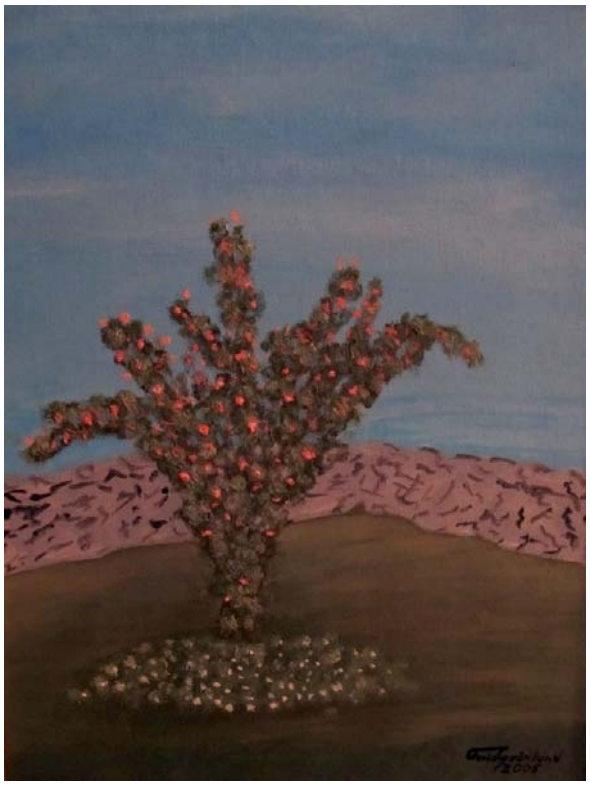
STUDIES IN SPORT, PHYSICAL EDUCATION AND HEALTH
173

Markku Paatelma

Orthopedic Manual Therapy on Low Back Pain with Working Adults

Clinical Tests, Subclassification and Clinical Trial of Low Back Pain



JYVÄSKYLÄN YLIOPISTO

STUDIES IN SPORT, PHYSICAL EDUCATION AND HEALTH 173

Markku Paatelma

Orthopedic Manual Therapy on Low Back Pain with Working Adults

Clinical Tests, Subclassification and Clinical Trial of Low Back Pain

Esitetään Jyväskylän yliopiston liikunta- ja terveystieteiden tiedekunnan suostumuksella
julkisesti tarkastettavaksi yliopiston Agora-rakennuksen auditoriossa 1
lokakuun 1. päivänä 2011 kello 12.

Academic dissertation to be publicly discussed, by permission of
the Faculty of Sport and Health Sciences of the University of Jyväskylä,
in Agora-building, Auditorium 1, on October 1, 2011 at 12 o'clock noon.

UNIVERSITY OF JYVÄSKYLÄ

JYVÄSKYLÄ 2011

Orthopedic Manual Therapy on Low Back Pain with Working Adults

Clinical Tests, Subclassification and Clinical Trial of Low Back Pain

STUDIES IN SPORT, PHYSICAL EDUCATION AND HEALTH 173

Markku Paatelma

Orthopedic Manual Therapy on Low Back Pain with Working Adults

Clinical Tests, Subclassification and Clinical Trial of Low Back Pain

UNIVERSITY OF JYVÄSKYLÄ

JYVÄSKYLÄ 2011

Editors
Harri Suominen
Department of Health Sciences, University of Jyväskylä
Pekka Olsbo, Ville Korkiakangas
Publishing Unit, University Library of Jyväskylä

Cover picture: Carelian Maid by Terho Grönlund, 2005

URN:ISBN:978-951-39-4436-0
ISBN 978-951-39-4436-0 (PDF)

ISBN 978-951-39-4428-5 (nid.)
ISSN 0356-1070

Jyväskylä University Printing House, Jyväskylä 2011

Dedicated to my grandchildren Arttu and Leo

“Let us spend more time together.”

- Markku -

ABSTRACT

Paatelma, Markku
Orthopedic manual therapy on low back pain with working adults; clinical tests, subclassification and clinical trial of low back pain
Jyväskylä: University of Jyväskylä, 2011, 98 p.
(Studies in Sport, Physical Education and Health
ISSN 0356-1070; 173)
ISBN 978-951-39-4428-5 (nid.)
ISBN 978-951-39-4436-0 (PDF)

Although the natural history of low back pain (LBP) is considered to be good, most sufferers have further episodes during the first year. Debate continues on the comparative effectiveness of advice on bed rest and staying active as part of the primary care management to avoid a vicious circle, preventing LBP from becoming chronic in its early phase (of < 12-week duration). LBP in its early phase is also a common reason to consult a physiotherapist (PT).

The objective of this dissertation was to evaluate the intertester and intratester reliability of selected clinical tests between PTs to assess LBP patients in the early phase, and to discover the sensitivity and specificity of these tests in acute/subacute and chronic LBP patients, and in a group of no "patient status" subjects. A further objective was to evaluate the intertester reliability of a pathoanatomical/pathophysiological classification by general practitioners in primary care physiotherapy compared to findings classified by a specialist in orthopedic manual therapy (OMT) and the findings of two OMT specialists when examining patients with early-phase LBP using selected clinical tests. The objective was also to compare effectiveness against low back and leg pain, disability, and sick leave of OMT physiotherapy using a pathoanatomical/pathophysiological classification to the effectiveness of the McKenzie method and "Advice" only to stay active.

The inter- and intratester reliability of clinical tests seemed to be at an acceptable level globally, although tests had both high and low reliability among test categories. Intratester reliability was slightly better than intertester reliability.

Of all the selected clinical tests, few tests were moderately sufficient in sorting the chronic low back pain patients (CLBP) from subacute low back pain patients (SLBP), and distinguishing patient groups from controls. These 31 clinical tests quite poorly sorted the CLBP or SLBP patients from the controls. It may be possible that a combination of tests may enhance the sensitivity and specificity in sorting the CLBP or SLBP patients from the controls.

Subclassification into clinical subgroups is reliable in the two most common subgroups also with PTs without specialization in OMT after a short post-graduate training. No further conclusion could be drawn from other subgroups because of the low number of subjects in these categories.

The OMT and McKenzie methods, compared to Advice-only to stay active for low back and leg pain, and disability showed no significant difference in effectiveness during a one-year follow-up of working adults. However, the OMT and McKenzie groups showed positive treatment effects compared to the Advice-only group in form of the smaller number of days of sick leave because of LBP. Regarding sick leave days, there was also a statistically significant difference between OMT and McKenzie groups in favor of OMT-physiotherapy.

Keywords: orthopedic manual therapy, McKenzie method, clinical tests, sub-classification, low back pain in early phase, randomized controlled trial

Author`s address

Markku Paatelma
Department of Health Sciences
University of Jyväskylä
P.O. Box 35
40014 Jyväskylä, Finland
markku.paatelma@auron.fi

Supervisors

Professor Ari Heinonen
Department of Health Sciences
University of Jyväskylä
P.O. Box 35
40014 Jyväskylä, Finland

Professor Tapio Videman
Faculty of Rehabilitation Medicine
University of Alberta
Edmonton, Canada

Reviewers

Docent Jari Ylinen
Physical Medicine and Rehabilitation
Central Hospital of Central Finland
Jyväskylä, Finland

Professor Jenny Keating
Faculty of Medicine Nursing
Department of Physiotherapy
Victoria, Australia

Opponent

Docent Markku Kankaanpää
Physical Medicine and Rehabilitation
Tampere University Hospital
Tampere, Finland

ACKNOWLEDGEMENTS

This dissertation was carried out in the Department of Health Sciences, University of Jyväskylä during the years 1998-2011. A large number of people have contributed to this work in many different ways and I wish to express my sincere gratitude to all of them.

I first like to thank my teachers in Norway, physiotherapists Freddy Kaltenborn and Olaf Evjenth, founders and creators of Orthopedic Manipulative Therapy. They showed me a way in physiotherapy already in early 80th and inspired me to start this thesis. Without their support, Orthopedic Manipulative Therapy would have raised the head much later.

I would like to thank my supervisors, professors Ari Heinonen, PT, PhD, Tapio Videman, MD, PhD. I am deeply grateful to my principal supervisor Ari Heinonen for his constant encouragement and advice during the course of this project, being available almost always.

I want to express my gratitude to official reviewers of this thesis, Docent Jari Ylinen, MD, PhD, and professor Jenny Keating, PT, PhD. They have had very clear comments and constructive criticism. I am truly indebted and thankful for Jari Ylinen for his time to go through the thesis from page by page. Through his advice this thesis improved a lot.

I owe sincere and earnest thankfulness to my co-authors Markku Alèn, MD, PhD, Riitta Simonen, PT, PhD, and especially to Sinikka Kilpikoski, PT, PhD, who finished her thesis last year, for her tireless input for preparing the fourth article. In the second article, Jouko Heiskanen, PT, MD, was a strong generator not only in that paper, but during this whole process, putting a theory to practice.

The dissertation would not have been possible without my beloved wife Eira Karvonen, PT, MSc, who was not only one of the co-authors in I-III articles, but far more. She has helped me professionally and intellectualmentally, supported me not to stop, she has helped my financially, and she has taken care of our house, children and grandchildren for these years.

My special thanks go to my very good friend Kari Koskikare, Orthopedic Surgeon, MD, who has trained me physically every week in squash court, and given me extra knowledge of low back pain, how to examine, how to treat, and especially when to consult.

I am obliged to many of colleagues, especially to Martti Palo, Keijo Pitkämäki, and Mikael Falck, who supported me with taking my tasks, helped me with patients, and divided supervision with me in OMT class. I would like to show my gratitude also to physiotherapists in Palokka Health Care center, who participate in study III, and Jari Rautiainen, PT, for making possible to carry out the study I in his premises.

I would like to thank Jukka-Pekka Kesonen, MSc, and Hannu Kautiainen, MSc, for analyzing data statistically. I also thank the Physiotherapist Association in Finland for funding part of my statistical analysis in study IV,

and travel costs to WCPT Congress in Vancouver to present one of the articles, and also to Finnish Association for Manual Therapy for sharing several travel costs for international seminars. Thanks go to Finnish Work Environment Fund, whose support helped me to take some weeks free to prepare these articles.

Further, I warmly thank my children Liisa and Vesa and their spouses for never-ending support, and grandchildren Arttu and Leo for their understanding of their grand-father's travels and the time spent in front of computer during week-ends. I feel also sorry that my mother, who had so long awaited the final input, passed away last November, not having possibility to shear my happiness.

Last but not least my grate gratitude goes to Carolyn Norris, PhD, for checking the content and author-editing my language, and for given me extra-ordinary important tips for 20 years on how to write.

Lahti 11th July 2011 Markku Paatelma

ABBREVIATIONS

ASLR	Active straight leg raise
CI	Confidence interval
CLBP	Chronic low back pain
ES	Effect size
ITT	Intension-to-treat-analysis
K	Kappa coefficient
LBP	Low back pain
MCIC	Minimally clinically important change
McK	McKenzie Classification
MSI	Movement System Impairments
OMT	Orthopedic manual therapy
OR	Odds ratio
PAP	Patho-anatomical / -physiological Classification
PKB	Prone knee bend
PPPP	Posterior pelvic pain provocation
QTFC	Quebec Task Force Classification
RCT	Randomized controlled trial
RMQ	Roland-Morris disability index
ROM	Range of motion
SD	Standard deviation
SLBP	Subacute low back pain
SLR	Straight leg raise
SLUMP	Neurodynamic lumbar spine provocation test
TBC	Treatment-Based Classification
TrA	Transverses abdominis
VAS	Visual analogue scale

LIST OF ORIGINAL PUBLICATIONS

The thesis is based on the following articles, which are referred to in the text by their Roman numerals:

I Paatelma M, Karvonen E, Heinonen A. 2010. Inter- and intratester reliability of selected clinical tests in examining patients with early phase lumbar spine and sacroiliac joint pain/dysfunction and dysfunction. Advances in Physiotherapy 12, 74-80.

II Paatelma M, Karvonen E, Heiskanen J. 2009. How do clinical test results differentiate chronic and subacute low back pain patients from ”non-patients”? Journal of Manual & Manipulative Therapy 17, 11-19.

III Paatelma M, Karvonen E, Heinonen A. 2009. Intertester reliability in classifying acute and subacute low back pain patients into clinical subgroups. A comparison of specialists and non-specialists, a pilot study. Journal of Manual & Manipulative Therapy 17, 221-229.

IV Paatelma M, Kilpikoski S, Simonen R, Heinonen A, Àlen M, Videman T. 2008. Orthopaedic manual therapy, McKenzie method or Advice only for low back pain in working adults: a randomized controlled trial with one year follow-up. Journal of Rehabilitation Medicine 40, 858-863.

CONTENTS

1 INTRODUCTION

In Finland, low back pain (LBP) and mental depression are the two most common complaints leading to work loss and early retirement (Kääriä et al. 2005, 2009, Näslindh-Ylispangar et al. 2008). LBP is a common condition in adults even in their 30s (Shiri et al. 2010). While the natural history of LBP is often considered to be good, many patients suffer recurrent episodes with consequences for their well-being as well as for their quality of life (Breivik et al. 2006, Waddell and Burton 2007).

According to recent guidelines, the primary goal in LBP rehabilitation is an active approach, resuming normal activities and restoring function (Hayden et al. 2005, Henchoz_and Kai-Lik. 2008), although trials of different treatment modalities have failed to determine what strategy is optimal, and no single intervention is likely to be effective in treating LBP, thus far (Airaksinen et al. 2006). European guidelines recommend in early phase to take a case history and make a clinical examination. If history-taking indicates possible serious spinal pathology or nerve root syndrome, carrying out a more extensive physical examination including neurological screening is recommended (van Tulder et al. 2006). Review of psychosocial factors, and reassessment of those patients who are not resolving in few weeks, or those who are on a worsening course, are indicated (van Tulder et al. 2006).

A common theme is that acute LBP should be managed in primary care because it is generally benign: recovery tends to be both rapid and complete, and the few cases of serious disease can be readily detected with a clinical assessment (Koes et al. 2001). It has been shown that subacute LBP patients can be treated successfully with an approach that includes a clinical examination and patient information concerning the nature of the problem to reduce fear and to motivate them to resume light activity. This has an effect on work absenteeism (Indahl et al. 1995, Karjalainen et al. 2004). Similar effects have not been reported in treating chronic LBP patients. But the notion that acute LBP has a favorable prognosis, a view common to all guidelines, should be reconsidered because of the inconsistency in the outcomes reported and the lack

of long-term follow-up data. Acute LBP may not be a benign, self-limiting condition (Henschke et al. 2008).

Physiotherapy has long been the choice for non-surgical management for LBP patients. According to the World Confederation of Physical Therapy (WCPT), physiotherapy is concerned with identifying and maximizing movement potential, with regard to prevention, rehabilitation, and treatment (TCSO. P, 2002). When physiotherapists and patients with LBP are interacting, clinical evaluation and expertise (clinical reasoning) should be used together with the best current evidence, in finding a treatment strategy that provides effective pain treatment and good function in the musculoskeletal system. The aim of that treatment strategy is not only to cure current episodes, but also to prevent future recurrences of disabling pain. (Abenhaim et al. 2000, Krismer and van Tulder 2007, Waddell et al. 2007).

Recommendations for assessment of LBP emphasize the importance of ruling out potentially serious spinal pathology, finding specific causes of LBP, and tracing neurologic involvement. Recommendations include also identifying risk factors for chronicity and measuring the severity of symptoms and functional limitations, through the history, and physical and neurologic examination. Recommendations for management of acute LBP emphasize patient education, with short-term use of acetaminophen, non-steroidal anti-inflammatory drugs, or diclofenac, paracetamol, or spinal manipulation therapy (Hancock et al. 2007, Dagenais et al. 2010). For chronic LBP, the addition of back exercises, behavioral therapy, and short-term opioid analgesics is suggested. Clinicians who care for patients with LBP should endeavor to adopt these recommendations to improve patient care (Dagenais et al. 2010).

The validity and reliability of clinical tests, which are the basis for subgroup classification, and further the treatment strategies of LBP patients, have still conflicting evidence. Identification of subgroups of LBP has been a focus of major research. Several authors suggest that because LBP is a benign problem, emphasis should be on clinical tests and assessments. LBP should not be viewed as a homogenous condition, and treatment outcomes can be improved when sub-grouping is used to guide treatment decision-making (McKenzie and May 2003, O'Sullivan 2005, Brennan et al. 2006, Fritz et al. 2007)

Very few studies address cultural issues, highlighting the lack of information on the impact of specific cultural factors on LBP classification procedures. The main purpose of the present thesis was to find which tests are reliable, and which tests/test batteries can be recommend when sub-classifying LBP in its early phase (i.e., <12 week duration) for working adults in primary care in Finnish population, with aim of preventing new LBP episodes, which was tested in randomized controlled trial (RCT).

2 REVIEW OF THE LITERATURE

2.1 Epidemiology and etiology of low back pain

Low back pain (LBP) is defined as pain and discomfort localized below the costal margin and above the inferior gluteal folds, with or without leg pain. The term LBP refers to a large number of heterogeneous groups of clinical and etiological entities. The lifetime prevalence of spinal pain has been reported as 54% to 80% (Gerdle et al. 2004, Ihlebaek et al. 2006). Studies of the prevalence of low back pain and neck pain and its impact in general have shown 23% of patients reporting low back pain (high pain intensity with disability) versus 15% with neck pain (Manchikanti et al. 2009). LBP has been shown to have a clear association with neck pain (NP), in particular radiating LBP is associated with radiating NP (Kääriä et al. 2010). The prevalence of chronic LBP in the Finnish population in the Terveys 2000 study was for men 10% and for women 11%, which was less than in the Mini-Suomi study 20 years earlier with its prevalence being for men 18% and women 16% (Riihimäki et al. 2002) Further, age-related prevalence of persistent pain appears to be much higher in the elderly associated with functional limitations and difficulty in performing daily life activities (Manchikanti et al. 2009). LBP is a major cause of disability in the adult working population, but afflicts all ages, from children to the elderly (Hill and Keating, 2010). Risk factors for developing spine pain are multidimensional; physical attributes, socioeconomic status, general medical health and psychological state, and occupational environmental factors are all thought to contribute to the risk for experiencing pain (Rubin 2007).

Many factors play a role in the etiology of LBP. Risk factors can be classified into physical load factors, individual factors, and psychosocial and psychological factors. It has been postulated that heavy physical work, frequent bending, twisting, lifting, pulling and pushing, repetitive work, static postures and vibration are the most common mechanical/physical load factors (Andersson 1997). Individual factors include age and gender, weight-related factors, smoking, and physical exercise or inactivity (Viikari-Juntura 2008, Shiri et al.

2010). Some studies show an increased risk for lumbar radicular pain in smokers with a long smoking history or in those with a high level of physical activity (Shiri et al. 2007). Psychosocial and psychological risk factors include stress, distress, anxiety, depression, cognitive dysfunction, pain behavior, job dissatisfaction, and mental stress at work. (Hoogendoorn et al. 2000, Linton 2005).

Several studies and also reviews have addressed the cardiovascular risk factors, such as showing an association between smoking and obesity and LBP (Shiri et al. 2010a, 2010b). When examining the interaction between physical and psychosocial demands of work associated with LBP, Fernandes et al. (2009) found these factors as being independent of each other. Of physically demanding material handling for LBP, the risk was 2.35 higher. Current smoking was associated with increased prevalence of LBP in the past month; OR 1.30, LBP in the past 12 months with OR 1.33, seeking care for LBP with OR 1.49, chronic LBP with OR 1.79, and disabling LBP with OR 2.14 (Shiri et al. 2010a). Obesity was associated with increased prevalence of LBP in the past 12 months (OR 1.33), seeking care for LBP (OR 1.56), and chronic LBP (OR 1.43).

Acute LBP is usually defined as the duration of an episode of LBP persisting for less than 6 weeks; sub-acute LBP as low back pain persisting between 6 and 12 weeks; chronic LBP as low back pain persisting for 12 weeks or more (van Tulder et al. 2006). Recurrent LBP is defined as a new episode after a symptom-free period of 6 months, but not an exacerbation of chronic LBP. In the Quebeck Task Force Classification (QTFC) acute LBP is defined as persisting for fewer than 7 days, subacute between 7 days and 7 weeks, and chronic more than 7 weeks (Spitzer et al. 1987). Non-specific LBP is defined as low back pain not attributed to any recognizable, known specific pathology (infection, tumour, osteoporosis, ankylosing spondylitis, fracture, inflammatory process, radicular syndrome, cauda equina syndrome) (van Tulder et al. 2006).

Acute LBP is usually self-limiting with an estimated recovery rate of 90% within 6 weeks (Croft et al. 2006). After an initial episode of LBP, 44 to 78% suffer relapses of pain, and 26 to 37%, relapses of work absence (Fritz et al. 2003). A common clinical finding in acute LBP patients is decreased range of motion of the spine with increased spinal activity. Disturbances in neuromuscular control have frequently been connected at least with chronic LBP and considered a possible linkage between pain and disability (Hodges and Moseley 2003).

Evidence suggests that prevention of various consequences of LBP is feasible. However, for those interventions with acceptable evidence, effect sizes are rather modest. The most promising approaches seem to involve physical activity/exercise and appropriate (biopsychosocial) education, at least for adults. Owing to its multidimensional nature, no single intervention is likely to be effective at preventing the overall problem of back pain, although benefit is likely from getting all the players onside (Burton et al. 2005).

LBP symptoms, pathology, and radiological appearance are poorly associated. It has been estimated that 5 to 15% of occurrences have a clear patho-anatomical diagnosis. The rest are considered to have non-specific LBP, a

variety of pathological and patho-physiological conditions (Deyo and Phillips 1996, Leboeuf-Yde 2004, Iles et al. 2008,). Pain is attributable neither to pathology nor to neurological encroachment in about 85% of cases. About 4% seen with low back pain in primary care have compression fractures, and about 1% have a neoplasm (Deyo et al. 1992). Ankylosing spondylitis and spinal infections are rarer. The prevalence of prolapsed intervertebral disc is about 1% to 3% in those of working age (Deyo et al. 1990). However, most recent studies have shown an association between lumbar disc degeneration and LBP. According to de Schepper et al. (2010), disc space narrowing at two or more levels appeared more strongly associated with LBP than did other radiographic features. Genetic and environmental influences on disc degeneration seem to be of similar importance; conversely, genetic and environmental influences differed substantially for upper versus lower lumbar levels (disc degeneration only at lower lumbar levels could be an independent factor for environmental influences), emphasizing the importance of examining these levels separately in studies of associated genes, other constitutional factors, and environmental influences (Battié et al. 2008).

In principle, any of the structures of the lumbar spine that receives any innervation could be a source of back pain. Accordingly, back pain could arise from any of the ligaments, muscles, fasciae, joints or discs of the lumbar spine (Adams et al. 2004). Experimental studies in normal volunteers and patients have shown that noxious stimulation of the back muscles, interspinosus ligaments, dura mater, zygapophysial joints, or the sacroiliac joint can produce local and referred pain in a particular patient, (Fortin et al. 1994, Schwarzer et al. 1994, Indahl et al. 1999).

2.2 Clinical tests and test batteries of low back pain patients

Low back pain in terms of non-specific or mechanical back pain describes an entity in which the patho-anatomical etiology is unknown (AHCPR 1994, CSAG 1994, Deyo 2002). Although some studies suggest that specific structural pathology can be diagnosed with clinical examination Young et al 2003, Laslett et al 2005), such reports are unusual. However, clinicians usually base their management decisions about findings on the examination. The type of clinical examination items for the assessment of patients with back pain include functional and mobility tests, inspection, provocation and alleviation of symptoms, muscle tightness, stability, and neurological and neurodynamic tests. However, little is known about the relationships between clinical findings in the low back and LBP in the normal working population (Kääriä et al. 2009).

Red flags are recognized as indicators of possible serious spinal pathology, and their use in clinical examination is indicated by numerous guidelines (Ferguson et al. 2010). The identification of yellow flags (psychosocial

prognostic factors) through early screening in people with acute or SLBP and their influence for outcomes has not yet been shown (Nicholas et al. 2011)

Although each LBP patient`s history and clinical tests are the basis for treatment interventions, there exist a great number of clinical tests with low sensitivity and specificity, and tests producing results that are difficult to interpret (Van Dillen et al. 1998). Because most primary care clinicians believe LBP comprises a number of subgroups but share little agreement regarding the symptoms and signs that identify these subgroups, there is likely to be considerable variability in the methods they use to assess LBP and uncertainty regarding best practice (Kent and Keating 2005). According to a study by Kent et al. (2009), 100% of clinicians very frequently or often assess physical impairment, 99% assess pain; but many fewer assess activity limitation (21%) or psychosocial function (7%) when examining patients in the early phase of LBP (i.e., <12-week duration).

Reliability of physical examination procedures has been studied by many professionals (Hestbaek and Leboeuf-Yde 2000, van der Wurff et al. 2000, Essendrop et al. 2002, Seffinger et al. 2004, van Trijffel et al. 2005). Van der Wurff et al. (2000) reviewed 11 studies which investigated the repeatability of the tests for the sacroiliac joint (SIJ). The results of their review, however, could not demonstrate reliable outcomes, and therefore provided no evidence on which to base acceptance of mobility tests of the SIJ into daily clinical practice. The authors suggested the necessity of further research in this area with an emphasis on multiple test scores and pain provocation tests of the SIJ.

Hestbaek and Leboeuf-Yde (2000) reviewed chiropractic tests and concluded that only tests with palpation for pain had consistently acceptable results in intratester reliability. Motion palpation of the lumbar spine might be valid but showed poor repeatability, whereas motion palpation of the SIJ seemed to be somewhat reliable but was not shown to be valid. The objective of the study by Essendrop et al. (2002) was to make a systematic literature review with preset quality criteria concerning reproducibility of the tests of the low back regarding strength, endurance, and range of motion. These authors concluded that information is quite scarce as to the reproducibility of functional measures for the low back, and therefore any recommendation for consensus is difficult. However, most tests performed in the sagittal plane (flexion-extension) are reliable for use on groups. A review by Seffinger et al. (2004) evaluated the reliability of palpation procedures of the spine with the conclusion that pain-provocation tests are the most reliable, and soft-tissue paraspinal palpatory diagnostic tests are unreliable. Van Trijffel et al. (2005) evaluated the reliability of passive assessment of intervertebral motion. They found that assessment of motion segments in the upper cervical spine almost consistently reached at least fair reliability, but for the lumbar spine, inter-examiner reliability was poor to fair. However, most studies were of poor methodological quality.

All in all, the literature shows the poor reliability of different types of clinical examination procedures commonly used for non-specific LBP; it shows

an obvious need to develop, establish, and enforce valid and reliable test procedures. In addition, clinical examination of LBP requires validation. The validation is typically divided into categories to test functions (walking, squatting, dressing); the shape and size of lumbar lordosis; and structural abnormalities (posture, length differences in lower limbs); mobility (total range of lumbar motion, segmental mobility); and different structures (nerves, muscles).

2.3 Reliability of clinical tests

The World Health Organization's Classification of Functioning, Disability, and Health (WHO-ICF) model was developed to describe, classify, and measure function in health care practice and research (Rundell et al. 2009). Because a relevant body structure for explaining pain or dysfunction might be difficult to evaluate, especially with chronic LBP, physiotherapists can focus on functional impairments by using functional tests such as analyzing walking, squatting, and bending.

In only a few studies has the reliability of functional tests been evaluated with LBP patients in their early phase. The Tidstrand and Horneij (2009) study showed inter-rater kappa values ranging from 0.47 to 1.0 of three standardized functional tests with non-specific LBP patients. The Weiss and Werkman (2009) study reported kappa values ranging from 0.47 to 1.0 of two functional tests with unspecific chronic LBP patients. Movement impairment and movement control tests described by O`Sullivan (2005) could also be classified as functional tests, although their study population comprised chronic LBP patients. The intertester kappa of these tests has shown a range between 0.24 and 0.82) (Luomajoki et al. 2007, Vibe Fersum et al. 2009), when clinical relevance can be considered as Kappa being >0.40 (Landis and Koch, 1977).

Diagnostic accuracy and reliability of muscle strength and endurance measurements in patients with CLBP has been tested (Gruther et al. 2009), but not with acute or SLBP patients.

Inspection

Spinal posture is commonly a focus in the assessment and clinical management of LBP patients. Inspection or observation of body alignment, posture, leg-length differences and others are typical procedures for examining LBP patients. In situations where additional loads or complex postures are anticipated, the muscle recruitment strategy may need alteration, with the temporary goal of enhancing spinal stability beyond the normal requirements, producing, for example, low back pain (Panjabi 1992). Lumbo-pelvic neutral postures may have a positive influence on spinal stability compared to that of equivalent poor postures (slouched sitting and sway-back standing) through the recruitment of the transverses abdominis muscle (Reeve and Dilley 2009).

The link between spinal posture and LBP is, however, not yet fully understood (French at al. 2010. Clinical assessment of lumbar lordosis has not been found to be in accordance with radiological assessments (Tüzün et al. 1999). Results of the study of Schneider et al. (2007) indicate that two clinicians show good reliability in determining the side of the short leg, but show poor reliability when determining the precise amount of that leg-length difference. Knowledge is still sparse regarding the precision of the inspection in LBP patients, if clinical relevance is considered as Kappa being >0.40 (Landis and Koch, 1977).

Mobility

According to nine systematic reviews, risk factors for the occurrence of and prognostic factors for the persistence of non-specific musculoskeletal pain are increased mobility of the lumbar spine in LBP patients (Lakke et al. 2009). Therefore, mobility tests seem to be an essential LBP assessment tool.

Mobility tests can be performed with or without specific equipment. Studies show the positive value of the inclinometer (goniometer) when evaluating spinal curvature and anatomical movements of the lumbar spine in LBP patients (Saur et. al 1999, Fritz et al. 2005). The lumbar range of motion can be determined in degrees by evaluation of radiographs and use of the inclinometer technique. The non-invasive inclinometer technique proves to be highly reliable and valid, but the measurement technique for extension needs further refinement. Lumbar range of motion measurements taken with and without radiologic determination have showed a very close association ($r = 0.93$; $P < 0.001$) (Saur et. al 1999). Flexion alone also has demonstrated a close correlation ($r = 0.95$; $P < 0.001$), whereas extension has shown a somewhat smaller correlation ($r = 0.82$; $P < 0.001$) with radiographs. Total lumbar range of motion ($r = 0.94$; $P < 0.001$) and flexion ($r = 0.88$; $P < 0.001$) have been closely related, as indicated by interrater correlation, whereas extension ($r = 0.42$; $P < 0.05$) has shown a lower correlation. Measurements taken radiographically and by inclinometer have demonstrated an almost linear correlation in measurement of the total lumbar range of motion (ROM) ($r = 0.97$; $P < 0.001$) and flexion ($r = 0.98$; $P < 0.001$), whereas extension ($r = 0.75$; $P < 0.001$) failed to correlate as well (Saur et. al 1999). Lower reliability scores when measuring extension ROM compared to that in flexion can be explained by physiological extension being only one-third of that in flexion.

Measurement of the physiological range of motion using instruments has proven to be more reliable than using vision, and in order to make reliable decisions about joint restrictions in clinical practice, van de Pol et al. (2010) recommend that clinicians measure the passive physiological range of motion by used of goniometers or inclinometers.

Testing subjects passively in the supine or side-lying position, postero-anterior (PA) mobility or segmental dysfunction/facet pain have been studied with LBP patients (Boline et al. 1993, Binkley et al. 1995, Hicks et al. 2003). Results show low kappa- or ICC values with LBP patients, but higher reliability when testing an asymptomatic population (Maher et al. 1998). Results of the

study by Landel et al. (2008) revealed that two examiners applying a PA force to a lumbar spinous process could agree on the lumbar segment that they perceived to be the least mobile segment, but they were less reliable in judging the segment deemed to be the most mobile segment.

Pain provocation

Evidence is moderate as to the high reliability of pain provocation tests of patients with LBP or SIJ dysfunction (Laslett and Williams 1994, van Dillen et al. 1998, Horneij et al. 2002, Laslett et al. 2005a, Haswell 2008). Laslett and Williams recommend combining multiple tests. Five of seven tests to detect a sacroiliac source of low back pain in their study were reliable showing a positive correlation: the percent agreement and the Kappa statistic ranged in value from 78% and 0.52 ($P < 0.001$) to 94% and 0.88 ($P < 0.001$); the other two were less reliable. Posterior pelvic pain provocation and ASLR tests have been relevant when evaluating this affliction in pregnant women with possible pelvic girdle pain (Robinson et al. 2010). Recent European guidelines for the diagnosis and treatment of pelvic girdle (sacroiliac joint) pain recommend using a combination of six different tests, with some of them being provocation tests (Vleeming et al. 2008).

Muscle tightness

Despite limited scientific knowledge, stretching of human skeletal muscle to improve flexibility is a widespread practice among patients with LBP (Magnusson 1998). The effectiveness of different stretching techniques is attributed to a change in stretch tolerance rather than to passive properties (Magnusson 1998).

Clinical signs associated with LBP may be associated with an inability to rotate the trunk about the hips because of weakness or tightness of muscles around the hip (McGregor and Hukins_ 2009). Testing the muscles around the hip in 11 asymptomatic individuals, including short hip flexors and extensors, and external and internal rotators, has been reliable within examiner, its ICC ranging from 0.87 to 0.99 (Bullock-Saxton and Bullock 1994). Holmich et al. (2004) tested 18 athletes with groin pain by a combination of 14 tests including muscle flexibility. The kappa values for the intra-tester agreement were above 0.60 in 11 of 14 tests, and those for inter-tester agreement in the pain tests were above 0.60 in 8 of 10 tests.

