking (m)	386.59 (81.21)	389.02 (93.59)	53	0.89 (0.81 - 0.93)
Inverse fifty-foot fast walking (1/sec)	0.098 (0.024)	0.101 (0.024)	52	0.76 (0.61 - 0.85)
Untransformed (sec)	10.92 (3.30)	10.68 (3.93)		
Inverse five times sit to stand (1/sec)	0.053 (0.019)	0.056 (0.020)	53	0.91 (0.84 - 0.94)
Untransformed (sec)	21.96 (10.01)	21.27 (12.17)		
Loaded forward reach (cm)	50.62 (13.09)	53.37 (13.27)	52	0.74 (0.59 - 0.84)
One-minute stair climbing (steps)	74.23 (26.89)	75.89 (27.02)	53	0.96 (0.93 - 0.98)
Pile test (lifting stages)	4.27 (2.80)	4.10 (2.61)	50	0.92 (0.87 - 0.96)

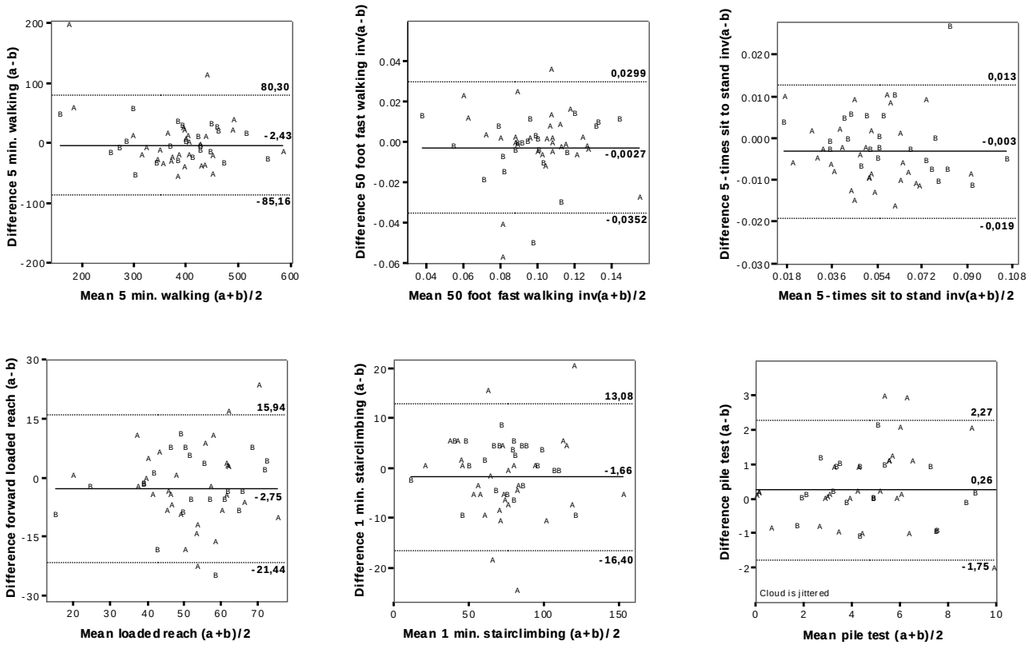

Fig.1: Bland-Altman plots, difference test-retest plotted against mean test-retest in six performance tasks
○: group A, △: group B. Reference line in the plots: —: mean, … : limits of agreement (±1.96 SD).

Acknowledgements

This study is supported by Zorgonderzoek Nederland/Medische Wetenschappen (ZonMw) grant number 014-32-007 and the Rehabilitation Centre Blixembosch.

References

- Altman DG. Some common problems in medical research. In: DG Altmans. Practical statistics for medical research. London: Chapman & Hall, 1991, pp. 396-403.
- Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986;1:307-10.
- Brouwer S, Kuijer W, Dijkstra PU, Goeken LN, Groothoff JW, Geertzen JH. Reliability and stability of the Roland Morris Disability Questionnaire: intra class correlation and limits of agreement. Disabil and Rehabil 2004;26:162-5.
- Brouwer S, Reneman MF, Dijkstra PU, Groothoff JW, Schellekens JM, Goeken LN. Test-retest reliability of the Isernhagen Work Systems Functional Capacity Evaluation in patients with chronic low back pain. J Occup Rehabil 2003;13:207-18.
- Dworkin RH, Turk DC, Farrar JT, Haythornthwaite JA, Jensen MP, Katz NP, Kerns RD, Stucki G, Allen RR, Bellamy N, Carr DB, Chandler J, Cowan P, Dionne R, Galer BS, Hertz S, Jadad AR, Kramer LD, Manning DC, Martin S, McCormick CG, McDermott MP, McGrath P, Quessy S, Rappaport BA, Robbins W, Robinson JP, Rothman M, Royal MA, Simon L, Stauffer JW, Stein W, Tollett J, Wernicke J, Witter J. Core outcome measures for chronic pain clinical trials: IMMPACT recommendations. Pain 2005;113:9-19.
- Fleiss J. Reliability of measurement. In: J Fleiss. Design and analysis of clinical experiments. New York: John Wiley & Sons, 1986, pp. 1-32.
- Gibbons LE, Videman T, Crites Battié M. Determinants of isokinetic and psychophysical lifting strength and static back muscle endurance: A study of male monozygotic twins. Spine 1997;23:2412-21.
- Harding VR, Williams AC, Richardson PH, Nicholas MK, Jackson JL, Richardson IH, Pither CE. The development of a battery of measures for assessing physical functioning of chronic pain patients. Pain 1994;58:367-75.
- Hopkins WG. Measures of reliability in sports medicine and science. Sports Med 2000;30:1-15.
- Kankaanpaa M, Taimela S, Laaksonen D, Hanninen O, Airaksinen O. Back and hip extensor fatigability in chronic low back pain patients and controls. Arch Phys Med Rehabil 1998;79:412-7.
- Keller A, Hellesnes J, Brox JI. Reliability of the isokinetic trunk extensor test, Biering-Sorensen test, and Åstrand bicycle test: assessment of intraclass correlation coefficient and critical difference in patients with chronic low back pain and healthy individuals. Spine 2001;26:771-7.
- Lee CE, Simmonds MJ, Novy DM, Jones S. Self-reports and clinician-measured physical function among patients with low back pain: a comparison. Arch Phys Med Rehabil 2001;82:227-31.
- Ljungquist T, Harms Ringdahl K, Nygren A, Jensen I. Intra- and inter-rater reliability of an 11-test package for assessing dysfunction due to back or neck pain. Physiother Res Int 1999;4:214-32.
- Lygren H, Dragesund T, Joensen J, Ask T, Moe Nilssen R. Test-retest reliability of the Progressive Isoinertial Lifting Evaluation (PILE). Spine 2005;30:1070-4.
- Magnussen LMPT, Strand LIPPT, Lygren HMPT. Reliability and Validity of the Back Performance Scale: Observing Activity Limitation in Patients with Back Pain. Spine 2004;29:903-7.
- Mayer TG, Barnes D, Kishino ND, Nichols G, Gatchel RJ, Mayer H, Mooney V. Progressive isoinertial lifting evaluation. I. A standardized protocol and normative database. Spine 1988a;13:993-7.
- Mayer TG, Barnes D, Nichols G, Kishino ND, Coval K, Piel B, Hoshino D, Gatchel RJ. Progressive isoinertial lifting evaluation. II. A comparison with isokinetic lifting in a disabled chronic low-back pain industrial population. Spine 1988b;13:998-1002.
- Mayer TG, Gatchel RJ. Functional restoration for spinal disorders: the sports medicine approach. Philadelphia: Lea & Febiger, 1988.

References

- Mayer TG, Gatchel RJ, Barnes D, Mayer H, Mooney V. Progressive isoinertial lifting evaluation. Erratum notice. Spine 1990;15:5.
- Newton M, Thow M, Somerville D, Henderson I, Waddell G. Trunk strength testing with isokinetic lifting-machines. Part 2: Experimental evaluation of the Cybec II back testing system in normal subjects and patients with chronic low back pain. Spine 1993;18:812-24.
- Novy DM, Simmonds MJ, Lee CE. Physical performance tasks: what are the underlying constructs? Arch Phys Med Rehabil 2002;83:44-7.
- Novy DM, Simmonds MJ, Olson SL, Lee CE, Jones SC. Physical performance: differences in men and women with and without low back pain. Arch Phys Med Rehabil 1999;80:195-8.
- Parks KA, Crichton KS, Goldford RJ, McGill SM. A comparison of lumbar range of motion and functional ability scores in patients with low back pain: assessment for range of motion validity. Spine 2003;28:380-4.
- Rankin G, Stokes M. Reliability of assessment tools in rehabilitation: an illustration of appropriate statistical analyses. Clin Rehabil 1998;12:187-99.
- Reneman MF, Dijkstra PU, Westmaas M, Goeken LN. Test-retest reliability of lifting and carrying in a 2-day functional capacity evaluation. J Occup Rehabil 2002;12:269-75.
- Roland M, Morris R. A study of the natural history of back pain. Part I: development of a reliable and sensitive measure of disability in low-back pain. Spine 1983;8:141-4.
- Shrout PE, Fleiss JL. Intraclass correlations: Uses in assessing rater reliability. Psychol Bull 1979;86:420-8.
- Simmonds MJ, Olson SL, Jones S, Hussein T, Lee CE, Novy D, Radwan H. Psychometric characteristics and clinical usefulness of physical performance tests in patients with low back pain. Spine 1998;23:2412-21.
- Smeets RJ, Vlaeyen JW, Hidding A, Kester AD, van der Heijden GJ, van Geel AC, Knottnerus JA. Active rehabilitation for chronic low back pain: Cognitive-behavioral, physical, or both? First direct post-treatment results from a randomized controlled trial [ISRCTN22714229]. BMC Musculoskelet Disord 2006;7:5.
- Strand LI, Moe Nilssen R, Ljunggren AE. Back Performance Scale for the assessment of mobility-related activities in people with back pain. Phys Ther 2002;82:1213-23.
- Waddell G. A new clinical model of low back pain and disability. The back pain revolution. London: Churchill Livingstone, 1998.
- Watson PJ. Non-psychological determinants of physical performance in musculoskeletal pain. In: M Mitchell. Pain 1999- an updated review. Refresher course syllabus 9th World Congress on Pain. Seattle: IASP Press, 1999, pp. 153-8.
- WHO. The International Classification of Functioning, Disability and Health (ICF), second edition. Geneva: WHO, 2001.
- Wittink H. Functional capacity testing in patients with chronic pain. Clin J Pain 2005;21:197-9.

Physical capacity tasks in chronic low back pain: What is the contributing role of cardiovascular capacity, pain and psychological factors?

Rob J.E.M. Smeets, Tineke C.M. van Geel, Arnold D.M. Kester, J. André Knottnerus

Published as:
Smeets RJEM, van Geel ACM, Kester ADM, Knottnerus JA. Physical capacity tasks in chronic low back pain: What is the contributing role of cardiovascular capacity, pain and psychological factors? Disability and Rehabilitation 2006: in press.

Chapter 9

Abstract

Purpose: To explore the association of personal and body functions with physical capacity tasks used in the evaluation of chronic low back pain (CLBP) treatment.
Method: Cross-sectional study in which 221 patients with nonspecific CLBP participated. Physical capacity was assessed by six capacity tasks, and several personal and body functions were assessed by questionnaires (age, gender, pain intensity, duration of pain, radiating pain to leg, fear of injury/movement, depression, pain catastrophizing and internal control of pain). Cardiovascular capacity was measured by a modified Åstrand submaximal bicycle test. The association of these hypothetically influential personal and body function factors with capacity tasks was examined with multiple linear regression analyses.
Results: The total explained variance was low to moderate (9% to 19%), except for stair climbing for which the variance just reached 30%. Many putative factors did not contribute significantly to the level of outcome of several different tasks.
Conclusions: Cardiovascular capacity, pain intensity, fear of injury/movement, cognitions, and depression had statistically significant but clinically minor effects on several, but not all capacity tasks. Radiating pain, age and duration of complaints had no significant influence at all. Due to anthropometric differences men outperformed women on most tasks. Apparently the influence of many personal, physical, but especially psychological factors on the selected capacity tasks is not high at all. This might indicate that these tasks are more objectively measuring physical capacity than expected.

Introduction

Most industrialized countries are confronted with a dramatic increase in people dealing with low back pain (LBP) (Pfingsten et al. 1997). The resulting disability has become an epidemic health and socioeconomic problem (van Tulder et al. 1995). A relatively small group of chronic low back pain (CLBP) sufferers accounts for up to 90% of all medical and social costs for LBP (Nachemson 1992).

In the clinical rehabilitation assessment of CLBP patients, mostly self-reported questionnaires are used. However, to obtain more insight into the patients' ability of functioning, more objective measures such as physical capacity tasks have been advocated (Wittink 2005), and tasks such as walking, climbing stairs and lifting are becoming common practice in the evaluation of CLBP treatment (Harding et al. 1994; Ljungquist et al. 1999; Mayer et al. 1988a; Mayer et al. 1988b; Simmonds et al. 1998). Although physical capacity tasks and self-reported disability show low to moderate correlations (Filho et al. 2002; Lee et al. 2001; Novy et al. 2002; Reneman et al. 2002; Simmonds et al. 1998; Wittink 2005), these correlations are higher than those between self-reported disability and the more physical tests such as mobility of the spine or muscle strength (Parks et al. 2003; Simmonds et al. 1998).

Self-reported disability and physical capacity tasks seem to measure different aspects of activity as defined by the International Classification of Functioning, Disability and Health (ICF). The ICF is similar to the biopsychosocial model, identifying three concepts described from the perspective of body systems (body functions and structures), the individual (activity) and society (participation). Activity is defined as the execution of specific tasks or actions by the individual (e.g. walking, bending), whereas participation is characterized as the individual functioning in his societal context (e.g. work, leisure time) (WHO 2001). According to the ICF, self-reported disability is a performance qualifier of the activity domain describing the perceived inability of a person to perform an activity in his actual environment. On the other hand, a physical capacity task can be defined as a capacity qualifier indicating the highest probable level of functioning of a person at a given moment and in a standardized environment. The score of physical capacity tasks is not only influenced by body functions and structures, but also by personal factors (e.g. coping style, fear, age, education) and environmental factors (solicitous spouse, societal factors) (Watson 1999).

Until now, several body functions and structures e.g. muscle strength, pain level, leg pain and cardiovascular capacity (VO_2max) have been identified as contributing factors to physical capacity tasks scores (Matheson et al. 2002; Simmonds 1999). Likewise, many personal factors, like age (Ljungquist et al. 2003; Mayer et al. 1988a), gender (Lee et al. 2001; Ljungquist et al. 2003; Novy et al. 1999; Simmonds 1999), duration of complaints (Geisser et al. 2000b), depression (Kaplan et al. 1996; Novy et al. 2002) and pain cognitions (Kaplan et al. 1996; Lackner and Carosella 1999; Lackner et al. 1996; Lee et al. 2001; Mayer et al. 1988a; Novy et al. 2002) have been reported. Especially fear of injury/movement has been postulated to be a very influential factor (Crombez et al. 1998; Crombez et al. 1999; McCracken et

al. 1992; Vlaeyen et al. 1995a). Based on the fear-avoidance model it is assumed that due to catastrophizing, the patient experiencing pain might become fearful, start to avoid activities, and finally develop more disability, disuse and depression (Vlaeyen and Linton 2000). Similarly, it might be expected that this patient will encounter difficulties in performing activities of daily life that can be assessed by capacity tasks (Verbunt et al. 2003; Vlaeyen et al. 1995b). Furthermore, based on the anthropometric differences, it is expected that men will outperform women on these tasks, and younger patients will outperform older ones.

Although many factors seem to influence the physical capacity in CLBP, the evidence regarding putatively influential factors is conflicting and often confusing. For example, many studies reported only a weak or even absent association between pain intensity and physical capacity (Filho et al. 2002; Novy et al. 2002; Rainville et al. 2004; Reneman et al. 2003), while other studies reported inverse associations (Gronblad et al. 1997; Lackner and Carosella 1999; Lackner et al. 1996; Ljungquist et al. 2003; Watson 1999). The same applies to the influence of fear of injury/movement (Crombez et al. 1998; Crombez et al. 1999; Geisser et al. 2000a; Lackner et al. 1996; McCracken et al. 1992; Reneman et al. 2003; Vlaeyen et al. 1995a). This contradiction is partly due to the fact that most studies only calculated univariate correlations between a putatively influential factor and the physical capacity tasks (Filho et al. 2002; Lee et al. 2001; Mayer et al. 1988a; Novy et al. 2002; Novy et al. 1999; Simmonds et al. 1998). Only a few studies, mainly concerning lifting tasks, examined the influence of a limited number of potentially contributing factors simultaneously (Crombez et al. 1999; Geisser et al. 2000a; Lackner et al. 1996; Ljungquist et al. 2003; Reneman et al. 2003).

It can be concluded that data about the simultaneous influence of hypothetically influential personal factors and body functions on physical capacity tasks commonly used in the evaluation of CLBP treatment, are scarce. The aim of this study was to further explore the simultaneous influence of these factors. The results and consequences of this exploratory study will be further discussed.

Methods

The results of this study are part of a larger randomized controlled trial (RCT) (ISRCTN22714229). In this RCT, a physical treatment, a cognitive-behavioral, a combination of both treatments and a waiting list control group are being compared. The methods of the RCT are extensively described elsewhere (Smeets et al. 2006a), and a description of participants and measurements is presented here with special emphasis on the potentially influential factors.

Participants

Between April 2002 – December 2004, patients who were referred by their general practitioner or medical specialist for treatment to an outpatient unit of three rehabilitation centres in the province of Noord-Brabant, The Netherlands, were selected. Of the eligible 309 patients, 223 patients agreed to participate in the RCT. The inclusion criteria were: (1) first referral to a rehabilitation centre to reduce disability due to nonspecific LBP existing for at least 3 months, (2) age between 18 and 65 years and (3) ability to walk at least 100 meters. Exclusion criteria were: (1) lumbar disc herniation with neurological symptoms, (2) inflammatory or neoplastic disease, (3) fracture of the spine, (4) spondylolisthesis or spondylolysis, (5) lumbar spondylodesis, (6) major psychiatric problems, (7) cardiac and/or pulmonary conditions, (8) severe addiction to drugs, narcotics or alcohol, and (9) pregnancy. A total of 82 patients initially referred, did not participate in the RCT. The reasons for not participating were: Not willing to participate in this research (n = 14), not meeting criteria (n = 31), preference for a particular treatment (n = 20), logistic problems (n = 15), and waiting time too long (n = 2).

Procedures

All data presented in this study were obtained during a pre-randomization testing session. Patients completed questionnaires, performed a modified Åstrand submaximal bicycle test and several capacity tasks. All patients gave written informed consent to participate. The medical ethics committee of the Rehabilitation Foundation Limburg and Institute for Rehabilitation Research at Hoensbroek in The Netherlands approved the study protocol.

Demographic data and disease characteristics

Information regarding age, gender, duration of complaints and level of radiation of pain to leg was collected.

Body functions

Pain intensity

Current pain intensity was measured by using the ten centimeters visual analogue scale (VAS) of the McGill Pain Questionnaire with on the left side ‘no pain’ and on the right side ‘unbearable pain’. Relevance, validity and reliability have been sufficiently tested for patients with LBP (Carlsson 1983; Melzack and Katz 1992).

Cardiovascular capacity

To predict maximum oxygen consumption (VO_2max in mL/kg LBM.min^{-1}), a modified submaximal Åstrand bicycle test was performed. In this test, the workload was gradually increased. Sufficient test-retest reliability (r = 0.87) and validity (r = 0.84, when compared with maximal VO_2-uptake measured with a maximum exercise test) have been established (Hodselmans et al. 2001).

First, patients' lean body mass (LBM) was measured according to the protocol of Durnin and Womersley (1974) by using a skinfold caliper. Next, the participants performed the test on a calibrated cycle ergometer. A monitor placed on the patients' chest monitored his heart rate (HR). The patient started cycling with a workload of 0.5 W/kg LBM at a constant rate of 60 rpm. After two minutes the workload was increased to 1.5 W/kg LBM. The increase with 0.5 W/kg LBM was continued until the HR reached 120 beats/min. When the HR exceeded 120 beats/min the patient cycled six minutes with a fixed workload in order to reach a steady state phase, meaning the HR did not vary more than five beats/min during the last minute of exercise. The average HR during the last minute was calculated. VO_2max was estimated by using the Åstrand's nomogram (Åstrand and Rohdahl 1986) based on the linear association between HR and increase in oxygen uptake. If the HR during the last minute varied more than five beats/min, no VO_2max could be estimated. The test was stopped when the patient did not reach a HR of at least 120 beats/min, the HR exceeded the predefined maximum HR ([220 – age] * 0.85), the blood pressure reached the level of 220/115 mm Hg, or if the patient showed signs of serious cardiovascular or pulmonary difficulties.

Personal Factors

Fear of injury/movement

The Tampa Scale for Kinesiophobia (TSK), which measures fear of injury and movement and consists of 17 items with a total score ranging from 17 to 68 was used. A higher score on the TSK indicates more fear. The questionnaire is considered reliable and valid in CLBP (Roelofs et al. 2004; Vlaeyen et al. 1995a).

Depression

Depression was measured by the Beck Depression Inventory (Beck et al. 1979), which is a reliable, valid and widely used questionnaire (Beck et al. 1988).

Cognitions

The subscales pain catastrophizing and internal control of pain of the Pain Cognition List (PCL) were used to measure pain catastrophizing and internal control of pain. The reliability, stability and validity of this questionnaire have been proven to be sufficient in CLBP patients (Vlaeyen et al. 1990; Vlaeyen et al. 2003). The PCL is a 39-item self-report questionnaire with a five-point Likert scale answering categories ranging from 'completely disagree' to 'completely agree'. The pain catastrophizing subscale consists of 16 items and the score ranges from 16 to 80. Some items are: "My thoughts are always concentrating on the pain", "I feel like an unlucky person", "I think that fate has struck me". The higher the score the more the person is catastrophizing. The internal control of pain subscale consists of five items. Some typical items are: "I know a way to decrease my pain a little", "I think that I can influence my pain positively". The total score of these five items is calculated and ranges from 5 to 25, with a higher score indicating increased control of pain.

Capacity tasks

Six capacity tests were selected from task batteries described in detail in several studies (Harding et al. 1994; Lee et al. 2001; Mayer et al. 1988a; Mayer et al. 1988b; Novy et al. 2002; Novy et al. 1999; Simmonds et al. 1998), based on their ability to measure activities that are normally hampered in CLBP patients. The tasks were all well described. The research assistants who did the assessments were carefully instructed not to react to the patients' behavior, and only to use the verbal instruction as written in the protocol. Thus solicitous behavior was prevented as much as possible. All research assistants received a special training and attended regular refresher courses two to three times a year to ensure reliable testing results. Testing was done to assess the pre-treatment level of performance on the capacity tasks and was not used to assess the work capacity or for matters concerning litigation. Due to the strict testing protocol, the environmental factors were equal for all patients, and therefore their influence was not further assessed. The physical capacity task battery consisted of:

Five-minute walking test

Participants were asked to walk as fast as possible, no running, for a five-minute period over a 30 meters long and eight-shaped circuit. The distance covered was recorded in meters. The test-retest reliability has been reported to be 0.92 (Novy et al. 2002; Simmonds et al. 1998), and a day-to-day reliability of 0.87 in persons with LBP (Lee et al. 2001; Simmonds et al. 1998).

Fifty-foot walking test

Participants had to walk as fast as possible over a 50 feet long and eight-shaped circuit. The time was recorded in seconds. The interrater and test-retest reliability of this test are 0.99 and 0.80, respectively, in LBP subjects (Novy et al. 2002; Simmonds et al. 1998).

Sit to stand test

Participants were asked to perform five times a sit to stand movement from a chair without arms as fast as possible. The task was performed twice and the average time in seconds needed to perform the task was calculated. The interrater and test-retest reliability have been reported to be 0.99 and 0.89, respectively, for people suffering from LBP (Novy et al. 2002; Simmonds et al. 1998).

Loaded forward reach

Participants were instructed to hold a stick with a weight of 4.5 kg at shoulder height and width. They had to reach forward as far as possible without lifting the heels off the floor. The distance reached was recorded in centimeters. The test-retest, day-to-day, and interrater reliability coefficients have been reported to range from 0.91 to 0.99 (Novy et al. 2002; Simmonds et al. 1998).

One-minute stair climbing test
Patients were asked to walk a stair up and down for one minute. The circuit was five steps high and eight-shaped. The number of steps climbed was recorded. The test-retest, and interrater reliability coefficients were calculated for a two-minute stair climbing test and were 0.94 and 0.999 respectively, and correlated highly with a one-minute test ($r = 0.98$; $P < 0.001$) (Harding et al. 1994).

PILE-test weight lifting from floor to waist
Patients had to lift a box with a weight four times within 20 seconds from floor up to a 75 cm high table. The weight was increased in a standardized manner. The test stopped when the patient could not lift the box four times within 20 seconds, the patient decided to stop because of tiredness or pain, the HR exceeded 85% of the maximal HR (0.85 * [220 – age]), maximal weight that could safely be lifted was reached, or the research assistant considered the lifting unsafe (Mayer et al. 1988a; Mayer et al. 1988b). The number of fully completed lifting stages was recorded. By recording the number of stages instead of the total amount of weight lifted, the difference in increase of weight between sexes was compensated for and a more normally distributed test outcome was obtained. The test-retest reliability has been reported to be 0.87 in a healthy population (Mayer et al. 1988b) and 0.91 in patients with long-lasting low back pain (Lygren et al. 2005).

Data Analysis
To determine the relative contribution of each of the predictors (age, gender, duration of complaints, level of radiating pain to leg, pain intensity, VO_2max, TSK, BDI, PCL-pain catastrophizing, PCL-internal control of pain) to the variance in each physical capacity task, multiple regression analysis was used. In case of non-normal distribution of the dependent variable (score of a capacity task), a transformation was applied to obtain a normal distribution. Unstandardized and standardized beta weights and tests of significance for each of the predictors were calculated.

For performing the multiple linear regression analysis with ten independent variables, the number of the variables times ten (100 patients) are needed, as recommended for multiple regression analysis by Dawson-Saunders and Trap (1998). Multicollinearity was checked by calculating the variance inflation factor (VIF); a VIF > 3 was judged as a problem of collinearity.

When performing regression analyses, only those patients with complete datasets can be incorporated. It is known from our previous research that for a sizable proportion of patients, no VO_2max could be calculated (Smeets et al. 2006b). Excluding this group from analysis would bias our results since significant differences between the patients with and those without VO_2max were found while using *t*-tests. We decided to replace the missing values using multiple imputation of the VO_2max. The data set of the patients for whom the VO_2max could be calculated was used and generated predicted values for VO_2max based on a linear regression model. The initial VO_2max prediction model

consisted of the remaining nine predictive variables, the level of disability and the level of activity during sport time, during work time and during leisure time. These last four variables were added because we hypothesized that they are indicators of the level of activity displayed by the patient, and hence, would correlate highly with and predict the actual level of aerobic capacity. Of these variables, only age, level of disability (measured by the RDQ) (Roland and Morris 1983), and level of activity during leisure time (measured by the Baecke Physical Activity Questionnaire) (Baecke et al. 1982) appeared to be significant predictors of the VO_2max. Next, we imputed the predicted values of VO_2max with the addition of a normally distributed random error term, whose standard deviation (SD) was taken equal to the residual SD from the prediction model. Imputation and analyses of the resulting completed data was repeated ten times, and the results were then combined using the methods of Rubin (1987).
All statistical analyses were performed with SPSS software, version 12.0.

Results

Of the 223 patients who were eligible for inclusion, two were excluded from further analysis because BDI data were not registered. Of the remaining 221 patients, 175 (79.2%) had a complete dataset, and 46 patients had missing VO_2max data (20.8%). Table 1 represents the characteristics of the total sample of 221 patients, and by group according to the availability of the VO_2max score. The total group had moderate to severe functional limitations and quite a high percentage of the patients was on sick leave or disability pension because of their low back pain.

Comparison between both groups showed that age, current pain, level of depression, internal control, five-minute walking task, fast walking task, sit to stand task, one-minute stair climbing task and the lifting task, differed significantly. Therefore we used the imputation technique described in the method section. Furthermore, the dependent variables fast walking and sit to stand task were not normally distributed. The inverse scores of these tasks showed normal distributions and were used in further analyses.

The results of multiple regression analyses are presented in table 2. In none of the regression analyses a VIF >3 was found, indicating that there was no problem with collinearity.
For each capacity task, the explained variance (adjusted R^2 and R^2) of all putative predictors simultaneously, and only the significantly contributing predictors (after correction for the other predictors) and their unstandardized, as well as the standardized beta weights are shown.

The results can be summarized as follows: (1) In the five-minute walking task male patients performed better, and patients with more pain and higher level of depression worse. (2) In the fast walking task, male patients and those experiencing a higher level of control of pain walked faster and those having more pain walked slower. (3) In sit to stand task, the patients with a higher level of VO_2max performed faster and those with higher level of depression slower.

Table 1: Baseline variables for the total study population and the samples with and without VO_2max

Variables	Total study population (n=221)	Sample with known VO_2max (n=175)	Sample without known VO_2max (n=46)
Age (years)	41.58 ± 9.97	40.10 ± 9.95*	47.22 ± 7.85*
Gender (% male)	52.5	52.6	52.2
Education (%)			
low	62.0	58.9	73.9
middle	37.0	40.0	26.1
high	0.9	1.1	0.0
Work status (%)			
full time	29.9	32.6	19.6
partial sick leave / disability pension	24.0	22.9	28.3
full sick leave / disability pension	37.6	37.1	39.1
no job / retired	8.6	7.4	13.0
Duration of LBP (mo)	56.67 ± 72.31	55.90 ± 71.54	59.59 ± 75.92
Radiation of pain (%)			
no radiation	14.0	15.4	8.7
above knee	37.1	38.3	32.6
below knee	48.9	46.3	58.7
RDQ	13.79 ± 3.76	13.49 ± 3.84*	14.91 ± 3.26*
Work activity (BPAQ-work)	1.90 ± 1.59	1.87 ± 1.57	2.00 ± 1.64
Sport activity (BPAQ-sport)	2.08 ± 0.69	2.15 ± 0.71*	1.83 ± 0.55*
Leisure activity (BPAQ-leisure)	2.84 ± 0.67	2.89 ± 0.69	2.67 ± 0.61
TSK	38.85 ± 6.88	38.77 ± 6.97	39.17 ± 6.57
VO_2max (mL/kg/min)	28.73 ± 7.21	28.73 ± 7.21*	-
Current pain	49.24 ± 24.81	47.77 ± 24.32	54.83 ± 26.09
BDI	10.09 ± 7.20	9.77 ± 6.91	11.30 ± 8.15
PCL pain catastrophizing	39.63 ± 12.47	39.85 ± 12.20	38.78 ± 13.54
PCL internal control	16.19 ± 3.73	16.62 ± 3.61*	14.57 ± 3.77*
Walking (m)	380.42 ± 83.16	390.40 ± 97.58*	340.73 ± 97.58*
Fast walking (sec)	10.13 ± 2.49	9.93 ± 2.36*	10.91 ± 2.87*
Sit to stand (sec)	21.37 ± 8.53	20.52 ± 7.55*	24.71 ± 11.07*
Loaded forward reach (cm)	54.11 ± 13.50	54.31 ± 13.02	53.36 ± 15.32
Stair climbing (number of stairs)	74.27 ± 23.14	77.26 ± 22.25*	62.89 ± 23.14*
Lifting (stages)	4.22 ± 2.88	4.43 ± 2.86*	3.43 ± 2.85*

Values presented as means and standard deviation or percentage. * Significant differences between the groups with known and unknown VO_2max ($P < 0.05$).
* Significant differences between the groups with known and unknown VO_2max ($P < 0.05$).
LBP = low back pain; RDQ = Roland Disability Questionnaire; BPAQ = Baecke Physical Activity Questionnaire; TSK = Tampa Scale for Kinesiophobia;
BDI = Beck Depression Inventory; PCL = Pain Cognition List.

(4) In the loaded forward reach task male patients reached further. (5) In the stair climbing task the patients with a higher level of VO_2max, catastrophizing and internal control climbed more stairs while those with higher level of pain and depression climbed less. (6) Finally, male patients performed more lifting cycles in the PILE-test, while patients with higher levels of fear of injury/movement and depression performed less cycles.

Table 2: Summary of multiple regression analyses of the performance tasks as a dependent variable with ten independent variables

Dependent	adj. R^2	R^2	Independents*	Standardized Beta	Unstandardized Beta	Confidence interval (CI 95%)**
Walking	0.155	0.209	gender	0.16	26.7	4.85 to 48.55
			current pain	-0.20	-0.67	-1.13 to -0.20
			depression	-0.18	-2.12	-4.18 to -0.07
Inverse fast walking	0.186	0.223	gender	0.24	0.0107	0.0049 to 0.0165
			current pain	-0.22	-0.0002	-0.0003 to -0.0001
			internal control	0.17	0.0010	0.0002 to 0.0018
Inverse sit to stand	0.129	0.168	VO^2max	0.16	0.0004	0.0000 to 0.0008
			depression	-0.29	-0.0007	-0.0011 to -0.0003
Forward reach	0.093	0.134	gender	0.22	6.03	2.34 to 9.73
Stair-climbing	0.303	0.335	VO_2max	0.17	0.57	0.09 to 1.05
			current pain	-0.29	-0.27	-0.39 to -0.15
			depression	-0.29	-0.92	-1.44 to -0.39
			pain catastrophizing	0.28	0.52	0.19 to 0.85
			internal control	0.17	1.07	0.29 to 0.19
Lifting	0.147	0.187	gender	0.14	0.82	0.05 to 1.60
			fear of injury/movement	-0.23	-0.10	-0.16 to -0.03
			depression	-0.25	-0.10	-0.17 to -0.03

* The ten independent variables are: gender, age, VO_2max, current pain, duration of complaints, fear of injury/movement, depression, radiation, catastrophizing and internal control. Only the significant independent variables after correction for other independent variables are displayed.
** Confidence interval of the unstandardized beta.

Discussion

The main aim of this study was to explore which hypothetically influential factors, personal and body functions, play a role in several capacity tasks, performed by patients with nonspecific CLBP. This is the first study we know of, in which the simultaneous influence of many putatively influential factors on different capacity tasks was evaluated in a large number of CLBP patients. The tasks resembled daily activities that are mostly restricted in patients with CLBP. The total explained variance was low to moderate (9% to 19%) for all capacity tasks, except for stair climbing for which the total explained variance just reached 30%. Part of the differential and low influence of the putative factors can be explained by the fact that each task reflects different aspects of function such as speed, coordination, strength, flexibility of spine and/or limbs and aerobic capacity. Nevertheless, several putative factors did not contribute significantly at all to the level of outcome of any capacity task.

As expected, the level of pain contributed significantly and inversely to the walking and stair climbing tasks. Several researchers, using similar walking tasks did not find this association, however, they only used one-way correlations (Filho et al. 2002; Novy et al. 2002). This significant association between pain and walking tasks might be explained by the fact that the alternate moving of both legs and body oscillation during walking and stair climbing, urges for more coordination of the trunk which probably is hampered, causing faster fatigue and thus a lower level of performance. The lifting task can be regarded as a straining and coordinative task for the lower back. Therefore, it is remarkable that the level of pain contributed significantly in the univariate model only, and not in the simultaneous model. Apparently, other factors such as gender, fear of injury/movement and depression were stronger predictors. Nevertheless, this last finding is in accordance with Reneman et al. (2003), but was not reported by others (Lackner and Carosella 1999; Lackner et al. 1996; Ljungquist et al. 2003).

Although the level of radiating pain was expected to influence physical capacity (Crombez et al. 1998; Simmonds 1999), we did not find any significant contribution, not even in the walking and stair climbing tasks.

According to our results, VO_2max only had a significant influence on the stair climbing and sit to stand task, but not on tasks such as five-minute walking and lifting, which are considered to be equally or even more aerobically demanding tasks. The lack of significant contribution to the lifting task is especially remarkable since Matheson et al. (2002) reported that VO_2max was responsible for 27% of the variance in their study, but they only controlled for spine strength and for none of the other variables. Our findings might be caused by the fact that we had to impute the level of VO_2max for 46 patients (20.8%), introducing some amount of uncertainty. Geisser et al. (2000a), also used imputation for 48 of the 133 CLBP patients (36.1%). They simply imputed the mean value of the calculated cardiovascular capacity of the respondents. Using that technique or not imputing at all, would have seriously weakened our results since the patients for whom no VO_2max could be calculated, scored significantly worse on several putative factors and

capacity tasks. By using the imputation method as described in the method section, we tried to approach the missing VO_2max data as accurately as possible.

Based on the fear-avoidance model, the lack of influence of the fear of injury/movement (fear-avoidance) as measured by the TSK on most capacity tasks is surprising. The capacity tasks were selected because they measure activities that are often hampered in patients with CLBP and it was expected that fear of injury/movement would significantly contribute to all capacity tasks. Only in the lifting task, a higher score on the TSK was significantly associated with a lower amount of lifting cycles, although this influence was not high since a patient having a 10-point higher TSK score performed only one lifting cycle less. This influence is much lower than described by Crombez et al. (1999), using a back holding task and controlling for several of the same variables (age, gender, radiating pain, catastrophizing), and in a lifting task described by Geisser et al. (2000a), controlling for cardiovascular capacity, pain and depression. In contrast, Reneman et al. (2003) found no significant contributing role of the TSK in a lifting task after adjusting for a gender, pain, duration of pain, and sick leave. This lack of influence on the other capacity tasks was not expected since the TSK is not measuring fear for a particular activity but a more global fear of injury/movement. On the other hand, the other tasks, although eliciting body movements, could have been too little fear evoking. Maybe it should be more specifically determined whether the activity being tested is fear evoking for that particular patient.

Pain catastrophizing and internal control of pain have often been mentioned as important mediating factors in the treatment of CLBP, meaning that they have to be changed in order to reduce the level of disability or better increase the level of activities (Burns et al. 2003; Jensen et al. 2001). According to the present study, these factors hardly contributed significantly to the level of performance. Only in the fast walking and the stair climbing task a higher level of internal control was associated with a better performance. Although pain catastrophizing, as expected on the basis of the fear-avoidance model (Vlaeyen and Linton 2000), showed a significant negative correlation with the number of steps climbed, the multiple regression analysis showed a reversed relationship: a higher level of pain catastrophizing resulted in higher number of stairs climbed. This is probably due to the high correlation between pain catastrophizing and the level of depression ($r = 0.72$) and fear of injury/movement ($r = 0.47$), and the fact that pain catastrophizing without correction for the other variables, explained only 1% of the variance in stair climbing.

The level of depression significantly contributed to four tasks which is in accordance with the fact that patients who get depressed are less active. Otherwise, it should be taken into account that the mean level of depression score was very low (10.1 ± 7.2). Only 12 patients had a score of ≥ 21 indicating a probable clinical depression (Geisser et al. 2000b). Therefore it can be debated whether the contribution of depression to capacity task scores in this population is really relevant. Otherwise, the results are in concordance with others (Kaplan et al. 1996; Novy et al. 2002). Only Geisser et al. (2000a) reported that, after correcting for several physical and psychological factors, the level of depression did not significantly contribute to the PILE-test.

Due to anthropometric differences between men and women, it was expected that men would perform better on most of the capacity tasks. Instead, men and women performed equally on the stair climbing and sit to stand task. Furthermore, it is remarkable that gender contributed significantly to the lifting task since the starting weight and amount of weight increment are higher for men (5.85 kg and 4.5 kg increment for each lifting cycle) than for women (3.6 kg and 2.25 kg increment for each lifting cycle). Apparently the construction of the PILE-test is not proportional and the increment of weight for males should have been higher or lower for women.
Although Mayer et al. (1988a) reported a low correlation between age and the PILE-test, age did not contribute significantly to the PILE-test or any other of the capacity tasks in our study.

By collecting data of 221 patients, we ensured sufficient power to perform the regression analyses with ten independent variables. It seems reasonable to state that the measurement of activity by using physical capacity tasks is only slightly influenced by the selected body functions and personal factors. Otherwise, the relatively low level of explained variance might also be due to the variance of the capacity tasks themselves and/or the questionnaires used. However, even taking these possible explanations into account, the level of explained variance is still quite low.

Our results were obtained from patients who agreed to participate in a RCT and therefore unconditional generalization to daily rehabilitation practice is not possible. Otherwise, the baseline characteristics (for example level of disability and fear of injury/movement) of the study population are similar to the population that is normally being treated in outpatient rehabilitation clinics (Crombez et al. 1999; Vlaeyen et al. 1995a; Vlaeyen et al. 1995b; VRIN/VRA 2000).