Stability

Stability of the lumbar spine and pelvic girdle can be evaluated actively (muscles) or passively (non-contractile structures such as ligaments, and joint capsules). When testing the ability to actively stabilize the pelvic girdle muscles, De Groot et al. (2008) found the active straight leg raise test (ASLR) to have a good differentiation capacity to distinguish between pregnant women with and

without LBP. Intertester reliability coefficients (kappa) were greater than 0.70 for the ASLR in a study by Roussel et al. (2007). Mens et al. (2001) reported for the ASRL test an intra-tester reliability correlation of r= 0.87, and inter-tester of r=0.83. However, Kankaanpaa et al. (2005) could not find in CLBP group impaired paraspinal muscle activity compared to controls in the 90s dynamic back endurance test.

The study by Hicks et al. (2003) showed that segmental mobility testing is unreliable (kappa ranging from -0.25 to 0.26). When testing the inter-tester reliability of common clinical examination procedures expected to identify patients with lumbar segmental instability, the prone instability test, generalized laxity test, and aberrant motion with trunk ROM demonstrated higher levels of reliability. Kappa values for the prone instability test (kappa=0.87) showed greater reliability than did the posterior shear test (kappa=0.22). The Beighton Ligamentous Laxity Scale (LLS) for generalized ligamentous laxity showed high reliability (intraclass correlation coefficient=0.79). Judgments of pain provocation (kappa range 0.25 to 0.55) were generally more reliable than were judgments of segmental mobility (kappa range, -0.02 to 0.26) during passive intervertebral motion testing between two testers (Hicks et al. 2003).

Neurology and neurodynamic tests

Clinical tests are done for LBP patients mainly to detect lumbosacral radiculopathy. The aim of Kerr et al. (1988) was to relate history and clinical signs to the myelograms and surgical findings. The percentage of patients presenting with sciatica reached 99%. The sign most frequently found in patients with a disk protrusion was reduction in SLR. The three signs that, when present, particularly indicated a disk protrusion were "crossed straight leg raising" (pain on contralateral straight leg raising), measured calf wasting, and impaired ankle reflex. A common clinical sign with sciatic patients is decreased strength of the extensor hallucis longus muscle, but this test lacks validity and reliability studies. However, evidence is sufficient regarding the accuracy of other specific tests for identifying sciatica or radiculopathy (Rubinstein and van Tulder, 2008).

The most common way to evaluate irritability of the sciatic nerve is by the SRL test. Intratester results have shown substantial reliability (Hunt et al. 2001), and a negative test outcome may be of greater diagnostic value than a positive one. Biomechanical devices, such as goniometers, by improving intra- and inter-tester reliability, have thus enhanced test reproducibility (Rebain et al. 2002). Kappa scores for agreement between raters for manual palpation of the sciatic nerve have ranged, during SLR and slump tests, from 0.70 to 0.80 (Walsh et al. 2009). Clinical tests which evaluate increased nerve mechanosensitivity and afferent/efferent nerve function show comparable moderate-to-substantial reliability (Schmid et al. 2009). One study shows the slump test to be even more sensitive than the SLR test, but conversely, SLR has a higher specificity (Majlesi et al. 2008).

Evidence for neurological and neurodynamic testing is still conflicting for LBP, and the strength of the associations identified and the extent of confounding between the prognostic investigated factors remain uncertain (Chorti et al. 2009).

Measures of low back pain and function are thought to be of great importance for clinicians and low-back researchers in general, providing information for subclassification and further treatment strategies (Ferguson et al. 2009). However, the usefulness of many tests should be questioned because of the lack of studies on their level of reproducibility, and test sensitivity and specificity.

2.4 Subgroup classification of low back patients

The importance of classifying LBP patients into homogeneous subgroups has been emphasized, and this has been called one of the biggest challenges in physiotherapy (Dankaerts et al. 2006). Because classifying patients with nonspecific LBP into meaningful subgroups should aid in clinical management and to increase the power of outcome assessments, it has been targeted as an important research priority. Use of homogeneous subgroups of LBP patients is considered by many experts to be essential for the improvement of clinical trials related to patient management and clinical outcomes (Delitto et al. 1993, Werneke et al. 1999, Fritz and George 2000, Dankaerts et al. 2009, Hall et al. 2009).

Several classification systems have been designed to categorize patients with low back pain into such homogeneous subgroups. These could guide clinical management decisions or predict pain and disability (McKenzie 1981, Spitzer 1987, Delitto et al. 1995, Riddle 1998, Wilson et al. 1999, Werneke et al 2001, Petersen et al. 2004, O`Sullivan 2005). Common features of subclassification systems are described in Table 1.

TABLE 1 Common features of subclassification systems of LBP

Classification	Quebec Task Force Classification (QTFC)	Treatment-Based Classification (TBC)	McKenzie classification (McK)	Movement System Impairments (MSI)	Patho-anatomical / -physiological Classification (PAP)
Conceptual model/ basis for LBP pain	Patho-anatomic diagnosis if possible.	Matched interventions for patients with LBP through key history and clinical findings	Mechanical deformation of periarticular tissue and /or disc with prolonged postures and repeated movements of the spine.	Repeated movements and sustained postures of the spine in specific direction(s) resulting in strategies. Continual use of strategies contributes to	Symptomatic site located by history and clinical examination.

				impairments, accelerated cumulative tissue stress, micro- and macro traumata.	
Majority of LBP are proposed to be associated	No preference	Results from excessive movement of lumbar spine or immobilization .	Lumbar flexion position.	Lumbar extension or rotation or both positions.	Lumbar disc degeneration and resulting segmental instability.
Examination includes	1. Anatomical location of pain 2.Possible presence of neurologic signs, findings from radiological imaging techniques 3. Subdivided according to pain duration and patients working status	1. Symptom duration 2. Symptom location 3. Fear-avoidance beliefs 4. Lumbar mobility 5. hip rotation range of motion	1. Symptom provocation tests with single and repeated end-range spinal movements and sustained end range spinal postures 2. Clinical evaluation of spinal and pelvic movements and alignments.	1. Primary tests of symptoms with movements in 7 different positions; Symptomatic tests followed by test to eliminate symptoms 2. Clinical evaluation of spine and pelvic movements and alignments with trunk and limb movements	1. Functions 2. Inspection 3. Mobility 4. Stability 5. Muscle length, 6. Pain provocation / alleviation tests 7. Neurologic tests of lower extremities
System specifics/ diagnostic categories	1 of 11 diagnostic categories from local pain to radiating pain	1. Specific exercise 2. Mobilization / manipulation 3. Immobilization 4. Traction	1. postural, 2.dysfunction, 3.derangement	1.extension 2.flexion 3.rotation 4.rotation with flexion 5.rotation with extension	1.disc pain, 2.instability 3.stenosis, 4.facet pain, 5.s-i joint pain
Treatment guidelines	Educational program to train PTs and physicians in a LBP assessment procedure based on the recommendation of The Clinical Practice Guidelines on Acute Low Back Pain Problems	1. In acute phase symptom relief 2 In subacute symptom relief and quick return to normal function 3. For selected patients who must return to activities requiring high physical demands and who demonstrate a lack of physical conditioning necessary to perform the desired activities safely	1. Instructions of sitting, standing and lying postures 2. Mobilization, manipulation or active repeated end-range spinal movement 3. Derangement corrected by directional preference movements	1.Education in tissue injury and healing 2. Instruction in direction-specific movement and alignment strategies with symptomatic functional activities 3. Active exercise to modify direction-specific strategies used with symptomatic impairment tests	1. Pain treatment, 2. Mobilization, manipulation, and / or traction 3. Active self-exercise 4. Immobilization: stabilization with exercise or external support or both. 5. Activation and motivation for self-care, to inform, instruct and train.

2.4.1 Quebec Task Force Classification (QTFC)

Of these classification systems, the Quebec Task Force Classification (QTFC) system has received the widest review. Health care professionals using the QTFC procedure classify patients into 11 diagnostic categories according to presence of pain, anatomical location of pain, presence of neurologic signs, findings from radiological imaging, and surgical history. (Table 1)

These 11 categories are further subdivided according to pain duration and patient`s work status. Simpler versions of the QTFC system have been recommended for use by primary care practitioners, emphasizing anatomical location of pain and results from clinical neurological assessments (Spitzer 1987, Atlas et al. 1996, O`Hearn 1997, Loisel et al. 2002). Studies have shown that QTFC for acute work-related LBP has good discriminant validity between acute nonspecific and radicular LBP (Werneke and Hart 2004). It is concluded also that the QTFC is a helpful descriptor and related to both physical and psychological disability and handicap in employment (Frank et al. 2000). Despite studies to evaluate the reliability and validity, no studies show its effectiveness when treating LBP patients according to their classification.

2.4.2 Treatment-Based Classification (TBC)

In 1995, Delitto et al. described a treatment-based classification (TBC) system that allowed physical therapists to systematically classify patients for physical therapy intervention. The system described subgroup classification of patients with LBP into manipulation, stabilization, specific exercise, and traction. The underlying premise of TBC is that subgroups of patients with acute LBP can be identified from key history and clinical examination findings (Table 1). Furthermore, the creators of TBC hypothesized that each subgroup would respond favorably to a specific intervention, but only when the intervention applied matched the subgroup's clinical presentation. In their study George and Delitto (2005) provided evidence supporting the discriminant validity of TBC with acute LBP patients.

A substantial amount of research has emerged in the years since the introduction of this classification system, including development of clinical prediction rules (CPR) providing new evidence for examination criteria useful in placing a patient into a subgroup and for each subgroup the optimal intervention strategies (Fritz et al. 2007, Cleland et al. 2009). These results also provide support for the assumption that the TBC can be generalized to settings different from those for which it was derived and validated.

The clinical prediction rule in TBC has been questioned by Stanton el al. (2010). These authors state that there exist, at present, little evidence that CPRs can be used to predict effects of treatment for musculoskeletal conditions. The principal problem is that most studies use designs that cannot differentiate between predictors of response to treatment and general predictors of outcome.

2.4.3 McKenzie Classification (McK)

The centralization phenomenon has been reported to be a key physical examination finding in the classification, evaluation, and management of patients with spinal impairments (Long 1995, Donelson et al. 1997, Werneke et al. 1999, Razmjou et al. 2000, Werneke and Hart 2001, 2003, Long et al. 2004). McKenzie originally defined centralization as "a situation in which pain arising from the spine and felt laterally from the midline or distally is reduced and transferred to a more central or near midline position when certain movements are performed" (McKenzie 1985). Patients who respond according to centralization phenomenon then receive directional-specific exercises (Long et al. 2008) (Table 1).

The reliability and validity of the McKenzie classification system has been tested (Kilpikoski et al 2002, Clare et al 2005). This system was shown to be reliable in LBP sub-grouping classification for suitably trained examiners, but not for minimally trained or untrained assessor. The effectiveness of classification-based treatment according to McKenzie is yet to be established (Machado et al. 2006).

2.4.4 Movement Impairment Classification (MIC)

The movement impairment classification is based on findings derived from a standardized examination that includes a history and physical examination (Van Dillen et al. 2003). During the examination, the clinician attempts to identify spinal motions or alignments that provoke symptoms (reproduce the patient's symptoms of pain or paresthesia) (Maluf et al. 2000, Van Dillen et al. 2003) (Table 1).

Once the subgroup has been identified, then treatment strategies can be implemented that restrict the symptom-provoking spinal motions or alignment during everyday activity (Van Dillen et al. 2003). O`Sullivan (2005) proposed that these patients present with either movement impairments (characterized by pain-avoidance behavior) or control impairments (characterized by pain-provocation behavior). These pain disorders are predominantly mechanically induced, and patients typically present with mal-adaptive primary physical and secondary cognitive compensations for their disorders, and that these become a mechanism for ongoing pain. The reliability of this classification system ranged in two studies from almost perfect agreement (kappa-coefficient 0.96; %-of-agreement 97%) in the first study to the second study`s kappa-coefficients ranging 0.47 to 0.80 and %-of-agreement ranging 60 to 84% (Dankaetrs et al. 2006 and 2009). The latest study supports the validation of the proposed classification system for chronic LBP patients, but no validation with patients in the early phase of LBP has been done. However, Luomajoki et al. (2008) demonstrated a significant difference between patients with LBP and subjects without back pain regarding their ability to actively control the movements of the low back, when comparing the ability of 108 patients with non-specific LBP

to 102 control subjects without back pain to control their movements in the lumbar spine by means of a set of six tests.

Ferreira et al. (2007) conducted an RCT in patients with chronic LBP to compare effects of general exercise, motor control exercise based on movement-impairment classification, and manipulative therapy on function and the perceived effect of intervention. In patients with chronic non-specific LBP, motor control exercise and spinal manipulative therapy produced slightly better short-term function and perceptions of effect than did general exercise, but not better long-term effects. A prospective study was carried out in two outpatient physiotherapy practices that involved 38 patients suffering from chronic LBP and movement control impairment; these treatment sessions occurred on average nine times (Luomajoki et al. 2010). Movement control, LBP (VAS) and disability (RMQ) showed a statistical improvement ($p<0.001$). But as there was no control group, the effectiveness of the treatment intervention cannot be judged on the basis of that study. No RCT studies comparing exercise based movement impairment classification have been done with LBP patients in the early phase.

2.4.5 Patho-anatomical/patho-physiological Classification (PAP)

The largest percentage of physiotherapists utilizes a general patho-anatomical/patho-physiological classification system for LBP (Petersen et al. 2003, Spoto and Collins 2008). This classification is also common among physicians dealing with LBP patients (Billis et al 2007). The aim of this patho-anatomical/patho-physiological classification is to provide matched interventions for patients with acute and sub-acute LBP through key history and clinical findings (Table 1). A clinical examination used in this system is non-invasive and widely available. However, the validity of this system has not yet been convincingly confirmed by objective methods. Clinical experience suggests that even a mere idea of the origin of the symptoms may aid the therapist as to the best choice of pain relief, treatment methods, and individual advice for self-care (Rothwell 2005, Brennan et al 2006).

Comparison of blinded clinical diagnoses with diagnoses based on available examination methods such as discography, facet-, sacroiliac- or hip-joint blocks, epidural injections, radiologic imaging studies, or any combination of these, diagnostic agreement on the six most common patho-anatomic categories (disc, facet joint, sacroiliac joint, hip joint, nerve root and spinal stenosis) produced a kappa of 0.31 (Laslett et al. 2005b). The authors concluded that clinical diagnoses agree with reference standards diagnoses more often than is predictable by chance.

2.4.6 Other sub-classification systems of low back pain patients

Sub-classification by duration is commonly applied: acute pain for 0 to 6 weeks with pain and disability from the onset, sub-acute pain for 6 to 12 weeks, and more than 12 weeks for chronic or persistent pain (von Korff 1994, Jonsson and

Nachemson 2000). Recurrent pain has been defined in several ways: patients seeking help after at least one month, going on sick-leave after at least one month of still working, or having a new episode after being symptom-free for 6 months. (McGorry et al. 2000, van Tulder and Koes 2006, Dunn et al. 2008).

The European guidelines for the management of LBP recommend a diagnostic triage, to exclude specific spinal pathology, and an assessment of prognostic factors, to maximize the benefits of treatment and to avoid unnecessary over- or under-treatment (Weiser and Rossignol 2006).

2.5 Exercise therapy, physical conditioning programs and patient education for LBP

Low back pain is most commonly treated in primary health care settings. Among health care providers, clinical management of acute as well as chronic LBP varies substantially. Moreover, many different primary health care professionals are involved in the management of LBP, such as general practitioners, physiotherapists and manual therapists like chiropractors, naprapaths, osteopaths and masseurs. There is a need to improve consistency in the management of LBP across professions. At present, an increasing international trend is towards evidence-based health care. Within the framework of evidence-based health care, clinicians should conscientiously, explicitly, and judiciously use the best current evidence in making decisions on the care of individual patients (Sackett and Wennberg, 1997). The field of LBP research in primary care is an excellent example of evidence-based health care because of a huge body from evidence of randomized trials. At present, more than 500 randomized controlled trials (RCTs) have been published, evaluating all types of conservative and alternative treatments for LBP, ones commonly used in primary care. These trials have been summarized in a large number of systematic reviews (van Tulder at al. 2003). The Cochrane Back Review Group (CBRG) offers a framework for conducting and publishing systematic reviews in the fields of back and neck pain (Bouter et al. 2003). The CBRG has also developed and published method guidelines to improve the quality of reviews in this field and to facilitate comparison across reviews and enhance consistency among reviewers (van Tulder at al. 2003).

Exercise therapy appears to be slightly effective at decreasing pain and improving function in adults with chronic LBP (Hayden et al. 2005). In subacute LBP, some evidence indicates that a graded activity program improves absenteeism outcomes. In acute LBP, exercise therapy is as effective as other conservative treatments (Hayden et al. 2005). Evidence of moderate quality indicates that post-treatment exercise programs can prevent recurrences of back pain (Choi et al 2010). Kankaanpää et al. (1999) showed the active progressive exercise program as being successful in reducing pain and self-experienced disability and also in improving lumbar endurance, more than was the passive

control treatment. However, the group difference in lumbar endurance tended to diminish at the 1-year follow-up.

The effectiveness of *physical conditioning programs* in reducing sick leave when compared to usual care or other exercises in workers with acute back pain shows no effect on sick leave, but a positive effect on sick leave may apply to workers with subacute and chronic back pain (Schaafsma et al. 2011).

For patients with acute or subacute LBP, *patient education* (e.g back-school) seems effective. For patients with chronic LBP, individual education has not been shown to be effective (Engers et al. 2008). Combined *respondent-cognitive therapy and progressive relaxation* therapy are more effective merely remaining on a waiting list for prediction of short-term pain relief (Ostelo et al. 2005).

Multidisciplinary biopsychosocial rehabilitation, back-school, and bed rest

There is moderate evidence as to the positive effectiveness of *multidisciplinary rehabilitation* for subacute LBP and to the fact that a workplace visit improves its effectiveness in reducng LBP-related costs (Karjalainen et al. 2004). Moderate evidence suggests that for patients with chronic and recurrent LBP, *back schools*, in an occupational setting, reduce pain and improve function and return-to-work status, compared to exercises, manipulation, myofascial therapy, advice, placebo or waiting list followup (Heymans et al. 2004). For those with acute LBP, advice to *rest in bed* is less effective than advice to stay active. For patients with sciatica, little or no difference exists in pain [SMD -0.03 (95% CI: -0.24, 0.18)] or functional status [SMD 0.19 (95% CI: -0.02, 0.41)] between advice to rest in bed and advice to stay active (Hagen et al. 2004).

2.6 Orthopedic manual therapy (OMT)

Manual therapy/Orthopedic manipulative therapy/Orthopedic manual therapy (OMT) is defined as a specialized area of physiotherapy/physical therapy for the management of neuro-musculo-skeletal conditions. It is based on clinical reasoning, using highly specific treatment approaches including manual techniques and therapeutic exercises. OMT also encompasses, and is driven by, the available scientific and clinical evidence and the bio-psychosocial framework of each individual patient (IFOMPT 2004).(Figure 1).

OMT is also defined as a form of manual therapy which involves movement of a joint past its usual end-range of motion, but not past its anatomic range of motion with mobilization or manipulation. Mobilization is usually considered movement of long-lever, low-velocity, as opposed to short-lever, high-velocity specific thrust - manipulation. (Assendelft et al 2003). Potential explanations for hypotheses for the working mechanism of spinal manipulative therapy (SMT) are: (1) release for the entrapped synovial folds, (2) relaxation of hypertonic muscle, (3) disruption of articular or periarticular adhesions, (4) unbuckling of motion segments that have undergone

disproportionate displacement, (5) reduction of disc bulge, (6) repositioning of miniscule structures within the articular surface, (7) mechanical stimulation of nociceptive joint fibres, (8) change in neurophysiological function, and (9) reduction in muscle spasm (Assendelft et al 2003).

Manual treatments in OMT are commonly used in combination with specific and functional exercises. The effectiveness of OMT is often summarized in reviews. These reviews present moderate-to-strong evidence that OMT can be effective for the relief of pain and improvement of function at least in the short term (Ferreira et al. 2002, Assendelft et al. 2004, Bronfort et al. 2004, Ernst and Canter 2006). Despite the many published RCTs, a substantial number of reviews and several national clinical guidelines, controversy still remains as to the efficacy of OMT for low back pain.

The present thesis concerns patients with non-specific LBP in its early phase (< 12 weeks), the majority of them at work, seeking physiotherapy treatment due to functional limitations and pain.

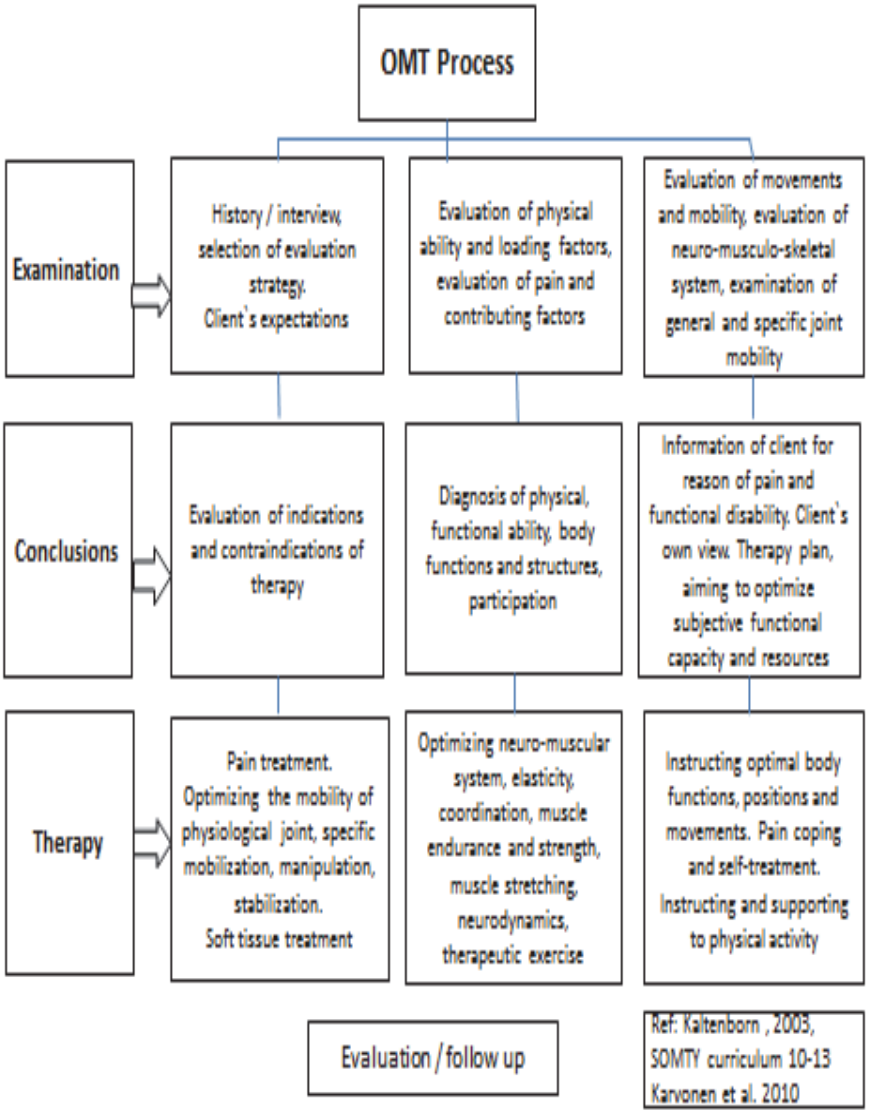


FIGURE 1 Framework of Orthopedic Manual Therapy

2.6.1 Effectiveness of orthopedic manual therapy (OMT) for LBP

During the last decade, four systematic reviews appeared on spinal manipulation for LBP. Ferreira et al. (2002) concluded that spinal manipulative therapy (SMT) produces slightly better outcomes than does placebo therapy, no treatment, massage, and short-wave therapy for nonspecific LBP of less than 3 months duration. SMT, exercise, the standard physiotherapy, and medical care appear to produce similar outcomes in the first 4 weeks of treatment. Ernst and Canter (2003) concluded that SMT was considered superior to sham manipulation but not better than conventional treatments. According to a systematic review by Assendelft et al. (2004), no evidence shows SMT to be superior to other standard treatments for patients with acute or chronic LBP. Comparison treatments were classified into the following seven categories: sham, conventional general practitioner care, analgesics, physical therapy, exercises, back school, or a collection of therapies judged to be ineffective (traction, corset, home care, topical gel, no treatment, diathermy, and minimal massage) or even harmful **(**bed rest). Bronfort et al. (2004), who included in SMT not only high-velocity thrust but also mobilization (MOB), suggest that these treatments can be recommended for the treatment of LBP.

However, the central element in the current debate about best practice management of NSLBP is the efficacy of targeted versus generic (non-targeted) treatment. Many clinicians and researchers believe that tailoring treatment to NSLBP subgroups positively impacts patient outcomes. Despite this, few systematic reviews compare the efficacy of targeted versus non-targeted manual therapy or exercise or both, in which classification and matched interventions have been utilized.

Childs et al. (2004) compared targeted SMT plus range-of-motion exercises with the control treatment of guidelines-based exercise for *acute* NSLBP. A test of interaction in this study indicated that the ability of the prediction rule (in TBC) to identify those who respond to this targeted treatment was statistically significant, but considering the statistical methods in this review, the size of that treatment modifier effect was not statistically significant.

Hancock et al (2008) compared the results of spinal mobilization with the control treatment of detuned ultrasound for *acute* NSLBP. Results showed a treatment-modifier effect size of 8.4% in baseline scores for short-term pain and 0.6% for intermediate-term pain, but neither was statistically significant, in contrast to the findings of Childs et al (2004). One factor for these differing results may be that Childs et al. in RCT used a special spinal manipulation technique, whereas RCT by Hancock et al. involved a pragmatic study in which most patients underwent spinal mobilization techniques and only 5% received manipulation. In summary, Hancock et al. (2008) suggested that the ability of the prediction rule to identify those who respond to a targeted treatment is not statistically significant, and no test of the size of any treatment modifier effect is statistically significant at any outcome time point.

The study by Hallegraeff et al (2009) comprised 64 participants, with *acute* NSLBP randomly assigned to two groups: an experimental group

(manipulative therapy plus physical therapy) and a control group (only physical therapy). The results showed a statistically significant effect for disability, but no statistically significant benefit from additional manipulative therapy over physical therapy for pain and mobility within four treatments. In a study by Jüni et al (2009), 104 patients with *acute* LBP were randomly assigned to SMT in addition to standard care (n = 52) or standard care alone (n = 52). Standard care consisted of general advice and paracetamol, diclofenac, or dihydrocodeine as required. The authors concluded that SMT is unlikely to result in relevant early pain reduction in patients with *acute* LBP.

Fritz et al. (2007) investigated the Delitto TBC method for treatment of *subacute* LBP patients. Results of this RCT showed matched treatment effects of 12.9% of baseline scores for short-term activity limitation and 8.5% for short-term pain, but neither was statistically significant. This study indicated that the ability of the prediction rule to identify those who respond to a matched treatment was statistically significant, but a test of the size of that matched treatment effect was not statistically significant.

A single high-quality study investigated McKenzie directional preference-based exercise with *chronic* LBP patients (Long et al. 2004). It was a multi-arm subgroup system RCT that showed statistically significant improvements in short-term activity and short-term pain limitation due to the matched treatment effect ranging from 23% to 34% of baseline scores compared to other groups: either exercises directionally "opposite" to their directional preference, or "nondirectional" exercises. As this study included only those with a directional preference, these results are applicable only to those who display a directional preference. The analysis used in this review compared the effect of directional preference exercises with the mean of both comparison groups (the opposite-direction exercise group and the non-directional exercise group), but an alternative would have been to use only the opposite-direction exercise group as the comparison. As clinicians and patients in this trial were not blind to treatment group allocation, treatment expectation may have inflated the subgroup system effect size in this RCT. The size of the matched treatment effect was statistically significant for both short-term activity and short-term pain limitation.

A systematic review with a meta-analysis by Fersum et al. (2009) was undertaken to determine the integration of sub-classification strategies with matched interventions in RCTs evaluating manual therapy treatment and exercise therapy for *NSCLBP*. These authors found in 68 studies only five (7.4%) to sub-classify patients after applying general in- and exclusion criteria. In these few studies utilizing classification and matched interventions, meta-analysis showed a statistically significant difference in favor of the classification-based intervention both for short- and long-term pain reduction and disability.

2.7 McKenzie method

The McKenzie Method is a comprehensive approach to the spine based on principles and fundamentals that when understood and followed correctly are very successful in pain relief according to McKenzie and May, (2003). Unique to the McKenzie Method is an algorithm that leads to the simple classification of spinal-related disorders. It is based on a consistent "cause and effect" relationship between historical pain behavior as well as the pain response to repeated test movements, positions, and activities during the assessment process. (McKenzie and May, 2003)

A systematic progression of applied mechanical forces (the cause) utilizes pain response (the effect) to monitor changes in movement and function. The underlying disorder can then be quickly identified through objective findings for each individual patient. The McKenzie classification of spinal pain provides reproducible means of separating patients with apparently similar presentations into definable sub-groups (syndromes) to determine appropriate treatment. McKenzie classified three mechanical syndromes: Postural, end-range stress of normal structures; Dysfunction, end-range stress of shortened structures (scarring, fibrosis, nerve root adherence); and Derangement, anatomical disruption or disc displacement within the motion segment.

Each distinct syndrome is addressed according to its unique nature, with mechanical procedures utilizing movement and positions. The Derangement syndrome, where the phenomenon of "centralization" occurs, is most common. Centralization of pain is defined as ''the abolition of distal limb symptoms in response to the deliberate application of repeated movements or sustained postures.'' ''Directional preference'' is closely related to pain centralization, and indicates the direction of force required to centralize the pain (McKenzie and May, 2003).

2.7.1 Effectiveness of the McKenzie method for LBP

The effectiveness of the McKenzie method in addition to first-line care for *acute* LBP was evaluated by a randomized controlled trial, where eligible participants were assigned to receive a treatment program based on the McKenzie method and first-line care (advice, reassurance, and time-contingent acetaminophen) or first-line care alone. Machado et al. (2010) found that when added to the currently recommended first-line care of *acute* low back pain for 3 weeks, a treatment program based on the McKenzie method failed to produce appreciable additional short-term improvement in pain, disability, function or global perceived effect. However, the McKenzie method seems to reduce health utilization due to low back pain although it does not reduce patient's risk of developing persistent symptoms.

Petersen et al. (2007) compared the effectiveness of McKenzie treatment to strength training for patients with *chronic* LBP. A total of 260 patients with *chronic* LBP were included in a randomized controlled trial of McKenzie

therapy versus strengthening training. Outcome variables were functional status, pain level, work status, and use of healthcare services during follow-up. No differences in outcomes emerged between the treatment groups at 14 months of follow-up.

The key element, the centralization phenomenon, has been studied by several authors. Werneke and Hart (2005) and Werneke et al. (2011) analyzed data from 177 consecutive patients with *acute* work-related low back syndromes referred to physical therapy and the association between centralization and non-centralization and baseline behavioral signs. The physical sign of non-centralization was associated with non-organic signs, overt pain behaviors, fear of work activities, and somatization. Berthelot et al. (2007) concluded that even the available data have failed to establish that centralization is sufficiently specific for discogenic pain as to obviate the need for investigations, particularly in patients considered for surgical treatment (e.g., fusion or implant). Nevertheless, centralization may indicate the high likelihood of discogenic pain and may provide therapeutic guidance.

Although, centralization correlates strongly with a positive discography (Donelson et al. 1997) the value of this sign as an indicator that surgery is needed remains highly controversial (Berthelot et al. 2007). But, because centralization is associated with better outcomes after nonsurgical treatment, even in patients with nerve root pain, its presence may constitute an argument against surgical treatment like discectomy (Berthelot et al. 2007).

2.8 Advice to stay active

European guidelines for *acute* LBP advise patients to stay active and continue normal daily activities including work if possible (van Tulder et al. 2005). Advice and exercise are widely recommended also for *subacute* low back pain, but the effectiveness of these interventions is unclear (Pengel et al. 2007). However, advice to stay active and continue normal daily activities has not been specified in guidelines (Chou et al. 2007).

Most *acute* LBP, with no red flags signs, is primarily managed in general practice. Williams et al. (2010) investigated how recommendations for best practice in international evidence-based guidelines for the management of *acute* LBP are followed by general practitioners (GPs) in primary care. The usual care provided by GPs for LBP does not match the care endorsed in international evidence-based guidelines and thus may not provide the best outcomes for patients; this situation has not improved over time.

2.8.1 Effectiveness of advice to stay active during LBP

Pengel et al. (2007) investigated the effectiveness of physiotherapist-prescribed exercise, advice, or both for subacute LBP (>6 weeks and <3 months in duration) with 259 patients. Primary outcomes were average pain over the past

week (scale, 0 to 10), function (Patient-Specific Functional Scale), and global perceived effect (11-point scale) at 6 weeks and 12 months. Secondary outcomes were disability (Roland-Morris Disability). They found that physiotherapist-directed exercise and advice were each more effective: -0.8 point 95%CI, -1.3 to -0.3 point; (P = 0.004) than placebo at 6 weeks. This effect was greatest when the interventions were combined. At 12 months, the only effect that persisted was a small effect on participant-reported function.