Although we examined ten different potentially influential factors on the capacity level of nonspecific CLBP patients, there are still several other factors which we did not assess in this study, some of which are difficult to test reliably (e.g. muscle strength and coordination) (Lariviere et al. 2002; Simmonds et al. 1998). One of the most interesting and promising factors seems to be functional self-efficacy as described by Lackner et al. (1996; 1999), and this should be investigated in future research.

Conclusions

Cardiovascular capacity, pain intensity and several personal factors like fear of injury/movement, cognitions, and depression had limited influence on the outcome of different, easy-to-administer capacity tasks in patients with CLBP. The level of radiation of pain to the leg, age and duration of complaints had no significant influence at all on any of these tasks. Due to anthropometric differences men outperformed women on most tasks. Apparently the influence of many personal, physical, but especially psychological factors on the selected capacity tasks is not high at all. This might indicate that these tasks are more objectively measuring physical capacity than expected. Otherwise, the influence of other, maybe even more influential psychological factors such as functional self-efficacy has not been investigated.

Acknowledgements

The authors would like to thank the rehabilitation centres Breda, Leijpark and Blixembosch and the many people who contributed to the successful completion of this study, particularly the participating patients, the rehabilitation centre contact persons, therapists, research assistants, Paula Rinkens for her help with the analyses and Harriët Wittink for critically reviewing the manuscript.

References

- Åstrand P, Rohdahl K. Textbook of work physiology, physiological bases of exercise. New York: McGraw-Hill, 1986, pp. 354-91.
- Baecke JA, Burema J, Frijters JE. A short questionnaire for the measurement of habitual physical activity in epidemiological studies. Am J Clin Nutr 1982;36:936-42.
- Beck AT, Rush AJ, Shaw BF, Emmery G. Cognitive therapy of depression. New York: Wiley and sons, 1979.
- Beck AT, Steer RA, Garbin MG. Psychometric properties of the Beck Depression Inventory: Twenty-five years of evaluation. Clin Psychol Rev 1988;8:77-100.
- Burns JW, Kubilus A, Bruehl S, Harden RN, Lofland K. Do changes in cognitive factors influence outcome following multidisciplinary treatment for chronic pain? A cross-lagged panel analysis. J Consult Clin Psychol 2003;71:81-91.
- Carlsson AM. Assessment of chronic pain. I. Aspects of the reliability and validity of the visual analog scale. Pain 1983;16:87-101.
- Crombez G, Vervaet L, Lysens R, Baeyens F, Eelen P. Avoidance and confrontation of painful, back-straining movements in chronic back pain patients. Behav Modif 1998;22:62-77.
- Crombez G, Vlaeyen JW, Heuts PH, Lysens R. Pain-related fear is more disabling than pain itself: evidence on the role of pain-related fear in chronic back pain disability. Pain 1999;80:329-39.
- Dawson-Saunders B, Trapp RG. Basic and clinical biostatistics. Norwalk: Appleton and Lange, 1998, pp. 207-29.
- Durnin JVGA, Womersley J. Body fat assessed from total body density and its estimation from skinfold thickness: measurements on 481 men and women aged 16 to 72 years. Br J Nutr 1974;32:77-97.
- Filho ITC, Simmonds MJ, Protas EJ, Jones S. Back pain, physical function, and estimates of aerobic capacity: what are the relationships among methods and measures? Am J Phys Med Rehabil 2002;81:913-20.
- Geisser ME, Haig AJ, Theisen ME. Activity avoidance and function in persons with chronic back pain. J Occup Rehabil 2000a;10:215-27.
- Geisser ME, Roth RS, Theisen ME, Robinson ME, Riley JL, 3rd. Negative affect, self-report of depressive symptoms, and clinical depression: relation to the experience of chronic pain. Clin J Pain 2000b;16:110-20.
- Gronblad M, Hurri H, Kouri JP. Relationships between spinal mobility, physical performance tests, pain intensity and disability assessments in chronic low back pain patients. Scand J Rehabil Med 1997;29:17-24.
- Harding VR, Williams AC, Richardson PH, Nicholas MK, Jackson JL, Richardson IH, Pither CE. The development of a battery of measures for assessing physical functioning of chronic pain patients. Pain 1994;58:367-75.
- Hodselmans AP, Jaegers SM, Goeken LN. Short-term outcomes of a back school program for chronic low back pain. Arch Phys Med Rehabil 2001;82:1099-105.
- Jensen MP, Turner JA, Romano JM. Changes in beliefs, catastrophizing, and coping are associated with improvement in multidisciplinary pain treatment. J Consult Clin Psychol 2001;69:655-62.
- Kaplan GM, Wurtele SK, Gillis D. Maximal effort during functional capacity evaluations: an examination of psychological factors. Arch Phys Med Rehabil 1996;77:161-4.
- Lackner JM, Carosella AM. The relative influence of perceived pain control, anxiety, and functional self efficacy on spinal function among patients with chronic low back pain. Spine 1999;24:2254-60.
- Lackner JM, Carosella AM, Feuerstein M. Pain expectancies, pain, and functional self-efficacy expectancies as determinants of disability in patients with chronic low back disorders. J Consult Clin Psychol 1996;64:212-20.
- Lariviere C, Arsenault B, Gravel D, Gagnon D, Loisel P, Vadeboncoeur R. Electromyographic assessment of back muscle weakness and muscle composition: reliability and validity issues. Arch Phys Med Rehabil 2002;83:1206-14.

References

- Lee CE, Simmonds MJ, Novy DM, Jones S. Self-reports and clinician-measured physical function among patients with low back pain: a comparison. Arch Phys Med Rehabil 2001;82:227-31.
- Ljungquist T, Harms Ringdahl K, Nygren A, Jensen I. Intra- and inter-rater reliability of an 11-test package for assessing dysfunction due to back or neck pain. Physiother Res Int 1999;4:214-32.
- Ljungquist T, Jensen IB, Nygren A, Harms Ringdahl K. Physical performance tests for people with long-term spinal pain: aspects of construct validity. J Rehabil Med 2003;35:69-75.
- Lygren H, Dragesund T, Joensen J, Ask T, Moe Nilssen R. Test-retest reliability of the Progressive Isoinertial Lifting Evaluation (PILE). Spine 2005;30:1070-4.
- Matheson LN, Leggett S, Mooney V, Schneider K, Mayer J. The contribution of aerobic fitness and back strength to lift capacity. Spine 2002;27:1208-12.
- Mayer TG, Barnes D, Kishino ND, Nichols G, Gatchel RJ, Mayer H, Mooney V. Progressive isoinertial lifting evaluation. I. A standardized protocol and normative database. Spine 1988a;13:993-7.
- Mayer TG, Barnes D, Nichols G, Kishino ND, Coval K, Piel B, Hoshino D, Gatchel RJ. Progressive isoinertial lifting evaluation. II. A comparison with isokinetic lifting in a disabled chronic low-back pain industrial population. Spine 1988b;13:998-1002.
- McCracken LM, Zayfert C, Gross RT. The Pain Anxiety Symptoms Scale: development and validation of a scale to measure fear of pain. Pain 1992;50:67-73.
- Melzack R, Katz J. The McGill Pain Questionnaire: appraisal and current status. In: DC Turk, R Melzack. Handbook of pain assessment. New York: Guilford Press, 1992, pp. 152-68.
- Nachemson AL. Newest knowledge of low back pain. A critical look. Clin Orthop 1992:8-20.
- Novy DM, Simmonds MJ, Lee CE. Physical performance tasks: what are the underlying constructs? Arch Phys Med Rehabil 2002;83:44-7.
- Novy DM, Simmonds MJ, Olson SL, Lee CE, Jones SC. Physical performance: differences in men and women with and without low back pain. Arch Phys Med Rehabil 1999;80:195-8.
- Parks KA, Crichton KS, Goldford RJ, McGill SM. A comparison of lumbar range of motion and functional ability scores in patients with low back pain: assessment for range of motion validity. Spine 2003;28:380-4.
- Pfingsten M, Hildebrandt J, Leibing E, Franz C, Saur P. Effectiveness of a multimodal treatment program for chronic low-back pain. Pain 1997;73:77-85.
- Rainville J, Hartigan C, Jouve C, Martinez E. The influence of intense exercise-based physical therapy program on back pain anticipated before and induced by physical activities. Spine J 2004;4:176-83.
- Reneman MF, Jorritsma W, Dijkstra SJ, Dijkstra PU. Relationship between kinesiophobia and performance in a functional capacity evaluation. J Occup Rehabil 2003;13:277-85.
- Reneman MF, Jorritsma W, Schellekens JM, Goeken LN. Concurrent validity of questionnaire and performance-based disability measurements in patients with chronic nonspecific low back pain. J Occup Rehabil 2002;12:119-29.
- Roelofs J, Goubert L, Peters ML, Vlaeyen JW, Crombez G. The Tampa Scale for Kinesiophobia: further examination of psychometric properties in patients with chronic low back pain and fibromyalgia. Eur J Pain 2004;8:495-502.
- Roland M, Morris R. A study of the natural history of back pain. Part I: development of a reliable and sensitive measure of disability in low-back pain. Spine 1983;8:141-4.
- Rubin DB. Multiple imputation for nonresponse in surveys. New York: Wiley, 1987, p. 78.

References

- Simmonds MJ. Physical function and physical performance in patients with pain: what are the measures and what do they mean? In: M Mitchell. Pain 1999- an updated review: Refresher course syllabus 9th World Congress on Pain. Seattle: IASP Press, 1999, pp. 127-36.
- Simmonds MJ, Olson SL, Jones S, Hussein T, Lee CE, Novy D, Radwan H. Psychometric characteristics and clinical usefulness of physical performance tests in patients with low back pain. Spine 1998;23:2412-21.
- Smeets RJ, Vlaeyen JW, Hidding A, Kester AD, van der Heijden GJ, van Geel AC, Knottnerus JA. Active rehabilitation for chronic low back pain: Cognitive-behavioral, physical, or both? First direct post-treatment results from a randomized controlled trial [ISRCTN22714229]. BMC Musculoskelet Disord 2006a;7:5.
- Smeets RJ, Wittink H, Hidding A, Knottnerus JA. Do patients with chronic low back pain have a lower level of aerobic fitness than healthy controls? Are pain, disability, fear of injury, working status, or level of leisure time activity associated with the difference in aerobic fitness level? Spine 2006b;31:90-7.
- van Tulder MW, Koes BW, Bouter LM. A cost-of-illness study of back pain in The Netherlands. Pain 1995;62:233-40.
- Verbunt JA, Seelen HA, Vlaeyen JWS, van der Heijden GJ, Heuts PH, Pons K, Knottnerus JA. Disuse and deconditioning in chronic low back pain: concepts and hypothesis on contributing mechanisms. Eur J Pain 2003;7:9-21.
- Vlaeyen JWS, Geurts SM, Kole-Snijders AMJ, Schuerman JA, Groenman NH, van Eek H. What do chronic pain patients think of their pain? Towards a pain cognition questionnaire. Br J Clin Psychol 1990;29:383-94.
- Vlaeyen JWS, Kole-Snijders AM, Boeren RG, van Eek H. Fear of movement/(re)injury in chronic low back pain and its relation to behavioral performance. Pain 1995a;62:363-72.
- Vlaeyen JWS, Kole-Snijders AM, Rotteveel AM, Ruesink R, Heuts PHTG. The role of fear of movement/(re)injury in pain disability. J Occup Rehabil 1995b;5:235-52.
- Vlaeyen JWS, Linton SJ. Fear-avoidance and its consequences in chronic musculoskeletal pain: a state of the art. Pain 2000;85:317-32.
- Vlaeyen JWS, van Breukelen G, Nooyen-Haazen IWCJ, Stomp-van den Berg SGM, Kole-Snijders AMJ. De Pijn Cognitie Lijst. Maastricht: Pijn Kennis Centrum academisch ziekenhuis Maastricht, 2003.
- VRIN/VRA. Productdefiniëring Revalidatiezorg: registratiepilot revalidatie-instellingen januari 1999 - juli 2000. Utrecht: VRIN, 2000.
- Watson PJ. Non-psychological determinants of physical performance in musculoskeletal pain. In: M Mitchell. Pain 1999- an updated review: Refresher course syllabus 9th World Congress on Pain. Seattle: IASP Press, 1999, pp. 153-8.
- WHO. The International Classification of Functioning, Disability and Health (ICF). Geneva: WHO, 2001.
- Wittink H. Functional capacity testing in patients with chronic pain. Clin J Pain 2005;21:197-9.

General discussion.

Chapter 10

General discussion

The aim of our study was multifarious; first we wanted to find out whether signs of physical deconditioning (loss of aerobic capacity and muscular atrophy and fibre changes) are present in chronic low back pain (CLBP) patients compared with matched healthy subjects and whether treatments specifically aimed at restoring these symptoms are effective. The main aim of our study was to assess the effectiveness of model-based treatments and whether a combination treatment is more effective than the single treatments. Finally, we wanted to test the reliability and usability of several easy-to-use physical performance/capacity tasks and to determine the association between several personal variables and body functions and the level of performance of CLBP patients on these tasks.
The ambition of this general discussion is to give an overview of the findings and consequences of our studies. First a summary of the main findings will be presented, followed by methodological issues, implications for clinical practice and implementation aspects, and finally we will provide several recommendations for future research.

Summary of findings

Do patients with CLBP have signs of physical deconditioning?

Based on our review, there is conflicting evidence that patients with CLBP have a lower aerobic capacity than healthy controls. Just recently, Brox et al. (2005) reported a lower aerobic capacity in patients with low back pain (LBP) who were sick-listed for 8-12 weeks, but not in CLBP patients waiting for lumbar instrumented fusion, compared with healthy persons matched for gender and age. Unfortunately, they only used four age cohorts, and did not match for the level of sport activity.
The results of our own research (chapter 3), showed that CLBP patients, and especially men, have a significantly lower level of aerobic capacity compared with healthy persons matched for age, gender and level of sport activity. The suggestion that this lower aerobic capacity level might be caused by the level of work activity (Verbunt et al. 2003), could not be confirmed, not even when we included the level of household activities. Also, the levels of disability and pain were not associated with the reduction of VO_2max. However, since our findings are based on cross-sectional research, the question whether the loss of VO_2max is a cause or effect of CLBP, remains unanswered.
Our review also showed that there is only limited evidence for the idea of the wasting of low back muscles, and the deeper multi-segmental multifidus muscles in particular, as these muscles are responsible for segmental stability.

Is treatment aimed at restoring cardiovascular and/or muscle strength and endurance effective?

On the basis of our review (chapter 3) no reliable answer can be given to the question whether cardiovascular training is effective, although it is often incorporated in rehabilitation treatments. There seems to be moderate evidence that specific training of the multifidus muscles of sufficient intensity and duration to fulfill the physical training principles, is more effective than less intensive strengthening exercises for CLBP patients with low to moderate levels of disability.

After we finished our literature search (December 2004), two interesting and relevant papers have been published. Hayden et al. (2005b) published a systematic review, showing that supervised strengthening exercises of high-dose or high-intensity and adjusted to the individual need, are indeed more effective in reducing disability than other exercise treatments. Koumantakis et al. (2005) performed a RCT in patients with recurrent LBP with a duration of six months or less and showed that the addition of stabilization enhancing exercises to lumbar and abdominal strengthening exercises of sufficient intensity did not provide additional benefit. Nevertheless, both treatments showed clinically relevant improvements in disability and pain, three months post-treatment. These two studies further confirm our earlier suggestion that it is very relevant to examine the effectiveness of intensive training of the multifidus muscles in patients with CLBP of longer duration and more severe level of perceived disability.

The recommendations presented in our review were taken into account while designing our trial. We investigated the effectiveness of the dynamic-static exercise protocol for back extensor muscles as described by Danneels et al. (2001) in a population of patients with a much higher level of perceived disability using a methodologically stronger design. Furthermore, we included exercises that were grounded on sound theoretical principles concerning exercise physiology and dose response. We not only reported on the intensity, but also the duration, frequency and progression of the aerobic training as recommended by Protas et al. (1996).
The results of our trial showed that the strength and endurance enhancing exercises for the multifidus combined with aerobic capacity training are indeed effective, even one year post-treatment. This physical treatment (APT) showed a trend to be more effective than the combination treatment (CT). Only patients with a high level of disability at the start of treatment reported significant more satisfaction immediately post-treatment in CT compared with APT.

Is the combination treatment more effective than the two single component treatments?

The finding that CT is not more effective than the single treatments is surprising and puzzling. Several possible explanations have already been discussed in chapters 4 and 5, and will be briefly summarized. First, the high effort required from patients while they were attending the outpatient combination treatment resulting in a lower compliance rate in comparison to the single treatments. Second, the lower total treatment intensity than the 100 hours of multidisciplinary treatment with functional restoration showing additional effect when compared with non-multidisciplinary treatment (Guzman et al. 2002). However, the conclusion of the Cochrane meta-analysis that treatments of more than 100 hours are more effective, might also be attributable to the fact that those attending such an intensive program were very motivated, their employer as well as insurance company advised them to participate, and last but not least, most programs were inpatient programs, meaning that the patients were invited to stay for at least five days in a row in the training facility.
On the other hand, it could not be ruled out that the mixture of APT and CBT even had an counteractive effect. The increase of exercise load in APT was based on training physiology using the 1-Repetition Maximum for each strengthening exercise and maximum heart rate for the cardiovascular training. The increase of activity in CBT was

based on a time-contingent increase of patient relevant activities irrespective of the level of pain. We used a clear and, in our opinion, logical treatment rationale for CT. The patients were told that they first have to increase their aerobic capacity, muscle strength and endurance and then start to increase important activities. Meanwhile, they learn how to improve their problem solving skills in order to cope better with problems and stress. In this way they save energy that can be used to exercise and cope with possible pain flare-ups. Nevertheless, this could have confused patients, resulting in insufficient generalization of the learned principles during follow-up and finally obscuring the synergetic effect of both treatments.
Furthermore, APT was more effective, especially at one year post-treatment, than originally expected (Hayden et al. 2005a; Liddle et al. 2004). APT seems to have resulted in both attitudinal and behavioral changes similar to those achieved with CBT, making the addition of CBT to CT probably redundant. This is further enhanced by the results presented in chapter 6, showing that APT, although not using cognitive-behavioral techniques, resulted in a similar decrease of pain catastrophizing which significantly mediated outcome, as is the case in both CBT and CT (Smeets et al. 2006).
At the same time CBT might not have been strong enough in reducing catastrophizing beliefs in the highly fearful patients. Exposure-in-vivo has shown to be one of the strongest CBT interventions in reducing pain-related fear and pain catastrophizing, and not problem solving training (Boersma et al. 2004; de Jong et al. 2005; Vlaeyen et al. 2002).

These findings urge us to further investigate whether several treatment modalities should be offered simultaneously since a synergetic effect of physical and cognitive-behavioral treatment could not be proven in this study. On the contrary, there is a trend in the direction of adverse effects when combining both approaches.

Did the treatment elements work as they were intended to?
The aerobic capacity (VO_2max) increased significantly in those patients who were treated with the physical training (APT and CT), and not in CBT, indicating that the cardiovascular training had a sufficient level of intensity. No reliable measure to check the strengthening exercises was available, but based on the analysis of the records regarding the increase of workload during training, we concluded that the patients attending the strengthening exercises showed improved muscle strength and endurance, or at least increased performance.
Analysis of the patient records indicated that in most patients participating in the graded activity training, a clear increase of patient relevant activities was achieved.
As we already discussed in chapter 4, the patients who received problem solving training did not show a larger reduction of their negative problem solving orientation than the patients who did not receive this training. The lack of additional improvement in negative problem solving after completion of problem solving training was also reported by van den Hout et al. (2003), although they reported improvement of outcome on the long term. We also expected that the addition of problem solving training would lead to better long term outcome since patients would have learned how to cope with their problems and also would have learned to keep on using the graded activity technique to increase or

sustain their level of activity. Unfortunately, we did not assess the negative problem orientation at the 6 and 12 month follow-up. Nevertheless, the results of CBT sustained and even improved a bit more until one year post-treatment, indicating that a number of sustained cognitive and/or behavioral changes seemes to have occurred. We already discussed that the problem solving training probably might have exerted its effect otherwise, and for future research it would be useful to assess other potential processes, for example self-efficacy.

In addition to our own research, an independent researcher made several observations of active treatment sessions (on-site, audio and video recordings) and interviewed patients (focus groups), as well as therapists involved in the trial (focus groups and personal interviews). After systematic analysis of the observations and interviews, he concluded that the therapists delivering the physical treatment did not use additional decatastrophizing methods, and the therapist responsible for the graded activity and problem solving training did not incorporate physiologic training principles (Smeets et al. 2006). However, since only one independent researcher did all the observations, interviewing, as well as the interpretation, no data regarding reliability of his conclusions are available.