A study by Pengel et al. (2007) concluded that a combination of exercise and advice was better than the effect of exercise and advice alone for *subacute* LBP. Cherkin et al. (1998) investigated the effectiveness and costs of treatments for *acute* LBP with the McKenzie method of physical therapy, with chiropractic manipulation, or with advice (provision of an educational booklet). Patients receiving the McKenzie method of physical therapy or chiropractic manipulation had only marginally better outcomes than did those receiving advice with the minimal intervention of an educational booklet. For all outcomes, no significant differences emerged between the physical-therapy and chiropractic groups and no significant differences among the groups in the numbers of days of reduced activity or missed work or in recurrences of back pain. Whether the limited benefits of these treatments are worth the additional costs is open to question. Malmivaara et al. (1995) conducted a study to compare bed rest, back-extension exercises or ordinary activity for patients with *acute* LBP. Outcomes and costs were assessed after 3 and 12 weeks. Differences were statistically significant in favor of the control group (ordinary activity) in duration of pain, pain intensity, lumbar flexion, ability to work as measured subjectively, the Oswestry back-disability index, and number of days absent from work. Recovery was slowest among the patients assigned to bed rest. Overall costs of care among the three groups did not differ significantly.

2.9 Summary

Relatively little is known about the accuracy of diagnostic procedures for LBP. Although most spinal conditions are benign and self-limiting, the real challenge to the clinician is to distinguish serious spinal pathology or nerve-root pain from non-specific LBP. The use of valid procedures can assist the clinician in this aim. A systematic review by Rubinstein and van Tulder (2008) has evaluated evidence for diagnostic procedures in the following categories: history, physical examination, and special studies, including diagnostic imaging, diagnostic blocks, and facet and sacroiliac joint injections. With regard to the physical examination, the straight-leg raise was the only sign consistently reported to be sensitive for sciatica due to disc herniation, but it is limited by its low specificity. The tests that authors reviewed led to this conclusion: "It is quite remarkable that while many named orthopedic tests of the neck and low back are often illustrated in orthopedic textbooks, there is little evidence to

support their diagnostic accuracy, therefore their use in clinical practice." The authors evaluated the tests based on sensitivity, specificity, and the ability of each to alter post-test probability.

Clinical tests are a mainstay of physical diagnosis. Within the context of the evidence-based practice paradigm, data on the diagnostic accuracy of these special tests are frequently used in the decision-making process when determining the diagnosis, prognosis, and selection of appropriate intervention strategies. However, the reported diagnostic utility of these tests is significantly affected by the study methodology of diagnostic accuracy studies. Methodological shortcomings can influence the outcome of such studies, and this in turn will affect the clinician's interpretation of diagnostic findings (Cook et al. 2007). Thus, the need is for further development of tests or test batteries with high sensitivity and specificity.

It is hypothesized that classification of LBP into homogenous groups and application of specific interventions tailored for these groups are likely to enhance treatment efficacy (Dankaerts et al. 2006). Early identification of high-risk populations that comprise those with weak psychological and physiological constitutions, who more easily than others may develop long-lasting back pain as well as other pain will allow for a selective primary and secondary preventive approach (Leboeuf-Yde 2004). It is also well established that LBP is a multi-dimensional problem consisting of pathoanatomical, neurophysiological, physical, and psychosocial factors (Borkan et al 2002, McCarthy et al 2007).

Choosing a classification is dependent on the time factor of LBP, on the education and the length of experience of the clinician, and on an understanding of the underlying mechanism driving the disorder (Laslett et al. 2005b). It is quite possible to combine different classifications, because none of the classifications seems to be superior to others. A lack of consensus among participating clinicians regarding LBP subgroups and a lack of evidence for the validity of LBP sub-grouping are a compelling argument for further research into this clinical practice (Kent and Keating 2005).

Targeting subgroups of LBP patients for treatments should provide improvement in outcome. Research shows that adequately powered controlled trials using designs capable of providing robust information on treatment effect modification are uncommon. (Kent et al. 2010).

3 PURPOSE OF THE STUDY

The object of this thesis was to evaluate the reliability of selected clinical tests, the reliability of sub-group classifications of LBP patients based on these selected tests, and to study the effect on LBP of orthopedic manual therapy, the McKenzie method, and advice only in working adults in the early phase of LBP.

3.1 Specific aims

Specific aims were

1. To evaluate intertester and intratester reliability of selected clinical tests between two physiotherapists in orthopedic manual therapy (OMT) to assess LBP patients in the early phase (< 12 week duration) (Study 1).
2. To investigate the sensitivity and specificity of selected clinical tests commonly used in examining acute, subacute, and chronic LBP patients and compare the results to a control group with "no patient status" (Study II).
3. To evaluate the intertester reliability of a pathoanatomical-pathophysiological classification between i) general practitioners in primary care physiotherapy and a specialist in OMT, and ii) a physiotherapist with long experience as a specialist in OMT and one with short experience when examining patients in the early phase of LBP (< 12 week duration)(Study III).
4. To compare the effectiveness of OMT and the McKenzie method with advice only to stay active on low back and leg pain, disability, and sick leave during a one-year follow-up of working adults (Study IV).

4 METHODS

4.1 Design

This thesis is based on three cross-sectional studies (Studies I, II, and III), and a randomized controlled follow-up trial (Study IV). The study designs are presented in Table 2.

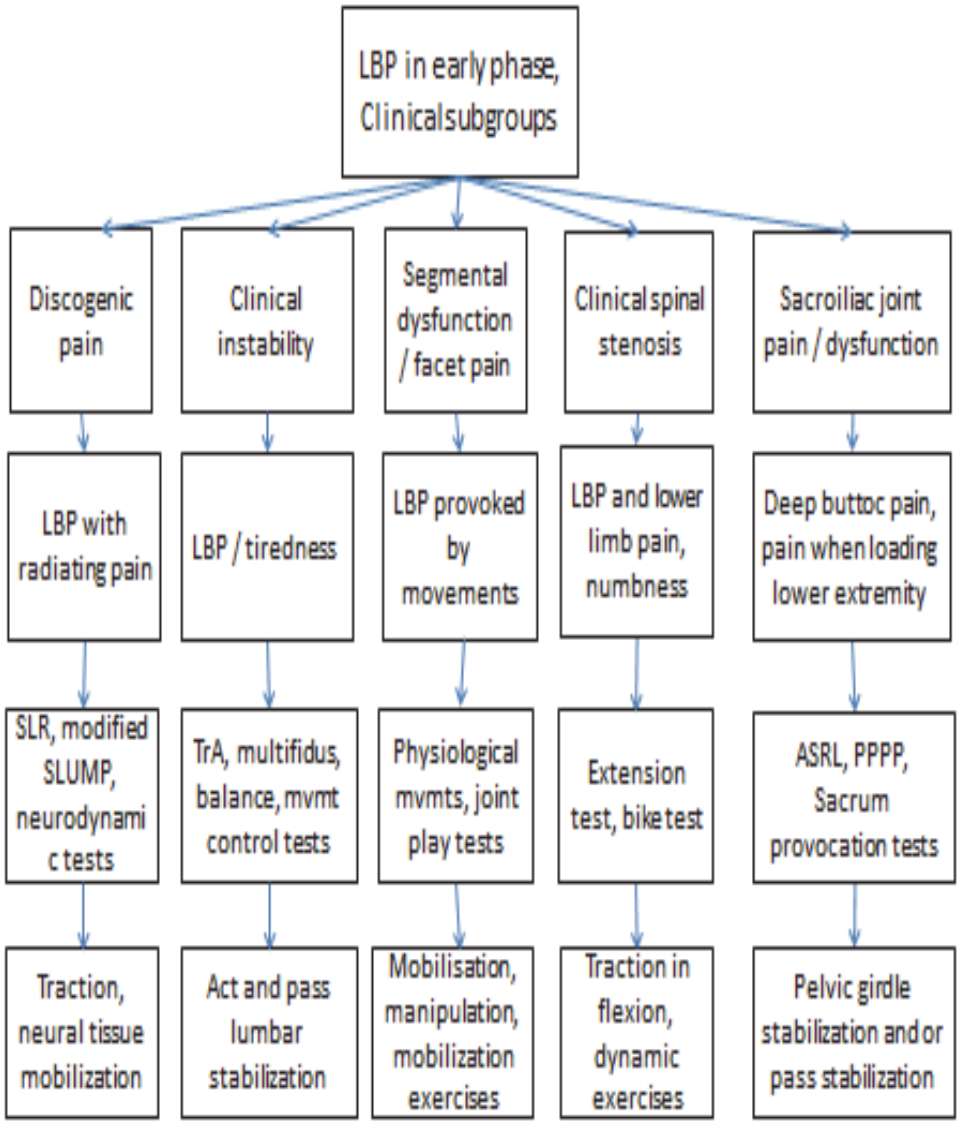
The clinical reasoning process when classifying patients into specific patho-anatomical/patho-physiological subgroups (Study I) is shown in Figure 2. Flow of patients through the randomized controlled trial (Study IV) is shown in Figure 3.

For all four studies, the subjects received written and oral information about the study and gave their informed consent before inclusion. Both confidentiality and the voluntary nature of a questionnaire and clinical measurement were stressed. The studies were approved by Ethics Committees of the University of Jyväskylä and Pirkanmaa Hospital District (formerly Tampere University Hospital).

TABLE 2 Studies and their design, number of subjects, group, measurements, and main outcomes

Study	N	Group	Measurements	Main outcome
I Cross-sectional	15	Early phase of LBP	Questionnaire, clinical assessment	Intra- and intertester agreement of clinical tests
II Cross-sectional	55 47 55	Early phase of LBP CLBP Controls	Questionnaire, clinical assessment	Differentiation capacity of clinical tests
III Cross-sectional	51	Early phase of LBP	Questionnaire, clinical assessment	Intertester reliability of subgroup classification between specialists, and between specialist and non-specialists
IV RCT follow-up	45 52 37	Working adults with LBP in early phase; OMT group McKenzie group Advice group	Questionnaire, clinical assessment	Pain Disability Sick leave because of LBP

LBP = Low back pain
CLBP = Chronic low back pain
OMT = Orthopedic manual therapy



LBP = Low-back pain
SLR = straight leg raise
SLUMP = lumbar spine provocation test in slump-sitting position
ASLR = active straight leg raise
TrA = transverses abdominis
PPPP = posterior pelvic pain provocation

FIGURE 2 Flow chart of decision rule for classifying patients into clinical subgroups.

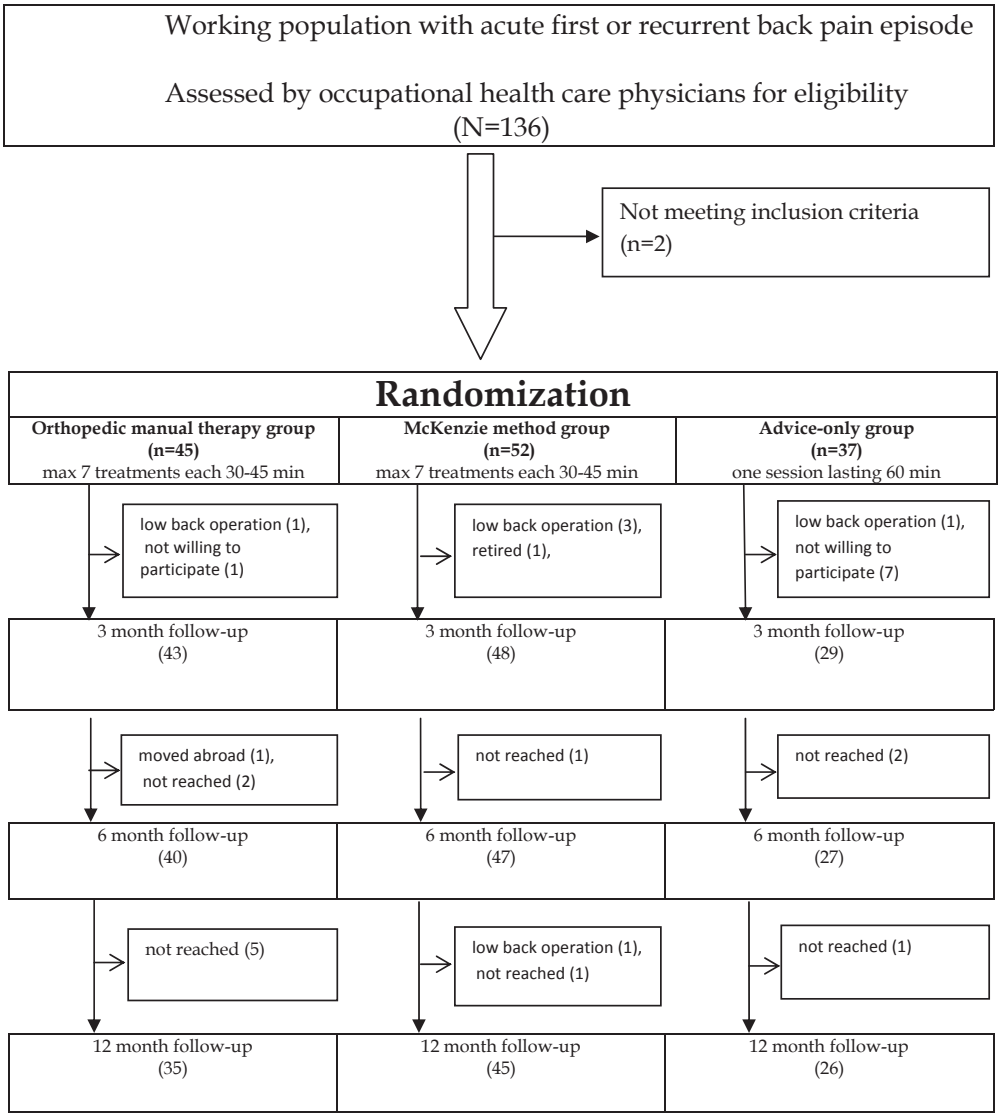


FIGURE 3 Flow of patients through trial (Study IV)

4.2 Subjects

The characteristics of subjects are presented in Table 3. *A total of 356* subjects participated in these studies, 301 LBP patients and 55 serving as a "non-LBP group". Altogether 171 women and 185 men, from 18 to 68 years of age, were recruited through hospital, occupational health or university departments.

In Study I, 15 eligible, consecutive, and voluntary patients with LBP lasting *less than 3 months* were recruited by an invitation letter from a private occupational health care center (Medivire) in the city of Jyvaskyla, Finland.

In Study II, the inclusion criteria for the 55 in the CLBP group: LBP lasting *more than 3 months* with or without radiating pain to one or both lower limbs. They were sent to an orthopedic surgeon's consultation at Tampere University Hospital in order to evaluate whether spinal surgery was needed. There were no exclusion criteria. For an SLBP group, 47 subjects were recruited from five occupational health care centers. The back pain episode could be the first or recurrent LBP, the last episode lasting *less than 3 months*. Their inclusion criteria were age 18 to 65, employed, and with current LBP with or without radiating pain to one or both lower legs. The back pain episode could be the first, or be recurrent with the last episode lasting *less than 3 months.* For controls in the "non-LBP" group, 55 subjects were recruited by invitation letter from the University of Jyväskylä. They had no medical low back diagnosis and had had no medical or physiotherapeutic treatment during the previous year.

For the first part of the Study III, 20 eligible, consecutive, and voluntary *acute or SLBP* patients were recruited by an invitation letter from four municipal health care centers in central Finland. For the second part of the study, 30 patients were recruited from a private occupational health care center (Medivire) in the city of Jyväskylä, Finland. The inclusion criteria were 18- to 65-year-old employees with current low back pain with or without radiating pain to one or both lower legs.

In Study IV, participants were selected according to the following inclusion criteria: 18- to 65-year-old employed people with current non-specific LBP with or without radiating pain to one or both lower legs, current back pain lasting *less than 3 months.* Patients who were randomized into one of these three groups started physiotherapy on average 7 days after a visit for their occupational health care. The 1 to 7 treatment sessions which patients normally used in these treatment groups lasted 4 to 6 weeks.

In all studies exclusion criteria were pregnancy, use of any psychogenic medication, back pain surgery less than 2 months previously, acute spinal trauma or serious pathology as evidenced by the popularity of red flag screening (age, cauda equine syndrome, significant and / or progressive neurological deficit, or other systemic illness, including cancer) (Bigos et al 1994, Bigos and Davis 1996), except for the CLBP group in Study IV. The participants` recruitment process and other subjects are described in more detail in the original papers.

TABLE 3 Characteristics of the subjects

	Study I	Study II			Study III		Study IV		
		CLBP	SLBP	Controls	Health care	Occupational health care	OMT	McKenzie	Advice
Number of subjects	15	55	47	55	20	30	45	52	37
Age (years, mean, sd)	38 (4.5)	42 (11.6)	45 (10.2)	38 (8.1)	40 (11.5)	38 (4.5)	44 (10.0)	44 (9.1)	44 (15.2)
Gender, (f/m)	10/5	31/24	18/29	33/22	13/7	20/10	19/26	15/37	13/24
Episodes of LBP									
0-1(%)	10	2	7	37	42	1	12	7	10
2-5 (%)	75	22	49	49	50	77	42	45	44
>6 (%)	15	76	44	14	8	28	46	48	46
On sick leave(%)	0	42	12	0	0	0	16	17	8
Duration of last LBP episode									
Acute (%)	27	0	32	x	35	30	48	56	50
Subacute (%)	73	0	68	x	65	70	52	44	50
Chronic (%)	0	100	0	x	0	0	0	0	0
VAS LBP (sd)	3 (2.9)	6 (5.1)	4 (3.1)	x	3 (3.5)	3 (2.8)	4 (3.3)	3 (3.0)	3 (4.1)
VAS Leg pain	x	x	x	x	x	x	2 (2.7)	2 (3.1)	2 (2.9)
R-M disability	x	x	x	x	x	x	9 (5.8)	9 (4.6)	8 (4.1)
Symptom location									
LBP only (%)	38	x	x	x	42	45	29	19	31
Above knee(%)	30	x	x	x	28	30	30	54	37
Below knee (%)	32	x	x	x	30	25	41	27	32
Type of work	67	77	89	94	65	67	89	94	86
Light (%)	33	23	11	6	35	33	11	6	14
Heavy (%)									

x = not recorded
f = female
m = male
CLBP = chronic low back pain
SLBP = subacute low back pain
OMT = orthopedic manual therapy
VAS = visual analogue scale
sd = standard deviation
R-M = Roland-Morris disability questionnaire

4.3 Measurements

Severity and symptoms of LBP were assessed in a short interview and the pain location by a pain drawing. Type of employment, absence from work because of LBP, and episodes of LBP during each subject's lifetime were all recorded. After having interviewed the subjects, physiotherapists performed selected clinical tests on all of them. This assessment consisted of a history and 34 different tests, divided into seven categories: **1) functions** of the lumbar spine and lower extremities (5 tests), **2) inspection** of posture (4), **3) mobility** tests of the lumbar spine, sacroiliac joints, and hip joints (5), **4) pain provocation** tests (7), **5) muscle tightness** tests (4), **6) stability** tests for the lumbar spine and pelvis (4 tests), and **7) neurological and neurodynamic** tests (5). (Table 4, Appendix 1; "Back examination form")

4.3.1 Visual analogue scale (VAS)

The visual analogue scale (VAS) was used to assess perceived pain in the low back and leg pain. The VAS used was a 100-mm horizontal line ranging on the left by "no pain" and on the right by "unbearable pain". Validity and reliability of VAS have been tested for patients with LBP (Scott and Huskisson 1978, Bijur et al 2001). For patients with LBP, the minimally clinically important change (MCIC) for pain on VAS should be at least 20 mm (Ostello and de Vet. 2005, Kovacs et al. 2007)

4.3.2 Roland- Morris Disability Questionnaire

The Roland-Morris Questionnaire (RMQ) is a self-administered disability measure in which greater levels of disability are reflected by higher numbers on a 24-point scale. The RMQ has been shown to yield reliable measurement, valid for inferring the level of disability, and to be sensitive to change over time for groups of patients with LBP (Roland and Fairbank, 2000, Stratford et al. 1996). For patients with LBP, the minimally clinically important change (MCIC) for functional disability measured with RMQ should be at least 3.5 points (Ostello and de Vet, 2005, Kovacs et al. 2007).

4.3.3 Sick-leave because of low back pain

The time and number of days of sick-leave because of LBP during the last 12 months were provided by a questionnaire prior to treatment interventions. During the 3, 6, and 12 months follow-up points were asked if a participant has been on sick-leave for current LBP, and the number of sick-leave days after the last follow-up.

TABLE 4 Classification of clinical tests. Test procedures for the decision are described in Appendix 1, and results of inter- and intratester reliability in Table 5. The result of the test was dichotomous: each test was either negative (normal) or positive.

1. Functions
Walking
Undressing
Walking on toes
Heel-walking
Squat and rise

2. Inspection
Posture of lumbar spine
Leg length difference
Posture of knee
Posture of feet

3. Mobility
Lumbar spine flexion
Lumbar spine extension
Lumbar spine lateral flexion
Hip rotation
Specific p-a mobility of T12-S1

4. Pain provocation
Extension with traction
Physiological movement
Posterior pelvic pain provocation
Kibler`s skin rolling
Sacroiliac joint provocation
L4, L5 rotation provocation

5. Muscle tightness
Hamstrings
Piriformis
Gluteus medius/minimus
Iliopsoas

6. Stability
One-leg standing
Active straight-leg raise
Isometric lumbar extension
Transverses abdominis activity

7. Neurology/neurodynamics
SLUMP in sitting
Straight-leg raise
Achilles reflex
Patella reflex
Ely`s (femoral nerve tension) test

4.4 Interventions

A randomized controlled trial (Study IV) with one-year follow-up was conducted at Jyväskylä University. Randomization of the participants into the treatment groups was by a stack of sealed envelopes, numbered in an order prepared from a random number table.

Orthopedic manual therapy. In the OMT group, the participants were clinically assessed and classified into five patho-anatomical/patho-physiological subgroups and treated according to OMT principles (Figure 2). Patients in this group underwent pain treatment, specific mobilization, spinal manipulation if indicated, and muscle-stretching techniques. Further, these patients were taught to perform one or a combination of self-mobilization, self-stabilization or stretching exercises at home once daily.

McKenzie Method of Mechanical Diagnosis and Therapy. In the McKenzie method group, the participants were clinically assessed and classified as to the

mechanical syndromes. If a non-mechanical syndrome was present, the subjects were transferred from conservative care for further investigation. If a mechanical syndrome was present, then one of the relevant treatment principles served as the management strategy. This consisted of an educational component, supported with the "Treat Your Own Back" book, and an active therapy component provided instructions in exercises repeated several times a day according to the principles of the approach.

Advice only. Subjects in the Advice-only group received 45 to 60 minutes of counseling from a physiotherapist concerning the good prognosis of LBP and concerning pain tolerance, medication, and early return to work. Patients in this group were told to avoid bed rest and advised to continue their routine as actively as possible, including exercise activities, within the limits permitted by their back pain. For support, a two-page educational back booklet was also supplied. Advice-only subjects were informed of the benign nature of LBP, and the importance of self-exercises and self-management.

Number of visits. The number of visits for subjects in the Advice-only group was one, and ranged from three to seven in the OMT and McKenzie groups.
Length of interventions. Subjects in the OMT and McKenzie groups visited their physiotherapist 1 to 2 times weekly.

The participants` recruitment process and other subjects are described in more detail in the original paper.

4.5 Subgroups and their definitions

Clinical subgroups were divided as follows: discogenic pain, clinical instability, clinical lumbar spinal stenosis, segmental dysfunction/facet pain, sacroiliac joint pain/dysfunction. These subgroups are based on clinical examination and there is no gold standard radiologic method to justify which anatomical structure is the origin of pain in the great majority of patients. Subgroup classification was based on clinical reasoning focusing on patients` history and clinical findings (Figure 2). Subgroup classification was not a medical diagnosis.

Briefly, *discogenic pain* was the diagnosis when a patient's pain (local or referred) could be provoked in modified slump test and when movement into extension was less painful or alleviated the same pain (centralized). Discogenic pain was recorded also when radiating pain was provoked by positive sciatic (SLR) or a femoral nerve tension test (PNB).

The construct for *clinical lumbar instability* involved assessment of three interdependent components: the passive, the active, and the neuromuscular subsystem. Clinical lumbar instability was recorded when the patient reported LBP and fatigue during prolonged sitting/standing/lying down, and when pain during extension was relieved and movement increased with traction. In addition, the classification was made if there were difficulties in a one-leg stance or active straight-leg raise (ASLR) or both, or inability to activate either

transverse abdominus or lumbar multifidi combined with local interspinal pain, or a combination of these problems.

Clinical spinal stenosis was recorded when the patient reported a clear pattern of intermittent claudication provoked by extension, which was relieved by sitting or a flexed spinal posture. Symptoms and signs could be combined with tightness of hip flexors or a positive sciatic (SLR) or a femoral nerve tension test (PNB). Diagnosis was recorded when the patient reported radiating pain with nerve tension tests and during extension/lateral flexion toward the symptomatic side or during transverse process provocation (passive foramina approach), or both.

Segmental dysfunction/facet pain was recorded when pain and movement restrictions were identified during physiological movements in standing and painful hypomobility while lying prone.

Sacroiliac joint pain/dysfunction was recorded if the patient's lower lumbar or buttock pain was provoked while standing on one leg and relieved with a sacroiliac joint belt, or provoked with sacral thrust and/or during posterior pelvic pain provocation (PPPP) or both, or if pain and difficulties occurred during an ASLR.

4.6 Statistics

Statistical analysis was carried out with appropriate programs for SPSS software, versions 14-16. In Studies I to III, percentage agreement and the kappa statistic served to test intra- and intertester agreement of clinical tests. The kappa statistic estimates the degree of agreement corrected for chance agreement. General agreement exists that kappa is one of the preferred statistics for estimation of the accuracy of nominal and ordinal data in clinical research (Haley and Osberg, 1989). When the prevalence of rating in the population is very high or low, the value of kappa may indicate poor reliability even with high observed proportion of agreement (Byrt et al. 1993). Agreement and kappa can be used together to uncover non-random examiner error (Hunt, 1986).

In Study II, to determine the best predictors of chronic and subacute LBP, as well as determining the best predictors among the five subgroups, a forward stepwise logistic model was applied. The odds ratio (OR), sensitivity, and specificity, with their confidence intervals (CIs) for all the tests were calculated.

In Study IV, randomization of the participants into the treatment groups was by a stack of sealed envelopes, numbered in an order prepared from a random number table. The data were analyzed by the intention-to-treat principle, including all randomized participants who provided follow-up data by post-hoc tests using ANOVA. Post-hoc between-groups comparisons were done with Sheffe`s adjustment for multiple comparison. An alternative analysis accounted for drop-outs at follow-up, whereby missing values were replaced with imputed values generated by a series of estimated marginal means of

measuring two-tailed equations; subjects' previous scores allowed determination of a predicted value that reduced the variance of the value for each variable. Baseline characteristics were summarized for descriptive purposes, with medians and quartiles used for continuous measures and percentages for categorical measures. Comparison of treatment effects among all groups were analyzed, calculating effect sizes according to the following categories: Effect size of 0.2 to 0.3 being a "small" effect, around 0.5 a "medium" effect, and 0.8 to infinity a "large" effect (Cohen 1969). Physiotherapists were not blinded as to the groups.

5 RESULTS

5.1 Clinical tests

5.1.1 Reliability of selected clinical tests between and within testers.

The values of the inter- and intratester reliability of the 34 clinical tests for in each test and test categories are shown in Table 5.

In all seven test categories, the mean *intertester* kappa was 0.5 (95% CI; -0.2 to 1.2) and agreement 82%. The mean reliability was at a moderate level in other test categories with broad confidence intervals. The inspection category was low. In functional tests, the between-testers` overall agreement was at a high level: kappa 0.9 and agreement 98%. Agreement was, however, poor (kappa 0.2 and agreement 51%) in the inspection test category overall, as well as in almost every single test in this category. In the other test categories, kappa and agreement% were moderate: mobility (kappa 0.5 and agreement 85%), pain provocation (0.5 and 78%), muscle tightness (0.4 and 79%), stability (0.5 and 80%), and by neurodynamic tests (0.5 and 82%). Even though overall intertester reliability was at an acceptable level in these test categories, in some single tests, reliability between testers ranged from poor to good, except in the functional test category. The decision in each test was dichotomous; tests were either negative (normal) or positive.

In all seven test categories, the mean *intratester* kappa was 0.5 (95% CI; 0.1 to 1.0) and agreement 85%. The mean reliability was at a moderate level in every test category, also in the inspection category with broad confidence intervals. In the functional test category, agreement was at a high level: kappa 0.9 and agreement 98%. In intratester reliability, kappa was fair and agreement moderate: kappa 0.5 and agreement 77% in the inspection test category as well as in all single tests in this category.

In other categories, kappa ranged from fair to good, and agreement% was good: mobility, kappa 0.5 and agreement 80%; pain provocation, kappa 0.4 and agreement 77%; muscle tightness: kappa 0.6 and agreement 90%; stability:

kappa 0.6 and agreement 83%; and by neurological and neurodynamic tests, kappa 0.5 and agreement 85% with broad or moderate CIs. Regarding intratester reliability, the overall agreement was at an acceptable level in these test categories. However, in some single tests in the mobility and provocation test categories, reliability between the test sessions was poor (kappa 0.10).