Association of the physical training treatment outcome: Is improvement of VO_2max important?
Patients who underwent the aerobic training showed a significant and clinically relevant improvement of their VO_2max. On the other hand, patients who were only offered CBT treatment showed no increase of VO_2max, although the level of disability decreased equally immediately post-treatment. How can this be explained and what does it mean?

Our review shows that there is scientific evidence for the use of reconditioning treatment when compared with doing nothing or passive treatments, but it also shows that this treatment is not more effective than other active treatments. It is hypothesized that the reactivation itself is the common and most important factor in the reduction of CLBP disability. Regular exercise, regardless of age, has the potential to favorably modify the psychological state of men and women. These modifications include an improved mood, decreased level of mild to moderate depression, change in state anxiety, reduction in neuroticism and improved self-esteem, self-concept, and general perception of personal worth (Mayer and Gatchel 1988; Simmonds et al. 1996). This mood elevating effect seems to be mediated by reducing psychosocial stress responses (Crews and Landers 1987) and the participating in a social event, rather than just the improved physical fitness (Thirlaway and Benton 1992). Perhaps this might be a consequence of modification in pain perception or simply because a feeling of well being and accomplishment associated with achieving an increase in physical fitness or socialization as well as their distractional properties (Abenhaim et al. 2000; Ferrell et al. 1997; Gurevich et al. 1994; Mannion et al. 1999). Other positive effects of increased physical fitness or activity participation might be: a) An increased sense of control that gives the patient a positive outlook, preventing the patient from taking the sick role (Campello et al. 1996). b) A decreased level of fatigue (Nutter 1988; Protas 1996) and pain behaviors (Fordyce 1976). For example, Mannion et al. (2001) showed that after a physical treatment, fear-avoidance beliefs about physical activity were significantly reduced in CLBP patients although no specific cognitive treatment was used to alter these beliefs.

These findings are in line with those of our own research described in chapter 6, showing that pain catastrophizing significantly mediates changes in perceived disability and pain severity in all three treatment conditions. These findings suggest that an increase in aerobic capacity might not be a necessary condition to obtain a significant and clinically relevant improvement of outcome (e.g. disability) in patients with CLBP.

Also striking was the finding that the effects of APT lasted at least one year post-treatment. Meta-analyses (Hayden et al. 2005a; Liddle et al. 2004) showed that the effect of exercise therapy on pain and disability diminishes slightly, 6 to 12 months post-treatment. Our long-term results (chapter 5) showed the same for pain but not for disability, which slightly improved in APT. This persistence of effect might also be mediated by the reduction of pain catastrophizing in APT despite the absence of cognitive-behavioral treatment (Smeets et al. 2006), and is in line with the suggestion that not the improved physical fitness but cognitive processes are responsible for the effectiveness of exercise training (Helmhout et al. 2004; Mannion et al. 1999).

How does each treatment exert its effect?
A previous study, using data from a Dutch RCT on several cognitive-behavioral treatments, showed that changes in several outcome measures were mediated by both pain catastrophizing and internal control of pain (Spinhoven et al. 2004). We were able to only partially corroborate these findings. In our study pain catastrophizing but not internal control of pain, mediated the outcome. One of the reasons might be that the previous study included an intervention that was aimed specifically at improving pain coping, while this was not the case in our study. Unfortunately, we could not carry out a cross-lagged analysis (Burns et al. 2003), but based on the available evidence, it seems pretty evident that all active treatments caused pain catastrophizing to diminish, which then resulted in improvement of several outcome measures, including disability. We advise that future treatments should try to focus even more on the systematic reduction of pain catastrophizing in patients with CLBP.

Is treatment expectancy and credibility important regarding the final effectiveness of treatment?
Treatment expectancy as well as credibility appeared to significantly influence several outcome measures for disability and satisfaction, beyond that predicted by pre-treatment outcome score, several demographic and disease specific variables, treatment centre and irrespective of the treatment offered. For global perceived effect, differential associations for treatments were found. Treatment expectancy showed a significant, moderate to high association in APT, and treatment credibility showed a significant, moderate association in CT.
Although the predictive power was mostly modest, these results underscore the importance of treatment expectancy and credibility for the outcome of different active treatments for CLBP. All three active treatments showed a similar level of treatment credibility and expectancy after the treatment rationale had been explained. We think that the explication of a clear treatment rationale for each separate treatment contributed to the moderate to high level of credibility of each of the treatments. Furthermore, treatment credibility and expectancy should be taken in account in future research as well as in clinical practice.

Usability of physical performance/capacity tasks
According to the ICF, the formerly used term physical performance task should be changed to physical capacity tasks (WHO 2001). Physical capacity tasks measure the highest probable level of functioning on these tasks of a person at a given moment and preferably in a standardized environment. To prevent further confusion, we will use both terms in this discussion. Despite a good to very high reliability, the limits of agreement for the lifting task (PILE) were that high that this task should better not be used in clinical practice. One explanation for the higher limits of agreement might be the fact that this task was performed at the end of the assessment, causing fatigue to influence the variability. On the other hand, Lygren et al. (2005) found the same limits of agreement and, besides the lifting task, they did not use other performance/capacity tasks.
Despite the moderate to good clinical usefulness of the other five performance/capacity tasks, for several patients a ceiling effect might have been applicable and for others the tasks were not specific enough as was already discussed in chapter 4. We propose to develop more patient specific performance/capacity tasks (see also Clinical implications and suggestions for clinically based research).

Influential factors of physical performance/capacity tasks
Although several authors suggested that one should evaluate pain-related fear (Verbunt et al. 2003; Vlaeyen et al. 1995; Vlaeyen and Linton 2000) and other psychological contributing factors (Curtis et al. 1994; Kaplan et al. 1996; Lackner and Carosella 1999; Lackner et al. 1996; Lee et al. 2001; Mayer et al. 1998; Moseley 2004; Novy et al. 2002) during the performance/capacity tasks in CLBP, we hardly found any cross-sectional associations between fear-avoidance or other psychological factors and the scores on physical performance/capacity tasks. No association between fear-avoidance and the performance on a modified submaximal bicycle test was found, indicating the absence of evidence of the influential role of pain-related fear on completing the test and the reached VO_2max. Only for the graded lifting task, higher pain-related fear was significantly, although only modestly, associated with a lower amount of lifting cycles. Apparently the influence of many personal, physical, but especially psychological factors on several easy-to-administer performance/capacity tasks that resemble daily life activities is not high at all. This indicates that these physical performance/capacity tasks are more objectively measuring physical performance/capacity of a person than we originally expected.

Role of fear-avoidance as an important variable for treatment outcome
The baseline pain-related fear appeared to be a significant prognostic variable for the primary outcome disability of all active treatments. It was also a significant contributor to treatment credibility and expectancy, which on their turn significantly influenced several outcome measures including disability. This indicates that patients having a higher level of pain-related fear not only expect less from an active treatment, but also tend to improve less when offered such an active treatment. Future research should investigate whether treatments more specifically aimed at reducing pain-related fear are worth the effort.

Methodological issues

Selection of patients

Three months after the start of the trial, we checked whether consultants in rehabilitation medicine responsible for including patients for our trial did use the predefined trial selection criteria. Of all patients who were not admitted to the trial, but who were offered regular rehabilitation treatment, the medical charts were collected. Next, the main researcher and the responsible physician discussed why he/she did not ask the patient to participate in the trial. It appeared that most physicians strictly followed the criteria. Nevertheless, two physicians used their own criteria, especially when they thought that psychosocial factors seemed to play an important role in the maintenance of disability. They preferred to offer such patients a biopsychosocial observation/intervention. The main researcher once again stressed the importance and necessity of not using one's own selection criteria, in order to be able to reliably assess the effectiveness of model-based treatments irrespective of additional, empirical selection criteria. After these sessions, we took a random sample, but did not encounter any problems. Based on these checks, we are quite confident that the participating patients were not selected otherwise than by using our trial selection criteria.

Cross-sectional associations

Cross-sectional associations were calculated for aerobic fitness (VO_2max) and putative influential variables (chapter 3), the scores on physical performance/capacity tasks and several putative influential factors (chapter 9), and finally the mediating role of pain catastrophizing and internal control of pain on several important outcome measures (chapter 6). Due to the trial design no longitudinal data collection was possible, although additional measurement of mediating variables and outcome measures during the treatment (for example half way the treatment) would have strengthened our results. Based on the high level of strain already put on the patients by the planned assessments (questionnaires and performance/capacity task pre-treatment, post-treatment and 6 and 12 months post-treatment, as well as weekly diaries), we decided not to add another assessment moment during the treatment.

Methods of measurement

Use of performance/capacity tasks

Although we incorporated 'more objective' performance/capacity tasks as secondary outcome measures, we encountered several logistic problems. Despite the many efforts made to facilitate patients to come to the rehabilitation centres for their 6 and 12 month follow-up assessment, the number of missing data was quite high. This makes interpretation of these outcome measures difficult, and conclusions can only be drawn with many reservations. Otherwise, this was one of the first trials in CLBP that reported on such outcome measures. Only one study using similar performance/capacity tasks, reported the total number of patients that dropped out during follow-up assessment, but not the results of the tasks itself. The number of dropouts appeared to be comparable to our study (Ljungquist et al. 2003). The requirement for patients to travel to the treatment centre in order to perform the

performance/capacity tasks appeared to be a big disadvantage compared with studies only using postal questionnaires. Furthermore, the reasons why patients did not show up for assessment were not solely caused by negative treatment results or increase of pain because of the performance/capacity tasks. An equal number of patients were not willing to stay away from work or household duties, or reported having no or only minor complaints anymore. Future studies should take these problems of missing data due to logistic problems and positive outcome into account, and for example visit patients at home or workplace.

Use of the modified Åstrand submaximal test

In daily practice we noticed that many CLBP patients could not finish the traditional Åstrand test because the initial workload was too high. Since reducing the workload is not allowed once the Åstrand test is started, we expected that we could not calculate the VO_2max for many patients, which would reduce the power of our study. However, a direct comparison of the modified Åstrand submaximal testing with the traditional test has still to be done. Furthermore, validation of this modified test against a maximal test has to be explored. Until now, only one study compared the modified test with maximal testing in healthy people (Hodselmans et al. 2001). However, maximal testing in CLBP patients will be extremely difficult, since Wittink et al. (2000) found that none of their CLBP patients at the end of maximal testing reached heart rate values indicating maximal effort. Therefore, future researchers should consider using the modified Åstrand submaximal testing instead of the maximal test.

Use of a new subscale of the Baecke Activity Questionnaire

We developed an alternative work index of the Baecke Activity Questionnaire. In the original questionnaire, household activities were not incorporated as contributing activities for the work index or any other activity index at all. Furthermore, the number of hours a patient worked during a week was not taken into account while calculating the work index. Therefore we used the same calculation method used for the sport index of the Baecke Activity Questionnaire, and multiplied the work intensity score by the hours of work and household activities in order to adjust for the amount of hours being active. This alternative work index seems to be more precisely measuring the level of activity during work and household, but further validation is needed.

Use of the Dutch version of the Credibility/Expectancy Questionnaire

We translated the Credibility/Expectancy Questionnaire to Dutch. By translating the words 'think' and 'feel' into the Dutch words 'denkt' en 'voelt', we were not sure whether credibility and expectancy could be accurately measured. The sub-scales treatment credibility and expectancy showed a moderate to high level of correlation. Nevertheless, with a confirmatory factor analysis, the original two factor structure could be confirmed and both sub-scales showed high internal consistency. Furthermore, we found proof for a low to moderate predictive validity of both constructs in a CLBP population treated with different types of rehabilitation treatment. Despite these findings, still more research on test-retest reliability, discriminative and predictive validity, as well as normative values in CLBP patients should be

established for the Dutch version. Furthermore, the scoring of all six items should be changed by using an identical scoring system for all items.

Power

The power of the study was originally calculated to detect a significant and clinically relevant difference regarding disability (2.5 point on the RDQ) between the three active treatments and the waiting list group (Roland and Fairbank 2000). As a consequence the found differences between the CT and the two single treatments (being smaller than the 2.5 points) did not reach significance, although these could be clinically relevant. No consensus is available on the magnitude of the clinically important difference between active treatments, although in ours and others opinion a ≥ 1 point additional reduction of the RDQ might be worth the effort (Moffett et al. 1999; UK BEAM Team 2004). Despite the possible lack of power we did find a trend for disability, and even a significant difference regarding self-perceived effect in favor of the single treatments compared with CT. As we already discussed in chapter 5, opposite to our hypothesis these results seem to indicate that APT and CBT are more promising than CT. In order to improve the power, an additional analysis in which both single treatments taken together are compared with CT could be done. However, since we did not use a factorial design, we did not perform this analysis.

Statistical problems

In the studies concerning the short and long term results (chapters 4 and 5) we primarily did not impute missing outcome data assessed by questionnaire, mainly because the number of missing data was rather low. The random coefficient analyses applied for the long term results already takes the results of the outcome measure on previous assessments into account, making imputation even less necessary (Twisk 2003). We did however, perform secondary analyses. We replaced the missing data by the last value carried forward and the worst case method, and reported whether important changes occurred.

The number of missing data on the performance/capacity tasks at the 6 and 12 month follow-up was so high that imputation was no option at all.

In the study concerning the influential role of personal, physical and psychological variables on the performance/capacity tasks, we were faced with the large amount of missing VO_2max data. Since the patients of whom the $VO_{2max\ data}$ were missing differed significantly on several other important independent and dependent variables (performance/capacity task outcome), we used the multiple imputation technique as described in detail in chapter 9.

Clinical implications and suggestions for clinically based research

What does this research add to the body of contemporary knowledge?

Our research contributed a small piece to the big puzzle of CLBP as it showed that model-based treatments are more effective than a waiting list control group. The effect sizes of all model-based treatments are comparable to other active

treatment trials, and even trials that used more extensive selection procedures (Guzman et al. 2002; Hayden et al. 2005a; Hayden et al. 2005b; Ostelo et al. 2005). Notable of the current study are:

a) The finding that the combination treatment is not more effective than the two single treatments immediately post-treatment. At follow-up the single treatments even showed a trend of being more effective by maintaining the positive results up to one year post-treatment as compared with the combination treatment. Based on these findings it can be concluded that we do not need to offer patients a treatment with all active treatment modalities studied in this trial.
b) The reduction of disability is mediated by pain catastrophizing (and not by internal control of pain), not only in the cognitive-behavioral condition, but in the physical treatment as well.
c) Treatment expectancy and credibility, are significant, although mostly modest predictors of the outcome of rehabilitation treatment.

Nevertheless, these conclusions and possible implications for daily rehabilitation practice should be regarded with caution since the selection procedure and treatments offered differ from daily practice, and an extensive cost-effectiveness analysis is still to be performed.

Characteristics of treatment program compared with daily practice (external validity)

Selection of participants

All patients were referred by general practitioners or medical specialists for disabling CLBP. Only patients for whom they thought that rehabilitation treatment would be meaningful, were referred. No patients were recruited by advertisements in newspapers or other media. Thus our study sample can be considered representative for the population usually seen at rehabilitation departments, at least in The Netherlands. Otherwise, the selection method used differed from normal clinical practice. Patients were included once the consultant in rehabilitation medicine decided that the patient was able and ready to participate in rehabilitation treatment. In daily clinical practice, after the rehabilitation specialist has seen the patient once or twice, mostly additional screening is used to decide whether and what treatment the patient should be offered. Sometimes a psychologist or social worker sees the patient for further interviewing, but mostly a multidisciplinary team consisting of physiotherapist, occupational therapist, social worker and/or psychologist, further assesses the patient before deciding what treatment the patient will be offered.

By following this procedure, it is imaginable that eventually fewer patients will be offered rehabilitation treatment, making comparison of the trial results with daily practice complicated. Unfortunately, clear criteria upon which clinicians and their multidisciplinary teams decide whether and what treatment should be offered are still not available or insufficiently reliable. Therefore, we urge clinicians and their rehabilitation teams to document their assessment procedures, the process of decision making and the final decision, more clearly.

Treatment
In daily practice, treatment modalities are selected based on the observation results as described above. This implicates that, at least in The Netherlands, patients are offered individual or group treatment which can contain biomedical (cardiovascular training, manual therapy, muscle strengthening and so on), cognitive-behavioral treatment elements varying from pacing, graded activity, graded exposure, counseling, Rational Emotive Therapy, assertiveness training, problem solving, counseling regarding relationships and work-related problems, body awareness, all kinds of relaxation training. Furthermore, information about anatomy, pain mechanisms and ergonomics, ergonomic advise, work place visits and job related functional training and so on, can be applied. However, clear definitions and selection criteria for all the abovementioned treatment modalities are lacking (Koke et al. 2005). Nevertheless it is obvious that when biomedical, cognitive as well as behavioral elements are incorporated in varying ways, the interaction and tuning of all these treatment modalities are vital for the final results.
For the patients being treated in the trial no such observation period and evolving selection of treatment modalities were used. The treatments were randomly offered and followed strict treatment protocols without prior assessment of the patient problems. The treatments were exclusively based on the deconditioning model (physical training), the cognitive-behavioral model (graded activity and problem solving training) or the biopsychosocial model (combination of physical and cognitive-behavioral treatment). Within these protocols, no other treatment modalities such as spouse training, work-related training and so on could be incorporated. Nevertheless, the trial treatments were individually adjusted. In the physical training, the training level was assessed by prior testing of the aerobic capacity and muscle strength (70% of the 1-Repetition Maximum). In the graded activity training, the activities that were time-contingently increased were selected on an individual basis. Furthermore, in the problem solving training, individual problems could be identified after which solutions were generated, selected, tested and evaluated by the patient himself. Summarizing, although the treatments offered in the trial were individually adjustable, less treatment modalities selected on the basis of patient specific identified problem areas could be offered than in daily rehabilitation practice.

Evaluation
A direct comparison between the trial results and daily practice is not possible due to the lack of sufficient reliable data on outcome measures of current rehabilitation treatment in the Netherlands. Only one Dutch trial described the short term effectiveness of a regular multidisciplinary rehabilitation treatment for moderately disabled CLBP patients. Vollenbroek et al. (2004) found a two-point reduction on the RDQ in 35% of the patients in the multimodal treatment group, but also in 23% of the waiting list patients immediately post-treatment. Our trial showed at least similar post-treatment figures: 44% (physical treatment), 47% (cognitive-behavioral treatment), 55% (combination treatment) and 24% (waiting list). Furthermore, we should not forget that for the trial population the intention-to-treat analysis was used, meaning that every patient, even when he/she did not show up or dropped out, contributed to the final results of the treatment he/she was randomized to. When we as clinicians are asked to discuss the results of our current rehabilitation treatment, we often tend to forget about the patients who did not show up or

quitted treatment before the intended end of treatment. The follow-up contacts with the rehabilitation physician mostly last until half a year post-treatment, and no reliable outcome data gathered during these visits are available. Furthermore, social desirable answers might overestimate the results. Taken all these reasons together, the expert-based opinion about the effectiveness of the current rehabilitation care might not be completely correct. To overcome this problem we urge to implement outcome assessment with a one year follow-up as a daily routine in our clinical practice. This assessment should preferably be done by blinded assistants, or at least by postal questionnaires in order to prevent social desirable behavior of the patient.

Group versus individual treatment
Group treatment can be an efficient tool and a stimulating learning environment, especially with cognitive-behavioral treatments (Keefe et al. 1996). The transfer of knowledge can be done more efficiently and patients can learn from each other especially by exchanging opinions, feedback and experiences. Furthermore, this kind of training is more cost reducing since you only need one or two therapists for four patients. Otherwise, this group treatment can also be a pitfall. Several times during the trial, the therapists reported that patients did not like the problem solving training and started to disorganize the training itself or even quitted the training causing the group to disintegrate. We therefore controlled for the cluster of four patients, the patients were randomized to (multilevel linear regression analyses), but did not find a significant influence on any of the outcome measures. Nevertheless, the therapists often mentioned negative or positive aspects of the group training which should be taken into account when we decide to provide group training in daily practice.

Implications for daily practice
For daily practice we should try to find out whether the extensive package we often offer to patients is really necessary. This is especially important since the long term trial results indicate that simply offering all treatment elements together has no additional value at all. Despite our careful design of CT, in which we had put much effort to perfect the treatment rationale and to ensure a logical building up of the treatment by first starting with the physical training and after four weeks starting the graded activity training, this treatment even showed a significant less positive result regarding self-perceived improvement, compared with the single treatments one year post-treatment.