TABLE 5 Kappa-values of inter- and intratester reliability of selected clinical tests in 7 test categories, and odds ratios (OR), sensitivities, and specificities between chronic LBP (CLBP) and controls, and subacute LBP (SLBP) and controls, (95% CI).

test categories and tests	intertester reliability	intratester reliability	CLBP vs Controls			SLBP vs. Controls		
	kappa	kappa	OR	Sensitivity	Specificity	OR	Sensitivity	Specificity
1. Functions								
Walking	0.8 (0.4 to 1,2)	0.6 (0.4to 1.2)	26 (3 to10)	33 (19 to 45)	100 (93to100)		2 (< 1 to 10)	100 (93 to 100)
Undressing	1.0 (1.0 to 1.0)	1.0 (1.0 to 1.0)	20 (3 to100)	27 (15 to 38)	100 (93to100)		4 (2 to 15)	100 (93 to 100)
Walking on toes	1.0 (1.0 to 1.0)	1.0 (1.0 to 1.0)	9 (1.to76)	15 (5 to 24)	100 (93to100)		2 (< 1 to 10)	100 (93 to 100)
Heel-walking	1.0 (1.0 to 1.0)	1.0 (1.0 to 1.0)	17 (2to100)	24 (5 to 28)	100 (93to100)		2 (< 1 to 10)	100 (93 to 100)
Squat and rise	1.0 (1.0 to 1.0)	1.0 (1.0 to 1.0)	6 (2to18)	38 (19 to 46)	87 (75 to 90)		15 (6 to 27)	69 (54 to 81)
Mean	**0.9 (0.4 to 1.2)**	**0.9 (0.4 to 1.2)**						
2. Inspection								
Posture of lumbar spine	0.4 (-0.1 to 0.5)	0.3 (0.1to 0.7)	3 (1 to 6)	72 (59 to 84)	58 (39 to 69)	3 (1 to 6)	72 (59 to 84)	58 (39 to 69)
Leg length difference	0.1 (-0.1 to 0.2)	0.5 (0.3to 0.9)		15 (8 to 29)	75 (54 to 81)		45 (15 to 39)	80 (60 to 86)
Posture of knee	0.1 (-0.1 to 0.3)	0.7 (0.4to 0.9)		15 (8 to 29)	93 (75 to 95)		16 (8 to 29)	93 (75 to 95)
Posture of feet	0.2 (-0.1 to 0.5)	0.6 (0.3to 0.9)		33 (15 to 39)	80 (60 to 86)		40 (8 to 29)	75 (54 to 81)
Mean	**0.2 (-0.1 to 0.6)**	**0.5 (0.1 to 0.9)**						
3. Mobility								
Lumbar spine flexion	0.5 (0.3to 0.9)	0.7 (0.4to 1.0)	27 (7 to100)	51 (37 to 65)	96 (89 to 100)	15 (3to69)	36 (37 to 65)	96 (89 to 100)
Lumbar spine extension	0.6 (0.3to 0,9)	0.1 (-0.2to0.4)	6 (3to13)	62 (46 to 75)	78 (71 to 93)	5 (2 to 10)	57 (48 to 75)	78 (70 to 93)
Lumbar spine lateral flexion	0.4 (0.1to 0.8)	0.6 (0.3to 0.9)	4 (2 to 9)	49 (35 to 63))	80 (67 to 91)	4 (2 to 9)	49 (35 to 63))	81 (67 to 91)
Hip rotation	0.6 (0.1 to 1.0)	0.6 (0.3to 0.9)		4 (3 to 20)	93 (77 to 97)		9 (3 to 20)	93 (77 to 97)
Specific p-a mobility	0.4(0.2 to 0,9)	0.7 (0.3to 1.0)	5 (2to11)	84 (71 to 92)	47 (37 to 67)	13 (6 to 47)	94 (71 to 99)	47 (33 to 67)
Mean	**0.5 (0.1 to 1.0)**	**0.5 (-0.2 to 1.0)**						
4. Pain provocation								
Extension with traction	0.7(0.4 to 0.9)	0.4 (-0.1to0.7)		31 (13 to 39)	80 (60 to 86)		23 (5 to 29)	80 (59 to 86)
Physiological movement	0.5 (0.2 to 0.8)	0.4 (0.1to 0.7)		35 (15 to 40)	87 (75 to 90)		23 (4 to 28)	87 (72 to 90)
Post pelvic pain provocation	0.8 (0.6 to 1.0)	0.6 (0.3to 1.0)		22 (5 to 28)	87 (754 to 90)		10 (6 to 28)	93 (82 to 99)
Interspinosus pain	0.5 (0.2 to 0,9)	0.7 (0.3to 1.0)	5 (2to11)	78 (65 to 88)	58 (48 to 78)	3 (1 to 7)	68 (65 to 88)	58 (48 to 78)
Kibler`s skin rolling	0.2 (-0,1to 0.5)	0.2(-0.1to0.5)	4 (2 to 9)	47 (34 to 64)	82 (70 to 93)		17 (34 to 64)	82 (70 to 93)
Sacroiliac joint	0.3 (-0.3to 0.7)	0.1 (-0.2to0.4)		13 (5 to 24)	93 (83 to 99)		2 (5 to 24)	93 (83 to 99)

provocation								
L4, L5 rotation provoc.	0.3(-0.1to 0.6)	0.6 (0.2 to 0.8)	5 (1to20)	24 (15 to 39)	95 (86 to 99)		9 (10 to 39)	76 (76 to 99)
Mean	**0.5 (-0.1 to 1.1)**	**0.4 (-0.1 to 1.0)**						
5. Muscle tightness								
Hamstrings	0.8 (0.5to 1.0)	0.8 (0.5to 1.0)		53 (36 to 66)	62 (49 to 78)		64 (36 to 66)	62 (49 to 78)
Piriformis	0.1 (-0.2 to 0.3)	0.7 (0.4to 1.0)		15 (6 to 27)	95 (39 to 69)		10 (6 to 27)	95 (39 to 69)
Gluteus medius/minimus	0.3 (-0.1to 0.7)	0.3 (-0.2to0.6)		15 (6 to 27)	95 (39 to 69)		16 (6 to 31)	95 (32 to 69)
Iliopsoas	0.2 (-0,1 to 0.5)	0.7(0.4to 1.0)	3 (1 to 6)	73 (59 to 84)	51 (39 to 69)	7 (3 to 19)	87 (59 to 84)	51 (39 to 69)
Mean	**0.4 (-0.1 to 1.1)**	**0.6 (-0.2 to 1.0)**						
6. Stability								
One-leg standing	0.7 (0.3 to 1.0)	0.6(0.1to 0.9)	x	x	x	x	x	x
Active SLR	0.6 (0.3to 0.9)	0.6 (0.3to 0.9)	x	x	x	x	x	x
Isometric lumbar extension	0.5 (0.1 to 0.8)	0.5 (0.2to 0.7)	12 (2 to 39)	75 (59 to 84)	87 (75 to 93)		32 (31 to 65)	80 (60 to 86)
Transverses abdominis	0.5 (0.1 to 0.8)	0.6 (0.3to 0.9)	6 (3to 13)	66 (46 to 75)	80 (67 to 91)		23 (6 to 29)	78 (70 to 93)
Mean	**0.5 (0.1 to 1.0)**	**0.6 (0.1 to 1.0)**						
7. **Neurol/neurodynam**								
SLUMP in sitting	0.3 (-0.2to 0.6)	0.2 (0.1to 0.4)	6 (2to16)	42 (29 to 56)	89 (80 to 98)	3 (1 to 8)	28 (29 to 56)	89 (80 to 98)
SLR	0.9 (0.4to 1.0)	0.8 (0.4to 1.0)	5 (2to17)	29 (19 to 45)	93 (86 to 99)	3 (1 to 9)	17 (16 to 45)	93 (86 to 99)
Achilles reflex	0.5 (0.2 to 0.9)	0.5 (0.2to 0.8)	3 (1to10)	16 (21 to 50)	93 (72 to 97)	2 (1 to 8)	15 (24 to 50)	95 (77 to 97)
Patella reflex	0.4 (0.1to 0.8)	0.4(0.1to 0.8)	3 (1 to 8)	24 (25 to 50)	95 (77 to 97)	3 (1to 12)	21 (20 to 50)	93 (71 to 97)
Ely`s test	0.7 (0.3 to 0.9)	0.7 (0.4to 1.0)		6 (2 to 18)	100 (93to100)		40 (8 to 59)	75 (54 to 81)
Mean	**0.5 (-0.2 to 1.2)**	**0.7(0.1 to 1.0)**						

x = not recorded, CLBP = chronic low back pain, SLBP = subacute low back pain, SLR = straight leg raise, ASLR = active straight leg raise, SLUMP = neurodynamic provocation test, Ely`s test = femoral nerve tension test, p-a =posterior-anterior, OR = odds ratio

5.1.2 Clinical tests to differentiate chronic and subacute LBP patients from controls (Table 5)

After analyzing the 34 clinical tests in seven categories, the isometric lumbar extension test was able *to differentiate CLBP from controls* at an acceptable level (OR 12), with sensitivity of 75% and specificity of 87%, and five tests at a moderate level, with significant odds ratio.

In the inspection test category, back posture was at an OR 3 with moderate sensitivity 72% and specificity 58%. High OR values also appeared in the mobility category in the lumbar extension test (OR 6) with sensitivity 62% and specificity 78% and in pain provocation tests interspinosus pain (OR 5) with sensitivity 75% and specificity 58%. In the muscle tightness category, the hip flexion test had the highest OR (3), and sensitivity 73% and specificity 51%. Despite high ORs in neurological and neurodynamic tests, no test had a sensitivity better than 42% (SLUMP-test).

Clinical differences between SLBP patients and controls were following:

Of the 34, 5 tests were able to differentiate SLBP from the controls at a moderate level. The tests were: back posture (OR 3) with sensitivity 72% and specificity 58%, lumbar spine extension (OR 5) with sensitivity 57% and specificity 78%, and the specific PA mobility test (OR 13) with sensitivity 94% and specificity 47%. For pain provocation tests, the interspinosus ligament pain test had the best values: OR (3), sensitivity 68%, and specificity 58%. In the muscle tightness category, the hip flexion test had the highest OR (7) and a sensitivity of 87% and specificity of 51%.

The tests which differentiate CLBP from controls were the same that also differentiated the SLBP group from controls, except in the isometric extension test.

5.2 Subgroup classification

Inter-tester reliability of clinicians' ability independently to classify patients with LBP, utilizing clinical tests and history-based classification methods

The reliability of classification of LBP patients in their early phase into five subgroups was tested in *Study III* with two cohorts of 20 subjects in municipal health care and 30 subjects in occupational health care *(Tables 6-7)*.

In the first cohort of the study, one physiotherapist with long experience in OMT and four experienced physiotherapists without OMT specialization examined 20 patients. LBP subgroup prevalence was as follows: Clinical instability and discogenic pain were the most common (35% and 30%), followed by segmental dysfunction/facet pain (18%), sacroiliac joint pain/dysfunction (10%), and clinical lumbar spinal stenosis (7%).

In the second cohort, where two OMT specialists examined 30 patients in an occupational health care, clinical instability and discogenic pain were also the most frequent subgroups (43% and 37%), followed by segmental dysfunction/facet pain (7%), sacroiliac joint pain/dysfunction(7%), and clinical lumbar spinal stenosis (6%).

Due to the few subjects in the segmental dysfunction/facet pain, sacroiliac joint pain/dysfunction, and clinical lumbar spinal stenosis subgroups, proper analysis could not be performed for these categories. Overall intertester agreement was 70% and overall Kappa coefficient 0.60 between the physiotherapist with long experience in OMT and the four physiotherapists without OMT specialization. Overall agreement between the OMT specialists was 77% and overall Kappa 0.65.

TABLE 6 Inter-examiner reliability in assessment of low back pain subgroup classification between a physiotherapist specialized in OMT (E.K.) and also by four physiotherapists (A.H., I.L., S.S., P.V.) without OMT specialization. Number of subjects was 20.

LBP classification	Number of positive and negative observations by examiner		Agree[1]	Disagree[2]	Agreement %	Kappa (95 % CI)
	Examiner EK	Examiners AH, IL, SS, PV				
Discogenic pain	7	5	18	2	90	0.76 (0.35 to 1.00)
Clinical instability	6	8	16	4	80	0.57 (0.14 to 0.90)
Clinical lumbar spinal stenosis	2	1	19	1	95	Φ
Segmental dysfunction/facet pain	3	4	15	5	75	Φ
Sacroiliac joint pain/dysfunction	2	2	20	0	100	Φ
Overall agreement and Kappa					70	0.60 (0.40 to 0.85)

[1]Examiners agree that a given number of patients had the classification
[2]Examiners disagree whether patients had the classification
Φ=calculation of Kappa impossible, due to low number of subjects in subgroups

TABLE 7 Inter-examiner reliability in assessment of low back pain subgroup classification between two physiotherapists specialized in OMT (expert M.P. vs. novice J.R.). Number of subjects was 30.

LBP classification	Number of positive and negative observations by examiner		Agree[1]	Disagree[2]	Agreement %	Kappa (95 % CI)
	Examiner M.P.	Examiner J.R.				
Discogenic pain	10	12	26	4	87	0.71 (0.40 to 0.93)
Clinical instability	13	13	24	6	80	0.59 (0.28 to 0.86)
Clinical lumbar spinal stenosis	2	2	30	0	100	Φ
Segmental dysfunction/facet pain	3	1	26	4	75	Φ
Sacroiliac joint pain/dysfunction	2	2	30	0	100	Φ
Overall agreement and Kappa					77	0.65(0.38 to 0.89)

[1]Examiners agree that a given number of patients had the classification
[2]Examiners disagree whether patients had the classification
Φ=calculation of Kappa impossible, due to low number of subjects in subgroups

5.3 Treatment effects

Effects of OMT compared to the McKenzie method and advice only to stay active for LBP patients among working adults.

Treatment effects were compared between two physiotherapy approaches (OMT and McKenzie) to one counseling (Advice) session with a physiotherapist concerning the good prognosis for LBP, pain tolerance, medication, and early return to work. This counseling was reinforced with a booklet including the same information which patients had received from the physiotherapist, and also self-treatment exercises. In the OMT and McKenzie groups, the number of visits ranged from three to seven. During OMT and McKenzie visits patients also received explanation and information about LBP including self-management at home-exercise.

The absolute values of pain and disability indices and percentages of patients on sick leave because of LBP at the 3-, 6-, and 12-month follow-up are

in Table 8. Mean changes from baseline at 3-, 6- , and 12-month follow-up points in leg pain, low back pain, and Roland-Morris disability index are in Figure 5.

TABLE 8 Outcome measures at baseline and 3-, 6-, and 12-month follow-up visits, mean and standard deviation (sd).

	OMT[1] (n=45)	McKenzie[2] (n=52)	Advice[3] (n=37)
Baseline values			
Leg pain (VAS, mm) γ	27 (26.2)	23 (25.4)	20 (21.6)
Low back pain (VAS, mm) γ	36 (19.5)	35 (19.7)	36 (2.8)
Roland-Morris (0-24) ε	9 (5.7)	9 (4.6)	7 (4.1)
On sick leave because of LBP (%)	16	17	8
Outcome measures at 3 months			
Leg pain (VAS, mm) γ	12 (18.2)	4 (8.4)	9 (12.9)
Low back pain (VAS, mm) γ	20 (16.6)	14 (15.4)	22 (17.9)
Roland-Morris (0-24) ε	3 (4.9)	3 (4.0)	2 (4.2)
On sick leave because of LBP (%)	5	26	11
Outcome measures at 6 months			
Leg pain (VAS, mm) γ	9(15.1)	4 (7.1)	16 (18.5)
Low back pain (VAS, mm) γ	19 (17.8)	12(10.3)	29 (20.2)
Roland-Morris (0-24) ε	2 (3.8)	1 (1.8)	3 (4.9)
On sick leave because of LBP (%)	6	7	19
Outcome measures at 12 months			
Leg pain (VAS, mm) γ	10 (17.6)	6 (10.5)	13 (16.7)
Low back pain (VAS, mm) γ	16 (17.3)	15 (19.4)	20(16.1)
Roland-Morris (0-24) ε	2 (3.7)	1 (2.1)	3 (4.5)
On sick leave because of LBP (%)	3	4	23

[1]Orthopedic manual therapy
[2]McKenzie method
[3]Advice-only group
γ Self-reported measures included a visual analogue scale (VAS).
ε 0-24 point scale on Roland-Morris Disability questionnaire.

Figure 4 gives the treatment effects with 95% confidence intervals in LBP, leg pain and disability index (Roland-Morris) at 3-, 6-, and 12- month follow-up points.

Low back pain (VAS)

At the 3-month follow-up point, mean back pain (VAS) decreased in all groups 12 to 21 mm from baseline, but no treatment effect appeared when comparing the OMT or McKenzie groups to the Advice-only group. A small effect-size (ES: 0.3¸ 95% CI -0.16 to 0.76) was evident in the McKenzie and an insubstantial one (0.04 95% CI -0.43 to 0.51) in the OMT group compared to the Advice-only group. The back pain difference was -1 mm between the OMT group and the

Advice-only group. The corresponding difference between the McKenzie group and the Advice-only group was -7 mm (Figure 4). Differences were not significant.

After the 6-month follow-up, mean back pain (VAS) decreased from baseline 20 to 24 mm in the therapy groups, and 4 mm in the Advice-only group. A small ES appeared between the OMT and the Advice-only group (ES: 0.42, 95% CI -0.08 to 0.91), and a medium ES between the McKenzie and the Advice-only group (ES: 0.76, 95% CI 0.28 to 1.26). The improvement in back pain was significantly (p=0.04) 15 mm better in the McKenzie group than in the Advice-only group. No significant difference emerged (mean difference 10 mm, p= 0.07) between the OMT and the Advice-only groups (Figure 4).

At the 12-month follow-up mean back pain (VAS) in all groups decreased 14 to 22 mm from baseline, but we saw no treatment effects when comparing the OMT or McKenzie groups with the Advice-only group. The smaller effect sizes occurred in the OMT (0.23; 95% CI -0.28 to 0.74) and McKenzie (ES: 0.17; 95% CI -0.31 to 0.65) groups compared with the Advice-only group. The difference in back pain from baseline in the OMT group compared to the Advice-only group was –4 mm (p=0.27), and the corresponding value with the McKenzie group was -4 mm (p=0.14) (Figure 4). Differences were not significant.

Leg pain (VAS)

At the 3-month follow-up point, mean leg pain (VAS) decreased in all groups 12 to 15 mm from baseline, but no treatment effect appeared when comparing the OMT or McKenzie groups to the Advice-only group. A small effect-size appeared in the McKenzie (ES: 0.3¸ 95% CI -0.1 to 0.77) and an insubstantial in the OMT group (ES: 0.1; 95% CI -0.29 to 0.59) compared to the Advice-only group. The leg pain difference was -4 mm (p=0. 96) between the OMT group and the Advice-only group. The corresponding difference between the McKenzie group and the Advice-only group was -5 mm (p=0.38) (Figure 4). No significant differences occurred.

At the 6-month follow-up mean leg pain decreased from baseline 15 to 16 mm in therapy groups and only 8 mm in the Advice-only group. A medium effect size was evident in the McKenzie (ES: 0.6; 95% CI 0.16 to 1.05) and OMT (ES: 0.5; 95% CI 0.07 to 0.96) groups compared to the Advice-only group. The non-significant change-difference (p=0.14) in leg pain in the OMT group compared to the Advice-only group was -8 mm, corresponding to a significant difference from the McKenzie group`s -7mm (p=0.01) (Figure 4).

At the 12-month follow-up mean leg pain had decreased 16 to 18 mm in therapy groups and 8 mm in the Advice-group from baseline, but no treatment effect emerged when comparing the OMT or McKenzie groups to the Advice-only group. A small effect size was evident in the McKenzie (ES: 0.4; 95% CI -0.03 to 0.85) and OMT (0.3; 95% CI -0.10 to 0.77) groups compared to the Advice-only group. Differences in leg pain in the OMT group compared to the Advice-only

group was -10 mm (p=0.27); the corresponding values with the McKenzie group was -8 mm (p=0.14) (Figure XI). No significant differences occurred.

Decrease in back pain (VAS) in the 1-year follow-up was more than 20 mm in all groups. Decrease in leg pain (VAS) was in the OMT group 18 mm and 16 mm in the McKenzie group, and 8 mm in the Advice-only group (Table 11). Decrease in back pain (VAS) of 20 mm or more occurred in the OMT group with 51% of the patients, in the McKenzie group with 44%, and in the Advice-only group with 40%.

Disability (R-M)

At the 3-month follow-up point, mean disability (R-M) had decreased in all groups by? 7 to 8 points, but when comparing the OMT or McKenzie groups to the Advice-only group, no treatment effect emerged. A small effect size (ES: 0.2; 95% CI -0.26 to 0.61) was evident in the McKenzie and (0.2; 95% CI --0.26 to 0.62) OMT groups compared to the Advice-only group. The non-significant difference in disability index was -1 point in the OMT (p=0.90) and McKenzie (p=0.75) groups compared to the Advice-only group (Figure 4).

At the 6-month follow-up mean disability decreased in the therapy groups 8 to 9 points and 7 points in the Advice-group from baseline. A medium effect size (ES: 0.5; 95% CI 0.09 to 0.99) occurred in the OMT group compared to the Advice-only group and (0.7; 95% CI 0.28 to 1.17) in the McKenzie group compared to the Advice-only group. The non-significant difference in disability index was -3 point (p=0.06) in the OMT group and the significant difference was -4 points (p=0.003) in the McKenzie group compared to the Advice-only group (Figure 4).

At the 12-month follow-up, mean disability was decreased in all groups 4 to 8 points from baseline. A medium effect size (ES: 0.8; 95% CI 0.34 to 1.24) was apparent in the McKenzie (0.6; 95% CI 0.14 to 1.04) and OMT groups compared to the Advice-only group. The non-significant change difference in disability index was -3 points (p=0.06) in the OMT group and a significant -4 points difference (p=0.03) in the McKenzie group compared to the Advice-only group (Figure 4).

No significant differences emerged between the OMT and McKenzie groups in pain and disability scores at any follow-up point. In addition, no inter-group differences emerged during follow-up in visits to physicians or other health care professionals or in use of pain-killers.

Sick leave (days)

At the 3-month follow-up point mean days of sick leave were decreased in the OMT group by 3 days, showed no change in the Advice-only group, but we saw an increase of 8 days in the McKenzie group from baseline. A medium effect size (ES: 0.5; 95% CI 0.05 to 0.94) occurred in the OMT group compared to the Advice-only group, and in the Advice-only group compared to the

McKenzie group (0.5; 95% CI -0.09 to 0.97). Change in days of sick leave from baseline in the OMT group was -3 days and in the McKenzie group +8 days compared to the Advice-only group. Between the OMT and McKenzie groups, the difference was significant.

At the 6-month follow-up mean days of sick leave were decreased from baseline in the OMT group -3 days, and -2 days in the McKenzie, and increased by +1 day in the Advice-only group. A medium effect size (ES: 0.6; 95% CI 0.14 to 1.04) occurred in the OMT group (0.6; 95% CI 0.15 to 1.03) and McKenzie groups compared to the Advice-only group. Change in days of sick leave from baseline in the OMT group was -4 days and in McKenzie group -3 days compared to days in the Advice-only group. A significant difference between OMT and McKenzie groups compared to Advice-only group appeared.

At the 12-month follow-up, mean days of sick leave decreased in the OMT group by -3 days, but an increase of +1 day occurred in the McKenzie and +4 days in the Advice-only group from baseline. A small effect size (ES: 0.4; 95% CI -0.06 to 0.83) was seen in the OMT group and insubstantial (0.1; 95% CI -0.31 to 0.55) in the McKenzie group compared to the Advice-only group. Change in days of sick leave from baseline in the OMT group was -7 days and in the McKenzie group -3 days compared to the Advice-only group. No significant differences occurred.

TABLE 9 Sick leave days at 12-month follow-up because of low back pain in OMT-, McKenzie- and Advice groups, mean (95%CI).

Variable	OMT	McK	Advice	P-value≠
Days of sick leave	1.4(03 to 5.3)	12.0 (6.0 to 25.6)	6.5 (1.5 to 30.9)	0.004(OMT/McK)

≠ confidence interval obtained by bias corrected and accelerated bootstrapping (5000 replications)
OMT = Orthopedic manual therapy
McK = McKenzie

There emerged fewer days of sick leave because of LBP in the Advice-only group prior to treatment than in the McKenzie and OMT groups, but during treatment periods and in the 12-month follow-up were more days of sick leave in the McKenzie group than the Advice-only and OMT groups. The difference between the OMT and McKenzie groups was statistically significant (Table 9). Despite the high percentage of patients on sick leave at the 6- and 12- month follow-up in the Advice-only group, the difference compared to other groups was not statistically significant.

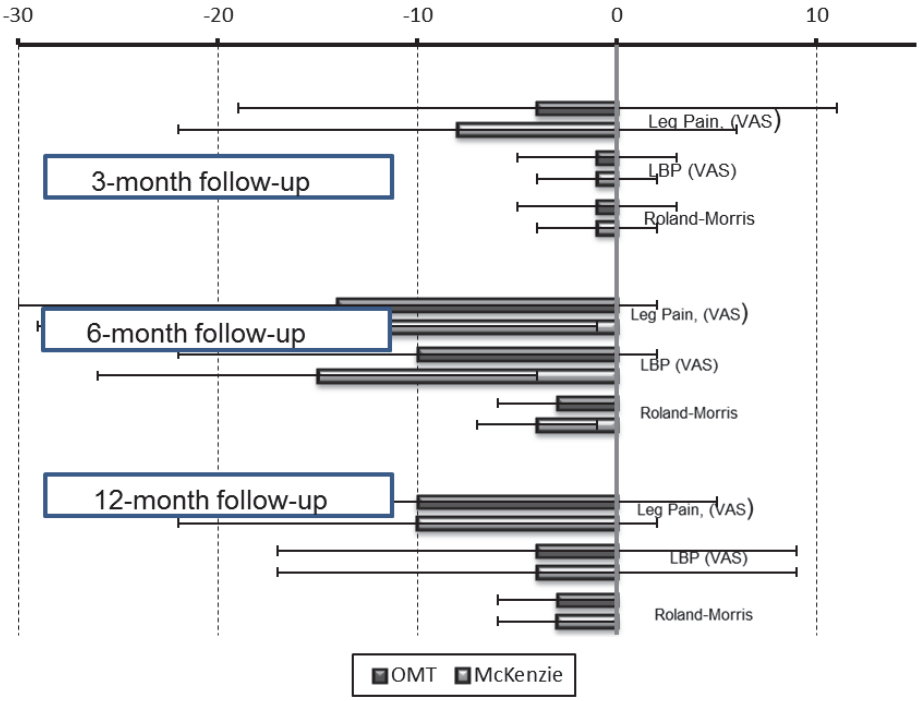


FIGURE 4 Treatment effects comparing OMT and McK groups to Advice only. Mean changes from baseline at 3-, 6-, and 12-month follow-up points in leg pain, low back pain VAS (0-10), and Roland-Morris (0-24) disability index among 26 participants in the Advice only, 35 in the OMT, and 45 in the McKenzie method group who completed the 12-month follow-up. Error bars represent 95% confidence intervals, and p-values indicate treatment effects in the OMT(#) and the McKenzie method(¤) groups compared with the Advice only group.

6 DISCUSSION

The objective of this thesis was to evaluate the reliability and predictive value of selected clinical tests and the reliability of subgroup classification of LBP patients based on these tests, and to compare OMT to the McKenzie method and to Advice-only.

6.1 Study sample

Participants in all studies comprised working-age LBP patients who visited their local health care or occupational health care centers, and whose current LBP episode had lasted less than 3 months. Only in Study II, did the participants have CLBP, and 42% of them were on sick leave. Subjects who remain in primary health care, as in the present study, except for the chronic LBP group, may be expected to be less disabled than patients who are referred to a specialized secondary setting (Denison et al 2007). The participants, most working, represented employed LBP patients quite well in the early phase, making the group more homogenous for interpreting of results.

In Study IV, the patient sample was adequate for statistical power. Our power analysis showed that the F-test will detect between-group differences equal to those implied by the sample size. The study population can be compared with that in studies published by Cherkin et al (1998), Wand et al. (2004), and Niemistö et al (2005), regarding average age, number of participants, level of pain and disability, and days of sick leave.

In Study IV, The aim was to gather 180 patients to achieve sufficient power for statistical measurements, but the number of subjects in Study IV remained at 136.

In these three groups, 8 to 17% of participants in these three groups were on sick leave. These subjects represent those typical patients who are usually seen in primary health care, such as in occupational health care, or who seek treatment from a variety of caregivers. It is postulated, even though this

hypothesis is unproven in larger RCTs, that if treatment strategies are not planned well enough, these patients will sooner or later have recurrences of LBP (Low Back Pain: Early Management of Persistent Non-specific Low Back Pain. National Collaborating Centre for Primary Care (UK). London: Royal College of General Practitioners (UK); 2009. National Institute for Health and Clinical Excellence: Guidance). For that reason, it is considered important to prevent first-time or recurrent LBP from becoming persistent and disabling. For early interventions to prevent new episodes and activate return to work depends on the target and source of the population, especially if patients are identified as being at high risk for developing chronic LBP (van Duijn et al. 2010, Whitfill et al. 2010). Looking at the results in our RCT in the 1-year follow-up, the recurrence rate for LBP episodes was low, especially with therapy groups, thus favoring this type of physiotherapy with these kinds of LBP patients. The recurrence rate was lower, and more patients had completely recovered in terms of pain and disability (back and leg VAS <10mm, and R-M <1point) compared to results by Croft et al. (1998).

6.2 Predictive factors of clinical tests

In analyzing all seven test categories, the mean *inter- and intratester* kappa and agreement were fair, ranging from 0.5 to 0.6 (Kappa) and 79 to 81 (agreement%). The inspection tests category showed the lowest reliability values and the functional tests the best values; in other categories, reliability was moderate. Even though overall inter- and intratester reliability was at an acceptable level in these test categories, in some single tests reliability between testers ranged from poor to good, except in the functional test category. Despite the reasonable Kappa – and agreement values, the confidence intervals were wide making interpretation unreliable. Such a situation offers a risk of misinterpretation of non-significant results of small studies, especially with a small number of participants (Altman 2005). If statistical significance is not reached, post hoc power analysis should be conducted in an attempt to rule in or out inadequate power (Onwuegbuzie and Leech, 2004).

The odds ratio of *Functional tests* for chronic LBP patients compared to other groups was high, indicating functional difficulties in this group. However, the differentiation capacity of clinical tests depends even more on these tests` sensitivities and specificities, which should both be at the level of 70% or even higher.

High OR was often accompanied by high specificity, meaning that for those not having LBP, the test was negative (normal value). But when the sensitivity was on some occasions high with high OR, the specificity was low, meaning that the differentiating capacity of those tests between LBP and non-LBP patients was not good, for time-related LBP. These tests might have a better differentiation capacity in patho-anatomical/patho-physiological

subgroups. Rubenstein and van Tulder (2008) have stated in their summary of findings for the diagnostic procedures for neck and LBP, that only a few tests and measures are valuable in ruling out or ruling in the condition.

The differentiation capacity of certain tests will be reduced by poor or fair inter- and or intratester reliabilities. Especially intratester reliability of lumbar spine posture, the lumbar spine extension test, physiological movement, and iliopsoas tests were low.

6.3 Subgroups

When using these 34 clinical tests to classify LBP patients in the early phase into five clinical subgroups, overall intertester agreement between the physiotherapist with long experience in OMT and the physiotherapists without OMT specialization, and between the two OMT specialists was moderate, and at the same level. The reliability of the McKenzie classification system has been shown to be reliable in the LBP sub-grouping classification for suitably trained examiners, but not for a minimally trained or untrained assessor (Kilpikoski et al 2002, Clare et al 2005). No similar studies with other classification systems comparing sub-group classification between experts and novice were found.

In Study III, the clinical lumbar spinal stenosis and sacroiliac joint pain/dysfunction subgroups were too small to reveal clinical implications for statistical evaluation, and thus, it seemed almost impossible to interpret their clinical usefulness. The low number of patients classified in the clinical lumbar spinal stenosis subgroup can be explained by the fact that there are often older than subjects in our study. Patients with sacroiliac joint pain/dysfunction are, in fact, often women having pain during or soon after childbirth or both. Those studies differ from our study subjects where no one pregnant included like in the study by Robinson et al. (2010). It could be also that these two subgroups are rarer than other subgroups. Since patients in discogenic pain and with a clinical lumbar spinal stenosis complaint of radiating pain, where the first aim in the early phase is to reduce that radiating pain with traction or with neural tissue mobilization or both, these subgroups could be combined in the same group. Patients in the sacroiliac joint pain/dysfunction subgroup have much in common with those with clinical instability: local back or pelvic pain, lack of stability and movement control. These two subgroups might be able to combinen (Vleeming et al, 1997). The third subgroup in future studies could be our segmental dysfunction/facet pain group.

In Study III, 30 to 37% of subjects were classified as in the discogenic pain subgroup if patients were complaining of radiating pain which could be provoked in flexion, and in SLR or SLUMP tests or both. In the McKenzie method this subgroup could be compared to the derangement subgroup, and radiating pain could be called “centralization phenomena” (Berthelot et al. 2007). Comparison between clinical tests to sub-classify patients into the

discogenic pain category in OMT classification or into the derangement syndrome category in the McKenzie method; comparison of treatment effects based on these classifications would be fruitful because both classifications were used in our Study IV in the early phase of LBP.

Arguments exist for and against subgrouping. One of the most compelling arguments is that clinicians believe in subgroups (Foster et al. 2011). Some have suggested that the current system for grouping patients in trials testing treatments for LBP is inadequate (McCarthy and Cairns, 2005). There exist arguments against subgrouping, stating that subgroups are as yet unsupported by data (Smeets et al. 2009). Although we found subgrouping as being quite reliable within and between testers, we did not randomize participants into subgroups in our RCT, rather into therapy groups with different philosophies underlying their methods.

6.4 Clinical trial

In Study IV, 136 LBP patients recruited from four occupational health care centers were allocated to OMT, McKenzie, and to Advice-only groups. At baseline, subjects in these subgroups did not differ from each other in disability, or in low back or leg pain. At baseline disability was at the same level as in studies by Niemistö et al. 2005, and Luomajoki et al. 2010, although pain (VAS) was 10mm higher in the Niemistö study.

Our results are in line with those of similar pragmatic studies (Cherkin et al. 1998, Brealey et al. 2003, Niemistö et al. 2005, Cairns et al. 2006, Ferreira et al. 2006) showing the effectiveness of physiotherapy at least in one-year follow-up. In the study by Niemistö, subjects were CLBP patients, and in other studies, patients were acute or subacute. We also found significant improvement in pain and disability in all groups with no differences between the groups at the 3-months visit. This may be due to good spontaneous recovery in the short term despite the fact that recurrence of LBP is frequent (Croft et al. 1998).

The recurrence rate was lower than in other epidemiological studies, with those showing recurrence at 1 year ranging from 24 to 80% (Hoy et al. 2010), when in our study recurrence was 15%. Of all patients 31% had completely recovered in terms of pain and disability at the 12-moth follow-up. Only three participants, who did not undergo spinal surgery, felt that their back pain was worse. No other negative effects were reported.

When comparing the effects of OMT and McKenzie to one counseling session with a physiotherapist (Advice-only) for treating low back pain/leg pain, and reducing disability and sick leave, no significant differences emerged between the OMT and McKenzie groups in pain and disability scores at any follow-up point, only in sick leave. Despite the broad variance of sick leave days, the difference between OMT and McKenzie groups was statistically significant (Table 9). The increase in days of sick leave occurred in McKenzie

group during the treatment period, in the first 3 months. Also two centralizers and two non-centralizers in that group were referred to orthopedic surgeon for discectomy.

Compared to the Advice-only group, the OMT and McKenzie methods seemed marginally more effective than was one session when patients were assessed for pain and disability. This difference occurred usually at the 6-month follow-up point.

Even absolute values between groups differed by several mm on the VAS scale or by several points on the RM scale, confidence intervals were quite large, and differences between groups were not statistically significant. After 12 months, all groups had only minimal pain and disability. Despite these large confidence intervals, in the Advice-only group those patients who had radiating pain (Leg VAS), showed less improvement than did other groups, and had increasing days of sick leave because of LBP after 12 months.

The minimally clinically important change (MCIC) for pain on VAS should be at least 20 mm (Ostello and de Vet, 2005, Kovacs et al. 2007). At the 12-month follow-up back pain (VAS) decreased in all groups 14 to 22 mm from baseline, and leg pain decreased 18 to 20 mm in the therapy groups and 5 mm in the Advice-only group. The mean decrease in VAS was in percentages high, even though the absolute values were moderate. Explanation for this is rather low baseline values. In therapy groups back pain (VAS) decreased 59 to 60% (VAS -20 to -21), and in Advice-only group up to 45% (VAS -16), and leg pain (VAS) in therapy groups 74 to 75% (VAS -18 to -20) and in the Advice-only group 26% (VAS -5). However, absolute values in changes of VAS and R-M are better related to clinical significance than are values in percentages

The minimally clinically important change (MCIC) for functional disability measured with R-M should be at least 3.5 points (Ostello and de Vet, 2005, Kovacs et al. 2007). Decrease in R-M was 8 to 9 points in all groups. The decrease of R-M was in percentages high, and the absolute values were moderate. In therapy groups R-M decreased 80 to 85% (R-M –7 to -8), and in Advice-only group 64% (R-M -5). Decrease in VAS and R-M were also clinically significant.

Whether LBP physiotherapy based on OMT or on McKenzie principles was superior to Advice-only could not be shown. There may be several reasons. Because of the randomization, the groups are supposed to be comparable. However, groups perhaps included the type of patient, who would have benefitted from participating in other groups. On the other hand, all groups improved very well, and no statistically meaningful differences emerged, only a trend toward the treatment groups` doing slightly better.

The definition of acute LBP differs among recent European guidelines (Koes et al. 2010), and the earlier Quebec Task force in Canada in 1987 and the Agency for Health Care Policy and Research guidelines in the United States in 1994 (Spitzer et al. 1987). In the earlier guidelines, acute LBP was defined as lasting from 0 to 7 days and subacute from 7 days to 7 weeks, and in recent guidelines acute LBP from 0 to 6 weeks and subacute from 6 weeks to 3 months,

thus making interpretation of the borderline between acute and subacute difficult (Spitzer et al.1987).

The first evaluation, which consisted of clinical and questionnaire assessments, was performed 3-months after study commencement. This was a point where differences in sick leave were largest between OMT and McKenzie groups (Table 9). At least one explanation for the higher number of sick leave days could be the ideology of the McKenzie method, that in the absence of "centralization phenomena," patients are referred back to a medical doctor for re-evaluation (Berthelot et al. 2007). The other possible explanation is that the waiting time for a back pain specialist is often several weeks, and patients may during that time be on sick leave.

Advice to stay active has for a long time been the recommendation of best practice in international evidence-based guidelines for the management of acute LBP, despite some conflicting evidence (Williams et al. 2010). Malmivaara et al. (1995) concluded that among patients with acute LBP, continuing ordinary activities within the limits permitted by the pain leads to more rapid recovery than either bed rest or back-mobilizing exercises. The McKenzie method and chiropractic manipulation in acute LBP had similar effects and costs, and patients receiving these treatments had only marginally better outcomes than did those receiving the minimal intervention of advice (Cherkin et al. 1998).

In a systematic review of randomized controlled trials involving advice for the management of LBP (Liddle et al. 2007), the authors concluded: "Advice to stay active is sufficient for acute LBP. No conclusions could be drawn as to the content and frequency of advice that is most effective for subacute LBP, due to the limited number or poor quality of RCTs. For chronic LBP evidence is strong to support the use of advice to remain active in addition to specific advice relating to the most appropriate exercise, and/or functional activities to promote active self-management."

Few studies have investigated whether prognostic indicators, ones which contribute to the transition from acute to chronic LBP, are also those which contribute to the persistence of chronic LBP. Grotle et al (2010) showed that the strongest prognostic indicators for disability at 12 months in both LBP groups were being unemployed, having widespread pain, and a high level Chronic Pain Grade, and catastrophizing. Liddle et al. (2007) stated that the effectiveness of treatment for acute and subacute symptoms will directly influence the development of chronicity, results suggesting that education and awareness of the causes and consequences of back pain may be a valuable treatment component for patient in the early phase of LBP. These educational components were also included in all our study groups in RCT.

All groups in our study improved in back and radiating pain, and in dysfunction to a level which was clinically significant. Improvement was evident during the first 3 months in all groups, and improvement occurred in therapy groups also in the 6- and 12-month follow-ups, which did not occur in the Advice-only group. However, only a small effect-size existed compared to that of the therapy groups.

Some studies are comparable to our RCT performed with subacute LBP patients (LBP lasting less than 3 months). In a study by Pengel et al. (2007) with 259 patients, physiotherapist-directed exercise and advice were each slightly more effective than placebo at 6 weeks. This effect was greatest when the interventions were combined. At 12 months, the only effect that persisted was a small effect on participant-reported function. In the study by Frost et al. (2004) with 286 LBP patients whose pain had lasted 6 weeks, these authors found no differences between routine physiotherapy (mobilization and exercise) and one session of assessment and advice.

In studies of acute LBP, Cherkin et al. (1998) investigated the effectiveness of the McKenzie method, chiropractic manipulation and advice. The McKenzie method and chiropractic had only marginally better outcomes than did a method providing advice with the minimal intervention of an educational booklet. Malmivaara et al. (1995) conducted a study comparing bed rest, back-extension exercises, and ordinary activity. After 3 and 12 weeks, "ordinary activity" patients had better recovery than those prescribed either bed rest or exercises.

The days of sick leave or percentage of patients in sick leave were low in therapy groups at the end of 1-year RCT compared to Advice-only. Total number of sick leave days in therapy groups was lower than in RCT for Karjalainen et al. (2004) working with SLBP patients. In RCT according to Faber et al. (2005), patients in the intervention group did not return to work earlier than did those in their control group.

Studies of cost-effectiveness of conservative treatment of acute or subacute LBP are sparse. Adherence to the recommendation for active care seems to improve the cost-effectiveness of care for acute LBP (Fritz et al. 2008). Despite the lack of any significant effect on intensity of LBP and perceived disability, a mini-intervention, including proper recommendations and advice, according to the "active approach," is able to reduce LBP-related costs (Karjalainen et al. 2004).

6.5 Limitations of the study

Studies I and III suffer from certain limitations because of small sample sizes, especially in regard to power. The cohorts investigated, even if small, were, however, considered homogenous. The study groups represent a normal outpatient population in any primary or occupational health care centers in Finland. However, our results do need confirmation in a larger sample, especially when such great numbers of examination findings are tested, in order to reduce the likelihood of spurious findings. The other earlier studies and ours are not fully comparable, because patient in our groups, acute and subacute patients, were not analyzed separately.

The prevalence of some of the subgroups in Study III was low, for example, sacroiliac joint pain/dysfunction and functional spinal stenosis; thus no calculation of kappa was possible. Combining these two subgroups with other subgroups would increase the feasibility of our results. Despite this limitation, a notable strength was that the patients were recruited from routine referrals, and because their care was provided for free, cost containment did not limit the physiotherapy examination. Although the time-limited examination lasted 30 minutes, some LBP patients may require longer assessment or even several assessments, with their responses to specific interventions and applications of home-exercises contributing to the final subgroup categorization.

Study II has some shortcomings. Its chronic LBP group may not be comparable to other chronic LBP groups, because there was suspicion of an indication for a spinal operation (the reason why the general practitioner involved wanted to consult the orthopedic surgeon). Moreover, we are aware of the Achilles heel of the study design, which was that the examiners were not blinded to group assignment. In this situation, however, it is likely that patients revealed their group unintentionally, even though the examiners were not allowed to discuss the length of their LBP problems. Blinding thus provided no additional value.

In Study II, 14 physiotherapists participated. All were taught by the one OMT specialist, with 22 years of experience in this field, to perform the specific clinical tests in the same way. Although, the experience each of those 14 PTs was similar, intertester reliability to perform those specific clinical tests was not measured. The strength of Study II was that the number of subjects was reasonable, and group division was accurate: Each group had a clearly differing length of LBP history. It is particularly important to study the sensitivity and specificity of a clinical test in a pain-duration-based group, which is very common in clinical medicine. Our results therefore provide new and important information as to the usefulness of these clinical tests for improving back pain physiotherapy.

In the first part of Study III, where intertester reliability and test-retest reproducibility of all clinical tests were analyzed in all test categories, only two physiotherapists specialized in OMT participated, which makes a generalization slightly complex. In the second part of Study III, were four PTs from healthcare centers, with similar experience. Even though they were taught to perform clinical tests by an experienced PT, specialized in OMT, their intertester reliability was not investigated prior to this study.

In Study IV, the drop-out rate was rather high and was highest in the Advice-only group. Reasons for this group`s high drop-out rate in included disappointment at having no pain treatment. There were also more subjects, unwilling to have follow-up measurements in that group. Those patients who suffered no more LBP during the follow-up were probably less willing to participate in follow-up measurements. One explanation could also be that the

locationor of some subjects was unreachable because of their work in a paper mill with much travelling required.

In our therapy groups, a physiotherapist controlled prescribed home-exercises during the treatment period, but no longer 3-, 6-, and 12-month follow-up points. Adherence to home-exercises was not assessed, which was also a limitation of this study. Whether the "adherence-outcome" interaction was mediated by improvements in pain and function related to the specific exercises, or by a more "global" effect of the program, remains to be examined.

The statistically significant difference in sick leave between McKenzie and OMT groups was seen during the first 3-months. At that point there were, however, no differences in LBP and leg pain, or in disability. This could be explained by acute and subacute pain being already relieved, also with those patients who had been on sick leave during this period, when follow-up status was recorded. The daily pain diary could have exposed more précised the level of pain and disability during sick leave.

However, the fact that drop-outs between groups did not differ as to outcome measures from those who completed the follow-ups strengthens validity. Further, sample size was quite small for a three-arm trial, which makes a type II error possible. For example, the confidence intervals in Figure 4 seem to indicate differences, especially between the therapy groups and the Advice-only group. A small number of subjects usually makes confidence intervals larger and thus weakens results**.** Still, this study had several strengths, for instance, its randomized controlled design and the fact that therapists in the OMT and McKenzie therapy were very experienced with over 20 and 10 years of experience as therapists as well as teachers. This could also be regarded as a weakness: the results cannot be generalized to those tests performed by novice physiotherapists.

The validity of the VAS and Roland-Morris was proved to be good. Our participants were recruited by routine referrals from occupational health care services, and interventions included commonly delivered treatments, so our results could be generalized with care because of the small number of subjects.

6.6 Clinical implications

Early information about and advice for LBP patients from a team including physicians and physiotherapists have been shown to be effective regarding the length of sick leave and prevention of LBP`s becoming chronic (Karjalainen et al. 2004). Direct access, also called early access, or the prompt access model tested in Great Britain has been acceptable for asking patients, physicians, and physiotherapists` opinion (Ferguson et al. 2010). In this model, the patient can directly contact a physiotherapist, or a patient can be referred to a physiotherapist directly by a physician.

A lower risk of subsequent medical service usage appears among patients who received physiotherapy early after an episode of acute LBP relative to those who receive physiotherapy at later times. Moreover, medical specialty variations exist regarding early use of physiotherapy, with potential underutilization among generalist specialties. (Gellhorn et al. 2010). Wand and O'Connell (2008) concluded that clinical trials offer little support for management of chronic LBP, and sub-grouping disappointing results from clinical trials in the chronic phase, as well. We found low subsequent medical service usage among patients in our RCT during 1-year follow-up, supporting the choice of early management of LBP. Our findings are in line with those in the study by Gellhorn et al (2010), a lower risk of subsequent medical service usage exists among patients who receive PT early after an episode of acute low back pain relative to those who receive PT at later times.

The subclassification based on selected clinical tests is moderately reliable also between physiotherapists who have not had a long specialization in OMT. This kind of subclassification could be a useful tool for recognizing the physiological background of LBP for targeted physiotherapy by physiotherapists working in primary contact with acute and subacute LBP patients in public health care centers and private physiotherapy clinics; here the classification criterion in terms of testing can be improved.

In our RCT, small but clinically important differences between therapy groups and Advice-only as to radiating pain and disability favor the OMT or McKenzie-type approach in the early phase of LBP. Our results can also encourage more study of how the inexpensive Advice-only method could be developed in general treatment in the early phase of LBP. The small number of sick leave days will reflect the effectiveness of the OMT-type of approach not only during the treatment sessions, but also during a 1-year follow up, if used along with a patient self-care program, which may facilitate the use of the OMT approach in the early phase of LBP.

6.7 Future research

Clinical examination, management, and treatment of LBP in its early phase are challenging. The aim of early management is to reduce recurrences of LBP, and prevent LBP`s becoming chronic, leading to bio-psycho-social problems both at the individual and community level. Further studies should evaluate the prognostic factors of clinical tests in the early phase of LBP with stronger study designs and validate the subgroup division with greater numbers of subjects. There is also a need to re-evaluate the reliability of subgroup classification with LBP patients in the early phase as well in patients with more pain and higher disability.

The most challenging area of LBP physiotherapy is prevention of LBP`s becoming chronic. Patients with LBP in most cases should be supported and

encouraged to be at work despite LBP (Lambeek et al. 2010). Work status has been shown to be associated with health status (Coste et al. 1994, Waddell et al. 2007). Patients in the early phase of LBP, who have localized back pain, have a good prognosis when returning to work as soon as possible (Coste et al. 1994). Although almost all patients of the therapy groups in our RCT had returned to work at the 1-year follow-up, the 2- to 3 year follow-up could be different, and future research should be challenged to include longer monitoring.

Two interesting factors were the lower adherence of the Advice-only group to self-treatment or their unwillingness to participate in follow-up visits in Study IV (Mannion et al. 2009). Seven patients were unwilling to carry on with the study (Figure 3). Improvement gained in the first 3 months was partly lost at the 6-month follow-up point among the Advice-only group (Figure 5). Similar lower adherence was detectable also in the study by Lamb et al. (2010), where the most frequent reason for participant withdrawal was unwillingness to complete questionnaires (the Advice-only group). Results by Rutten et al. (2010) also indicate that higher percentages of guideline adherence are related to better improvement in physical functioning and to a lower utilization of care. This means that proper assessment of the relationship between the process of physical therapy care and outcomes may require a comprehensive set of process indicators to measure guideline adherence.

Improvement in our Advice-only patients was good until the 3-month follow-up point, as did other treatment groups. But after that point, some decline occurred, when compared to other groups in leg pain, disability, and in sick leave.

Despite the fact that improvement in the therapy groups increased slightly, the effect-sizes compared to those of the Advice-group were only small or moderate.

The benefits of rehabilitation depend to a large extent on the patients` exercise behavior outside the formal physiotherapy sessions. Hence, more effort should be invested in finding ways to improve patients` motivation to take responsibility for the success of their own therapy, perhaps by increasing exercise self-efficacy (Mannion et al. 2009). Because adherence was lower in the Advice-only group, where patients had only one meeting with a PT, compared with 1 to 7 sessions with the OMT and McKenzie therapists, it might be fruitful to see the effect on the Advice-group of one or two sessions more with a PT giving only advice. Their deficient adherence could perhaps be avoided by determining whether subjects understood the information, and whether they had learned to perform their self-exercises. Their adherence to self-exercises could be monitored by telephone, as well.

Three psychosocial profile groups questionnaires in the study by Riipinen et al. (2005) as a predictor for LBP`s becoming chronic could be used to avoid chronicity. The so-called "dysfunctional" patients in particular might benefit more from a follow-up visit to a physiotherapist as would "interpersonally distressed" or "adaptive copers." The validity and usability of these kinds of questionnaires should be studied for LBP in its early phase, when patients pay a

visit to their physiotherapist in primary care. We need to discover predictive factors showing which type of patients might benefit more from self-exercise, and who might benefit from hands-on physiotherapy. This could be fruitful if we were to perform subgrouping of patients with the help of questionnaires and clinical tests into manual therapy treatment, and then randomize patients into one group having manual therapy and a second group having advice only.

Considering how central is the notion of targeted treatment to OMT principles, further studies should also explore the interplay between biological, social, and psychosocial factors in the early phase of LBP.

7 CONCLUSIONS

1. Some clinical tests had high inter- and intratester reliability in most of the test categories in the early phase of LBP with subjects having only moderate pain on the VAS scale, and having no severe functional limitations. Although these tests` inter- and intratester reliability was at an acceptable level globally, some tests had both high and low inter- and intratester reliability in different categories. Intratester reliability`s being slightly better than intertester reliability is an important factor in each individual patient`s follow-up.

2. Overall, only a few tests among all the selected clinical tests were moderately sufficient to sort the CLBP from SLBP, and patient groups from controls. These 31 selected clinical tests quite poorly sorted the CLBP or SLBP patients from the controls. It could be possible that combinations of different tests may enhance the sensitivity and specificity to sort such CLBP or the SLBP patients from their controls. Further studies need to reveal which tests in combination best sort the LBP patients from asymptomatic ones. In addition, whether these tests are sufficiently sensitive to classify a more specific diagnostic or clinical subgroup remains untested, and further data must help to enhance the effectiveness of clinical tests in differentiating among pathological conditions.

3. Subclassification into clinical subgroups in the early phase of LBP was reliable in the two most common classifications also with physiotherapists without specialization in OMT after short post-graduate training. This clinically relevant and clearly defined pain pattern system uses key elements of the history and examination to classify patients with LBP. However, larger trials using those tests in every subgroup, including those which have high odd ratios, are necessary before we can make general statements about the reliability of subgroup classification in the early phase of low back pain.

4. Concerning leg pain and disability, OMT and McKenzie methods, compared to Advice-only to stay active, showed no significant difference at 12 months of follow-up. However, the small number of sick leave days because of LBP in the intervention groups at the end of the study suggests the positive effects of these therapies. The low number of sick leave days and patients` early return to work reflects the effectiveness of OMT-physiotherapy.

TIIVISTELMÄ

Vaikka alaselkäkipua pidetään luonteeltaan hyvänlaatuisena, on selkäkivun uusiutuminen jo ensimmäisen vuoden aikana yleistä. Keskustelua käydään siitä paraneeko selkäkipu ohjeistuksella ”pysy aktiivisena”, ja onko kyseinen ohjeistus riittävä selkäkivun varhaisvaiheessa (<3kk) ehkäisemään selkäkivun pitkittymistä. Fysioterapeutti on usein selkäkipupotilaan ensimmäinen kontakti terveydenhuollossa tässä selkäkivun varhaisvaiheessa ohjeistamassa näitä potilaita.

Tämän neljästä osajulkaisusta koostuvan tutkimuksen tarkoitus oli arvioida fysioterapeuttien sisäistä ja välistä luotettavuutta kliinisten testien toistettavuudessa selkäkipupotilailla selkäkivun varhaisvaiheessa, ja arvioida näiden testien herkkyyttä ja tarkkuutta selkäkivun varhaisvaiheessa, sekä kroonisessa selkäkivussa, ja niiden erottelukykyä verrattaessa näitä ryhmiä kontrolliryhmään, jolla ei ole lääketieteellistä selkäkipudiagnoosia. Tutkimuksen tarkoitus oli myös arvioida kahden fysioterapeutin, jotka olivat erikoistuneet ortopediseen manuaaliseen terapiaan (OMT), luotettavuutta luokitella selkäkipupotilaat patoanatomisen/pato-fysiologisen luokittelun pohjalta viiteen eri alaluokkaan. Samaan tutkimukseen kuului myös arviointi onko luokittelu OMT-fysioterapeutin ja ei-erikoistuneen fysioterapeutin välillä verrattavissa näihin kahteen OMT-fysioterapeuttiin. Yhtenä tarkoituksena oli myös arvioida OMT-fysioterapian vaikuttavuutta verrattaessa sitä McKenzie-menetelmään ja itsehoito-ohjeistukseen ”pysy aktiivisena” selkäkipupotilailla randomisoidussa hoitotutkimuksessa. Mittareina olivat selkäkipu, alaraajaan säteilevä kipu, toimintakyky, sekä selkäkivun aiheuttamat sairauslomat.

Tutkimustulokset osoittavat että yleisesti ottaen kliiniset testit ovat sekä testaajan sisäistä että testaajien välistä toistettavuutta arvioitaessa hyväksyttävällä tasolla, vaikka eri testiluokissa olikin testejä jotka olivat hyvin toistettavia, tai testejä joiden toistettavuus oli huono. Sen sijaan testien erottelukyky kroonisten ja varhaisvaiheessa olevien selkäpotilaiden välillä oli huono, mutta jonkin verran parempi erottelemaan nämä potilasryhmät verrokkiryhmästä.

Luokittelu viiteen kliiniseen alaryhmään oli luotettavaa kahden yleisimmän alaryhmän välillä (välilevyperäinen kipu ja kliininen instabiliteetti) sekä kahden OMT-fysioterapeutin välillä että OMT-fysioterapeutin ja ei-erikoistuneen työpaikkakoulutuksen saaneen fysioterapeutin välillä. Muiden kolmen alaryhmän (kliininen stenoosi, segmentaarinen toimintahäiriö/fasettikipu, sacroiliaca-nivelen kipu/toimintahäiriö) luokittelun luotettavuutta ei tulosten valossa voida tarkastella johtuen näiden ryhmien pienestä koosta.

Kun verrattiin OMT-fysioterapian ja McKenzie-menetelmän vaikuttavuutta itsehoito-ohjeistukseen ”pysy aktiivisena”, todettiin alaraajan säteilevän kivun ja toimintakyvyn osalta 6-kuukauden seurannassa tilastollisia eroja, mutta ei selkäkivun osalta. Sairauspoissaoloja OMT-ryhmässä oli tilastollisesti merkitsevästi vähemmän kuin McKenzie-ryhmässä. Sairauspoissaolojen pienempi

määrä OMT-ryhmässä viittaa sen olevan tehokkaampi hoitomuoto tämän mittarin mukaan. Kuitenkin vuoden seurannassa näiden kahden ryhmän osalta sairauspoissaolot olivat vähentyneet lähes kokonaan, kun ne itsehoitoryhmässä olivat lisääntyneet, asettaen yhden fysioterapiakäynnin vaikuttavuuden kyseenalaiseksi.

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 J. 2000;25 Suppl:1S-33S.

Adams MA. Biomechanics of back pain. Acupunct Med. 2004;22:178-88.

Airaksinen O, Brox JI, Cedraschi C, Hildebrandt J, Klaber-Moffett J, Kovacs F, Mannion AF, Reis S, Staal JB, Ursin H, Zanoli G; COST B13 Working Group on Guidelines for Chronic Low Back Pain. Chapter 4. European guidelines for the management of chronic nonspecific low back pain. Eur Spine J. 2006;Suppl 2:S192-300.

Altman DG. Why we need confidence intervals. World J Surg. 2005;29:554-6

Andersson GB. Low back pain. J Rehabil Res Dev. 1997;34:202-12.

Assendelft WJ, Morton SC, Yu EI, Suttorp MJ, Shekelle PG. Spinal manipulative therapy for low back pain. A meta-analysis of effectiveness relative to other therapies. Ann Intern Med. 2003;138:871-81.

Atlas SJ, Deyo RA, Patrick DL, Convery K, Keller RB, Singer DE. The Quebec Task Force classification for Spinal Disorders and the severity, treatment, and outcomes of sciatica and lumbar spinal stenosis. Spine. 1996;21:2885-92.

Battié MC, Videman T, Levälahti E, Gill K, Kaprio J. Genetic and environmental effects on disc degeneration by phenotype and spinal level: a multivariate twin study. Spine J. 2008;33:2801-8.

Berthelot JM, Delecrin J, Maugars Y, Passuti N. Contribution of centralization phenomenon to the diagnosis, prognosis, and treatment of diskogenic low back pain. Joint Bone Spine. 2007;74:319-23.

Berthelot JM, Delecrin J, Maugars Y, Passuti N. Contribution of centralization phenomenon to the diagnosis, prognosis, and treatment of diskogenic low back pain. Joint Bone Spine. 2007;74:319-23.

Bigos SJ, Davis GE. Scientific application of sports medicine principles for acute low back problems. The Agency for Health Care Policy and Research Low Back Guideline Panel (AHCPR, Guideline #14). J Orthop Sports Phys Ther. 1996;24:192-207.

Bijur PE, Silver W, Gallagher EJ. Reliability of the visual analog scale for measurement of acute pain. Acad Emerg Med. 2001;8:1153-7.

Billis EV, McCarthy CJ, Oldham JA. Subclassification of low back pain: a cross-country comparison. Eur Spine J. 2007;16:865-79.

Binkley J, Stratford PW, Gill C. Interrater reliability of lumbar accessory motion mobility testing. Phys Ther. 1995;75:786-92.

Boline PD, Haas M, Meyer JJ, Kassak K, Nelson C, Keating JC Jr. Interexaminer reliability of eight evaluative dimensions of lumbar segmental abnormality: Part II. J Manipulative Physiol Ther. 1993;16:363-74.

Borkan J, Van Tulder M, Reis S, Schoene ML, Croft P, Hermoni D. Advances in the field of low back pain in primary care: a report from the fourth international forum. Spine J. 2002;27:E128-32.

Bouter LM, Pennick V, Bombardier C; Editorial Board of the Back Review Group. Cochrane back review group. Spine J. 2003;28:1215-8.

Brealey S, Burton K, Coulton S, Farrin A, Garratt A, Harvey E, Letley L, Martin J, Klaber MJ, Russell I, Torgerson D, Underwood M, Vickers M, Whyte K, Williams M; UK Back pain Exercise And Manipulation (UK BEAM) trial--national randomised trial of physical treatments for back pain in primary care: objectives, design and interventions. BMC Health Serv Res. 2003;3:16.

Breivik H, Collett B, Ventafridda V, Cohen R, Gallacher D. Survey of chronic pain in Europe: prevalence, impact on daily life, and treatment. Eur J Pain. 2006;10:287-333.

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 J. 2006;31:623-31.

Bronfort G, Haas M, Evans RL, Bouter LM. Efficacy of spinal manipulation and mobilization for low back pain and neck pain: a systematic review and best evidence synthesis. Spine J. 2004;3:335-56.

Bullock-Saxton JE. Local sensation changes and altered hip muscle function following severe ankle sprain. Phys Ther. 1994;74:17-28.

Burton AK, Balagué F, Cardon G, Eriksen HR, Henrotin Y, Lahad A, Leclerc A, Müller G, van der Beek AJ; COST B13 Working Group on European Guidelines for Prevention in Low Back Pain. How to prevent low back pain. Best Pract Res Clin Rheumatol. 2005;19:541-55.

Byrt T, Bishop J, Carlin JB. Bias, prevalence and kappa. J Clin Epidemiol. 1993;46:423-9.

Cairns MC, Foster NE, Wright C. Randomized controlled trial of specific spinal stabilization exercises and conventional physiotherapy for recurrent low back pain. Spine. 2006;31:E670-81.

Cherkin DC, Deyo RA, Battié M, Street J, Barlow W. A comparison of physical therapy, chiropractic manipulation, and provision of an educational booklet for the treatment of patients with low back pain. N Engl J Med. 1998;339:1021-9.

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.

Choi BK, Verbeek JH, Tam WW, Jiang JY. Exercises for prevention of recurrences of low-back pain. Occup Environ Med. 2010;67:795-6.

Chorti AG, Chortis AG, Strimpakos N, McCarthy CJ, Lamb SE. The prognostic value of symptom responses in the conservative management of spinal pain: a systematic review. Spine (Phila Pa 1976). 2009;34:2686-99.

Clare HA, Adams R, Maher CG.Reliability of McKenzie classification of patients with cervical or lumbar pain. J Manipulative Physiol Ther. 2005;28:122-7.

Cleland JA, Fritz JM, Kulig K, Davenport TE, Eberhart S, Magel J, Childs JD. Comparison of the effectiveness of three manual physical therapy techniques in a subgroup of patients with low back pain who satisfy a clinical prediction rule: a randomized clinical trial. Spine J. 2009;34:2720-9.

Cohen, J. Statistical Power Analysis for the Behavioral Sciences. NY: Academic Press. 1969.

Cook C, Cleland J, Huijbregts P. Creation and critique of studies of diagnostic accuracy: Use of the STARD and QUADAS methodological quality assessment tools. J Man Manip Ther. 2007;15:93-102.

Coste J, Delecoeuillerie G, Cohen de Lara A, Le Parc JM, Paolaggi JB. Clinical course and prognostic factors in acute low back pain: an inception cohort study in primary care practice. BMJ. 1994;308:577-80.

Croft PR, Dunn KM, Raspe H. Course and prognosis of back pain in primary care: the epidemiological perspective. Pain. 2006;122:1-3.

Dagenais S, Tricco AC, Haldeman S. Synthesis of recommendations for the assessment and management of low back pain from recent clinical practice guidelines. Spine J. 2010;10:514-29.

Dankaerts W, O'Sullivan P, Burnett A, Straker L, Davey P, Gupta R. Discriminating healthy controls and two clinical subgroups of nonspecific chronic low back pain patients using trunk muscle activation and lumbosacral kinematics of postures and movements: a statistical classification model. Spine J. 2009;34:1610-8.

Dankaerts W, O'Sullivan PB, Straker LM, Burnett AF, Skouen JS. The inter-examiner reliability of a classification method for non-specific chronic low back pain patients with motor control impairment. Man Ther. 2006;11:28-39.

de Groot M, Pool-Goudzwaard AL, Spoor CW, Snijders CJ. The active straight leg raising test (ASLR) in pregnant women: differences in muscle activity and force between patients and healthy subjects. Man Ther. 2008;13:68-74.

de Schepper EI, Damen J, van Meurs JB, Ginai AZ, Popham M, Hofman A, Koes BW, Bierma-Zeinstra SM. The association between lumbar disc degeneration and low back pain: the influence of age, gender, and individual radiographic features. Spine J. 2010;35:531-6.

Delitto A, Cibulka MT, Erhard RE, Bowling RW, Tenhula JA. Evidence for use of an extension-mobilization category in acute low back syndrome: a prescriptive validation pilot study. Phys Ther. 1993;73:216-22.

Delitto A, Erhard RE, Bowling RW. A treatment-based classification approach to low back syndrome: identifying and staging patients for conservative treatment. Phys Ther. 1995;75:470-85.

Denison E, Asenlöf P, Sandborgh M, Lindberg P. Musculoskeletal pain in primary health care: subgroups based on pain intensity, disability, self-efficacy, and fear-avoidance variables. J Pain. 2007;8:67-74.

Deshpande A, Furlan A, Mailis-Gagnon A, Atlas S, Turk D. Opioids for chronic low-back pain. Cochrane Database Syst Rev. 2007;3:CD004959.

Deyo RA, Phillips WR. Low back pain. A primary care challenge. Spine (Phila Pa 1976). 1996;21:2826-32.

Deyo RA, Loeser JD, Bigos SJ. Herniated lumbar intervertebral disk. Ann Intern Med. 1990;112:598-603.

Dunn KM, Croft PR, Main CJ, Von Korff M. A prognostic approach to defining chronic pain: replication in a UK primary care low back pain population. Pain. 2008;135:48-54.

Engers A, Jellema P, Wensing M, van der Windt DA, Grol R, van Tulder MW. Individual patient education for low back pain. Cochrane Database Syst Rev. 2008;1:CD004057.

Ernst E, Canter PH. A systematic review of systematic reviews of spinal manipulation. J R Soc Med. 2006;99:192-6.

Essendrop M, Maul I, Läubli T, Riihimäki H, Schibye B. Measures of low back function: a review of reproducibility studies. Clin Biomech. 2002;17:235-49.

Faber E, Bierma-Zeinstra SM, Burdorf A, Nauta AP, Hulshof CT, Overzier PM, Miedema HS, Koes BW. In a controlled trial training general practitioners and occupational physicians to collaborate did not influence sick leave of patients with low back pain. J Clin Epidemiol. 2005;58:75-82.

Ferguson F, Holdsworth L, Rafferty D. A national framework for supporting improvements in the physiotherapy assessment and management of low back pain: the Scottish experience. Physiotherapy. 2010;96:198-205.

Ferguson SA, Marras WS, Burr DL, Woods S, Mendel E, Gupta P. Quantification of a meaningful change in low back functional impairment. Spine. 2009;34:2060-5.

Ferguson F, Holdsworth L, Rafferty D. Low back pain and physiotherapy use of red flags: the evidence from Scotland. Physiotherapy. 2010;96:282-8.

Fernandes Rde C, Carvalho FM, Assunção AA, Silvany Neto AM. Interactions between physical and psychosocial demands of work associated to low back pain. Rev Saude Publica. 2009;43:326-34.

Ferreira ML, Ferreira PH, Latimer J, Herbert R, Maher CG. Does spinal manipulative therapy help people with chronic low back pain? Aust J Physiother. 2002;48:277-84.

Ferreira ML, Ferreira PH, Latimer J, Herbert RD, Hodges PW, Jennings MD, Maher CG, Refshauge KM. Comparison of general exercise, motor control exercise and spinal manipulative therapy for chronic low back pain: A randomized trial. Pain. 2007;131:31-7.

Ferreira PH, Ferreira ML, Maher CG, Herbert RD, Refshauge K. Specific stabilisation exercise for spinal and pelvic pain: a systematic review. Aust J Physiother. 2006;52:79-88.

Fersum KV, Dankaerts W, O'Sullivan PB, Maes J, Skouen JS, Bjordal JM, Kvåle A. Integration of sub-classification strategies in RCTs evaluating manual therapy treatment and exercise therapy for non-specific chronic low back pain (NSCLBP): a systematic review. Br J Sports Med. 2009;44:1054-62.

Fersum KV, Dankaerts W, O'Sullivan PB, Maes J, Skouen JS, Bjordal JM, Kvåle A. Integration of sub-classification strategies in RCTs evaluating manual

therapy treatment and exercise therapy for non-specific chronic low back pain (NSCLBP): a systematic review. Br J Sports Med. 2010;44:1054-62.