Starting the treatment with the explanation of the treatment rationale was positively judged by the therapists and patients. In each rationale the necessity for the patient to take his/her own responsibility for his/her treatment, to increase his/her activity level and to make plans to keep on active in the future (generalization) was stressed. Despite prior unsuccessful treatment(s), patients found the treatment sufficiently credible. Each rationale stresses that rehabilitation treatment is a kind of self-management course. This is also in agreement with the current opinion that rehabilitation treatment should be more patient-centreed with goal-setting and goal-attainment. We think that a clear treatment rationale which is explained to the patient in depth, is obligatory to improve treatment outcome, however,

additional research is necessary to confirm this hypothesis. For daily practice we advice to use a clear treatment rationale and to check the treatment credibility and expectancy. In case of low treatment credibility and expectancy it seems better to first invest in motivating the patient for this kind of rehabilitation treatment, e.g. by using motivational interviewing and counseling before the start of treatment (Jensen 1996; Jensen et al. 2003; Kerns et al. 2005).

Attitude of physicians and therapists

During the presentation of the assessment protocol and especially the physical training protocol, many physicians and therapists expressed their concern that patients would drop out because of the high intensity of the assessment procedure and treatments offered. Several therapists even thought that the physical training would cause negative side-effects. The research team itself had some doubts about the training intensity, but since the treatment had to be based on physiologic training principles we decided not to change the protocols. However, after several pilot assessments and treatments, most patients performed the submaximal bicycle tests, performance/capacity tasks and physical training without any side-effects. Furthermore, we developed clear criteria for adjustment of the treatment protocol, which ensured that therapists did not change the treatment intensity and training load based on their empirical experience and attitudes/beliefs regarding the exercise load or level of activities CLBP patients can cope with. During the trial we also did not encounter serious side-effects, and actually, in the physical training group the drop-out rate was the lowest. This indicates that the patients with CLBP are capable of performing better than we as clinicians/therapists think is possible or safe. For clinical practice, we advise clinicians and therapists not to be too careful in asking patients to be active or increase their level of activity. Many of us, who participated in the trial, changed their attitudes/beliefs about the intensity of training either by physical treatment or increasing activity by using cognitive-behavioral training principles, and assessment procedures such as submaximal testing and performance/capacity tasks.

Role of satisfaction for choosing treatment modalities

Based on our short-term trial results it might be better to provide a physical treatment to patients who have a low level of disability. For patients with moderate to high level of disability, cognitive-behavioral treatment elements seem to result in higher overall treatment satisfaction. Otherwise, it should be kept in mind that treatment satisfaction was only a secondary outcome measure and could have been influenced by other factors such as interaction between patients and therapists, although we did not find any significant contribution of the cluster of patients the patients were training in. In daily practice we should also use satisfaction as an outcome measure and collaborate with patients to find out whether this construct can be more elaborated than the rather simple visual analogue scale that we used in our trial.

Performance/capacity tasks

Based on the findings presented in chapters 8 and 9, we think that the five-minute walking task and the one-minute stair climbing task immediately can be used in clinical practice. Furthermore, several patients reported ceiling effects

and others reported that for example, the walking distance and not the speed as assessed with the five-minute walking task, was their main problem. This illustrates the need for the development of more patient specific performance/capacity tasks that should not be time-consuming, easy-to-administer, and maybe even assessed at their home or work place, in order to decrease the number of non-responders.

Implementation issues

Implementation of treatments

Before starting to implement the most effective treatments, an extensive cost-effectiveness analysis is still to be done before finally concluding which treatment is the best choice. However, based on the amount of treatment hours and the long-term results we already can state that CBT and probably also APT are worth the effort, and can be used in daily care. It seems not logical to combine both treatments since this combination did not show additional effectiveness. We think that this study shows that the graded activity technique is definitely worth implementing in daily care. This is also in agreement with the trend of more patient-centreed treatment. Protocols are available and clinicians as well as therapists can be quite easily and quickly trained in delivering this treatment. The problem solving training was designed by a researcher/psychologist and applied by psychologists and social workers with several years of experience in treating CLBP disability. During the pilot-treatments, we encountered several practical problems and the protocol was not easy to follow. We therefore adjusted the lay-out of the protocol and designed easy-to-use audiovisual materials. Most therapists who worked with the protocol feel that this treatment is valuable for daily practice, but should only be offered to patients who really encounter difficulties in handling their problems in daily life and are sufficiently motivated. Several patients participating in the trial said that they already could solve their problems quite good or experienced no problems at all. Others simply could not understand the rationale of the training. Despite these observations, we think that this problem solving training has to be considered for its usefulness in clinical practice.

Clinicians' and therapists' attitudes

As described earlier, we and others adjusted our attitudes regarding the exercise load, level of activity and treatment frequency most CLBP patients can handle without serious side-effects. This also led to changes in our daily practice since we started to use more intensive assessment tools (submaximal bicycle testing, performance/capacity tasks), and the increment of the physical load as well as daily activities is also higher than before the start of the trial. Another important change is that many therapists now use cognitive-behavioral techniques while treating patients. The pain-contingent adjustment is increasingly being replaced by more time-contingent increase or pacing of activities. Still, this change of attitude has to be monitored and maintained by continuous education and evaluation of our daily care. This evaluation can be done by using the Pain Attitudes and Beliefs Scale for Physiotherapists (PBAS-PT) (Houben et al. 2005).

Informing general public

The role of cognitive-behavioral treatment in advising and treating CLBP patients should be further implemented by the use of the media. We have already one good example how a simple press release regarding our short-term results article can result in much media attention. It was striking to note that many people still think that CLBP is an exclusive biomedical problem. Many journalists stated that, after reading the conclusions of our article, especially that cognitive-behavioral treatment works as good as physical training, this was a very interesting conclusion and new information for them. We think that it is time for the government to start a public campaign in order to change the beliefs about LBP in both people and health care workers as well. Already one scientific study showed that a population based media-campaign can be successful in reducing the burden and costs of LBP associated disability (Buchbinder et al. 2001).

Recommendations for future research

First of all, a detailed cost-effectiveness analysis in which the total health care and work-related costs are taken into account, is necessary before firmly recommending single treatments over combination treatment. This analysis, which is based on the collection of weekly cost diaries, is currently being done.

Our data will also be used to calculate the minimal clinical important change of the Roland Disability Questionnaire (RDQ), Quebec Back Pain Disability Scale (QBPDS) and Patient-specific Main Complaints by using anchor based and distribution based methods (Crosby et al. 2003). Furthermore, the limits of agreement of the RDQ in the trial population before the start of treatment will be calculated so that these results can be used for daily rehabilitation practice.

The identification of subgroups of patients requiring specific treatment is a promising future research priority. Recently, for the treatment of patients with acute low back pain in primary care, a group of investigators developed a clinical prediction rule based on history taking and physical examination for the use of a brief spinal manipulation (Childs et al. 2004; Fritz et al. 2005), and specific exercises and stabilization exercises (Brennan et al. 2006). However, this algorithm is only applicable for patients with less than 90 days of back pain and not for CLBP patients (Brennan et al. 2006; Fritz et al. 2006). For these patients, subgrouping might be more effectively performed on the basis of psychosocial factors (Pincus et al. 2002; van der Hulst et al. 2005).
The data of our trial will be used to investigate whether prognostic factors for the differing treatments can be identified, although the number of patients per treatment group will limit the number of potential prognostic factors that can be investigated. Based on the interesting findings regarding the influential role of treatment expectancy and credibility, these factors should be further investigated with special emphasis on the potential role of a clear treatment rationale and motivational interviewing before offering the patient a treatment reduce LBP associated disability.

Since the treatment in daily rehabilitation practice is fine-tuned by selecting treatment modalities based on patient's characteristics, problems and needs, a next step could be comparing our highly structured treatments to these individually designed treatments. Since we did not find a clear relationship between the improvement of cardiovascular capacity (VO_2max) and the reduction of disability, it would be interesting whether physical training without addition of the aerobic training is equally effective. Furthermore, it is recommended to use reliable measurements regarding muscle strength/endurance and coordination in order to find out whether the reduction of disability is mediated by neuromuscular changes alongside cognitive and behavioral changes.

For the improvement of the selection procedure and the matching of treatments not only time and cost consuming randomized controlled trials, but especially (replicated) single case experimental designs with frequently repeated measurements should be used. These designs can also be used to study the effects of novel interventions that systematically focus on modifying pain catastrophizing and other putative mediating variables such as self-efficacy. Single case experimental designs can facilitate cross-lagged comparisons which can be used to confirm our findings regarding the mediating role of pain catastrophizing.

The recruitment of patients differs in CLBP trials. Patients selectively recruited by advertisement in local newspapers do not share the characteristics of those referred to a rehabilitation centre or pain clinic (Hayden et al. 2005a). Furthermore, mostly different LBP disability scales are used that cannot be compared with each other, and some studies do not use a disability scale at all. We therefore recommend that future trials should describe the population of patients extensively, and especially restrict the selection to patients with chronic nonspecific LBP who are referred for treatment. Furthermore, researchers should use at least one similar LBP disability scale (Dworkin et al. 2005), and consensus should be reached on the magnitude of the clinically important difference between active treatments.

Another major finding of our research is the lack of high quality research focusing on physical deconditioning in CLBP. Specifically, only cross-sectional studies have been used to investigate the link between deconditioning, back pain and disability. Future research should use prospective designs to study this relationship.

In order to improve the percentage of CLBP patients being able to perform a submaximal test, especially for the older and more disabled patients, the modified submaximal Åstrand bicycle test should be further adjusted. This might be possible by increasing the workload in smaller steps according to the mean increase of the heart rate in the preceding 2 minutes.

Beside the development of more patient-specific performance/capacity tasks, the development and use of ambulatory devices, preferably combined with methods to measure the total energy consumption, should be further encouraged in

order to provide answers on how active a patient with CLBP really is, and to get some more insight into the real decline in physical activities after the onset of back pain (Bussmann et al. 2001; Verbunt et al. 2001).

The increasing urge to involve patients in the selection of treatment and especially treatment goals, should be accompanied by the use of reliable and valid assessment tools. In our opinion the use of the Patients-specific Main Complaints is certainly advisable. The Canadian Occupational Performance Measure (COPM) is another instrument that can be of great value and should be further investigated (Walsh et al. 2004).

In summary, we think that future research should be directed on defining subgroups of patients, and treatments specifically aimed at reducing pain catastrophizing. Beside the often used randomized controlled trials, single case experimental designs are very promising and less time and cost consuming.

References

- Abenhaim L, Rossignol M, Valat JP, Nordin M, Avouac B, Blotman F, Charlot J, Dreiser RL, Legrand E, Rozenberg S, Vautravers P. The role of activity in the therapeutic management of back pain. Report of the International Paris Task Force on Back Pain. Spine 2000;25:1S-33S.
- Boersma K, Linton S, Overmeer T, Jansson M, Vlaeyen J, de Jong J. Lowering fear-avoidance and enhancing function through exposure in vivo. A multiple baseline study across six patients with back pain. Pain 2004;108:8-16.
- Brennan GP, Fritz JM, Hunter SJ, Thackeray A, Delitto A, Erhard RE. Identifying subgroups of patients with acute/subacute "nonspecific" low back pain: results of a randomized clinical trial. Spine 2006;31:623-31.
- Brox JI, Storheim K, Holm I, Friis A, Reikeras O. Disability, pain, psychological factors and physical performance in healthy controls, patients with sub-acute and chronic low back pain: a case-control study. J Rehabil Med 2005;37:95-9.
- Buchbinder R, Jolley D, Wyatt M. 2001 Volvo Award Winner in Clinical Studies: Effects of a media campaign on back pain beliefs and its potential influence on management of low back pain in general practice. Spine 2001;26:2535-42.
- Burns JW, Kubilus A, Bruehl S, Harden RN, Lofland K. Do changes in cognitive factors influence outcome following multidisciplinary treatment for chronic pain? A cross-lagged panel analysis. J Consult Clin Psychol 2003;71:81-91.
- Bussmann JB, Martens WL, Tulen JH, Schasfoort FC, van den Berg Emons HJ, Stam HJ. Measuring daily behavior using ambulatory accelerometry: the Activity Monitor. Behav Res Methods Instrum Comput 2001;33:349-56.
- Campello M, Nordin M, Weiser S. Physical exercise and low back pain. Scand J Med Sci Sports 1996;6:63-72.
- Childs JD, Fritz JM, Flynn TW, Irrgang JJ, Johnson KK, Majkowski GR, Delitto A. A clinical prediction rule to identify patients with low back pain most likely to benefit from spinal manipulation: a validation study. Ann Intern Med 2004;141:920-8.
- Crews DJ, Landers DM. A meta-analytic review of aerobic fitness and reactivity to psychosocial stressors. Med Sci Sports Exerc 1987;19:S114-S120.
- Crosby RD, Kolotkin RL, Williams GR. Defining clinically meaningful change in health-related quality of life. J Clin Epidemiol 2003;56:395-407.
- Curtis L, Mayer TG, Gatchel RJ. Physical progress and residual impairment quantification after functional restoration. Part III: Isokinetic and isoinertial lifting capacity. Spine 1994;19:401-5.
- Danneels LA. Evaluation and rehabilitation of functional spinal stability. Ghent: Ghent University, 2001.
- Dworkin RH, Turk DC, Farrar JT, Haythornthwaite JA, Jensen MP, Katz NP, Kerns RD, Stucki G, Allen RR, Bellamy N, Carr DB, Chandler J, Cowan P, Dionne R, Galer BS, Hertz S, Jadad AR, Kramer LD, Manning DC, Martin S, McCormick CG, McDermott MP, McGrath P, Quessy S, Rappaport BA, Robbins W, Robinson JP, Rothman M, Royal MA, Simon L, Stauffer JW, Stein W, Tollett J, Wernicke J, Witter J. Core outcome measures for chronic pain clinical trials: IMMPACT recommendations. Pain 2005;113:9-19.
- Ferrell BA, Josephson KR, Pollan AM, Loy S, Ferrell BR. A randomized trial of walking versus physical methods for chronic pain management. Aging 1997;9:99-105.
- Fordyce WE. Behavioral methods for chronic pain and illness. St. Louis: Mosby, 1976.
- Fritz JM, Brennan GP, Clifford SN, Hunter SJ, Thackeray A. An examination of the reliability of a classification algorithm for subgrouping patients with low back pain. Spine 2006;31:77-82.
- Fritz JM, Childs JD, Flynn TW. Pragmatic application of a clinical prediction rule in primary care to identify patients with low back pain with a good prognosis following a brief spinal manipulation intervention. BMC Fam Pract 2005;6:29.
- Gurevich M, Kohn PM, Davis C. Exercise-induced analgesia and the role of reactivity in pain sensitivity. J Sports Sci 1994;12:549-59.

References

- Guzman J, Esmail R, Karjalainen K, Malmivaara A, Irvin E, Bombardier C. Multidisciplinary bio-psycho-social rehabilitation for chronic low back pain. Cochrane Database Syst Rev 2002:CD000963.
- Hayden JA, van Tulder MW, Malmivaara AV, Koes BW. Meta-analysis: exercise therapy for nonspecific low back pain. Ann Intern Med 2005a;142:765-75.
- Hayden JA, van Tulder MW, Tomlinson G. Systematic review: strategies for using exercise therapy to improve outcomes in chronic low back pain. Ann Intern Med 2005b;142:776-85.
- Helmhout PH, Harts CC, Staal JB, Candel MJ, de Bie RA. Comparison of a high-intensity and a low-intensity lumbar extensor training program as minimal intervention treatment in low back pain: a randomized trial. Eur Spine J 2004;13:537-47.
- Hodselmans AP, Jaegers SM, Goeken LN. Short-term outcomes of a back school program for chronic low back pain. Arch Phys Med Rehabil 2001;82:1099-105.
- Houben RM, Ostelo RW, Vlaeyen JW, Wolters PM, Peters M, Stomp-van den Berg SG. Health care providers' orientations towards common low back pain predict perceived harmfulness of physical activities and recommendations regarding return to normal activity. Eur J Pain 2005;9:173-83.
- Jensen MP. Enhancing motivation to change in pain treatment. In: RJ Gatchel, DC Turk. Psychological approaches to pain management. A practitioner's handbook. New York, NY: The Guilford Press, 1996, pp. 78-111.
- Jensen MP, Nielson WR, Kerns RD. Toward the development of a motivational model of pain self-management. J Pain 2003;4:477-92.
- de Jong JR, Vlaeyen JW, Onghena P, Goossens ME, Geilen M, Mulder H. Fear of movement/(re)injury in chronic low back pain: education or exposure in vivo as mediator to fear reduction? Clin J Pain 2005;21:9-17.
- Kaplan GM, Wurtele SK, Gillis D. Maximal effort during functional capacity evaluations: an examination of psychological factors. Arch Phys Med Rehabil 1996;77:161-4.
- Keefe FJ, Beaupre PM, Gil KM. Group therapy for patients with chronic pain. In: RJ Gatchel, DC Turk. Psychological approaches to pain management. A practitioner's handbook. New York: The Guilford Press, 1996, pp. 259-82.
- Kerns RD, Jensen MP, Nielson WR. Motivational issues in pain self-management. In: H Flor, E Kalso, JO Dostrovsky. Proceedings of the 11th World Congress on Pain. Seattle: IASP Press, 2006, pp. 555-65.
- Koke A, Brouwer M, Heuts PH, Schiphorst-Preuper R, Smeets RJ, Swaan L, Vlaeyen JW, Patijn J. Consensus rapport pijnrevalidatie Nederland. Maastricht: Pijn Kennis Centrum Maastricht, 2005.
- Koumantakis GA, Watson PJ, Oldham JA. Trunk muscle stabilization training plus general exercise versus general exercise only randomized controlled trial of patients with recurrent low back pain. Phys Ther 2005;85:209-25.
- Lackner JM, Carosella AM. The relative influence of perceived pain control, anxiety, and functional self efficacy on spinal function among patients with chronic low back pain. Spine 1999;24:2254-60.
- Lackner JM, Carosella AM, Feuerstein M. Pain expectancies, pain, and functional self-efficacy expectancies as determinants of disability in patients with chronic low back disorders. J Consult Clin Psychol 1996;64:212-20.
- Lee CE, Simmonds MJ, Novy DM, Jones S. Self-reports and clinician-measured physical function among patients with low back pain: a comparison. Arch Phys Med Rehabil 2001;82:227-31.
- Liddle SD, Baxter GD, Gracey JH. Exercise and chronic low back pain: what works? Pain 2004;107:176-90.
- Ljungquist T, Jensen IB, Nygren A, Harms Ringdahl K. Physical performance tests for people with long-term spinal pain: aspects of construct validity. J Rehabil Med 2003;35:69-75.

References

- Lygren H, Dragesund T, Joensen J, Ask T, Moe Nilssen R. Test-retest reliability of the Progressive Isoinertial Lifting Evaluation (PILE). Spine 2005;30:1070-4.
- Mannion AF, Muntener M, Taimela S, Dvorak J. A randomized clinical trial of three active therapies for chronic low back pain. Spine 1999;24:2435-48.
- Mayer TG, Gatchel RJ. Functional restoration for spinal disorders: the sports medicine approach. Philadelphia: Lea & Febiger,1988.
- Mayer TG, Polatin PB, Gatchel RJ. Functional restoration and other rehabilitation approaches to chronic musculoskeletal pain disability syndromes. Crit Rev Physical Rehab Med 1998;10:209-21.
- Moffett JK, Torgerson D, Bell-Syer S, Jackson D, Llewlyn-Phillips H, Farrin A, Barber J. Randomised controlled trial of exercise for low back pain: clinical outcomes, costs, and preferences. BMJ 1999;319:279-83.
- Moseley GL. Evidence for a direct relationship between cognitive and physical change during an education intervention in people with chronic low back pain. Eur J Pain 2004;8:39-45.
- Novy DM, Simmonds MJ, Lee CE. Physical performance tasks: what are the underlying constructs? Arch Phys Med Rehabil 2002;83:44-7.
- Nutter P. Aerobic exercise in the treatment and prevention of low back pain. Occup Med 1988;3:137-45.
- Ostelo RW, Tulder M, Vlaeyen J, Linton S, Morley S, Assendelft W. Behavioural treatment for chronic low-back pain. Cochrane Database Syst Rev 2005:Cd002014.
- Pincus T, Burton AK, Vogel S, Field AP. A systematic review of psychological factors as predictors of chronicity/disability in prospective cohorts of low back pain. Spine 2002;27:E109-20.
- Protas EJ. Aerobic exercise in the rehabilitation of individuals with chronic low back pain: a review. Crit Rev Physical Rehab Med 1996;8:283-95.
- Roland M, Fairbank J. The Roland-Morris Disability Questionnaire and the Oswestry Disability Questionnaire. Spine 2000;25:3115-24.
- Simmonds MJ, Kumar A, Lechelt E. Psychosocial factors in low back pain: Causes or consequences? Disabil and Rehabil 1996;18:161-8.
- Smeets RJEM, Vlaeyen JWS, Kester ADM, Knottnerus JA. Reduction of pain catastrophizing mediates the outcome of both physical and cognitive-behavioral treatment in chronic low back pain. J Pain 2006;7:261-71.
- Spinhoven P, Ter Kuile M, Kole-Snijders AM, Hutten-Mansfeld M, Den Ouden DJ, Vlaeyen JW. Catastrophizing and internal pain control as mediators of outcome in the multidisciplinary treatment of chronic low back pain. Eur J Pain 2004;8:211-9.
- UK BEAM Team. United Kingdom back pain exercise and manipulation (UK BEAM) randomised trial: effectiveness of physical treatments for back pain in primary care. BMJ 2004;329:1377-81.
- Thirlaway K, Benton D. Participation in physical activity and cardiovascular fitness have different effects on mental health and mood. J Psychosom Res 1992;36:657-65.
- Twisk JWR. Applied longitudinal data analysis for epidemiology. A practical guide. Cambridge: Cambridge University Press, 2003.
- van den Hout JH, Vlaeyen JW, Heuts PH, Zijlema JH, Wijnen JA. Secondary prevention of work-related disability in nonspecific low back pain: does problem-solving therapy help? A randomized clinical trial. Clin J Pain 2003;19:87-96.

References

- van der Hulst M, Vollenbroek-Hutten MM, IJzerman MJ. A systematic review of sociodemographic, physical, and psychological predictors of multidisciplinary rehabilitation - or, back school treatment outcome in patients with chronic low back pain. Spine 2005;30:813-25.
- Verbunt JA, Seelen HA, Vlaeyen JWS, van der Heijden GJ, Heuts PH, Pons K, Knottnerus JA. Disuse and deconditioning in chronic low back pain: concepts and hypothesis on contributing mechanisms. Eur J Pain 2003;7:9-21.
- Verbunt JA, Westerterp KR, van der Heijden GJ, Seelen HA, Vlaeyen JW, Knottnerus JA. Physical activity in daily life in patients with chronic low back pain. Arch Phys Med Rehabil 2001;82:726-30.
- Vlaeyen JWS, de Jong J, Geilen M, Heuts PHTG, van Breukelen G. The treatment of fear of movement/(re)injury in chronic low back pain: further evidence on the effectiveness of exposure in vivo. Clin J Pain 2002;18:251-61.
- Vlaeyen JWS, Kole-Snijders AM, Boeren RG, van Eek H. Fear of movement/(re)injury in chronic low back pain and its relation to behavioral performance. Pain 1995;62:363-72.
- Vlaeyen JWS, Linton SJ. Fear-avoidance and its consequences in chronic musculoskeletal pain: a state of the art. Pain 2000;85:317-32.
- Vollenbroek-Hutten MM, Hermens HJ, Wever D, Gorter M, Rinket J, IJzerman MJ. Differences in outcome of a multidisciplinary treatment between subgroups of chronic low back pain patients defined using two multiaxial assessment instruments: the multidimensional pain inventory and lumbar dynamometry. Clin Rehabil 2004;18:566-79.
- Walsh AD, Kelly JS, Johnson SP, Rajkumar S, Bennetts K. Performance problems of patients with chronic low-back pain and the measurement of patient-centreed outcome. Spine 2004;29:87-93.
- WHO. The International Classification of Functioning, Disability and Health (ICF). Geneva: WHO, 2001.
- Wittink H, Michel TH, Kulich R, Wagner A, Sukiennik A, Maciewicz R, Rogers W. Aerobic fitness testing in patients with chronic low back pain: which test is best? Spine 2000;25:1704-10.

Summary.

Summary

Summary

Chapter 1

Introduction

Low back pain is a highly prevalent health disorder, especially in industrialized countries, causing a huge personal as well as socio-economic burden. Three frequently used models for the explanation of the development and maintenance of functional limitations due to chronic low back pain (CLBP) are the deconditioning model, the cognitive-behavioral model and the biopsychosocial model. Similarly, three major rehabilitation strategies for the reduction of CLBP disability can be distinguished: a biomedical approach aimed at restoring aerobic capacity, muscular strength and endurance, especially of the lower back, a cognitive-behavioral one aimed at increasing the patients activity level and modifying dysfunctional beliefs and the acquisition of cognitive and behavioral coping skills, and finally, a combination of both approaches. There is substantial evidence that these strategies are more effective than no treatment.
However, much controversy exists whether one strategy is more effective than the other, and there is no consensus about the content, intensity and frequency of the different treatment elements. This might also be attributable to the fact that many treatments are not solely based on one of these three models, or do not fulfill, for example, the physiologic training principles to be classified as reconditioning or strengthening treatment.
The overall aim of the research reported upon in this thesis was to assess the effectiveness of three model-based treatments by performing a randomized controlled trial (RCT). The second aim was to assess the putative mediating role of pain catastrophizing and internal control of pain, and whether treatment credibility and expectancy influenced the final treatment outcome. Additionally, a systematic review, as well as a cross-sectional study was performed in order to assess the association of physical deconditioning and CLBP. Finally, the usability of physical performance/capacity tasks and the association of personal and body functions and the score on these tasks were assessed.

Chapter 2

Physical deconditioning and chronic low back pain

This chapter provides a systematic review regarding the association of physical deconditioning and CLBP. The physical deconditioning theory is described and features of physical deconditioning are defined. Systematic analysis of observational studies regarding the loss of cardiovascular capacity assessed by calculating the VO_2max, showed that there is conflicting evidence regarding the existence of aerobic deconditioning in CLBP patients. The loss of lower back muscle strength and endurance assessed by cross sectional area, amount of fatty infiltration, and fibre-type changes, respectively, provided limited evidence for wasting of the multifidus muscle. However, the results based on muscle fibre-changes are inconclusive.
On the basis of high quality RCTs, identified by using the methodology scoring list of the Cochrane Collaboration, no study regarding the effectiveness of cardiovascular enhancing treatments was found. Moderate evidence was found that intensive low back extensor muscle strengthening is more effective than less intensive strengthening.

Reactivation caused by active treatment and not the reconditioning itself might be the important factor in the reduction of disability, and further prospective and evaluative research into the role of physical deconditioning is necessary. The effectiveness of intensive low back extensor muscle strengthening should be assessed in a population of patients with a moderate to high level of perceived disability, using a methodologically strong design.

Chapter 3

Aerobic fitness and chronic low back pain

As appeared from the review, there were still many uncertainties regarding the existence of aerobic deconditioning in CLBP patients. Many researchers, reporting a reduced aerobic fitness level hypothesized that the activity level, especially regarding sport might be a very important contributing factor. Surprisingly, a sound comparison of CLBP patients, and healthy controls matched for their sport activity had never been carried out. Furthermore, several putative influential variables for the reduction of aerobic capacity were postulated, but most studies only calculated univariate correlations.
We conducted a cross-sectional study to compare the aerobic fitness level of 108 CLBP patients, assessed by a modified Åstrand submaximal bicycle test, with data from a database of healthy Dutch controls matched for gender, age and level of sport activity.
Furthermore, we evaluated the association of the difference in aerobic fitness level (observed VO_2max minus expected VO_2max) with pain intensity, duration and degree of disability, fear of injury, and level of activity during work including household, and leisure time. VO_2max could be calculated in 78% of the patients. The patients who stopped the test prematurely, were older and more disabled. Both men and women with CLBP had a significant and clinically relevant lower VO_2max than the healthy referents (10 mL/kg LBM.min^{-1} and 5.6 mL/kg LBM.min^{-1}, respectively), and this difference was significantly greater in men. Multiple regression analysis showed that the level of aerobic fitness was not associated with any of the presumed variables. The finding that men showed a lower level of observed versus expected aerobic fitness level than women had been reported earlier. However, the assumption that this might be caused by women being more engaged in childcare and various household duties, could not be confirmed.

Chapter 4

Short term results of a randomized controlled trial

First, three different active treatments, each based on one of three frequently used models regarding the development and maintenance of CLBP were developed; 10 weeks of aerobic training and muscle strengthening of back extensors (active physical treatment; APT), 10 weeks of gradual assumption of patient relevant activities based on operant behavioral principles and problem solving training (cognitive-behavioral treatment; CBT), and APT combined with CBT (combination treatment; CT). Next a randomized controlled trial was conducted to assess the effectiveness of these treatments compared with a waiting list control group (WL). Furthermore, we tested the hypothesis that CT is more effective than APT and CBT, respectively. Of 309 CLBP patients screened, 223 patients were randomized to APT, CBT, CT, or WL. There were only minor treatment protocol deviations with no relevant differences between treatments.

The treatment adherence was highest in APT, followed by CBT and, although only slightly lower, lowest in CT. Hardly any adverse effects were reported. For 212 patients, the primary outcome data were available for analysis. After treatment, significant and mostly clinically relevant reductions were observed in the level of disability, patient's main complaints, self-perceived improvement, and pain intensity for all three active treatments compared with WL. For satisfaction an interaction effect between type of intervention and baseline level of disability was found. Patients allocated to APT and a low to moderate disability level, were significantly more satisfied compared with WL than those with a high level of disability. For the patients allocated to CBT and CT, the level of satisfaction was significantly higher compared with WL, and the higher the baseline RDQ score, the greater this difference became. For patients with a high disability level, CT showed a significantly higher level of satisfaction than APT. Several physical performance/capacity tasks improved in APT and CT, but not in CBT. However, for all outcome measures no clinically relevant differences were found between the CT and APT, or between CT and CBT.

Chapter 5

One year post-treatment results of a randomized controlled trial

Of 156 of the 172 patients, who were allocated to one of the three active treatments (91%), primary outcome data were available one year post-treatment. At one year post-treatment, both APT and CBT when compared with CT, showed a higher, but statistically and clinically non-significant reduction of RDQ (level of disability): 1.16 [CI -0.52 to 2.84], and 1.62 [CI -0.06 to 3.31], respectively. Two alternative analyses (last value carried forward and worst case scenario) confirmed our findings. Furthermore, consecutive post-treatment measurements showed an increasing small to moderate difference in pain in favor of the single treatments, and the improvement of main complaints favored CBT. Self-perceived improvement reached statistical significance one year post-treatment for CBT and APT compared with CT. No relevant differences were found regarding depression and performance tasks, although 33% of the performance/capacity data were missing. Given that CT implies a higher burden for patients, our study does not support CT as a more useful treatment option.

Several explanations were offered for not finding CT more effective than APT and CBT, respectively, including the potential counteractive effect of combining physiologic and operant behavioral training principles, the unexpected persistent effectiveness of APT, as well as the lower compliance rate in CT.

Chapter 6

Mediating role of pain catastrophizing in chronic low back pain

In order to improve the effectiveness/efficacy of treatment we need insight into the mechanisms that are responsible for the desired outcomes. This process is called mediation. Reduction of pain catastrophizing and increase of internal control of pain have been postulated to be important mediating factors in CBT. However, their role has never been investigated in physical treatment. In order to examine whether treatments based on different models change pain catastrophizing and internal control of pain, and whether changes in these factors mediate treatment outcome, an

additional analysis of 211 patients with nonspecific CLBP participating in our RCT was conducted. Patients attended APT (n = 52), CBT (n = 55), CT (n = 55), or WL (n = 49). Immediately post-treatment, pain catastrophizing appeared to have decreased in all three active treatment groups, but not in the WL. There was no difference in the change in internal control across all four groups. Change in pain catastrophizing mediated the reduction of disability, main complaints and pain intensity. However, it should be kept in mind that we were not able to collect data at mid-treatment and could not perform a cross-lagged correlation analysis. Therefore, the direction of causality among the variables remains ambiguous. Nevertheless, the results might contribute to the development of more effective interventions, specifically aiming at the reduction of pain catastrophizing. This study shows that treatment elements that do not deliberately target cognitive factors, but instead concentrate on letting the patients experience that performing physical exercises or normally hampered activities in a controlled environment is still possible, are responsible for a substantial cognitive change.

Chapter 7

Treatment expectancy and credibility in chronic low back pain

Patients' initial beliefs about the success of a given pain treatment are shown to affect final treatment outcome. The Credibility/Expectancy Questionnaire (CEQ) has recently been shown to measure treatment credibility and expectancy reliably. The objectives of this study were 1) to investigate the factor structure of the CEQ in our RCT sample of CLBP patients by means of a confirmatory factor analysis (CFA), 2) to examine the association between treatment credibility and expectancy and patient characteristics and 3) to assess the extent to which treatment expectancy and credibility are associated with outcome of CLBP rehabilitation treatment. 167 patients randomized to either APT (n = 51), CBT (n = 57) or CT (n = 59) completed the CEQ after the treatment rationale was explained. CFA supported the two-factor structure (credibility/expectancy). Lower credibility was associated with higher pain-related fear and lower internal control of pain and lower expectancy with higher levels of pain-related fear and no radiating pain. Multiple linear regression analyses revealed that after controlling for age, gender, intervention, treatment centre, pain intensity and duration of disability, expectancy was significantly, although slightly to modestly, associated with disability and satisfaction for all treatments. Credibility was significantly associated with patient-specific complaints and satisfaction. For the outcome variable global perceived effect modification was found. Treatment expectancy was predictive in APT only, and treatment credibility showed a significant, but moderate association in CT only. Expectancy and credibility did not predict change in pain in any of the treatment conditions. When treating CLBP patients it might be promising to use (pre-)treatments that focus more specifically on increasing both treatment credibility and expectancy.

Chapter 8

Usability of physical performance/capacity tasks

In our RCT, six physical performance/capacity tasks were used to more 'objectively' assess physical capacity in CLBP patients. To improve objectivity and reliability of the assessment within the RCT, the testing protocols of the six selected tasks had to be adjusted. As a result, the existing psychometric data were not valid anymore. Furthermore, no data about the influence of task experience were available. Finally, no data were available regarding the limits of agreement, defining the minimal change needed to be confident that the observed change between two measurements reflects real change in an individual patient, and not merely a measurement error. We therefore performed a test-retest study (5 to 9 day interval) within a subgroup of our trial population and included 30 patients with no task experience and 23 patients already having undertaken the tasks on at least two occasions.

Task experience did not significantly influence test-retest differences. All tasks showed high to very high reliability (ICC varied from 0.74 to 0.99). The limits of agreement expressed as percentage of the mean score was low to moderate for five-minute walking and one-minute stair climbing (21% and 20%, respectively), moderate for 50-foot fast walking, sit to stand and forward reach (33%, 29% and 36%, respectively), and high for the PILE (48%). We concluded that the five-minute walking and stair climbing task, and to a lesser degree the 50-foot walking, sit to stand and loaded forward reach, seem clinically useful. However, there are major concerns about the usability of the PILE.

Chapter 9

Contributing factors of capacity tasks in chronic low back pain

Although the performance/capacity tasks are thought to be more objective than questionnaires, these tasks are also influenced by cognitive and emotional processes. In this cross-sectional study, for 221 patients participating in our RCT, the association of personal and body functions with the pre-treatment score of each of the six performance/capacity tasks was assessed by multiple linear regression analysis. At baseline, the independent variables age, gender, pain intensity, duration of pain, radiating pain to leg, VO_2max, pain-related fear, depression, internal control of pain and pain catastrophizing were collected. Because of a relatively high percentage of missing VO_2max data (20.8%) a multiple imputation technique had to be used.

Despite the inclusion of ten covariates, the total explained variance was low to moderate (9% to 19%) except for stair climbing, for which the variance just reached 30%. Cardiovascular capacity, pain intensity fear of injury/movement, cognitions and depression had statistically significant, but clinically minor effects on several but not all performance/capacity tasks. Radiating pain, age and duration of complaints had no significant influence at all. Due to anthropometric differences men outperformed women on most tasks. Apparently the influence of many personal, physical, but especially psychological factors on the selected capacity tasks is not high at all. This might indicate that these tasks measure physical capacity more objectively than expected.

Chapter 10

General discussion

This final chapter provides an extensive summary of findings and conclusions, and tries to integrate the results of all the studies presented in this thesis. The most important finding of these studies is that CT is not more effective than the single treatment components. Furthermore, the effectiveness of the APT is mediated by the reduction of pain catastrophizing. Despite the existence of deconditioning (reduced VO_2max) in our sample of CLBP patients, reduction of disability can be achieved without improving the cardiovascular capacity. Treatment credibility and expectancy can be important influential variables in the treatment of CLBP patients, although their association with treatment outcome in our sample is only low to modest. Several performance/capacity tasks are clinically useful, and the influence of several personal and body functions is rather limited.

Next, methodological aspects, clinical implications and suggestions for clinically based research, aspects of implementation and finally, suggestions for future research are addressed.

Samenvatting

Hoofdstuk 1

Inleiding

Lage rugpijn is met name in de geïndustrialiseerde landen een veel voorkomend gezondheidsprobleem en veroorzaakt hoge persoonlijke en sociaal-economische lasten. Drie veel gebruikte verklaringsmodellen voor het ontstaan en in stand houden van beperkingen in het dagelijks functioneren ten gevolge van chronische lage rugpijn (CLRP) zijn het deconditionering model, het cognitief-gedragstherapeutisch model en het biopsychosociaal model. Hoewel deze drie modellen niet geheel exclusief zijn, kunnen ook drie belangrijke vormen van revalidatiebehandeling ter vermindering van met CLRP samenhangende beperkingen worden onderscheiden: een biomedische benadering gericht op het herstellen van de aërobe capaciteit en het verbeteren van kracht en duurvermogen van met name van de lage rugspieren, een cognitief-gedragstherapeutische benadering gericht op het vergroten van het activiteitenniveau van de patiënt en het veranderen van disfunctionele gedachten (attributies en verwachtingen) en het aanleren van cognitieve en gedragsmatige copingvaardigheden en tot slot, de biopsychosociale benadering die bestaat uit een combinatie van beide eerder genoemde behandelstrategieën. Er is substantieel bewijs voorhanden dat deze drie behandelstrategieën effectiever zijn dan het geven van geen behandeling. Anderzijds bestaat er nogal wat controverse of de ene behandelstrategie effectiever is dan de andere. Verder bestaat er geen overeenstemming over de exacte inhoud, intensiteit en frequentie van de verschillende behandelonderdelen. Deze controverse is waarschijnlijk te wijten aan het feit dat vele behandelvormen onvoldoende of vaak op meerdere verklaringsmodellen tegelijk zijn gebaseerd. Een voorbeeld hiervan is een fysieke behandeling die niet voldoet aan de fysiologische trainingsprincipes om als aërobe conditietraining of spierkracht en –vermogen verbeterende behandeling te mogen worden geclassificeerd.

Het belangrijkste doel van het onderzoek waarover in dit proefschrift wordt gerapporteerd is het bepalen van de effectiviteit van drie modelgebaseerde behandelingen door middel van een gerandomiseerde gecontroleerde studie (randomized controlled trial; RCT). Het tweede doel is het onderzoeken of pijn-catastroferende gedachten en de ervaren interne controle over pijn het effect van de verschillende behandelvormen mediëren. Het derde doel is te bepalen of de geloofwaardigheid van en de verwachting ten aanzien van de aangeboden behandeling het uiteindelijk behandelresultaat beïnvloeden.

Aanvullend werd een systematisch literatuuronderzoek, alsook een cross-sectionele studie uitgevoerd om de relatie tussen fysieke deconditionering en CLRP nader te onderzoeken. Tot slot werd de klinische bruikbaarheid van enkele fysieke performance/capaciteit taken, alsmede de relatie tussen de scores op deze taken en meerdere persoonsgebonden en fysieke functies bepaald.