Fortin JD, Dwyer AP, West S, Pier J. Sacroiliac joint: pain referral maps upon applying a new injection/arthrography technique. Part I: Asymptomatic volunteers. Spine (Phila Pa 1976). 1994;19:1475-82.

Fritz JM, Delitto A, Erhard RE. Comparison of classification-based physical therapy with therapy based on clinical practice guidelines for patient with acute low back pain: a randomized clinical trial. Spine. 2003;28:136371.

Foster N, Hill J, Hay E. Subgrouping patients with low back pain in primar care: are we getting any better at it? Man Ther. 2011;16:3-8.

Frank AO, De Souza LH, McAuley JH, Sharma V, Main CJ. A cross-sectional survey of the clinical and psychological features of low back pain and consequent work handicap: use of the Quebec Task Force classification. Int J Clin Pract. 2000;54:639-44.

French SD, Green S, Buchbinder R, Barnes H. Interventions for improving the appropriate use of imaging in people with musculoskeletal conditions. Cochrane Database Syst Rev. 2010:CD006094.

Fritz JM, Cleland JA, Childs JD. Subgrouping patients with low back pain: evolution of a classification approach to physical therapy. J Orthop Sports Phys Ther. 2007;37:290-302.

Fritz JM, Cleland JA, Speckman M, Brennan GP, Hunter SJ. Physical therapy for acute low back pain: associations with subsequent healthcare costs. Spine. 2008;33:1800-5.

Fritz JM, George S. The use of a classification approach to identify subgroups of patients with acute low back pain. Interrater reliability and short-term treatment outcomes. Spine J. 2000;25:106-14.

Fritz JM, Lindsay W, Matheson JW, Brennan GP, Hunter SJ, Moffit SD, Swalberg A, Rodriquez B. Is there a subgroup of patients with low back pain likely to benefit from mechanical traction? Results of a randomized clinical trial and subgrouping analysis. Spine J. 2007;32:E793-800.

Fritz JM, Piva SR, Childs JD. Accuracy of the clinical examination to predict radiographic instability of the lumbar spine. Eur Spine J. 2005;14:743-50.

Gellhorn AC, Chan L, Martin B, Friedly J. Management Patterns in Acute Low Back Pain: The Role of Physical Therapy. Spine J. 2010 Nov 19.

George SZ, Delitto A. Clinical examination variables discriminate among treatment-based classification groups: a study of construct validity in patients with acute low back pain. Phys Ther. 2005;85:306-14.

Gerdle B, Björk J, Henriksson C, Bengtsson A. Prevalence of current and chronic pain and their influences upon work and healthcare-seeking: a population study. J Rheumatol. 2004;31:1399-406.

Croft PR, Macfarlane GJ, Papageorgiou AC, Thomas E, Silman AJ. Outcome of low back pain in general practice: a prospective study. BMJ. 1998;316:1356-9.

Grotle M, Foster N, DunnK, Croft P. Are prognostic indicators for porr outcome different for acute and chronic low back pain consulters in primary care? Pain. 2010151:790-7.

Gruther W, Wick F, Paul B, Leitner C, Posch M, Matzner M, Crevenna R, Ebenbichler G. Diagnostic accuracy and reliability of muscle strength and endurance measurements in patients with chronic low back pain. J Rehabil Med. 2009;41:613-9.

Hagen KB, Hilde G, Jamtvedt G, Winnem M. Bed rest for acute low-back pain and sciatica. Cochrane Database Syst Rev. 2004;4:CD001254.

Haley SM, Osberg JS. Kappa coefficient calculation using multiple ratings per subject: a special communication. Phys Ther. 1989;69:970-4.

Hall H, McIntosh G, Boyle C. Effectiveness of a low back pain classification system. Spine J. 2009;9:648-57.

Hallegraeff JM, de Greef M, Winters JC, Lucas C. Manipulative therapy and clinical prediction criteria in treatment of acute nonspecific low back pain. Percept Mot Skills. 2009;108:196-208.

Hancock MJ, Maher CG, Latimer J, Herbert RD, McAuley JH. Independent evaluation of a clinical prediction rule for spinal manipulative therapy: a randomised controlled trial. Eur Spine J. 2008;17:936-43.

Hancock MJ, Maher CG, Latimer J, McLachlan AJ, Cooper CW, Day RO, Spindler MF, McAuley JH. Assessment of diclofenac or spinal manipulative therapy, or both, in addition to recommended first-line treatment for acute low back pain: a randomised controlled trial. Lancet. 2007;370:1638-43.

Haswell K, Gilmour J, Moore B. Clinical decision rules for identification of low back pain patients with neurologic involvement in primary care. Spine J. 2008;33:68-73.

Hayden JA, van Tulder MW, Malmivaara AV, Koes BW. Meta-analysis: exercise therapy for nonspecific low back pain. Ann Intern Med. 2005;142:I71.

Hayden JA, van Tulder MW, Malmivaara AV, Koes BW. Meta-analysis: exercise therapy for nonspecific low back pain. Ann Intern Med.2003;142:9:I71.

Henchoz Y, Kai-Lik So A. Exercise and nonspecific low back pain: a literature review. Joint Bone Spine. 2008;75:533-9.

Henschke N, Maher CG, Refshauge KM, Herbert RD, Cumming RG, Bleasel J, York J, Das A, McAuley JH. Prognosis in patients with recent onset low back pain in Australian primary care: inception cohort study. BMJ. 2008;337:a171.

Hestbaek L, Leboeuf-Yde C. Are chiropractic tests for the lumbo-pelvic spine reliable and valid? A systematic critical literature review. J Manipulative Physiol Ther. 2000;23:258-75.

Heymans M, Tulder M, Esmail R, Bombardier C, Koes B. Back schools for non-specific low-back pain. Cochrane Database Syst Rev. 2004;4:CD000261.

Hicks GE, Fritz JM, Delitto A, Mishock J. Interrater reliability of clinical examination measures for identification of lumbar segmental instability. Arch Phys Med Rehabil. 2003;84:1858-64.

Hill J, Keating J. Risk factors for the first episode of low back pain in children are infrequently validated across samples and conditions: a systematic review. J Physiother. 2010;56:237-44.

Hodges PW, Moseley GL. Pain and motor control of the lumbopelvic region: effect and possible mechanisms. J Electromyogr Kinesiol. 2003;13:361-70.

Holmich P, Holmich LR, Bjerg AM. Clinical examination of athletes with groin pain: an intraobserver and interobserver reliability study. Br J Sports Med. 2004;38:446-51.

Hoogendoorn WE, van Poppel MN, Bongers PM, Koes BW, Bouter LM. Systematic review of psychosocial factors at work and private life as risk factors for back pain. Spine J. 2000;25:2114-25.

Horneij E, Hemborg B, Johnsson B, Ekdahl C. Clinical tests on impairment level related to low back pain: a study of test reliability. J Rehabil Med. 2002;34:176-82.

Hoy D, Brooks P, Blyth F, Buchbinder R. The Epidemiology of low back pain. Best Pract Res Clin Rheumatol. 2010;24:769-81.

Hunt DG, Zuberbier OA, Kozlowski AJ, Robinson J, Berkowitz J, Schultz IZ, Milner RA, Crook JM, Turk DC.Reliability of the lumbar flexion, lumbar extension, and passive straight leg raise test in normal populations embedded within a complete physical examination. Spine. 2001;26:2714-8.

Hunt RJ. Percent agreement, Pearson's correlation, and kappa as measures of inter-examiner reliability. J Dent Res. 1986;65:128-30.

Ihlebaek C, Hansson TH, Laerum E, Brage S, Eriksen HR, Holm SH, Svendsrød R, Indahl A. Prevalence of low back pain and sickness absence: a "borderline" study in Norway and Sweden. Scand J Public Health. 2006;34:555-8.

Iles RA, Davidson M, Taylor NF. Psychosocial predictors of failure to return to work in non-chronic non-specific low back pain: a systematic review. Occup Environ Med. 2008;65:507-17.

Indahl A, Velund L, Reikeraas O. Good prognosis for low back pain when left untampered. A randomized clinical trial. Spine J. 1995;20:473-7.

Jüni P, Battaglia M, Nüesch E, Hämmerle G, Eser P, van Beers R, Vils D, Bernhard J, Ziswiler HR, Dähler M, Reichenbach S, Villiger PM. A randomised controlled trial of spinal manipulative therapy in acute low back pain. Ann Rheum Dis. 2009;68:1420-7.

Kääriä S, Kaila-Kangas L, Kirjonen J, Riihimäki H, Luukkonen R, Leino-Arjas P.Low back pain, work absenteeism, chronic back disorders, and clinical findings in the low back as predictors of hospitalization due to low back disorders: a 28-year follow-up of industrial employees. 2005;30:1211-8.

Kääriä S, Solovieva S, Leino-Arjas P.Associations of low back pain with neck pain: a study of industrial employees with 5-, 10-, and 28-year follow-ups. Eur J Pain. 2009;13:406-11.

Kääriä SM, Mälkiä EA, Luukkonen RA, Leino-Arjas PI. Pain and clinical findings in the low back: A study of industrial employees with 5-, 10-, and 28-year follow-ups. Eur J Pain. 2010;14:759-63.

Kankaanpää M, Taimela S, Airaksinen O, Hänninen 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.

Kankaanpaa M, Colier WN, Taimela S, Anders C, Airaksinen O, Kokko-Aro SM, Hanninen O. Back extensor muscle oxygenation and fatigability in health subjects and low back pain patients during dynamic back extension exertion. Pathophysiology. 2005 Nov 7.

Karjalainen K, Malmivaara A, Mutanen P, Roine R, Hurri H, Pohjolainen T. Mini-intervention for subacute low back pain: two-year follow-up and modifiers of effectiveness. Spine J. 2004;29:1069-76.

Karjalainen K, Malmivaara A, Mutanen P, Roine R, Hurri H, Pohjolainen T. Mini-intervention for subacute low back pain: two-year follow-up and modifiers of effectiveness. Spine J. 2004;29:1069-76.

Kent P, Keating JL. Classification in nonspecific low back pain: what methods do primary care clinicians currently use? Spine J. 2005;30:1433-40.

Kent P, Keating JL. Classification in nonspecific low back pain: what methods do primary care clinicians currently use? Spine. 2005;30:1433-40.

Kent P, Mjosund HL, Petersen DH. Does targeting manual therapy and/or exercise improve patient outcomes in nonspecific low back pain? A systematic review. BMC Med. 2010;8:22.

Kent PM, Keating JL, Taylor NF. Primary care clinicians use variable methods to assess acute nonspecific low back pain and usually focus on impairments. Man Ther. 2009;14:88-100.

Kerr RS, Cadoux-Hudson TA, Adams CB. The value of accurate clinical assessment in the surgical management of the lumbar disc protrusion. J Neurol Neurosurg Psychiatry. 1988;51:169-73.

Kilpikoski S, Airaksinen O, Kankaanpää M, Leminen P, Videman T, Alen M. Interexaminer reliability of low back pain assessment using the McKenzie method. Spine J. 2002;27:E207-14.

Koes BW, van Tulder MW, Ostelo R, Kim Burton A, Waddell G. Clinical guidelines for the management of low back pain in primary care: an international comparison. Spine J. 2001;26:2504-13.

Kovacs FM, Abraira V, Royuela A, Corcoll J, Alegre L, Cano A, Muriel A, Zamora J, del Real MT, Gestoso M, Mufraggi N. Minimal clinically important change for pain intensity and disability in patients with nonspecific low back pain. Spine J. 2007;32:2915-20.

Krismer M, van Tulder M; Low Back Pain Group of the Bone and Joint Health Strategies for Europe Project. Strategies for prevention and management of musculoskeletal conditions. Low back pain (non-specific). Best Pract Res Clin Rheumatol. 2007;21:77-91.

Lakke SE, Soer R, Takken T, Reneman MF. Risk and prognostic factors for non-specific musculoskeletal pain: a synthesis of evidence from systematic reviews classified into ICF dimensions. Pain. 2009;147:153-64.

Lamb SE, Hansen Z, Lall R, Castelnuovo E, Withers EJ, Nichols V, Potter R, Underwood MR; Back Skills Training Trial investigators. Group cognitive

behavioural treatment for low-back pain in primary care: a randomised controlled trial and cost-effectiveness analysis. Lancet. 2010;375:916-23.

Lambeek LC, Bosmans JE, Van Royen BJ, Van Tulder MW, Van Mechelen W, Anema JR. Effect of integrated care for sick listed patients with chronic low back pain: economic evaluation alongside a randomised controlled trial. BMJ. 2010;341:c6414. doi: 10.1136/bmj.c6414.

Landel R, Kulig K, Fredericson M, Li B, Powers CM. Intertester reliability and validity of motion assessments during lumbar spine accessory motion testing. Phys Ther. 2008;88:43-9.

Landis J, Koch G. The measurement of observer agreement for categorical data. Biometrics 1977;33: 159-174.

Laslett M, Aprill CN, McDonald B, Young SB. Diagnosis of sacroiliac joint pain/dysfunction: validity of individual provocation tests and composites of tests. Man Ther. 2005;10:207-18.

Laslett M, McDonald B, Tropp H, Aprill CN, Oberg B. Agreement between diagnoses reached by clinical examination and available reference standards: a prospective study of 216 patients with lumbopelvic pain. BMC Musculoskelet Disord. 2005;6:28. ??

Laslett M, Williams M. The reliability of selected pain provocation tests for sacroiliac joint pathology. Spine J. 1994;19:1243-9.

Leboeuf-Yde C. Back pain--individual and genetic factors. J Electromyogr Kinesiol. 2004;14:129-33.

Leboeuf-Yde C. Back pain--individual and genetic factors. J Electromyogr Kinesiol. 2004;14:129-33.

Liddle SD, Gracey JH, Baxter GD. Advice for the management of low back pain: a systematic review of randomised controlled trials. Man Ther. 2007;12:310-27.

Linton SJ, Andersson T. Can chronic disability be prevented? A randomized trial of a cognitive-behavior intervention and two forms of information for patients with spinal pain. Spine (Phila Pa 1976). 2000;25:2825-31.

Linton SJ, Gross D, Schultz IZ, Main C, Côté P, Pransky G, Johnson W. Prognosis and the identification of workers risking disability: research issues and directions for future research. J Occup Rehabil. 2005;15:459-74.

Loisel P, Vachon B, Lemaire J, Durand MJ, Poitras S, Stock S, Tremblay C. Discriminative and predictive validity assessment of the quebec task force classification. Spine (Phila Pa 1976). 2002;27:851-7.

Long, A; Donelson, R. Does it Matter Which Exercise?: A Randomized Control Trial of Exercise for Low Back Pain. Spine J. 2004;29:2593-2602.

Luomajoki H, Kool J, de Bruin ED, Airaksinen O. Improvement in low back movement control, decreased pain and disability, resulting from specific exercise intervention. Sports Med Arthrosc Rehabil Ther Technol. 2010;2:11.

Luomajoki H, Kool J, de Bruin ED, Airaksinen O. Improvement in low back movement control, decreased pain and disability, resulting from specific

exercise intervention. Sports Med Arthrosc Rehabil Ther Technol. 2010;2:11.

Luomajoki H, Kool J, de Bruin ED, Airaksinen O. Movement control tests of the low back; evaluation of the difference between patients with low back pain and healthy controls. BMC Musculoskelet Disord. 2008;9:170.

Luomajoki H, Kool J, de Bruin ED, Airaksinen O. Reliability of movement control tests in the lumbar spine. BMC Musculoskelet Disord. 2007;8:90

Machado LA, de Souza MS, Ferreira PH, Ferreira ML. The McKenzie method for low back pain: a systematic review of the literature with a meta-analysis approach. Spine. 2006;31:E254-62.

Machado LA, de Souza MS, Ferreira PH, Ferreira ML. The McKenzie method for low back pain: a systematic review of the literature with a meta-analysis approach. Spine J. 2006;31:E254-62.

Magnusson SP. Passive properties of human skeletal muscle during stretch maneuvers. A review. Scand J Med Sci Sports. 1998;8:65-77.

Maher CG, Latimer J, Adams R. An investigation of the reliability and validity of posteroanterior spinal stiffness judgments made using a reference-based protocol. Phys Ther. 1998;78:829-37.

Majlesi J, Togay H, Unalan H, Toprak S. The sensitivity and specificity of the Slump and the Straight Leg Raising tests in patients with lumbar disc herniations. J Clin Rheumatol. 2008;14:87-91.

Malmivaara A, Häkkinen U, Aro T, Heinrichs ML, Koskenniemi L, Kuosma E, Lappi S, Paloheimo R, Servo C, Vaaranen V, et al.The treatment of acute low back pain--bed rest, exercises, or ordinary activity? N Engl J Med. 1995;332:351-5.

Maluf KS, Sahrmann SA, Van Dillen LR. Use of a classification system to guide nonsurgical management of a patient with chronic low back pain. Phys Ther. 2000;80:1097-111.

Manchikanti L, Derby R, Wolfer L, Singh V, Datta S, Hirsch JA. Evidence-based medicine, systematic reviews, and guidelines in interventional pain management: Part 7: systematic reviews and meta-analyses of diagnostic accuracy studies. Pain Physician. 2009;12:929-63.

Mannion AF, Helbling D, Pulkovski N, Sprott H. Spinal segmental stabilisation exercises for chronic low back pain: programme adherence and its influence on clinical outcome. Eur Spine J. 2009;18:1881-91.

McCarthy CJ, Gittins M, Roberts C, Oldham JA. The reliability of the clinical tests and questions recommended in international guidelines for low back pain. Spine J. 2007;32:921-6.

McGorry RW, Webster BS, Snook SH, Hsiang SM.The relation between pain intensity, disability, and the episodic nature of chronic and recurrent low back pain. Spine J. 2000;25:834-41.

McGregor AH, Hukins DW. Lower limb involvement in spinal function and low back pain. J Back Musculoskelet Rehabil. 2009;22:219-22.

McKenzie RA, May S, editors. 2003. The lumbar spine mechanical diagnosis and therapy. Waikanae: Spinal Publication Ltd.

Mens JM, Vleeming A, Snijders CJ, Koes BW, Stam HJ. Reliability and validity of the active straight leg raise test in posterior pelvic pain since pregnancy. Spine (Phila Pa 1976). 2001;26:1167-71.

Nicholas MK, Linton SJ, Watson PJ, Main CJ; "Decade of the Flags" Working Group. Early identification and management of psychological risk factors ("yellow flags") in patients with low back pain: a reappraisal. Phys Ther. 2011;91:737-53.

Niemisto L, Rissanen P, Sarna S, Lahtinen-Suopanki T, Lindgren KA, Hurri H.Cost-effectiveness of combined manipulation, stabilizing exercises, and physician consultation compared to physician consultation alone for chronic low back pain: a prospective randomized trial with 2-year follow-up. Spine. 2005;30:1109-15.

No authors listed. Acute low back problems in adults: assessment and treatment. Agency for Health Care Policy and Research. Clin Pract Guidel Quick Ref Guide Clin. 1994;14:iii-iv, 1-25.

Näslindh-Ylispangar A, Sihvonen M, Sarna S, Kekki P. Health status, symptoms and health counselling among middle-aged men: comparison of men at low and high risk. Scand J Caring Sci. 2008;22:529-35.

Ostelo RW, de Vet HC. Clinically important outcomes in low back pain. Best Pract Res Clin Rheumatol. 2005;19:593-607.

Ostelo RW, de Vet HC. Clinically important outcomes in low back pain. Best Pract Res Clin Rheumatol. 2005;19:593-607.

O'Sullivan P. Diagnosis and classification of chronic low back pain disorders: maladaptive movement and motor control impairments as underlying mechanism. Man Ther. 2005;10:242-55.

O'Sullivan P. Diagnosis and classification of chronic low back pain disorders: maladaptive movement and motor control impairments as underlying mechanism. Man Ther. 2005;10:242-55.

Panjabi MM. The stabilizing system of the spine. Part I. Function, dysfunction, adaptation, and enhancement. J Spinal Disord. 1992;5:383-9.

Pengel LH, Refshauge KM, Maher CG, Nicholas MK, Herbert RD, McNair P. Physiotherapist-directed exercise, advice, or both for subacute low back pain: a randomized trial. Ann Intern Med. 2007;146:787-96.

Petersen T, Larsen K, Jacobsen S. One-year follow-up comparison of the effectiveness of McKenzie treatment and strengthening training for patients with chronic low back pain: outcome and prognostic factors. Spine. 2007;32:2948-56.

Petersen T, Olsen S, Laslett M, Thorsen H, Manniche C, Ekdahl C, Jacobsen S. Inter-tester reliability of a new diagnostic classification system for patients with non-specific low back pain. Aust J Physiother. 2004;50:85-94.

Petersen T, Olsen S, Laslett M, Thorsen H, Manniche C, Ekdahl C, Jacobsen S. Inter-tester reliability of a new diagnostic classification system for patients with non-specific low back pain. Aust J Physiother. 2004;50(2):85-94.

Razmjou H, Kramer JF, Yamada R. Intertester reliability of the McKenzie evaluation in assessing patients with mechanical low-back pain. J Orthop Sports Phys Ther. 2000;30:368-83.

Rebain R, Baxter GD, McDonough S. A systematic review of the passive straight leg raising test as a diagnostic aid for low back pain (1989 to 2000). Spine J. 2002;27:E388-95.

Reeve A, Dilley A. Effects of posture on the thickness of transversus abdominis in pain-free subjects. Man Ther. 2009;14:679-84.

Riddle DL. Classification and low back pain: a review of the literature and critical analysis of selected systems. Phys Ther. 1998;78:708-37.

Riipinen M, Niemisto L, Lindgren KA, Hurri H.Psychosocial differences as predictors for recovery from chronic low back pain following manipulation, stabilizing exercises and physician consultation or physician consultation alone.J Rehabil Med. 2005;37:152-8.

Robinson HS, Mengshoel AM, Bjelland EK, Vøllestad NK. Pelvic girdle pain, clinical tests and disability in late pregnancy. Man Ther. 2010;15:522-8.

Roelofs PD, Deyo RA, Koes BW, Scholten RJ, van Tulder MW. Nonsteroidal anti-inflammatory drugs for low back pain: an updated Cochrane review. Spine (Phila Pa 1976). 2008;33:1766-74.

Roland M, Fairbank J. The Roland-Morris Disability Questionnaire and the Oswestry Disability Questionnaire. Spine (Phila Pa 1976). 2000;25:3115-24.

Rothwell PM. Treating individuals 2. Subgroup analysis in randomised controlled trials: importance, indications, and interpretation Lancet. 2005;365:176-86.

Roussel NA, Nijs J, Truijen S, Smeuninx L, Stassijns G. Low back pain: clinimetric properties of the Trendelenburg test, active straight leg raise test, and breathing pattern during active straight leg raising. J Manipulative Physiol Ther. 2007;30:270-8.

Rubin DI. Epidemiology and risk factors for spine pain. Neurol Clin. 2007;25:353-71.

Rubinstein SM, van Tulder MA. Best-evidence review of diagnostic procedures for neck and low-back pain. Best Pract Res Clin Rheumatol. 2008;22:471-82

Rundell SD, Davenport TE, Wagner T. Physical therapist management of acute and chronic low back pain using the World Health Organization's International Classification of Functioning, Disability and Health. Phys Ther. 2009;89:82-90.

Rutten GM, Degen S, Hendriks EJ, Braspenning JC, Harting J, Oostendorp RA. Adherence to clinical practice guidelines for low back pain in physical therapy: do patients benefit? Phys Ther. 2010;90:1111-22.

Sackett DL, Wennberg JE. Choosing the best research design for each question. BMJ. 1997;315:1636.

Saur PM, Ensink FB, Frese K, Seeger D, Hildebrandt J. Lumbar range of motion: reliability and validity of the inclinometer technique in the clinical measurement of trunk flexibility. Spine J. 1996;21:1332-8.

Schaafsma F, Schonstein E, Ojajärvi A, Verbeek J. Physical conditioning programs for improving work outcomes among workers with back pain. Scand J Work Environ Health. 2011:37:1-5.

Schmid AB, Brunner F, Luomajoki H, Held U, Bachmann LM, Künzer S, Coppieters MW. Reliability of clinical tests to evaluate nerve function and mechanosensitivity of the upper limb peripheral nervous system. BMC Musculoskelet Disord. 2009;10:11.

Schneider M, Homonai R, Moreland B, Delitto A. Interexaminer reliability of the prone leg length analysis procedure. J Manipulative Physiol Ther. 2007;30:514-21.

Schwarzer AC, Derby R, Aprill CN, Fortin J, Kine G, Bogduk N. The value of the provocation response in lumbar zygapophyseal joint injections. Clin J Pain. 1994;10:309-13.

Scott J, Huskisson E. Vertical and horizontal visual analogue scales. Ann Rheum Dis 1978,38-40.

Seffinger MA, Najm WI, Mishra SI, Adams A, Dickerson VM, Murphy LS, Reinsch S. Reliability of spinal palpation for diagnosis of back and neck pain: a systematic review of the literature. Spine J. 2004;29:E413-25.

Shiri R, Karppinen J, Leino-Arjas P, Solovieva S, Varonen H, Kalso E, Ukkola O, Viikari-Juntura E. Cardiovascular and lifestyle risk factors in lumbar radicular pain or clinically defined sciatica: a systematic review. Eur Spine J. 2007;16:2043-54.

Shiri R, Karppinen J, Leino-Arjas P, Solovieva S, Viikari-Juntura E. The association between smoking and low back pain: a meta-analysis. Am J Med. 2010;123:87.e7-35.

Shiri R, Solovieva S, Husgafvel-Pursiainen K, Viikari J, Raitakari OT, Viikari-Juntura E. Incidence of nonspecific and radiating low back pain: followup of 24-39-year-old adults of the Young Finns Study. Arthritis Care Res. 2010;62:455-9.

Smeets R, Maher C, Nicholas M, Refshauge K, Herbert R. Do psychological characteristics predict response to exercise and advice for subacute low back pain? Arthritis Rheum. 2009;61:1202-9.

Spitzer et al. Scientific approach to the assessment and management of activity-related spinal disorders. A monograph for clinicians. Report of the Quebec Task Force on Spinal Disorders. Spine. 1987;7 Suppl:S1-59.

Spoto MM, Collins J. Physiotherapy diagnosis in clinical practice: a survey of orthopaedic certified specialists in the USA Physiother Res Int. 2008;13:31-41.

Stanton TR, Hancock MJ, Maher CG, Koes BW. Critical appraisal of clinical prediction rules that aim to optimize treatment selection for musculoskeletal conditions. Phys Ther. 2010;90:834-54.

Stratford PW, Binkley J, Solomon P, Finch E, Gill C, Moreland J. Defining the minimum level of detectable change for the Roland-Morris questionnaire. Phys Ther. 1996;76:359-65.

Tidstrand J, Horneij E. Inter-rater reliability of three standardized functional tests in patients with low back pain. BMC Musculoskelet Disord. 2009:2;10:58.

Tüzün C, Yorulmaz I, Cindaş A, Vatan S. Low back pain and posture. Clin Rheumatol. 1999;18:308-12.

Urquhart DM, Hoving JL, Assendelft WW, Roland M, van Tulder MW. Antidepressants for non-specific low back pain. Cochrane Database Syst Rev. 2008;1:CD001703.

van de Pol RJ, van Trijffel E, Lucas C. Inter-rater reliability for measurement of passive physiological range of motion of upper extremity joints is better if instruments are used: a systematic review. Aust J Physiother. 2010;56:7-17.

van der Wurff P, Hagmeijer RH, Meyne W. Clinical tests of the sacroiliac joint. A systematic methodological review. Part 1: Reliability. Man Ther. 2000;5:30-6.

Van Dillen LR, Sahrmann SA, Norton BJ, Caldwell CA, Fleming DA, McDonnell MK, Woolsey NB. Reliability of physical examination items used for classification of patients with low back pain. Phys Ther. 1998;78:979-88.

Van Dillen LR, Sahrmann SA, Norton BJ, Caldwell CA, McDonnell MK, Bloom NJ. Movement system impairment-based categories for low back pain: stage 1 validation. J Orthop Sports Phys Ther.2003;33:126-42.

van Duijn M, Eijkemans MJ, Koes BW, Koopmanschap MA, Burton KA, Burdorf A. The effects of timing on the cost-effectiveness of interventions for workers on sick leave due to low back pain. Occup Environ Med. 2010;67:744-50.

van Trijffel E, Anderegg Q, Bossuyt PM, Lucas C. Inter-examiner reliability of passive assessment of intervertebral motion in the cervical and lumbar spine: a systematic review. Man Ther. 2005;10:256-69.

van Tulder M, Koes B. Chronic low back pain. Am Fam Physician. 2006;74:1577-9.

van Tulder MW, Koes B, Malmivaara. A Outcome of non-invasive treatment modalities on back pain: an evidence-based review Eur Spine J. 2006;Suppl 1:S64-81.

van Tulder MW, Touray T, Furlan AD, Solway S, Bouter LM; Cochrane Back Review Group. Muscle relaxants for nonspecific low back pain: a systematic review within the framework of the cochrane collaboration. Spine J. 2003;28:1978-92.

Viikari-Juntura E, Shiri R, Solovieva S, Karppinen J, Leino-Arjas P, Varonen H, Kalso E, Ukkola O. Risk factors of atherosclerosis and shoulder pain--is there an association? A systematic review. Eur J Pain. 2008;12:412-26.

Von Korff M. Studying the natural history of back pain. Spine (Phila Pa 1976). 1994;18 Suppl:2041S-2046S.

Waddell G, Burton K, Aylward M. Work and common health problems. J Insur Med. 2007;39:109-20.

Walsh J, Hall T. Reliability, validity and diagnostic accuracy of palpation of the sciatic, tibial and common peroneal nerves in the examination of low back related leg pain. Man Ther. 2009;14:623-9.

Wand BM, Bird C, McAuley JH, Doré CJ, MacDowell M, De Souza LH. Early intervention for the management of acute low back pain: a single-blind randomized controlled trial of biopsychosocial education, manual therapy, and exercise. Spine J. 2004;29:2350-6.

Wand BM, O'Connell NE. Chronic non-specific low back pain - sub-groups or a single mechanism? BMC Musculoskelet Disord. 2008;9:11.

Weiser S, Rossignol M. Triage for nonspecific lower-back pain. Clin Orthop Relat Res. 2006;443:147-55.

Weiss HR, Werkmann M. Unspecific chronic low back pain - a simple functional classification tested in a case series of patients with spinal deformities. Scoliosis. 2009:17;4:4.

Werneke M, Hart DL, Cook D. A descriptive study of the centralization phenomenon. A prospective analysis. Spine J. 1999;24:676-83.

Werneke M, Hart DL. Discriminant validity and relative precision for classifying patients with nonspecific neck and back pain by anatomic pain patterns. Spine J. 2003;28:161-6.

Werneke MW, Hart DL, Cutrone G, Oliver D, McGill T, Weinberg J, Grigsby D, Oswald W, Ward J. Association between directional preference and centralization in patients with low back pain. J Orthop Sports Phys Ther. 2011;40:801-10.

Werneke MW, Hart DL. Categorizing patients with occupational low back pain by use of the Quebec Task Force Classification system versus pain pattern classification procedures: discriminant and predictive validity. Phys Ther. 2004;84:243-54.

Werneke MW, Hart DL. Centralization: association between repeated end-range pain responses and behavioral signs in patients with acute non-specific low back pain. J Rehabil Med. 2005;37:286-90.

Whitfill T, Haggard R, Bierner SM, Pransky G, Hassett RG, Gatchel RJ. Early intervention options for acute low back pain patients: a randomized clinical trial with one-year follow-up outcomes. J Occup Rehabil. 2010;20:256-63.

Williams CM, Latimer J, Maher CG, McLachlan AJ, Cooper CW, Hancock MJ, Day RO, McAuley JH, Lin CW. PACE--the first placebo controlled trial of paracetamol for acute low back pain: design of a randomised controlled trial. BMC Musculoskelet Disord. 2010;11:169.

Williams CM, Maher CG, Hancock MJ, McAuley JH, McLachlan AJ, Britt H, Fahridin S, Harrison C, Latimer J. Low back pain and best practice care: A survey of general practice physicians. Arch Intern Med. 2010;170:271-7.

Wilson L, Hall H, McIntosh G, Melles T. Intertester reliability of a low back pain classification system. Spine. 1999;24:248-54.

Young S, Aprill C, Laslett M. Correlation of clinical examination characteristics with three sources of chronic low back pain. Spine J. 2003;3:460-5.