Hoofdstuk 2

Fysieke deconditionering en chronische lage rugpijn

In dit hoofdstuk wordt op basis van een systematisch onderzoek van de bestaande wetenschappelijke literatuur verslag gedaan van de relatie tussen fysieke deconditionering en CLRP. Allereerst wordt de theorie van deconditionering beschreven en worden de verschijnselen van fysieke deconditionering nader gedefinieerd. Een systematische analyse van observationele studies over het verlies van cardiovasculaire capaciteit, vastgesteld door middel van het berekenen van de VO_2max, laat zien dat er conflicterend bewijs is voor het bestaan van aërobe deconditionering bij mensen met CLRP. Studies die het verlies van spierkracht en duurvermogen van de lagerugspieren bestudeerden door middel van het meten van de omvang van deze spieren met behulp van MRI- of CT-scan opnames en de mate van vervetting alsmede veranderingen van spiervezeltypologie door middel van spierbiopten, werden verzameld en bestudeerd. Op basis van deze studies kon worden geconcludeerd dat er beperkt bewijs bestaat voor verminderde kracht en uithoudingsvermogen van de multifidus spieren. Echter, op basis van de spiervezel studies kunnen geen conclusies worden getrokken. Gebruikmakend van de methodologie scoringslijst van de Cochrane Collaboration werden vervolgens alleen RCTs van hoge kwaliteit geselecteerd. Er werd geen enkele studie gevonden waarin de effectiviteit van cardiovasculaire capaciteit verhogende behandelingen werd onderzocht. Verder bleek er matig bewijs te bestaan dat intensieve spierkrachttraining van de lagerugstrekkers effectiever is dan een minder intensieve krachttraining.

De reactivering die wordt veroorzaakt door de actieve behandelvormen- en niet de reconditionering sec, zou wel eens de belangrijke factor kunnen zijn bij het reduceren van beperkingen. Verder prospectief en evaluatief onderzoek naar de rol van fysieke deconditionering is noodzakelijk. De effectiviteit van een intensieve spierkrachttraining van de lagerugstrekkers dient door middel van een studie van methodologisch hoge kwaliteit te worden onderzocht in een populatie van patiënten met een matig tot ernstig niveau van ervaren beperkingen.

Hoofdstuk 3

Aërobe capaciteit en chronische lage rugpijn

Zoals bleek uit het systematisch literatuuroverzicht (hoofdstuk 2), bestaan er nog steeds vele onzekerheden en onduidelijkheden ten aanzien van het bestaan van aërobe deconditionering bij CLRP patiënten. Vele onderzoekers die een verlaagde aërobe capaciteit rapporteerden, suggereren dat het activiteitenniveau, vooral op sportief vlak, een zeer belangrijke bijdragende factor zou kunnen zijn. Het is echter opmerkelijk dat een degelijke vergelijking tussen CLRP patiënten en gezonde controle personen die een even hoog sportactiviteiten niveau hebben, nooit is uitgevoerd. Verder worden in de literatuur andere factoren voorondersteld die verantwoordelijk zouden kunnen zijn voor het verlies aan aërobe capaciteit, hoewel de meeste van deze studies alleen univariate correlaties berekenden.

Wij voerden een cross-sectionele studie uit, waarin we de aërobe capaciteit (VO_2max) van 108 CLRP patiënten bepaalden door middel van een gemodificeerde Åstrand submaximaal fietstest en vervolgens vergeleken met normgegevens uit een databestand van gezonde Nederlandse mensen van dezelfde leeftijd, geslacht en met een vergelijkbaar niveau van sportactiviteiten. Verder bepaalden we de relatie tussen het gemeten verschil in aërobe capaciteit (gevonden VO_2max van de CLRP patiënten verminderd met de normwaarde VO_2max uit het databestand) en de pijnintensiteit, duur en mate van beperkingen, bewegingsvrees en activiteitenniveau tijdens werk inclusief huishouden en tijdens vrije tijd. Voor 78% van de CLRP patiënten kon de VO_2max worden berekend. De patiënten die de test vroegtijdig moesten beëindigen, waren ouder en meer beperkt. Zowel mannen als vrouwen met CLRP hadden een significant en klinisch relevant lagere VO_2max dan de gezonde referentiegroep (respectievelijk 10 mL/kg LBM.min^{-1} en 5.6 mL/kg LBM.min^{-1}) en dit verschil was significant groter bij de mannen. Multipele regressie analyse liet zien dat het verschil tussen gemeten en verwachte aërobe capaciteit niet samenhing met een of meerdere van de tevoren invloedrijk geachte variabelen. De bevinding dat mannen met CLRP een lagere gemeten versus verwachte aërobe capaciteit laten zien dan vrouwen met CLRP werd al eerder in andere studies gerapporteerd. Echter, de hypothese dat dit veroorzaakt wordt doordat vrouwen meer actief zijn met de verzorging van kinderen en verschillende andere huishoudelijke taken dan mannen, kon in onze studie niet worden bevestigd.

Hoofdstuk 4

Korte termijn resultaten van een gerandomiseerd onderzoek

Allereerst werden drie verschillende actieve behandelingen, ieder gebaseerd op een van de drie frequent gebruikte modellen betreffende het ontstaan en instandhouden van CLRP, ontwikkeld: een 10 weken durende aërobe training en spierversterkende oefeningen van de lagerugstrekkers (active physical treatment; APT); een 10 weken durende graduele hervatting of uitbreiding van patiëntrelevante activiteiten gebaseerd op operant-gedragsmatige principes en probleem oplossende vaardigheidstraining (cognitive-behavioral treatment; CBT) en een 10 weken durende combinatie behandeling bestaande uit zowel de fysieke als de cognitief-gedragstherapeutische behandeling (combination treatment; CT). Vervolgens werd een gerandomiseerd gecontroleerde studie (RCT) uitgevoerd om de effectiviteit van deze behandelingen in vergelijking met een wachtlijst-controlegroep (WL) te bepalen. Tevens werd de hypothese getoetst dat de combinatiebehandeling (CT) effectiever is dan respectievelijk de fysieke behandeling (APT) en de cognitief-gedragstherapeutische behandeling (CBT).

Van de 309 patiënten die werden gescreend voor deelname aan de RCT werden uiteindelijk 223 patiënten gerandomiseerd naar de APT, CBT, CT of WL. Bij slechts een klein aantal patiënten werd een afwijking van het behandelprotocol vastgesteld en bovendien bleken er geen relevante verschillen ten aanzien van de geconstateerde protocolafwijkingen tussen de behandelingen onderling te bestaan. Binnen de APT was het percentage patiënten dat de behandeling voldoende trouw volgde het hoogst, gevolgd door CBT en, hoewel nauwelijks lager, het laagst in CT. Er werden vrijwel geen negatieve bijwerkingen gerapporteerd. Van 212 patiënten waren de data betreffende de primaire uitkomstmaat beschikbaar voor analyse. Meteen na het einde van de behandeling werd voor alle drie de actieve behandelingen, in vergelijking met de WL, een significante en meestal ook klinisch relevante reductie van het niveau van ervaren beperkingen, patiëntspecifieke klachten, ervaren algemene verbetering en pijnintensiteit gevonden. Voor de tevredenheid over de behandeling werd een interactie gevonden tussen het type behandeling en de aanvangsscore van de ervaren beperkingen. Patiënten die gerandomiseerd werden naar de APT en een laag tot matig niveau van beperkingen bij aanvang van de behandeling rapporteerden, bleken significant meer tevreden te zijn dan de WL-patiënten. Dit in tegenstelling tot diegenen die een hoog niveau van beperkingen bij aanvang van de behandeling rapporteerden. Patiënten die de CBT en CT kregen toegewezen waren significant meer tevreden over de behandeling dan de WL-patiënten en, hoe hoger de aanvangsscore van de ervaren beperkingen, hoe groter dit verschil in tevredenheid werd. Voor patiënten met een hoge aanvangscore van ervaren beperkingen, bleek dat diegenen die CT volgden een significant hogere tevredenheid rapporteerden dan diegenen die de APT volgden. Meerdere fysieke performance/capaciteit taken verbeterden in de APT en CT, maar niet in de CBT. Echter, voor geen van de uitkomstmaten werden klinisch relevante verschillen gevonden tussen de CT en APT, of tussen CT en CBT.

Hoofdstuk 5

Lange termijn resultaten van een gerandomiseerd onderzoek

Een jaar na het einde van de behandeling waren van 156 van de 172 patiënten, die werden gerandomiseerd naar een van de drie actieve behandelvormen (91%), de data van de primaire uitkomstmaat beschikbaar. Zowel APT als CBT lieten in vergelijking met CT, een grotere afname van beperkingen ten gevolge van de CLRP zien. Echter, deze verschillen waren zowel statistisch als klinisch niet relevant: respectievelijk 1.16 [95%-betrouwbaarheidsinterval -0.52 tot 2.84] en 1.62 [95%-betrouwbaarheidsinterval -0.06 tot 3.31]. Twee alternatieve statistische analysetechnieken (ontbrekende data vervangen door de laatst beschikbare waarde van de desbetreffende patiënt [last value carried forward] en de ontbrekende data vervangen de 10de percentiel score van de behandelgroep op het desbetreffende meetmoment [worst case scenario]) bevestigden onze bevindingen. Verder lieten achtereenvolgende metingen gedurende het jaar na het einde van de behandeling een toenemend klein tot matig verschil ten aanzien van pijnintensiteit ten faveure van de enkelvoudige behandelingen zien, en de verbetering van patiëntspecifieke klachten bleek in het voordeel van de CBT. Een jaar na einde van de behandeling bleek dat patiënten die CBT of APT hadden ondergaan een significant hogere mate van ervaren algemene verbetering rapporteerden dan patiënten die CT hadden doorlopen.

Er werden geen relevante verschillen gevonden voor depressie en de performance/capaciteit taken, hoewel moet worden aangetekend dat een jaar na het einde van de behandeling van 33% van de patiënten de data van de performance/capaciteit taken ontbraken. Gegeven het feit dat CT een hogere belasting voor de patiënten met zich meebrengt, levert onze studie geen bewijs dat CT een zinvollere behandeloptie is.

Meerdere verklaringen voor het feit dat de gecombineerde behandeling niet effectiever is dan de fysieke en cognitief-gedragstherapeutische behandeling afzonderlijk worden besproken: het mogelijk oppositionele effect van de combinatie van fysiologische en operant-gedragsmatige trainingsprincipes, de onverwachte effectiviteit op de langere termijn van APT, alsook het lagere percentage patiënten in CT dat een voldoende intensieve behandeling volgde.

Hoofdstuk 6

Pijn-catastroferen als mediator bij chronische lage rugpijn

Om de effectiviteit/efficiëntie van behandelingen te verbeteren hebben we inzicht nodig in de mechanismen die verantwoordelijk zijn voor de gewenste uitkomsten. Dit proces wordt mediatie genoemd. Afname van pijn-catastroferen en toename van ervaren interne controle over pijn worden genoemd als mogelijk belangrijke mediatoren binnen CBT, alhoewel hun rol binnen fysieke behandelvormen nooit is onderzocht. Om nu te onderzoeken of behandelingen, gebaseerd op verschillende theoretische modellen, pijn-catastroferen en ervaren interne controle over pijn veranderen en of deze veranderingen de uiteindelijk uitkomst van de behandeling mediëren werd een aanvullende analyse van 211 patiënten met aspecifieke lage rugpijn die deelnamen aan onze RCT, uitgevoerd. Patiënten namen deel aan APT (n = 52), CBT (n = 55) CT (n = 55) of WL (n = 49). Meteen na einde van de behandeling bleek de mate van pijn-catastroferen in alle drie de actieve behandelvormen te zijn afgenomen, terwijl in de WL geen verandering was opgetreden. In alle vier de behandelgroepen bleek geen verschil in de verandering van de ervaren interne controle over pijn te zijn opgetreden.

De afname in pijn-catastroferen medieerde de vermindering van ervaren beperkingen, patiëntspecifieke klachten en pijnintensiteit. Er moet wel in ogenschouw worden genomen dat we niet in staat waren om halverwege de behandelingen data te verzamelen en dus geen cross-lagged correlatie analyse konden uitvoeren. Derhalve bestaat er nog een zekere mate van ambiguïteit over de richting van causaliteit tussen pijn-catastroferen en de uitkomstmaten. Desalniettemin kunnen onze resultaten bijdragen tot de ontwikkeling van meer effectieve interventies, specifiek gericht op de vermindering van pijn-catastroferen.

Deze studie toont ook aan dat fysieke behandelvormen die niet specifiek gericht zijn op het beïnvloeden van cognitieve factoren toch verantwoordelijk zijn voor een substantiële cognitieve verandering. Mogelijk dat de ervaring dat activiteiten die normaal gesproken door de pijn achterwege werden gelaten, nog altijd mogelijk zijn, de interpretatie van pijn kan beïnvloeden.

Hoofdstuk 7

Geloofwaardigheid en verwachting van behandeling en chronische lage rugpijn

De initiële overtuigingen van een patiënt over het succes van de aangeboden behandeling hebben invloed of de uiteindelijke resultaten van die behandeling. Recentelijk werd aangetoond dat de Credibility/Expectancy Questionnaire (CEQ) een betrouwbaar instrument is om de geloofwaardigheid van en verwachting ten aanzien van een aangeboden behandeling te meten. De doelstellingen van deze studie waren: het bepalen van de factor structuur van de CEQ door middel van een confirmerende factor analyse (CFA) in de populatie van CLRP patiënten die deelnamen aan onze RCT; het onderzoeken van de relatie tussen geloofwaardigheid en verwachting en patiënt kenmerken; en het bepalen van de mate van invloed van geloofwaardigheid en verwachting op de uitkomst van behandelingen voor patiënten met CLRP. 167 patiënten die werden gerandomiseerd naar APT (n = 51), CBT (n = 57) of CT (n = 59), vulden de CEQ in nadat zij uitleg hadden gekregen over de achterliggende redenen (rationale) van de behandeling die ze hadden geloot. Door middel van de CFA werd de tweefactoren structuur (geloofwaardigheid/verwachting) van de CEQ bevestigd. Een lagere geloofwaardigheid bleek samen te hangen met hogere pijngerelateerde vrees en lagere ervaren interne controle over pijn. Een lagere verwachting bleek gerelateerd aan hogere pijngerelateerde vrees en het ontbreken van uitstralende pijn in been. Multipele lineaire regressie analyses lieten zien dat verwachting, gecontroleerd voor leeftijd, geslacht, type interventie, behandelcentrum, pijnintensiteit en duur van beperkingen voor alle drie de behandelingen significant van invloed was op de ervaren beperkingen en tevredenheid over behandeling. De omvang van deze invloed was echter gering tot matig. Geloofwaardigheid bleek significant van invloed te zijn op patiëntspecifieke klachten en tevredenheid over behandeling. Voor de uitkomstmaat ‘ervaren algemene verbetering’ bleek er sprake te zijn van effectmodificatie. Verwachting was alleen voorspellend voor APT en geloofwaardigheid was alhoewel matig, voorspellend voor CT. Geloofwaardigheid en verwachting voorspelden niet de verandering van pijn voor alle drie interventies.
Bij de behandeling van CLRP patiënten lijkt het veelbelovend om behandelvormen toe te passen die meer specifiek gericht zijn op het vergroten van zowel de geloofwaardigheid als de verwachting ten aanzien van de aangeboden behandeling.

Hoofdstuk 8

De bruikbaarheid van fysieke performance/capaciteit taken

Om meer ‘objectief’ de fysieke capaciteit van CLRP patiënten vast te stellen, gebruikten we in onze RCT zes fysieke performance/capaciteit taken. Om de objectiviteit en betrouwbaarheid van de metingen te vergroten, werden de protocollen van deze zes uit de literatuur geselecteerde taken verder aangepast en aangescherpt. Dit had tot gevolg dat de bestaande psychometrische data niet meer valide waren. Verder bleken er evenmin data beschikbaar over de invloed van de mate van ervaring met het uitvoeren van de taak. Tenslotte waren geen data beschikbaar over de mate van stabiliteit van de taken. Als maat van stabiliteit wordt gebruik gemaakt van de ‘limits of agreement’. Hiermee wordt de minimaal benodigde verandering gedefinieerd om er zeker van te zijn dat de waargenomen verandering tussen twee meetmomenten ook een daadwerkelijke verandering voor de individuele patiënt betreft en niet is toe te schrijven aan een meetfout. Derhalve voerden we een studie uit waarin een deel van onze RCT-populatie met een testinterval van 5 tot 9 dagen, twee maal de performance/capaciteit taken uitvoerde. In totaal werden 30 patiënten zonder enige ervaring met de taken en 23 patiënten die de taken al twee of drie maal hadden uitgevoerd, geïncludeerd.
De mate van ervaring met het uitvoeren van de taak bleek geen significante invloed te hebben op de test-hertest verschillen. Alle taken bleken een goede tot zeer goede test-hertest betrouwbaarheid te hebben (ICC variërend van 0.74 tot 0.99). De ‘limits of agreement’ uitgedrukt als het percentage van de gemiddelde score van de studiepopulatie, was laag tot matig voor vijf minuten lopen en een minuut traplopen (respectievelijk 21% en 20%), matig voor 50-foot snel lopen, zit-tot-stand en voorwaarts reiken (respectievelijk 33%, 29% en 36%) en hoog voor de tiltaak (Progressive Isoinertial Lifting Evaluation: 48%). We concludeerden derhalve dat de vijf minuten lopen en een minuut traplopen taak en in mindere mate de 50-foot snel lopen, zit-tot-stand en voorwaarts reiken taak bruikbaar zijn voor de dagelijkse praktijk. Echter, er bestaan grote bezwaren tegen de bruikbaarheid van de tiltaak.

Hoofdstuk 9

Invloed van persoonsgebonden en fysieke factoren op fysieke capaciteit

Alhoewel verondersteld wordt dat fysieke performance/capaciteit taken het functioneren van patiënten met CLRP meer objectief meten dan vragenlijsten, worden ook deze taken beïnvloed door cognitieve en emotionele processen. In een cross-sectionele studie werd voor 221 patiënten die deelnamen aan onze RCT de relatie tussen persoonsgebonden en fysieke functies en de score van elk van de zes fysieke performance/capaciteit taken voor de start van behandeling bepaald. Hiervoor werden multipele regressie analyses uitgevoerd. Meteen voor de start van de behandeling werden de gegevens over de onafhankelijke variabelen leeftijd, geslacht, pijnintensiteit, duur van pijn, mate van uitstralende pijn in been (3 categorieën), VO_2max, pijngerelateerde vrees, depressie, ervaren interne controle over pijn en pijn-catastroferen verzameld. Vanwege een hoog percentage van de patiënten waarvoor geen VO_2max kon worden berekend (20.8%), werd gebruik gemaakt van een multipele imputatie techniek.

Ondanks de inclusie van tien variabelen, bleek de totaal verklaarde variantie laag tot matig (9% tot 19%), met uitzondering voor traplopen dat een totaal verklaarde variantie van 30% werd gevonden. Cardiovasculaire capaciteit (VO_2max), pijnintensiteit, pijngerelateerde vrees, cognities en depressie bleken significant, maar klinisch nauwelijks relevant geassocieerd te zijn met enkele, maar niet alle performance/capaciteit taken. Uitstralende pijn in been, leeftijd en duur van klachten bleken niet significant geassocieerd te zijn met de prestatie op de performance/capaciteit taken. Ten gevolge van antropometrische verschillen presteerden mannen op de meeste taken beter dan vrouwen.

Blijkbaar is de invloed van vele persoonsgebonden, fysieke maar vooral psychologische factoren op de prestaties op de geselecteerde performance/capaciteit taken helemaal niet zo groot. Dit zou er op kunnen wijzen dat deze taken toch meer objectief de fysieke capaciteit meten dan verwacht.

Hoofdstuk 10

Algemene discussie

In dit laatste hoofdstuk wordt een uitgebreide samenvatting gegeven van de resultaten en de conclusies en wordt getracht de resultaten van alle in dit proefschrift gepresenteerde studies te integreren. De belangrijkste bevinding van dit proefschrift is dat de gecombineerde behandeling niet effectiever is dan de afzonderlijke behandelingen, zijnde de fysieke en de cognitief-gedragstherapeutische behandeling. Verder blijkt de effectiviteit van de fysieke behandeling te worden gemedieerd door een afname van pijn-catastroferen. Ondanks het bestaan van aërobe deconditionering (verlaagde VO_2max) in onze populatie van CLRP patiënten, kon een vermindering van beperkingen worden bereikt zonder dat de cardiovasculaire capaciteit verbeterde. De geloofwaardigheid van en verwachting ten aanzien van de aangeboden behandeling lijken belangrijke prognostische factoren bij de behandeling van patiënten met beperkingen ten gevolge van CLRP, hoewel de associatie met het uiteindelijke behandelresultaat gering tot matig is. Meerdere performance/capaciteit taken zijn bruikbaar voor de dagelijkse praktijk en de invloed van verschillende persoonsgebonden en fysieke functies is slechts beperkt.

Vervolgens worden methodologische aspecten, klinische implicaties en suggesties voor toegepast klinisch onderzoek, implementatie aspecten en tot slot, suggesties voor toekomstig onderzoek beschreven.

Dankwoord.

Dankwoord

Dankwoord

Het schrijven van een ogenschijnlijk gemakkelijk hoofdstuk kostte mij toch nog de nodige moeite. Vooral omdat ik niemand te kort wil doen. Zeker bij een multicentre studie zoals ik die heb uitgevoerd is dit niet ondenkbeeldig. Mijn welgemeende excuses voor diegene die ik onverhoopt vergeten mocht zijn!
Een zeer belangrijk project, zeg maar rustig levensproject, van bijna 6 jaar wordt nu afgesloten en ik heb dan ook zeer veel geleerd en niet in de laatste plaats over mijzelf. Velen hebben hier, ieder op hun manier aan bijgedragen.

Allereerst wil ik mijn José bedanken. José, jij staat al ruim 25 jaar aan mijn zijde en samen hebben we 6 jaar geleden besloten dat ik aan dit promotietraject zou beginnen. Dit project kan ik het beste vergelijken met het rijden van een Tour de France. Alleen voerde deze tocht veel verder dan enkele duizenden kilometers door het mooie Brabantse en Limburgse land. Het tempo varieerde sterk en enkele keren had ik de indruk dat ik geen meter verder kwam en in kringetjes bleef rijden. Soms ging ik door mist en regen, maar met name de laatste drie jaar brak de zon steeds vaker door. Samen reden we door enkele dalen maar we klommen altijd weer tot mooie hoogtepunten met dan nu de werkelijk prachtige top. José, je bent voor mij in de afgelopen jaren dan niet alleen de waterdrager, knecht (hield me uit de wind), soigneur, mecanicien maar ook mijn slapie, maatje en soms zelfs ploegleider geweest. Ik ben je dan ook zeer erkentelijk voor de gegeven steun en je onvoorwaardelijke liefde. Ik ben gelukkig niet een promovendus die nadat hij het beoogde doel heeft bereikt, opeens bemerkt dat hij het mooiste en allerdierbaarste in zijn leven is kwijtgeraakt.

Dit promotietraject kwam tot stand toen ik in 2000 een subsidieaankondiging zag voor klinisch werkzame revalidatieonderzoekers. Ik was reeds gestart met de SGO-wetenschappelijke vorming voor revalidatieartsen en wilde dolgraag zelf wetenschappelijk onderzoek beginnen. Na enkele verkennende gesprekken kwam ik terecht bij Geert van der Heijden, Johan Vlaeyen en Ludwig Goëken. Het idee om een multicentre trial op te starten om de effectiviteit van enkelvoudige behandelingen en een multidisciplinaire behandeling te vergelijken werd de leidraad voor de subsidieaanvraag. In die fase werd Alita Hidding bij het project betrokken. Door het ontbreken van een hoogleraar revalidatiegeneeskunde aan de UM, kwamen we al spoedig bij André Knottnerus terecht. Als hoogleraar Huisartsgeneeskunde draagt hij de revalidatie een zeer warm hart toe en toonde hij zich onvoorwaardelijk bereid om mijn promotor te zijn. Voortvarend werd de aanvraag geschreven en ook gehonoreerd. Vervolgens gaven mijn werkgever, het revalidatiecentrum Blixembosch in de persoon van directeur Vincent Buitendijk, alsmede mijn collega revalidatieartsen, mij volledige steun en werd ik voor een periode van 6 jaar gedurende de helft van mijn werktijd vrijgesteld voor het uitvoeren van dit onderzoek. Waarvoor veel dank!

Dan nu mijn beide promotores.
Beste André, “een beter promotor kan een promovendus zich niet wensen” was een veel gehoorde opmerking. Ik kan en wil dit alleen maar bevestigen. Je hebt mij op een zeer professionele en diplomatieke wijze en met scherp

oordeelsvermogen begeleid. Hierbij toonde je tevens een welgemeende belangstelling voor mij als mens. Nadat de trial was opgestart en de wetenschappelijke output meer en meer aan de orde kwam, nam jij zelfs de dagelijkse begeleiding op je. Dit deed je er zomaar even bij ondanks al je andere drukke werkzaamheden. Ook al had je een keer onvoldoende tijd om stukken te bestuderen, je toonde je een absolute meester in het binnen een zeer kort tijdbestek weer oppakken van de essentie van de voorliggende vraagstukken. Je coaching t.a.v. publicaties bleek van grote waarde. Zelfs nu het proefschrift klaar is ben je zeer actief betrokken bij het verder uitstippelen van mijn carrière. Ik hoop dan ook nog vele jaren met je te mogen samenwerken.

Beste Johan, ook jou ben ik veel dank verschuldigd. Onze samenwerking startte al in 1988, toen ik als student mocht meewerken aan jouw promotieonderzoek. Het is voor mij dan ook een zeer grote eer dat jij nu mijn tweede promotor bent. Nadat de ontwikkeling van de interventies en de keuze van meetinstrumenten waren afgerond, koos je ervoor om wat meer op de achtergrond aanwezig te zijn. Echter, met name bij de uitwerking van de trial resultaten en enkele belangrijke deelstudies, toonde jij je kwaliteiten van een zeer kundig onderzoeker en begeleider. Ook je redactionele vaardigheden waardeer ik zeer. Ik weet dat je recent voor de moeilijkste keuze t.a.v. je wetenschappelijke carrière hebt gestaan. Ik ben zeer verheugd dat je deels behouden blijft voor de UM. Ik ga er dan ook vanuit dat we even plezierig blijven samenwerken bij het in de praktijk toetsen van de door jou ontwikkelde theoriegebaseerde behandelvormen.

Dank ook aan mijn copromotor Alita. In de beginfase was onze samenwerking intensief. De geslaagde vormgeving van de behandelingen, trial, meetinstrumenten en strategie t.a.v. het werven van deelnemende revalidatiecentra en patiënten, is dan ook mede aan jou te danken. Samen hebben we een gedegen opleidingsplan samengesteld, waardoor ik een behoorlijke bagage als onderzoeker heb meegekregen. De laatste jaren was je niet meer als dagelijkse begeleider actief, maar wel betrokken bij het redigeren van enkele publicaties en zat je de projectgroepvergaderingen op zeer professionele wijze voor.

Ook wil ik de andere projectgroepleden bedanken: Ludwig Goëken, Geert van der Heijden, Arnold Kester, Frederieke Berendsen en Saskia Beelen.
Beste Ludwig, gedurende de beginfase van het project heb je een belangrijke bijdrage geleverd aan het vormgeven van de trial. Gezien je naderende emeritaat en je gezondheid besloot je geen promotor meer te zijn. Ik wil je bedanken voor je inzet en voor het op de kaart zetten van het rugonderzoek binnen de revalidatie. Zelfs nu, enkele jaren nadat je bent gaan genieten van je vrije tijd, blijkt hoe groot jouw invloed is geweest, ook al hebben velen dat nooit zo gezien en heb je jezelf daar ook nooit op laten voorstaan. Deze bescheidenheid is in mijn ogen niet terecht, maar siert jou als persoon. Je hebt je gedurende de afgelopen jaren steeds over de voortgang laten informeren. Ik respecteer je beslissing om in de luwte te blijven maar bij de promotieplechtigheid zul je wel in mijn gedachten aanwezig te zijn.
Beste Geert, je leek af en toe wel de ‘luis’ in mijn onderzoekerspels. Iedere keer als ik dacht dat we een beslissing hadden genomen over methodologische en statistische vraagstukken, slaagde jij erin twijfel te zaaien. Jouw opmerkingen

leidden ertoe dat ik steeds kritischer en gedrevener raakte om zaken nog beter uit te zoeken. Dit hele proces heeft er toe bijgedragen, dat ik nu veel beter kan beargumenteren waarom we bepaalde keuzes hebben gemaakt. Bedankt voor je bijdrage aan mijn methodologische vorming.
Beste Arnold, je raakte iets later bij de projectgroep betrokken doordat je collega Hubert Schouten ziek werd. Je hebt mij met je ontwapende nuchterheid en heerlijk 'droge' opmerkingen, zeer geholpen bij alle statistische problemen. Je deed je voorletters A.D.M, zoals je zelf zei: "Amsterdamse Droogdok Maatschappij", eer aan en trok mij af en toe op het droge. Geen enkele vraag was je teveel, zelfs niet in het weekend. Je opmerking "uiteindelijk wil je toch alles weten", bleek inderdaad zeer vaak van toepassing. Ik hoop je in de toekomst je nog regelmatig te mogen lastig vallen ook al ben je begonnen met de aanloop tot de vroege pensionering.

Dan mijn beide onderzoeksassistenten Frederieke Berendsen en Saskia Beelen. Zonder jullie was dit alles echt niet gelukt, jullie waren echt super!
Beste Frederieke, jij hebt in de beginfase van het onderzoek door je zeer kundige en uiterst tactvolle aanpak enorm veel krediet opgebouwd bij zowel de medewerkers in de centra als bij de deelnemende patiënten. Nog steeds hoor ik jouw naam wanneer ik met deze mensen spreek. Ik was dan ook zeer bevreesd dat de trial in gevaar zou komen toen je mij vertelde dat je als teamleider binnen Blixembosch was aangenomen. We hebben er bij de verantwoordelijk manager hard voor gepleit dat je voldoende tijd kreeg om de zaken over te dragen. Dit is uiteindelijk zeer goed gelukt. Het doet me veel plezier te zien dat je als teamleider uitermate goed functioneert. Gelukkig vond ik Saskia bereid jouw taken over te nemen.
Beste Saskia, aangezien het onderzoek geleidelijk aan in een andere fase terecht kwam werden andere eisen aan jou gesteld. Je hebt op voortreffelijke wijze het vertrouwen dat Frederieke reeds had gekweekt weten te behouden en soms nog verder uitgebouwd. Verder heb je zeer kritisch meegedacht over onderzoeksvragen en verwerking van data en heb je er hard aan getrokken om de dataset zo volledig mogelijk te maken. Je begeleidde op een prettige en professionele manier onze stagiaires. Ook bleek je te beschikken over goede redactionele kwaliteiten en had je een belangrijk aandeel in een deelstudie. Ik ben dan ook blij dat we in het kader van het project Zorgmonitoring en de afronding van Lobadis blijven samenwerken.

Zeer zeker wil ik Ria Laros niet vergeten. Beste Ria, gelukkig kon ik jou vrij snel na de start van mijn onderzoek interesseren voor een zeer belangrijke klus: de controle en invoer van data. Dit werk heb je geheel vrijwillig en met grote zorgvuldigheid en motivatie gedaan. Dagelijks mag ik hiervan nog de vruchten plukken. Echt super Ria! Verder bleek je zeer vaardig met het ontwerpen van bedank- en attentiekaartjes en verzorgde je de lay-out van onze mooie nieuwsbrief. Ook je echtgenoot Jan werd erbij betrokken en maakte professionele videobeelden van de behandelingen en metingen. Uiteraard mag de uitwisseling van onze vakantie-ervaringen niet worden vergeten.

Een speciaal woord van dank aan Albérè Köke en Anja van den Hout die bereid waren om respectievelijk de graded activity en probleemoplossende vaardigheidstraining verder vorm te geven en behandelaars te trainen.

Dank wil ik ook uitspreken naar de medeauteurs van buiten de begeleidingsgroep.
Harriet Wittink, jouw inbreng was van groot belang bij mijn eerste publicatie. Discussiëren met jou werkt stimulerend en je enthousiasme is waarlijk aanstekelijk.
Peter van Leeuwen, ook de collega met wie ik een goed functionerend behandelteam mag aansturen. Beste Peter, bedankt voor het uit de wind houden en het luisterend oor, je bent niet voor niets een van mijn paranimfen!
Mariëlle Goossens en Erik Schouten, voor jullie specifieke deskundigheid bij de deelstudie "verwachtingen en geloofwaardigheid van behandeling".
Helma Hijdra, in het kader van je opleiding tot revalidatiearts binnen ons opleidingscircuit HET nam je een groot deel van het test-hertest onderzoek van de performance taken voor je rekening. Minou Hitters, je droeg als hoofdopleider hieraan je steentje bij.
Tineke van Geel, in het kader van je opleiding Gezondheidswetenschappen voerde je jouw afstudeeropdracht binnen mijn onderzoek uit. Dit verliep zo voorspoedig dat je medeauteur van "het korte termijnresultaten"-artikel werd. Er bleef zelfs tijd over om de interessante deelstudie naar invloed van persoonsgebonden en fysieke factoren op fysieke capaciteit op te starten en te publiceren.
And last but not least, I want to thank Derick Wade. During this research project I have had the privilege to discuss the proceedings of my research with you. You advised me to write down my thoughts and findings about the literature regarding the effectiveness of exercise therapy and deconditioning. Eventually, this resulted in our hypothesis-oriented review. Furthermore, I am very grateful for you willingness to review several articles before submitting them to journals. I hope that we can keep on collaborating together!

Ik bedank de beoordelingscommissie bestaande uit Rob de Bie, Jan Geertzen, Stephen Morley, Trudy van der Wijden en Jeanine Verbunt voor de tijd die zij hebben genomen om mijn proefschrift te lezen en te beoordelen.

Een groot woord van dank wil ik richten aan de directies, medische staven en medewerkers van de participerende revalidatiecentra. De revalidatiecentra Breda, Leypark te Tilburg en Blixembosch te Eindhoven hebben hun nek durven uitsteken daar, waar meerdere revalidatiecentra het toch niet aandurfden. Deze drie Brabantse centra verdienen in mijn ogen een zéér groot compliment omdat zij, ondanks het gebrek aan ervaring met een dergelijk grote trial, het lef hebben gehad om samen met mij dit onderzoek aan te gaan. Dit getuigt van een duidelijke visie en toont overduidelijk aan dat er wel degelijk revalidatiegeneeskundig onderzoek buiten de geijkte academische setting mogelijk is.

Uiteraard wil ik alle behandelaars, testassistenten, planners en managers van de deelnemende revalidatiecentra die heel veel werk hebben uitgevoerd niet vergeten: Rudy, Monique, Sacha, Henk, Nick, Claire, Mark, Rien, Sabrina, Lisette, Loes, Anneke, Liesbeth, Theo, Jerome, Lia, Marjan, Maria, Marjolein, Annelieze, Rita, Iris, Rian, Jeanette, Mark, Gust, Herman, Jeroen, Paul, Wilbert, Leonie, Fenne, Nellie, Sjan, Hanneke, Jeanne, Aukje, Hanneke, Ideke, Annemiek, Kristel, Letty, Maartje, Stan, Mariken, Carlijn, Ilse, Marian, Esther, Dianne, Jeroen, Marieke, Christian en Willy-Jan. Gina, mijn medisch secretaresse, jou wil ik special bedanken omdat je mij tijdens mijn onderzoekstijd op voortreffelijke wijze afschermde van taken die daarin niet thuishoorden, geweldig!

Verder alle revalidatieartsen die patiënten verwezen voor deelname en speciaal Peter van Leeuwen, Mariëtte Schreibers, Gijs Kuijpers, Natascha Ringeling en Nieke Rijnders die als revalidatiearts eindverantwoordelijk waren voor de behandelingen.
Dank aan alle huisartsen die bereid waren mijn wervingscampagne aan te horen en vervolgens patiënten verwezen voor een revalidatie- en mogelijk onderzoekstraject. Speciaal wil ik, zonder andere collega huisartsen te kort te willen doen, Heinz Lövenig bedanken die Lobadis in de regio Tilburg, mede op de kaart heeft gezet.

Alle medewerkers van de facilitaire dienst, automatisering, medisch secretariaat, EAD en personeelszaken van Blixembosch, Harrie Kusters, Ben Eisermann, Robert Horbach, Jos Slangen en niet te vergeten de super secretaresses Hanny Prick en Kim Luyten. Bedankt voor de logistieke en praktische ondersteuning. Caroline van Heugten bedankt voor het uitwisselen van onderzoeksideeën en het soms gewoon, lekker samen even frustraties delen.

Een apart woord voor de onafhankelijk arts Hans Maarten Kruitbosch. Beste Hans Maarten, hoewel je geen enkele vraag van patiënten kreeg, ben jij voor mij van grote waarde geweest. Jij gaf mij in 2000, ten koste van meer werk voor jezelf, de ruimte om mijn onderzoeksaspiraties waar te maken. Bedankt voor de al jaren durende vriendschap. Ook jij bent niet voor niets een van mijn paranimfen!

De allerbelangrijkste partners voor het doen van onderzoek zijn echter de patiënten met chronisch lage rugpijn. Jullie wil ik speciaal bedanken voor de bereidheid om aan het onderzoek deel te nemen en het trouw verschijnen voor de metingen en het invullen van de kostendagboeken. De onderzoeksresultaten hebben geleid tot verdere kennisontwikkeling, hoewel ik mij realiseer dat ze meestal niet meteen van invloed zijn op jullie eigen situatie.
Ook wil ik de Stichting Patiëntenbelangen Orthopaedie in de persoon van Linda Riemens bedanken voor de prettige samenwerking.

Vele andere onderzoekers met wie ik tijdens dit traject zinvol heb mogen discussiëren. De Lobadis- en WIRR-leden: Jeanine, Sandra, Wietske, Bas, Maaike, Petra, Emanuel, Jan, Ant, Jos, Daan, Rita, Annemarijke, Miriam, Marli, Marije, Michiel, en de FADIS-gangers: Ruud, Jacques, Camiel, Judith, Jeroen, Ilse, Peter, Eric, Sita, Madelon, Jeffrey.

Paula Rinkens, voor de begeleiding bij het mij eigen maken van SPSS. Het ontwerpen en controleren van data en syntaxen bleek door de samenwerking met jou helemaal niet saai!

Hilde Sielhorst en Anita Legtenberg van het MEMIC voor de keurige verwerking van de vragenlijsten, hoewel een aantal vragenlijsten in een verhuiscontainer ons wel enkele slapeloze nachten bezorgden. Uiteindelijk is alles goed gekomen!

Dit project had nooit tot een goed einde gebracht kunnen worden als ik niet de nodige fysieke inspanning en sociale afleiding had mogen ervaren. Vooral de jongens en meiden van de mountain bike groep TIONA (Toon, Trees, Theo, Jos, Harrie, Henry, Hans, Yolanda, Marfa, Diana en John) en onze vele trouwe vrienden en familieleden (gelukkig te veel om allemaal op te noemen!) hebben dit mogelijk gemaakt.

Uitermate veel dank spreek ik uit naar mijn ouders. Pa en ma, ik vind het echt geweldig zoals jullie, ondanks de zware last die jullie toebedeeld kregen toen zowel Wil als ik een familiaire spastische paraplegie bleken te hebben, erin zijn geslaagd ons een zeer goede en vooral liefdevolle opvoeding te geven. Wij staan ons mannetje en jullie hebben voortreffelijk werk geleverd. Ik ben trots op jullie!

About the author.

About the author

About the author

Rob Smeets was born August 15th, 1964 in Sittard, the Netherlands. He grew up in Elsloo. He attended high school at the Scholengemeenschap St. Michiel in Geleen and obtained his Gymnasium ß diploma in 1982. He then started his medical training at the Maastricht University and graduated in February 1989 (met genoegen). From March 1989 till March 1990 he worked at the department of internal medicine of the Catharina Hospital in Eindhoven. From March 1990 until March 1991 he worked at the pain and rheumatology department of the Rehabilitation Centre Hoensbroek after which he started as a trainee in rehabilitation medicine at the Rehabilitation Foundation Limburg. He graduated in March 1995, and ever since has been working as a consultant in rehabilitation medicine at the Rehabilitation Centre Blixembosch in Eindhoven. His main topic is the multidisciplinary treatment of patients with chronic musculoskeletal pain disorders. In 2001 he finished his training as a researcher in rehabilitation medicine (VRA-SGO wetenschappelijke vorming voor revalidatieartsen). In 2001 he was also awarded funding from ZonMW (klinische onderzoeksplaats revalidatiegeneeskunde) and started his research, resulting in this thesis.
At the moment he is an active member of three committees concerning the multidisciplinary treatment of chronic (low back) pain: Dutch chapter of Consultants in Rehabilitation Medicine in Chronic Pain (WPN), Dutch collaboration for the implementation of low back pain rehabilitation research (Lobadis-WIRR), and the consensus group on core outcome measures for chronic pain rehabilitation (OPR-VRA klinimetrie werkgroep). From 2002 to 2004, he was a member of the committee of the Dutch consensus project Pain Rehabilitation (Consensus Pijnrevalidatie Nederland). He is a member of the Special Interest Group of Pain and Movement of the IASP. In 2005 he was involved as an advisor for the Dutch revised guidelines for General Practitioners regarding sciatica, and just recently has become a member of a task force for the Dutch government (Gezondheidsraad), designing guidelines for physicians responsible for awarding disability pensions to people suffering from sciatica. Since 1987 he is happily married to José Spronkmans.