Appendix 1

BACK EXAMINATION FORM:
Date:
Name:
Date of birth:

Test n:	**procedure**
1. Walking	Follow the patient's movements; if movement is difficult (limping, slow), then make a finding (2); otherwise normal (1).
2. Undressing	Follow the patient`s dressing / undressing during inspection, if the patient is having difficulties with clothes, socks or shoes, then make a finding (2); otherwise normal (1).
3. Walking on toes	Follow the patients walking 5 steps forward on toes. If walking is difficult, make a finding (2); otherwise normal (1).
4. Heel-walking	Have patient walk 5 steps backwards on his heels. If walking is difficult, make a finding (2); otherwise normal (1).
5. Squat and rise	The patient stands with feet apart and holds the therapist by the hands. The patient is asked to squat so that his heels are in contact with the ground at all times, and then the same when the patient is standing on his toes. If this is difficult, make a finding (2); otherwise normal (1).
6. Posture of lumbar spine	The patient stands with his back towards the therapist, feet slightly apart and pointing in the same direction. Note the position of the back from behind and the lateral aspects. If any scoliosis or abnormal

	lordosis, make a finding (2); otherwise normal (1).
7. Posture of knee	Knees observed from behind and in lateral aspects; if genu varus, valgus, or hyperextension, make a finding (2); otherwise normal (1).
8. Posture of feet	Check for position and possible lowering of the arch of each foot, hammertoe, or hallux valgus. If findings, mark them on the form.
9. Leg length difference	Estimate lower limb length difference while the patient is standing and bending forward. If asymmetry in PSIS on the same side in both positions, make a finding (2); otherwise normal (1).
10. Lumbar spine flexion	Patient bends forward as far as possible with palms together. If bending is painful or restricted (fingertips more than 10 cm from the ground), make a finding (2). If bending causes radiating pain to the gluteus area or lower limb, mention it. The therapist places an arm around the patient and causes traction to learn whether pain is relieved.
11. Lumbar spine extension	With therapist behind the patient, holding the pelvis motionless, the patient leans back and pushes hips forward as far as possible. Ensure that the patient does not extend his neck. If leaning is painful or restricted (with no skin fold between one or more spinous processes), mark finding (2); otherwise normal (1).
12. Lumbar spine lateral flexion	The patient laterally flexes to the right, sliding his right hand towards the back of

	the knee. Control this movement to avoid forward or backward bending. If movement painful or restricted, make a finding (2); otherwise normal (1). Repeat on the other side.
13. Specific p-a movement (joint-play)	The patient, lying prone, lays both hands on the table beside his body, if lumbar lordosis increases, place a pillow under the waist. Push with the thenar the lower spinous process and with the other hand's fingers compare movement of the spinous process above. Palpate first down then up until Th12-L1. Then up to down, changing hands to now push on the upper spinous process and compare it to the movement of the lower spinous process. Mark the finding (2) if one or more hypo- or hypermobile segments can be compared to the other; otherwise normal (1). Mark on form in which moving segment the problem is and whether hypo- or hypermobile.
14. Hip rotation	With the patient lying prone on the plinth, knees in 90 degrees flexion, the examiner takes the patient's right lower limb in his armpit and produces hip outward- and inward rotations. If rotation painful or limited (end-feel is hard), mark the finding (2); otherwise normal (1). Repeat for the left hip.
15. Lumbar spine extension with traction	If leaning forward is painful, the examiner puts a hand on the patient's lower back at the painful area and provide gentle traction. Pain relief, and movement-increase are signs of hypermobility.

16. Physiological movement	The patient moves from his maximal extension and side bending left to his flexion and side-bending right. If movement is restricted or painful, mark the finding (2); otherwise normal (1). Repeat on the other side.
17. Posterior pelvic pain provocation	With patient lying supine on the plinth, right knee in 90-degree flexion. Push down on the right knee. If pain occurs, mark the finding (2); otherwise normal (1). Repeat on the left.
18. Interspinous pain	Palpate with fingertips between lumbar spine processes; if painful, make finding (2); otherwise normal (1).
19. Kibler's skin rolling test	Skin roll from down to up, bilaterally, from the spinous processes. If the subcutaneous layer doesn't roll or is painful, make the finding (2); otherwise normal (1).
20. Sacroiliac joint provocation	The examiner, on right side, places the right thumb on patient's right top corner of the sacrum, pushing the thumb towards the floor. If pain occurs, make a finding (2); otherwise normal (1). Repeat on the left.
21. L4, L5 rotation provocation	Examiner with thumb on the L4 and L5 transverse processes on the right side, with his other hand's thenar presses on his own thumb towards the floor, rotating the vertebrae. If rotation provokes radiating symptoms to the right gluteus or distally, make a finding (2); otherwise normal (1). Repeat on the left.

22. Hamstring tightness	Examiner takes a grip above the patient's malleolus, lifting the lower limb in the sagittal plane towards the ceiling, knee extended. If movement is limited because of tightness (a contract-relax technique will increase range of movement), mark the finding (2); otherwise normal (1). Mark approximate angle at which the leg was lifted. Repeat on the left.
23. Piriformis tightness	Examiner takes a grip above the patient's malleolus, supporting at the same time with his body the patient's waist. The examiner will passively flex the patient's hip to 60 degrees, with adduction, and inward rotation. If the position produces radiating pain to the lower limb, flex the knee. If the symptoms still occur, mark the finding (2); otherwise normal (1). Repeat on the left.
24. Gluteus med/min tightness	With a grip above the patient's malleolus, supporting at the same time with his body the patient's waist, the examiner passively performs the patient's hip flexion, adduction, and medial rotation with the right leg in his armpit. If movement is limited because of tightness (a contract-relax technique will increase the range of movement), mark the finding (2); otherwise normal (1). Repeat on the left
25. Iliopsoas tightness	Fixing with his right arm the patient's waist from the right hip, the examiner flexes the patient's right knee, heel towards buttocks. If more than two fingers fit between heel and buttocks, make a finding (2); otherwise normal (1). Repeat on the left.

26. One-leg standing	Patient flexes his left hip and knee to 90 degrees and stands for 30 seconds. If the patient loses balance, make a finding (2); otherwise normal (1). Report also if a pelvic belt makes a difference. Repeat on the left.
27. Active straight leg raise	With patient lying supine on the plinth, ask the patient to lift his right leg about 5 cm from the plinth. If no difficulties, provide slight resistance above the malleoli. If pain occurs, or the patient feels weakness, mark the finding (2); otherwise normal (1). Repeat on left
28. Isometric lumbar extension	With the patient lying prone on the plinth, fixate the patient`s feet above the malleoli firmly against the plinth. Ask him to extend his back as high as possible with hands behind his neck, and hold that position for 60 seconds. Mark down the angle at which the back has been lifted with an inclinometer 10 cm above the line of both spina iliaca posterior superior. Mark the finding if the patient cannot sustain that position for 60 seconds or if the back drops 10 degrees or more (2); otherwise normal (1).
29. Transversus abdominis activity	Supine on the plinth with the stabilizer under the lumbar spine, if the patient can activate the abdominals without tilting the pelvis, increase pressure from 40 to 50 mmHg. If the patient cannot maintain 50 mmHg for 10 seconds, or cannot repeat it 10 times, mark the finding (2); otherwise normal (1).

30. SLR	The examiner takes a grip above the patient's malleolus, lifting the lower limb in the sagittal plane towards the ceiling, knee extended. If pain occurs in the buttock or more distally, mark the finding (2); otherwise normal (1). Mark down the approximate angle the leg has been lifted. Repeat on the left.
31. SLUMP in sitting	With the patient in a sitting position, back of the knee touching the plinth, arms behind the back, with the sacrum's posterior side as vertical as possible, the therapist then takea a grip above the malleolus and raises the foot with the knee in full extension. Next the patient flexes the spine forward, keeping the neck straight. Then the patient flexes the neck. The same procedure is done with the other leg. If pain occurs in the buttock or more distally, mark the finding (2); otherwise normal (1). If radiating pain occurs, mark on the form at which point of the test it occurred and to which lower limb.
32. Achilles reflex	With the patient supine on the plinth, strike the right achilles tendon with the reflex hammer. If no muscle contraction occurs, mark the finding (2); otherwise normal (1). Repeat on the left.
33. Patella reflex	With knee and hip in 90 degrees flexion, and calf in the examiner's armpit, strike the patella tendon with the reflex hammer. If no muscle contraction occurs, mark the finding (2), otherwise normal (1). Repeat on the left.

34. Ely's test	With the patient prone on the plinth, the examiner lifts the right leg above the knee while fixating the hip, knee in 90 degrees of flexion. If pain occurs in the area innervated by the femoral nerve and increases during passive movement of the right ankle, mark the finding (2); otherwise normal (1). Repeat on the left.

ORIGINAL PAPERS

I

Inter- and intra-tester reliability of selected clinical tests in examining patients in early phase lumbar spine and sacroiliac joint pain and dysfunction

By

Paatelma, M., Karvonen, E., & Heinonen, A.

Advances in Physiotherapy 2010, 12, 74-80.

Advances in Physiotherapy, 2010; 12: 74–80

informa
healthcare

ORIGINAL ARTICLE

Inter- and intra-tester reliability of selected clinical tests in examining patients with early phase lumbar spine and sacroiliac joint pain and dysfunction

MARKKU PAATELMA, EIRA KARVONEN & ARI HEINONEN

Department of Health Sciences, University of Jyvaskyla, Finland

Abstract
Of all patients with low back pain (LBP), 85% are diagnosed as "non-specific lumbar pain". It has been postulated that the only reliable test is a straight leg raise test for detecting sciatic pain, and for other diagnostic subgroups, tests are of varying value. Only a few standardized tests exist to test function, inspection, mobility, pain, muscle flexibility or stability for patients with early phase lumbar spine and sacroiliac joint pain and dysfunction in the non-laboratory setting. The aim of this study was therefore to examine the inter- and intra-tester reliability of selected standardized clinical tests. Fifteen eligible, consecutive and voluntary LBP patients (aged 18−56 years), whose pain had lasted less than 3 months were recruited by an invitation letter. Patients were examined by two physiotherapists specialized in orthopaedic manual therapy. These PTs examined the same patients at a 1-week interval, changing the examination order at the second session. The assessment consisted of seven categories: history, observation of posture, function of low back and lower extremities, stability tests, pain provocation and mobility tests, neurological and neurodynamic tests, and tests of muscle tightness. In analysing these seven test categories, the mean inter-tester kappa was 0.5 (95% CI −0.05 to 1.20) and intra-tester kappa 0.6 (95% CI −0.40 to 1.28). Overall inter- and intra-tester reliability was at an acceptable level, except for the inspection test category, where agreement was poor. However, the reliability of individual tests ranged from poor to very good. In conclusion, when assessing LBP patients in the early phase (<12-week duration), reliability of one or more tests were acceptable on a group level with inter- and intra-tester reliability in every test category. However, our results need confirmation in a larger sample, especially when this great number of examination findings is tested, to reduce the likelihood of spurious findings.

Key words: *acute and subacute LBP patients, clinical tests, OMT, orthopaedic manual therapy*

Introduction

Manual therapy techniques allow physiotherapists to assess and treat musculoskeletal disorders. However, methods to assess subacute (<12-week duration) low back pain (LBP) vary considerably (1,2). Adaptation of greater standardization accuracy of assessment by clinicians may require demonstration of the capacity of this standardization in order to improve patient outcome (1). An important step of this standardization and improvement of commonly used LBP patients' examination tests or test batteries is to develop inter- and intra-tester reliability to discover the most reproducible tests or test combinations to identify LBP patients' impairments. These classified impairments (e.g. altered spinal mobility) thus suggest specific treatment choices (2). To date, physical examination in LBP typically includes observation, palpation, motion testing, muscle force and neuro-vascular assessment.

For any clinical test to be useful, it must yield reliable data. In addition, the procedure must remain true to its clinical implementation and interpretation. Studies of clinical assessments of LBP patients have demonstrated varied validity and reliability (3−10). It has been postulated that the only reliable test is a straight leg raise test for detecting sciatic pain, and for other diagnostic subgroups, tests are of varying value (11). So far, few standardized tests exist to test reliability of function, inspection, mobility, pain provocation, muscle flexibility or stability tests for patients

Correspondence: Markku Paatelma, Harkalahdenkuja 9, FI-00850 Helsinki, Finland. E-mail: markku.paatelma@auron.fi

(Received 21 December 2009; accepted 22 December 2009)

ISSN 1403-8196 print/ISSN 1651-1948 online © 2010 Informa UK Ltd. (Informa Healthcare, Taylor & Francis AS)
DOI: 10.3109/14038190903582154

with early phase lumbar spine and sacroiliac joint pain and dysfunction in the non-laboratory setting. In addition, there are also clinical tests lacking inter- or intra-tester reliability data. Therefore, in order to develop reasonable test batteries for clinical use, the existing tests' inter- and intra-tester reliability need be studied.

The aim of the study was to evaluate inter-tester reliability of selected clinical tests between two physiotherapists in orthopaedic manual therapy (OMT), and intra-tester reliability of these physiotherapists to assess LBP patients in the early phase (<12-week duration).

Methods

Subjects

Fifteen eligible, consecutive and voluntary patients with LBP lasting less than 3 months were recruited by an invitation letter from a private occupational healthcare centre (Medivire) in the city of Jyvaskyla, Finland. Exclusion criteria were pregnancy, use of psychogenic medication and diagnosed osteoporosis (Table I).

Procedure

After receiving oral and written information as to requirements for participation, the participants signed an informed consent. Patients were examined by two physiotherapists (PT) specializing in OMT, one with long experience (20 years) in OMT and one with short experience in OMT (2 years). The first one had 5 and the second 6 years of clinical experience in the field of musculoskeletal physiotherapy prior to specialization in OMT. Before starting the study, both PTs participated in a 1-day training session during which they performed the tests to be used. This included training on both patients and healthy subjects to standardize the tests.

Table I. Characteristics for 15 subjects, values in percentages unless otherwise stated.

Age, years, mean±SD	37.9±4.5
Gender (female/male)	74/26
History of present LBP episode	
Acute: <6 weeks	27
Subacute >6 weeks but <12 weeks	73
On sick-leave because of LBP	0
Pain and symptom location during the first examination	
VAS from 0 to 10 over last 24 h (mean±SD)	3±2.9
Signs of pain drawing in low back (mean±SD)	3.7±8.5
Signs of pain drawing in lower leg (mean±SD)	2.9±4.8
Pain and symptom location during second examination	
VAS from 0 to 10 over last 24 h (mean±SD)	4±4.2
Signs of pain as drawn in low back (mean±SD)	3.7±7.8
Sighs of pain as drawn in lower leg (mean±SD)	3.4±2.8
Physical work	
Light	67
Heavy	33

LBP, low back pain; VAS, visual analogue scale.

The clinical assessment included 34 different tests; one assessment session lasted 30 min. When the first PT finished the examination session with patient A, the second PT started immediately with the same patient. Both PTs examined those 15 patients during the same day, blinded to each other's test results. All tests were done by both therapists in the same manner and in the same order (Appendix 1 – only available in the online version of the journal. Please find this material with the following direct link to the article: http://www.informahealthcare.com/[10.3109/14038190903582154]). Thereafter, the PTs examined the same patients at a 1-week interval, changing the examination order at the second session. In all tests, the decision was binary: the test was either negative (normal finding) or positive (pathological finding). The ethics committee of Jyvaskyla University approved the study protocol.

The data were collected by a lumbar spine assessment form. Patients filled in a questionnaire with pain drawings and a visual analogue scale (from 0 to 10). The assessment consisted of history and 34 different tests, which were divided into seven categories: (i) functions of the lumbar spine and lower extremities (five tests), (ii) inspection of posture (four tests), (iii) mobility tests of the lumbar spine, sacroiliac joints and hip joints (five tests), (iv) pain provocation tests (seven tests), (v) muscle tightness tests (four tests), (vi) stability tests for the lumbar spine and pelvis (four tests), and (vii) neurological and neurodynamic tests (five tests) (Table II).

Statistical analysis

Statistical analyses were performed with SPSS software, version 16.0. Percentage agreement and the kappa statistic served to test intra- and inter-tester agreement. The kappa statistic estimates the degree of agreement corrected for chance agreement. General agreement exists that kappa is the one of the preferred statistics for estimation of the accuracy of nominal and ordinal data in clinical research. Percentage agreement does not take into account any agreement related solely to chance (12).

All of the 34 clinical tests were repeated by both PTs at 1-week intervals, and the PTs' results were pooled for intra-tester analysis. In addition, both PTs performed all the tests on both test days, and thus, inter-tester analysis was calculated between two examiners comparing their pooled results of both test days. The kappa value can be interpreted as follows –

Table II. Classification of clinical tests.

1. Functions
Walking
Undressing
Walking on toes
Heel-walking
Squat and rise
2. Inspection
Posture of lumbar spine
Leg length difference
Posture of knee
Posture of feet
3. Mobility
Lumbar spine flexion
Lumbar spine extension
Lumbar spine lateral flexion
Hip rotation
Specific p-a mobility of T12-S1
4. Pain provocation
Extension with traction
Physiological movement
Posterior pelvic pain provocation
Interspinosus pain
Kibler's skin rolling
Sacroiliac joint provocation
L4, L5 rotation provocation
5. Muscle tightness
Hamstrings
Piriformis
Gluteus medius/minimus
Iliopsoas
6. Stability
One-leg standing
Active straight-leg raise
Isometric lumbar extension
Transverses abdominis activity
7. Neurology/neurodynamics
Slump in sitting
Straight-leg raise
Achilles reflex
Patella reflex
Ely's (femoral nerve tension) test

poor: 0−0.20; fair: 0.21−0.40; moderate: 0.41−0.60; good: 0.61−0.80; and very good: 0.81−1.00 (13).

However, when the prevalence of a rating in the population is very high or low, the value of kappa may indicate poor reliability even with a high observed proportion of agreement (14). A kappa coefficient above 0.40 is generally regarded as acceptable (15) and percentage of agreement as 70% (16). In addition, clinical relevance is considered acceptable when lower limit of CI (95%) in kappa is >0.40 (17).

Results

Inter- and intra-tester kappa and agreement values are given in Table III.

In analysing all seven test categories, the mean *inter-tester* kappa was 0.5 (95% CI −0.05 to 1.20) and agreement 79% (ranging from 51% to 98%). In functional tests, the between-testers' overall agreement was at a high level: kappa 0.9 (95% CI 0.4−1.2) and agreement 98%. However, agreement was poor: kappa 0.2 (95% CI −0.0 to 0.6) and agreement 51% in the inspection test category overall, as well as in almost all single tests in this category. However, in the other test categories, kappa was moderate and agreement % good: mobility − kappa 0.5 (95% CI 0.1−1.0) and agreement 85%; pain provocation − 0.5 (95% CI −0.1 to 1.1) and agreement 78%; muscle tightness − 0.4 (95% CI 0.1−1.1) and agreement 79%; stability − 0.5 (95% CI 0.1 −1.0) and agreement 80%; and by neurological and neurodynamic tests − 0.5 (95% CI −0.2 to 1.2) and agreement 82%. Even though overall inter-tester reliability was at an acceptable level in these test categories, in some single tests reliability between testers varied from poor to good, except in the functional test category (Table III).

Analysing all seven test categories, the mean *intra-tester* kappa was 0.6 (95% CI −0.20 to 1.28) and agreement 81% (ranging from 69% to 98%). In the functional test category, agreement was at a high level: kappa 0.9 (95% CI 0.4−1.2) and agreement 98%. In intra-tester reliability, kappa was fair and agreement moderate; kappa 0.5 (95% CI 0.1−0.9) and agreement 77% in the inspection test category as well as in all single tests in this category.

In other categories, kappa ranged from fair to good, and agreement % was good: mobility − kappa 0.5 (95% CI 0.1−1.0) and agreement 80%; pain provocation − kappa 0.4 (95% CI −0.1 to 1.0) and agreement 77%); muscle tightness − kappa 0.6 (95% CI −0.2 to 1.0) and agreement 90%; stability − kappa 0.6 (95% CI 0.1−1.0) and agreement 83%; and by neurological and neurodynamic tests − kappa 0.5 (95% CI 0.1−1.0) and agreement 85%. Regarding intra-tester reliability, the overall agreement was at an acceptable level in these test categories. However, in some single tests in the mobility and provocation test categories, reliability between the test sessions was poor (kappa 0.10) (Table III).

Discussion

On the basis of the kappa statistic, assessments of the lumbar spine clinical tests performed by the two physiotherapists were moderately reliable, i.e. at an acceptable group level (kappa >0.40) within and between testers in most of the test categories. They were not, however, reliable for assessing posture. The percentages of agreement between the two examiners were rather good and similar for both assessments at a group level. The high percentages of agreement were almost always accompanied by moderate or good kappa values. In a few cases, when the prevalence

Table III. Results of test battery.

	Inter-tester kappa (95% CI)	Inter-tester agreem. %	Intra-tester kappa (95% CI)	Inter-tester agreem. %
1. Functions				
Walking	0.78 (0.37−1,19)	93	0.63 (0.37−1.19)	93
Undressing	1.0 (1.0−1.0)	100	1.0 (1.0−1.0)	100
Walking on toes	1.0 (1.0−1.0)	100	1.0 (1.0−1.0)	100
Heel-walking	1.0 (1.0−1.0)	100	1.0 (1.0−1.0)	100
Squat and rise	1.0 (1.0−1.0)	100	1.0 (1.0−1.0)	1.0
Mean	0.9 (0.4−1.2)	98	0.9 (0.4−1.2)	98
2. Inspection				
Posture of lumbar spine	0.35 (−0.09 to 0.55)	60	0.32 (0.03−0.67)	67
Leg length difference	0.06 (−0.14 to 0.18)	33	0.59 (0.31−0.89)	77
Posture of knee	0.10 (−0.10 to 0.28)	47	0.66 (0.39−0.93)	83
Posture of feet	0.20 (−0.10 to 0.50)	63	0.57 (0.27−0.87)	83
Mean	0.2 (−0.1 to 0.55)	51	0.5 (0.1−0.93)	77
3. Mobility				
Lumbar spine flexion	0.53 (0.16−0.88)	84	0.71 (0.40−1.00)	90
Lumbar spine extension	0.62 (0.32−0,88)	80	0.10 (−0.20 to 0.41)	54
Lumbar spine lateral flexion	0.38 (0.10−0.78)	87	0.60 (0.27−0.93)	87
Hip rotation	0.60 (0.08−0.97)	87	0.59 (0.23−0.95)	84
Specific p-a mobility	0.43 (0.19−0,90)	87	0.71 (0.34−1.00)	84
Mean	0.5 (0.1−1.0)	85	0.5 (−0.2 to 1.0)	80
4. Pain provocation				
Extension with traction	0.65 (0.36−0.92)	83	0.40 (−0.10 to 0.69)	87
Physiological movement	0.52 (0.21−0.83)	77	0.39 (0.06−0.72)	67
Post pelvic pain provocation	0.82 (0.57−1.00)	94	0.63 (0.30−0.96)	87
Interspinosus pain	0.50 (0.19−0,90)	87	0.71 (0.33−1.00)	84
Kibler's skin rolling	0.23 (−0,10 to 0.54)	60	0.21 (−0.13 to 0.55)	67
Sacroiliac joint provocation	0.27 (−0.23 to 0.77)	87	0.10 (−0.26 to 0.40)	76
L4, L5 rotation provocation	0.31 (−0.10 to 0.60)	66	0.55 (0.24−0.87	80
Mean	0.5 (−0.1 to 1.07)	78	0.4 (−0.1 to 1.0)	77
5. Muscle tightness				
Hamstrings	0.79 (0.52−1.00)	94	0.79 (0.52−1.00)	94
Piriformis	0.10 (−0.17 to 0.26)	80	0.72 (0.37−1.00)	94
Gluteus medius/minimus	0.30 (−0.10 to 0.69)	73	0.33 (−0.20 to 0.69)	80
Iliopsoas	0.19 (−0,13 to 0.50)	70	0.71 (0.41−1.00)	90
Mean	0.4 (−0.1 to 1.1)	79	0.6 (−0.2 to 1.0)	90
6. Stability				
One-leg standing	0.67 (0.32−1.00)	84	0.59 (0.04−0.89)	90
Active SLR	0.62 (0.28−0.90)	83	0.61 (0.28−0.91)	83
Isometric lumbar extension	0.46 (0.13−0.78)	70	0.48 (0.16−0.77)	67
Transverses abdominis activity	0.45 (0.08−0.82)	84	0.63 (0.30−0.96)	90
Mean	0.5 (0.1−1.0)	80	0.6 (0.1−0.96)	83
7. Neurology/neurodynamics				
Slump in sitting	0.25 (−0.15 to 0.58)	73	0.24 (0.01−0.41)	70
SLR	0.78 (0.37−1.00)	96	0.78 (0.37−1.00)	96
Achilles reflex	0.54 (0.14−0.88)	84	0.51 (0.14−0.88)	84
Patella reflex	0.35 (0.10−0.80)	73	0.41 (0.10−0.80)	90
Ely's test	0.66 (0.33−0.93)	86	0.70 (0.40−1.00)	90
Mean	0.5 (−0.2 to 1.2)	82	0.5 (0.1−1.0)	85

SLR, straight-leg raise.

index was high, kappa was reduced accordingly (the prevalence of the positive rating being either very high or very low), the percentages of agreement and kappa were at different levels, as occurred between testers in sacrum provocation, piriformis, slump and iliopsoas tests, and within testers in sacrum provocation and slump tests. Because of the small sample size, the 95% confidence intervals were wide, which suggest the low-level individual precision of these tests, and thus, rather low clinical relevance.

As one can expect, inter- and intra-tester overall agreement and clinical relevance was very good in all functional tests. This is in line with the study by Weiss & Werkmann (18), who have shown that chronic non-specific LBP is also possible clearly to classify physically with functional tests. The high reliability in the functional tests can be explained because tests used in this study were easy to carry out and observe, and there was low variability in functional limitations. Despite the fact that the reliability was very good

in these tests, our previous study showed that the sensitivity of the functional tests is low in separating chronic or subacute LBP patients from controls. Because only one-third of LBP patients can be identified (19), the usability of functional tests in clinical work should be reassessed.

The intra-tester reliability in inspection was acceptable at the group level in leg length difference, and posture of knee and feet, but clearly below standard in inter-tester agreement. This is in accordance with the study by O'Haire and Gibbons (20) that also showed a low grade of reliability as an inter-tester observation. The poor inter-tester reliability can be explained by difficulties to localise certain anatomical landmarks. The literature suggests that inter- and intra-tester reliability can be enhanced in some of the inspection tests by use of a specific measurement device such as goniometry for measurement of the lumbar lordosis/lumbo-pelvic angle (21).

Inter- and intra-tester reliability of all mobility tests seemed to be at least moderate at the group level. The exception was the lumbar extension test, where intra-tester reliability was poor. The reason for these poor reliability scores might be the variation in pain between the first and second tests. Therefore, in order to improve test accuracy, instructions and interpretation scores of pain and quality of movement (22) should be added, so that the tests are not merely evaluation of the mobility.

In the segmental p-a mobility test, reliability between examiners was moderate. However, the study by Landel et al. (4) reported good kappa (0.71) and agreement (83%) scores for identifying the least mobile segment, but low scores for identifying the most mobile segment (kappa 0.29). Their findings are at least partly in line with ours, where examiners evaluated both hypo- and hypermobile segments, but because only a binary decision was made (the test was either normal or abnormal), it is difficult to identify whether the abnormal result was because of hypo- or hypermobility.

In the lumbar flexion test, which also turned out to be a clinical relevant test within examiners, our results are line with those of the studies that have been reporting good inter-tester scores (10,23,24). Moreover, it seems that the precision of spine mobility tests might be improved with special devices (25). Furthermore, pain assessment may improve mobility tests' precision (22), and thus it would be reasonable to include it in the test procedure. Overall, the tests in this category are acceptable at a group level except for the lumbar spine extension test.

In provocation tests, inter-tester reliability was little better than intra-tester reliability, which could also mean that pain could have changed between test procedures. The highest inter-tester agreement was in the posterior pelvic pain provocation (PPPP) test, which is also a clinically relevant test, and the lowest in Kibler's skin rolling test. The highest intra-tester value was in PPPP and the lowest in the sacroiliac joint provocation test. Since it has been shown that sacroiliac joint dysfunction tests have generally poor inter-tester reliability, it is our and others' recommendation (9) to use PPPP test instead of joint dysfunction or motion palpation tests. In the interspinosus provocation test, reliability ranged from moderate to good. In the provocation test category, our suggestion is in line with the recommendation by Laslett (26) as to the sensitivity and specificity of these tests; Paatelma et al. (19) also speaks in favour of them.

There occurred at least moderate overall intra-tester reliability at a group level regarding the tightness of the hamstrings, piriformes and iliopsoas muscles, with acceptable clinical relevance within testers. However, inter-tester reliability was good only in the hamstrings test, which corresponds only partly to the results by Holmich et al. (27) that showed both inter- and intra-tester reliability was at a good level in the iliopsoas test. These observations regarding reliability are at the same level as our previous findings regarding the sensitivity and specificity of muscle tightness tests (19), which can recognize only 50% of the LBP patients from among healthy controls. Therefore, the standardization of muscle tightness tests is necessary to enhance inter-tester reliability.

Reliability ranged from moderate to good in all four stability tests being acceptable at a group level. The validity and reliability of stability tests for sacroiliac joints are described by several authors: De Groot et al. (28) found the active straight leg raise test (ASLR) to have a good differentiation capacity to distinguish between pregnant women with and without LBP. Inter-tester reliability coefficients (kappa) were >0.70 for the ASLR in the study by Roussel et al. (29). Mens et al. (30) reported for the ASRL test an intra-tester reliability correlation of $r=0.87$ and inter-tester $r=0.83$. Devices like ultrasound images or EMG have proven to be more precise methods than are non-device methods for detecting the activity of segmental muscles like the transverses abdominis and multifidus (31–37). Their use is, however, expensive and time-consuming, and thus impossible for everyday clinical practice.

We examined five neurological and neurodynamic tests, for which inter- and intra-tester reliability ranged from fair to good. The most reliable both in inter- and intra-tester categories was a straight-leg raise (SLR) test and its reliability was almost clinically relevant. The slump test had the lowest values. These findings agree with our previous ones, showing that neural and neurodynamic tests can detect less than 40% of the LBP patients from among the healthy subjects (19).

The SRL test is the most common way to evaluate irritability of the sciatic nerve. Intra-tester results have shown good reliability (38), and a negative test outcome may be of greater diagnostic value than a positive one. Biomechanical devices have improved intra- and inter-tester reliability and thus have increased test reproducibility (39). Kappa scores for agreement between testers for manual palpation of the sciatic nerve have been shown ranging from 0.70 to 0.80 during the SLR and slump tests (40). Clinical tests, which evaluate increased nerve mechanosensitivity and afferent/efferent nerve function, show comparable moderate-to-good reliability (41). With regard to the physical examination, SLR is the only sign consistently reported as sensitive for sciatica related to disc herniations; the diagnostic accuracy of other neurological signs and tests is unclear (11). One study shows the slump test to be even more sensitive than the SLR test, but in contrast, SLR has a higher specificity (42). Accuracy and perhaps sensitivity and specificity can also be improved by goniometry.

Because two examiners performed the physical examination of each patient independently, each examiner could make an independent judgment about the history and clinical tests. Testing the same patient twice during 1 day allowed findings of a single test to vary, because of increased or decreased pain. In future, the level of pain on a VAS scale during the test procedure should be reported and should be included in the interpretation of the tests. It would also be reasonable to discuss and also study, if the test accuracy can be improved, for example, by a Likert-like scale instead of a binary decision. Between the first and second test trial, patients received no interventions but were advised to stay normally active according to international clinical guidelines (43).

All in all, these tests' inter- and intra-tester reliability ranged from poor to very good. Some tests showed high group level inter- and intra-tester reliability in all categories, but some tests also showed poor reliability in some of the categories. However, clinical relevance of tests was weak except in functional tests and the hamstring test. In addition, the PPPP test showed clinical relevance between examiners, and lumbar spine flexion, iliopsoas and Ely's test within examiners. Future research is needed for development of these clinical tests, which had high kappa values to determine their differentiate capacity in an LBP population. Before that, our results need also confirmation in a larger sample, especially when these great numbers of examination findings are tested, to reduce the likelihood of spurious findings.

Declaration of interest: The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

References

1. May S, Littlewood C, Bishop A. Reliability of procedures used in the physical examination of non-specific low back pain: A systematic review. Aust J Physiother. 2006;52: 91–102.
2. Zimny NJ. Diagnostic classification and orthopaedic physical therapy practice: What we can learn from medicine. J Orthop Sports Phys Ther. 2004;34:105–9.
3. Kent PM, Keating JL, Taylor NF. Primary care clinicians use variable methods to assess acute nonspecific low back pain and usually focus on impairments. Man Ther. 2009;14: 88–100.
4. Landel R, Kulig K, Fredericson M, Li B, Powers CM. Intertester reliability and validity of motion assessments during lumbar spine accessory motion testing. Phys Ther. 2008;88:43–9.
5. Arab AM, Abdollahi I, Joghataei MT, Golafshani Z, Kazemnejad A. Inter- and intra-examiner reliability of single and composites of selected motion palpation and pain provocation tests for sacroiliac joint. Man Ther. 2009;14:213–21.
6. Qvistgaard E, Rasmussen J, Laetgaard J, Hecksher-Sørensen S, Bliddal H. Intra-observer and inter-observer agreement of the manual examination of the lumbar spine in chronic low-back pain. Eur Spine J. 2007;16:277–82.
7. McCarthy CJ, Gittins M, Roberts C, Oldham JA. The reliability of the clinical tests and questions recommended in international guidelines for low back pain. Spine. 2007; 32:921–6.
8. Laslett M, McDonald B, Tropp H, Aprill CN, Oberg B. Agreement between diagnoses reached by clinical examination and available reference standards: A prospective study of 216 patients with lumbopelvic pain. BMC Musculoskelet Disord. 2005;6:28.
9. Riddle DL, Freburger JK. Evaluation of the presence of sacroiliac joint region dysfunction using a combination of tests: A multicenter intertester reliability study. Phys Ther. 2002;82:772–81.
10. Vincent-Smith B, Gibbons P. Inter-examiner and intra-examiner reliability of the standing flexion test. Man Ther. 1999;4:87–93.
11. Rubinstein SM, van Tulder M. A best-evidence review of diagnostic procedures for neck and low-back pain. Best Pract Res Clin Rheumatol. 2008;22:471–82.
12. Kramer M, Feinstein A. Clinical biostatistics LIV. The biostatistics of concordance. Clin Pharmacol Ther. 1981;29: 111–123.
13. Altman DG. Practical statistics for medical research. Boca Raton, FL: Chapman & Hall; 1991. pp. 403–409.
14. Byrt T, Bishop J, Carlin JB. Bias, prevalence and kappa. J Clin Epidemiol. 1993;46:423–9.
15. Landis J, Koch G. The measurement of observer agreement for categorical data. Biometrics. 1977;33:159–174.
16. Potter N, Rothstein J. Intertester reliability for selected clinical tests of the sacroiliac joint. Phys Ther. 1985;65: 1671–1675.
17. Sim J, Wright C. The Kappa statistics in reliability studies: Use, interpretation, and sample size requirements. Phys Ther. 2005;85:257–268.
18. Weiss HR, Werkmann M. Unspecific chronic low back pain – A simple functional classification tested in a case series of patients with spinal deformities. Scoliosis. 2009;4:4.
19. Paatelma M, Karvonen E, Heiskanen J. How do clinical test results differentiate chronic and subacute low back pain patients from "non-patients". J Man Manip Ther. 2009;17: 11–19.
20. O'Haire C, Gibbons P. Inter-examiner and intra-examiner agreement for assessing sacroiliac anatomical landmarks using

palpation and observation: Pilot study. Man Ther. 2000; 5:13–20.
21. Gilliam J, Brunt D, MacMillan M, Kinard RE, Montgomery WJ. Relationship of the pelvic angle to the sacral angle: Measurement of clinical reliability and validity. J Orthop Sports Phys Ther. 1994;20:193–9.
22. O'Sullivan P. Diagnosis and classification of chronic low back pain disorders: Maladaptive movement and motor control impairments as underlying mechanism. Man Ther. 2005; 10:242–55.
23. Luomajoki H, Kool J, de Bruin ED, Airaksinen O. Reliability of movement control tests in the lumbar spine. BMC Musculoskelet Disord. 2007;12;8:90.
24. Dankaerts W, O'Sullivan PB, Straker LM, Burnett AF, Skouen JS. The inter-examiner reliability of a classification method for non-specific chronic low back pain patients with motor control impairment. Man Ther. 2006;11:28–39.
25. Madson TJ, Youdas JW, Suman VJ. Reproducibility of lumbar spine range of motion measurements using the back range of motion device. J Orthop Sports Phys Ther. 1999;29: 470–7.
26. Laslett M. Evidence-based diagnosis and treatment of the painful sacroiliac joint. J Man Manip Ther. 2008;16: 142–52.
27. Holmich P, Holmich L, Bjerg A. Clinical examination of athletes with groin pain: An intraobserver and interobserver reliability study. Br J Sports Med.2004;38:446–451.
28. de Groot M, Pool-Goudzwaard AL, Spoor CW, Snijders CJ. The active straight leg raising test (ASLR) in pregnant women: Differences in muscle activity and force between patients and healthy subjects. Man Ther. 2008;13:68–74.
29. Roussel NA, Nijs J, Truijen S, Smeuninx L, Stassijns G. Low back pain: Clinimetric properties of the Trendelenburg test, active straight leg raise test, and breathing pattern during active straight leg raising. J Manip Physiol Ther. 2007; 30:270–8.
30. Mens JM, Vleeming A, Snijders CJ, Koes BW, Stam HJ. Reliability and validity of the active straight leg raise test in posterior pelvic pain since pregnancy. Spine. 2001;26: 1167–71.
31. Costa LO, Costa Lda C, Cançado RL, Oliveira Wde M, Ferreira PH. Short report: Intra-tester reliability of two clinical tests of transversus abdominis muscle recruitment. Physiother Res Int. 2006;11:48–50.
32. Hides JA, Miokovic T, Belavý DL, Stanton WR, Richardson CA. Ultrasound imaging assessment of abdominal muscle function during drawing-in of the abdominal wall: An intrarater reliability study. J Orthop Sports Phys Ther. 2007;37:480–6.
33. Evans K, Refshauge KM, Adams R. Trunk muscle endurance tests: Reliability, and gender differences in athletes J Sci Med Sport. 2007;10:447–55.
34. Wallwork TL, Hides JA, Stanton WR. Intrarater and interrater reliability of assessment of lumbar multifidus muscle thickness using rehabilitative ultrasound imaging. J Orthop Sports Phys Ther. 2007;37:608–12.
35. Mannion AF, Pulkovski N, Gubler D, Gorelick M, O'Riordan D, Loupas T, et al. Muscle thickness changes during abdominal hollowing: An assessment of between-day measurement error in controls and patients with chronic low back pain. Eur Spine J. 2008;17:494–501.
36. Coorevits P, Danneels L, Cambier D, Ramon H, Druyts H, Karlsson JS, et al. Test–retest reliability of wavelet- and Fourier-based EMG (instantaneous) median frequencies in the evaluation of back and hip muscle fatigue during isometric back extensions. J Electromyogr Kinesiol. 2008;18: 798–806.
37. Koppenhaver SL, Hebert JJ, Fritz JM, Parent EC, Teyhen DS, Magel JS. Reliability of rehabilitative ultrasound imaging of the transversus abdominis and lumbar multifidus muscles. Arch Phys Med Rehabil. 2009;90:87–94.
38. Hunt DG, Zuberbier OA, Kozlowski AJ, Robinson J, Berkowitz J, Schultz IZ, et al. Reliability of the lumbar flexion, lumbar extension, and passive straight leg raise test in normal populations embedded within a complete physical examination. Spine. 2001;26(24):2714–8.
39. Rebain R, Baxter G, McDonough S. A systematic review of the passive straight leg raising test as a diagnostic aid for low back pain. Spine. 2002;27:E388–395.
40. Walsh J, Hall T. Reliability, validity and diagnostic accuracy of palpation of the sciatic, tibial and common peroneal nerves in the examination of low back related leg pain. Man Ther. 2009;14:623–9.
41. Schmid AB, Brunner F, Luomajoki H, Held U, Bachmann LM, Kunzer S, et al. Reliability of clinical tests to evaluate nerve function and mechanosensitivity of the upper limb peripheral nervous system. BMC Musculoskelet Disord. 2009;10:11.
42. Majlesi J, Togay H, Unalan H, Toprak S. The sensitivity and specificity of the Slump and the Straight Leg Raising tests in patients with lumbar disc herniation. J Clin Rheumatol. 2008;14:87–91.
43. van Tulder M. Conservative non-pharmacological treatment for chronic low back pain. BMJ. 2008;337:a974.

Supplementary material available online

Appendix 1

ORIGINAL PAPERS

II

How do clinical test results differentiate chronic and subacute low back pain patients from "non-patients"?

By

Paatelma, M., Karvonen, E., & Heiskanen, J.

The Journal of Manual and Manipulative Therapy 2009, 17, 11-19

Clinical Perspective: How Do Clinical Test Results Differentiate Chronic and Subacute Low Back Pain Patients from "Non-Patients"?

Markku Paatelma, MSc, PT[1]; **Eira Karvonen**, MSc, PT[2]; **Jouko Heiskanen**, MD[3]

Low back pain (LBP) is one of the most common reasons for people to seek medical treatment in Western societies, with the majority of LBP sufferers having non-specific low back pain. This definition includes any type of back pain or referred leg pain or both that does not fall into the category of nerve root pain or serious spinal pathology[1]. Our limited understanding of underlying conditions for back pain is reflected in the common use of pain-duration-based groupings: LBP lasting for 6 weeks or less has been defined as *acute*, from 7 to 12 weeks as *subacute*, and for 12 weeks or more as *chronic* back pain[2,3]. Still, classification according to LBP duration is questionable, because of evidence that new, acute LBP is rare among adults and that most experiences of all LBP will reoccur[4].

There is a consensus about the duration of the symptoms approach toward the standardization of back pain definitions for use in prevalence studies[5]. Although inter-country differences also exist in the management of LBP sufferers from differing cultural backgrounds, clinical guidelines in different countries are mostly based on the pain-duration classification (acute, subacute, and chronic), or a classification into a diagnostic triage: serious pathology or sciatic syndrome or non-specific LBP[6].

Several physical characteristics have been associated with LBP development. Physiologic changes such as muscle dysfunction occur in the lumbar spine at the same time as do initial episodes of pain: changes that remain after the pain has subsided[7,8]. Knowing the degree of relationship between each of these factors and LBP will guide prevention treatment strategies[9]. Pengel et al[10] suggested placing more emphasis on change in pain and disability scores than on physical impairment. Substantial evidence also exists that psychosocial variables are strongly linked to the transition from acute to chronic LBP disability[11].

To date, there is only marginal literature that describes variations in clinical findings between duration-based classifi-

ABSTRACT: Our limited understanding of underlying conditions for back pain is reflected in the common use of pain-duration-based groupings. The aim of this paper was to investigate typical clinical tests used in examining low back pain (LBP) patients in order to discover how tests distinguish between chronic low back pain patients (CLBP) and subacute low back pain patients (SLBP) and if they distinguish these groups from those with no "patient status." CLBP patients in this study were from a university hospital and SLBP patients were from five occupational health care centers. Control subjects were recruited from a university. Determination of the best predictors between CLBP and SLBP patients and between CLBP and SLBP patients and non-patients was made by a forward stepwise logistic model. A total of 157 subjects were included in the study. Of all the clinical tests, several tests in each category had high odds ratio, differentiating CLBP patients from controls. Only a few tests differentiated between CLBP and SLBP patients. The only clinical differences between SLBP patients and controls were in the mobility test and in one test of muscle tightness. The best predictor for CLBP was the lumbar spine flexion test. SLBP patients seemed to differ from the control group in lumbar flexion, in a specific anterior-posterior mobility test, and in tightness of hip flexor muscles. CLBP patients differed from SLBP patients in functional tests, in the presence of sensation in the feet, and in different pain provocation tests. Whether these tests are sufficiently sensitive to classify a more specific diagnostic or clinical subgroup remains untested, and further studies with clinical tests to differentiate among pathological conditions are necessary.

KEYWORDS: Clinical Tests, Low Back Pain, Odds Ratios, Orthopedic Manipulative Therapy, Sensitivity, Specificity

[1]OMT-instructor, University of Jyvaskyla, Jyvaskyla, Finland
[2]OMT, University of Jyvaskyla, Jyvaskyla, Finland
[3]Medical Adviser in Rehabilitation, Metropolia Polytechnic, Helsinki, Finland
Address all correspondence and requests for reprints to: Markku Paatelma, markku.paatelma@tcr.fi

cations. Subsequently, the purpose of this study was to investigate typical clinical tests used in examining LBP patients in order to discover which tests distinguish between chronic low back pain patients (CLBP) and subacute low back pain patients (SLBP), and how clinical findings distinguish these groups from those with no "patient status".

Methods

Design

The prospective non-randomized comparative design was approved by the ethics committees of Pirkanmaa Hospital District and the University of Jyväskylä. All participants gave their informed consent.

Subjects

A total of 157 subjects were asked to participate, and all subjects expressed an interest. Eighty-two women and 75 men, from 18 to 68 years of age, were recruited through one hospital and a number of occupational health centers and university departments. Based on the inclusion criteria, 55 subjects were selected for the CLBP group. For this group, subjects were selected from a population of primary-care patients in the city of Tampere, Finland. The inclusion criteria were LBP lasting more than 3 months, with or without radiating pain to one or both legs. Each subject was sent for an orthopedic surgeon's consultation at Tampere University Hospital in order to evaluate whether spinal surgery was needed. If there was an acute need for spinal operation or other red flags, patients were not referred for this study.

For the SLBP group, 47 subjects were recruited from five occupational health care centers in the city of Jyväskylä, Finland. Inclusion criteria were employment, age 18 to 65, with current LBP with or without radiating pain to one or both lower legs. The back-pain episode could be the first or recurrent LBP, the last episode lasting less than 3 months. Exclusion criteria were pregnancy, back pain surgery less than 2 months earlier, or serious pathology as evidenced by the frequency of red-flag screening (age, cauda equina syndrome, cancer, fractures, spinal infections, ankylosing spondylitis, significant or progressive neurological deficit or both, or other systemic illness)[12].

For the control (non-patient) group, 55 subjects were recruited from the University of Jyväskylä in Finland. None had any medical low back diagnosis nor had a history of medical or physiotherapeutic low back treatment during the previous year. In this control group, exclusion criteria were the same as for the SLBP group: serious pathology as evidenced by the frequency of red-flag screening (age, cauda equina syndrome, cancer, fractures, spinal infections, ankylosing spondylitis, significant or progressive neurological deficit, or both, or other systemic illness) as based on the criteria of Bigos[12]. All subjects' characteristics at baseline measurement are given in Table 1.

Procedure

Severity, location of pain, and symptoms of the LBP were assessed in a short interview. Type of employment, absence from work, and episodes of LBP during each subject's lifetime were recorded. After having interviewed the subjects, physiotherapists specializing in orthopedic manual therapy performed 45 selected clinical tests on all these subjects, without having access to any medical records. These tests are commonly used during clinical examinations of LBP patients.

Overall, 14 physiotherapists were involved in examination of the subjects. Their years of clinical experience in management of LBP ranged from 5 to 16 years (mean 10.5). All physiotherapists were specialists in orthopedic manual

TABLE 1. Characteristics of the subjects

	CLBP	SLBP	Controls	Significance
Number of Subjects	55	47	55	NS
Age (yr.)				
Range	18–68	22–60	21–60	
Mean, SD	42.3 (11.6)	44.6 (10, 2)	37.5 (8, 1)	NS
Gender (female/male)	31/24	18/29	33/22	0.05 (CL[¤], Co/SL)
Number of LBP Episodes				
0	1 (2%)	3 (7%)	20 (37%)	0.05 (Co[¶], CL/SL)
1–5	12 (22%)	23 (49%)	27 (49%)	NS
6–10	22 (40%)	12 (25%)	4 (7%)	0.05 (CL, Co/SL[Å])
More Than 10	20 (36%)	9 (19%)	4 (7%)	0.05 (CL, Co/SL)
Type of Work				
Light	41 (77 %)	42 (89%)	52 (94%)	NS
Heavy	14 (23%)	5 (11%)	3 (6%)	NS
Sick Leave at Baseline				
Number of Subjects (%)	23 (42%)	6 (12%)	0	0.05 (CL, Co/SL)

CLBP = chronic low back pain; SLBP = subacute low back pain; NS = not significant
[¤] CLBP group; [Å] SLBP group; [¶] Control group; LBP = low back pain

therapy (OMT), which involves a postgraduate examination in Nordic countries. Prior to this study, all physiotherapists were taught by the OMT specialist, with 22 years of experience in this field, how to perform the specific clinical tests the same way. The physiotherapists were unaware of the ongoing study.

The clinical tests were divided into six classes as follows: functional tests, inspection, mobility assessment, muscle-tightness, neurological and neurodynamic, and pain-provocation tests (Table 2). In all tests, the decision was binary: results were either negative (normal finding) or positive (pathological finding). An identical assessment form was used for all patients, with categories always in the same sequence. The examination sequence lasted 30 minutes in total. These tests were performed with great care, so that the patients' underlying condition would not have changed during the examination. Those tests that were symmetrical (for instance, lumbar spine lateral flexion to right and left) were recorded as positive if flexion in one or both directions was abnormal. The number of tests analyzed was thus 31. The order in which the tests and measures were assessed within each category was based on the patient's positioning considerations (standing, supine, lying on one side, prone, or lying on the other side).

TABLE 2. Test categories

Type of Test	Specific Test
Functional tests	Walking
	Undressing
	Walking on toes
	Heel walking
	Squat and rise
Inspection tests	Posture of low back
	Posture of feet
	Posture of knees
	Length difference of lower limbs
	Distribution of hair on lower back
Mobility tests	Lumbar spine right lateral flexion
	Lumbar spine left lateral flexion
	Lumbar spine right rotation
	Lumbar spine left rotation
	Lumbar spine extension
	Lumbar spine flexion
	Specific PA (posterior-anterior) mobility
	Mobility of right hip
	Mobility of left hip
Muscle tightness	Hamstring tightness right
	Hamstring tightness left
	Piriformis tightness left
	Iliopsoas tightness right
	Iliopsoas tightness left
Neurological and neurodynamic tests	Slump (pain provocation test in sitting)
	SLR (left straight leg raise) (Laseque)
	SLR (right straight leg raise) (Laseque)
	Reflexes (right patella tendon)
	Reflexes (left patella tendon)
	Reflexes (right Achilles tendon)
	Reflexes (left Achilles tendon)
	Babinski right
	Babinski left
	Sensation in feet (right)
	Sensation in feet (left)
	Ely's test right (n. femoralis)
	Ely's test left (n. femoralis)
Pain provocation tests	Heel drop
	Kibler's skin-roll test
	Interspinosus ligament sensitivity
	Transverse process pressure right (unilat PA)
	Transverse process pressure left (unilat PA)
	Sacro-iliac joint pressure right
	Sacro-iliac joint pressure left

Test Reliability

Prior to this study, intertester reliability and test-retest reproducibility of all clinical tests were analyzed for in all test categories. Two physiotherapists specializing in OMT performed all the clinical tests independently during the same day/s by examining 15 volunteer LBP patients from occupational health; this tested intertester reliability. Test-retest reliability was evaluated by examining the same stable patients after one week. The 70% agreement level and 0.40 in Kappa were adopted as the minimum criteria for acceptable intra- and intertester reliability. As suggested by Potter and Rothstein, this level of agreement should be the minimum criterion for any test to be considered clinically meaningful[13]. The results of intertester reliability are presented in Table 3 and intratester reliability in Table 4. In all tests, intratester reliability exceeded 70% agreement and 0.40 Kappa level, but intertester reliability did not reach that level in inspection tests (Kappa and agreement) nor in muscle tightness tests (Kappa alone).

Data Analysis

Statistical analyses were performed with SPSS software, version 14.0. Percentage agreement and the Kappa statistic were used to test intra- and intertester reliability in clinical tests. The results were

TABLE 3. Intertester Kappa values and agreement percentages

Test	Intertester Kappa	Intertester Agreement %
Functional tests		
Mean (SD)	0.9 (0.13)	99.2 (2.37)
Inspection tests		
Mean (SD)	0.2 (0.19)	52.1 (17.57)
Mobility tests		
Mean (SD)	0.5 (0.15)	82.3 (8.71)
Provocation/alleviation tests		
Mean (SD)	0.4 (0.25)	74.4 (13.79)
Muscle tightness tests		
Mean (SD)	0.3 (0.32)	79.2 (12.06)
Stability tests		
Mean (SD)	0.5 (0.17)	80 (10.11)
Neural and neurodynamics tests		
Mean (SD)	0.4 (0.28)	82.2 (11.70)

SD = standard deviation

TABLE 4. Intratester Kappa-values and agreement percentages

Test	Intratester Kappa	Intratester Agreement %
Functional tests		
Mean (SD)	0.9 (0.13)	99.2 (2.37)
Inspection tests		
Mean (SD)	0.4 (0.27)	70.7 (14.96)
Mobility tests		
Mean (SD)	0.5 (0.22)	76.7 (15.40)
Provocation/alleviation tests		
Mean (SD)	0.4 (0.26)	85 (10.99)
Muscle tightness tests		
Mean (SD)	0.7 (0.23)	89.2 (8.68)
Stability tests		
Mean (SD)	0.6 (0.18)	82.5 (11.82)
Neural and neurodynamics tests		
Mean (SD)	0.5 (0.28)	85 (12.21)

SD = standard deviation

expressed as means and standard deviations (SD). Determination of the best predictors of chronic and subacute LBP was by a forward stepwise logistic model. The odds ratio (OR), sensitivity, specificity, and 95% confidence intervals (CIs) of all the tests were calculated.

Results

Significant differences were present in baseline characteristics. Patients in the CLPB group were 4.7 and the SLBP 7.1 years older than in the control group. Ratios between women and men were similar in the CLBP and control groups, but males predominated in the SLBP group. The CLBP group included more heavy laborers than did the SLBP or the control group. More subjects were on sick leave because of LBP in the CLBP group (Table 1).

The odds ratio, sensitivity, and specificity of all the tests are provided in Tables 5–7. Of the 31 clinical tests, several tests in every test category differentiated CLBP patients from controls, with high odds ratio, including all gross functional tests, mobility tests (lumbar spine flexion and extension, and specific posterior-anterior (PA) mobility), neurological and neurodynamic tests (slump, straight leg raise (SLR), and sensation in the feet), and pain provocation tests (heel drop, Kibler's skin roll, interspinosus ligament sensitivity, and transverse process pressure) (see Table 5).

Of the clinical tests, only gross functional tests (walking, undressing, walking on toes, and heel walking), one neurological test (sensation in the feet), and one pain provocation test (Kibler's skin roll) differentiated between CLBP and SLBP patients (see Table 6). In functional tests, one-third of the CLBP subjects had difficulty in squat and rise (31%), walking (33%), and undressing (24%), whereas the SLBP group showed only a few positive findings in this test category. The only positive finding in the control group was in the squat and rise test.

The only clinical differences between SLBP patients and controls were in mobility tests (lumbar spine flexion and extension, and specific PA mobility) and in one test of muscle tightness: tightness of hip flexors (see Table 7). Of the inspection tests, only back posture was entered in the forward stepwise logistic model, with a high odds ratio in comparison of CLBP and of SLBP with controls. In this test, proportions of positive findings were 72% in the CLBP, 72% in the SLBP, and 49% in the control group.

In the mobility test category, decreased or painful movement of the lumbar spine in extension and flexion differentiated between the CLBP and SLBP patients and the controls, especially in the flexion test. In this test, the likelihood of a positive finding was 28-fold greater for CLBP than controls. The proportions of positive findings in the flexion and extension tests were moderate, being 49% (in flexion) and 62% (in extension) in the CLBP subjects, 36% and 57% in the SLBP, and 4% and 22% in the controls, respectively. In a specific PA mobility test, the finding was positive in 84% of CLBP patients, in 93% of SLBP, and 54% for controls.

Among the tests that examine muscle tightness, the hip flexor test (ilio-

TABLE 5. Forward stepwise logistic models in six test classes: sensitivity and specificity of the selected clinical tests for chronic LBP compared to controls

Variable	Forward Stepwise Logistic Model † OR (95% CI ‡)	Sensitivity (95% CI)	Specificity (95% CI)
Functional tests:			
Walking	**26.36** (3.36 to 10)	33 (19 to 45)	100 (93 to 100)
Undressing	**20.25** (2.57 to 100)	27 (15 to 38)	100 (93 to 100)
Walking on toes	**9.19** (1.10 to 76.21)	15 (5 to 24)	100 (93 to 100)
Heel walking	**16.71** (2.10 to 100)	24 (5 to 28)	100 (93 to 100)
Squat and rise	**6.18** (2.12 to 17.97)	38 (19 to 46)	87 (75 to 95)
Inspection tests:			
Posture of low back	**2.77** (1.25 to 6.12)	72 (59 to 84)	58 (39 to 69)
Posture of feet		15 (8 to 29)	75 (54 to 81)
Posture of knees		15 (8 to 29)	93 (75 to 95)
Length difference of lower limbs		33 (15 to 39)	80 (60 to 86)
Distribution of hair on lower back		9 (3 to 20)	96 (89 to 100)
Mobility tests:			
Lumbar spine lateral flexion	**3.86** (1.66 to 8.99)	49 (35 to 63)	80 (67 to 91)
Lumbar spine rotation	**2.39** (1.03 to 5.51)	40 (27 to 54)	78 (67 to 91)
Lumbar spine extension	**5.80** (2.50 to 13.44)	62 (48 to 75)	78 (70 to 93)
Lumbar spine flexion	**27.48** (6.7 to100)	51 (37 to 65)	96 (89 to 100)
Mobility of hip		4 (3 to 20)	93 (77 to 97)
Specific PA mobility	**4.58** (1.88 to 11.55)	84 (71 to 92)	47 (37 to 67)
Muscle tightness tests:			
Tightness of hamstring		53 (36 to 66)	62 (49 to 78)
Tightness of piriformis		15 (6 to 27)	95 (39 to 69)
Tightness of hip flexors	**2.77** (1.25 to 6.12)	73 (59 to 84)	51 (39 to 69)
Neurological and neurodynamic tests:			
Slump	**5.87** (2.15 to 16.00)	42 (29 to 56)	89 (80 to 98)
SLR (straight leg raise)	**5.23** (1.62 to 16.89)	29 (19 to 45)	93 (86 to 99)
Reflexes (patella tendon)	**2.49** (0.72 to 8.65)	16 (21 to 50)	93 (72 to 97)
Reflexes (Achilles tendon)	**3.37** (1.13 to 10.08)	24 (25 to 50)	95 (77 to 97)
Babinski	-	-	-
Sensation in feet	**7.18** (2.26 to 18.97)	27 (16 to 41)	100 (93 to 100)
Ely's test		6 (2 to 18)	100 (93 to 100)
Pain provocation tests:			
Heel drop	**5.91** (1.59 to 21.99)	26 (17 to 43)	95 (86 to 99)
Kibler's skin-roll test	**4.03** (1.70 to 9.59)	47 (34 to 64)	82 (70 to 93)
Interspinosus ligament sensitivity	**4.99** (2.16 to 11.49)	78 (65 to 88)	58 (48 to 78)
Transverse process pressure	**5.37** (1.43 to 20.08)	24 (15 to 39)	95 (86 to 99)
Sacro-iliac joint presuure		13 (5 to 24)	93 (83 to 99)

† Only those variables are shown that were entered in the model
‡ Robust estimate of variance
LBP = low back pain, CI = confidence interval

TABLE 6. Forward stepwise logistic models in six test classes: sensitivity and specificity of the selected clinical tests for chronic LBP compared to subacute LBP

Variables	Forward Stepwise Logistic Model † OR (95% CI ‡)	Sensitivity (95% CI)	Specificity (95% CI)
Functional tests:			
Walking	**22.38** (2.85 to100)	33 (20 to 46)	98 (89 to 100)
Undressing	**8.44** (1.82 to 39.19)	27 (14 to 38)	96 (86 to 99)
Walking on toes	**7.80** (0.94 to 65.12)	15 (4 to 26)	98 (89 to 100)
Heel walking	**14.24** (1.79 to100)	24 (4 to 26)	98 (89 to 100)
Squat and rise	**3.53** (1.34 to 9.31)	38 (19 to 45)	94 (71 to 94)
Inspection tests:			
Posture of low back		73 (59 to 84)	28 (39 to 69)
Posture of feet		16 (8 to 29)	60 (54 to 81)
Posture of knees		16 (8 to 29)	83 (75 to 95)
Length difference in lower limbs		33 (15 to 39)	55 (60 to 86)
Distribution of hair on lower back		9 (3 to 20)	96 (89 to 100)
Mobility tests:			
Lumbar spine lateral flexion		49 (35 to 63)	51 (67 to 91)
Lumbar spine rotation	**2.05** (0.81 to 5.15)	40 (27 to 54)	81 (67 to 91)
Lumbar spine extension		62 (48 to 75)	43 (40 to 93)
Lumbar spine flexion		51 (37 to 65)	64 (89 to 100)
Mobility of hip		4 (3 to 20)	89 (77 to 97)
Specific PA mobility		84 (71 to 92)	6 (37 to 67)
Muscle tightness tests:			
Tightness of hamstring		53 (36 to 66)	36 (49 to 78)
Tightness of piriformis		15 (6 to 27)	89 (39 to 69)
Tightness of hip flexors		73 (59 to 84)	13 (39 to 69)
Neurological and neurodynamic tests:			
Slump	**2.08** (0.91 to 2.91)	42 (29 to 56)	72 (70 to 98)
SLR (straight leg raise)		29 (19 to 45)	83 (86 to 99)
Reflexes (Patella tendon)		16 (14 to 50)	79 (77 to 97)
Reflexes (Achilles tendon)		24 (24 to 50)	85 (77 to 97)
Babinski	-	-	-
Sensation in feet	**8.42** (1.54 to 26.60)	27 (16 to 41)	96 (93 to 100)
Ely's test	-	7 (2 to 18)	100 (93 to 100)
Pain provocation tests:			
Heel drop	**3.69** (1.06 to 8.47)	26 (16 to 45)	92 (71 to 94)
Kibler's skin-roll test	**4.41** (2.39 to 6.54)	47 (34 to 64)	83 (70 to 93)
Interspinosus ligament sensitivity		78 (65 to 88)	31 (38 to 78)
Transverse process pressure	**3.34** (0.97 to 7.97)	24 (15 to 39)	98 (86 to 99)
Sacro-iliac joint pressure	**5.91** (1.81 to 46.87)	13 (5 to 24)	94 (83 to 99)

† Only those variables showed that were entered in the model
‡ Robust estimate of variance
LBP = low back pain, CI = confidence interval

TABLE 7. Forward stepwise logistic models in six test classes: sensitivity and specificity of the selected clinical tests for subacute LBP compared to controls

Variable	Forward Stepwise Logistic Model † OR (95% CI ‡)	Sensitivity (95% CI)	Specificity (95% CI)
Functional tests:			
Walking		2 (< 1 to 10)	100 (93 to 100)
Undressing		4 (2 to 15)	100 (93 to 100)
Walking on toes		2 (< 1 to 10)	100 (93 to 100)
Heel walking		2 (< 1 to 10)	100 (93 to 100)
Squat and rise		15 (6 to 27)	69 (54 to 81)
Inspection tests:			
Posture of low back	**2.71** (1.18 to 6.22)	72 (59 to 84)	58 (39 to 69)
Posture of feet		40 (8 to 29)	75 (54 to 81)
Posture of knees		16 (8 to 29)	93 (75 to 95)
Length difference of lower limbs		45 (15 to 39)	80 (60 to 86)
Distribution of hair on lower back		4 (3 to 20)	96 (89 to 100)
Mobility tests:			
Lumbar spine lateral flexion	**3.83** (1.60 to 9.19)	49 (35 to 63)	81 (67 to 91)
Lumbar spine rotation		19 (27 to 54)	78 (67 to 91)
Lumbar spine extension	**4.84** (2.04 to 11.46)	57 (48 to 75)	78 (70 to 93)
Lumbar spine flexion	**15.02** (3.25 to 69,49)	36 (37 to 65)	96 (89 to 100)
Mobility of hip		9 (3 to 20)	93 (77 to 97)
Specific PA mobility	**13.15** (6.42 to 47.47)	94 (71 to 92)	47 (37 to 67)
Muscle tightness tests:			
Tightness of hamstring		64 (36 to 66)	62 (49 to 78)
Tightness of piriformis		10 (6 to 27)	95 (39 to 69)
Tightness of hip flexors	**7.09** (2.59 to 19.39)	87 (59 to 84)	51 (39 to 69)
Neurological and neurodynamic tests:			
Slump	**2.80** (0.96 to 8.18)	28 (29 to 56)	89 (80 to 98)
SLR (straight leg raise)	**2.62** (0.73 to 9.32)	17 (16 to 45)	93 (86 to 99)
Reflexes (Patella tendon)	**2.49** (0.44 to 8.47)	21 (20 to 50)	93 (71 to 97)
Reflexes (Achilles tendon)	**3.03** (0.74 to 12.47)	15 (24 to 50)	95 (77 to 97)
Babinski	-	-	-
Sensation in feet		4 (6 to 41)	100 (93 to 100)
Ely's test	-	-	-
Pain provocation tests:			
Heel drop		9 (3 to 20)	95 (86 to 99)
Kibler's skin-roll test		17 (34 to 64)	82 (70 to 93)
Interspinosus ligament sensitivity	**2.97** (1.32 to 6.70)	68 (65 to 88)	58 (48 to 78)
Transverse process pressure		9 (10 to 39)	76 (76 to 99)
Sacro-iliac joint pressure		2 (5 to 24)	93 (83 to 99)

† Only those variables are shown that were entered in the model
‡ Robust estimate of variance
LBP = low back pain, CI = confidence interval

psoas) was the only one entered into the model and was 3-fold to 7-fold greater for the CLBPs and SLBPs than for the controls, with the proportion of positive test findings 67% in the CLBP group, 87% in the SLBP group, and 49% in the control group.

Among neurological and neurodynamic tests, slump, SLR, and sensation in the feet were those that differentiated between CLBPs and controls. Positive findings for CLBP were 42% in the slump test, 28% in SLR tests, 27% in sensation in the feet; respectively; for SLBP, 24%, 18%, and 5%, respectively; and for controls 11%, 7%, and 0%, respectively.

For pain provocation tests, many tests differentiated between CLBP patients and controls with high odds ratios (heel drop, Kibler's skin roll, interspinal ligament sensitivity and tranverse process pressure). Heel drop, Kibler's skin roll, and transverse process pressure differentiated between CLBP patients and SLBP as well. Positive findings occurred in CLBP for heel drop 25%, for Kibler's test 50%, for interspinosus ligament sensitivity 78%, and for tranverse process pressure (a unilateral PA) 24%, while for SLBP the corresponding figures were 10%, 17%, 68%, and 10%, respectively, and for controls 6%, 18%, 40%, and 5%, respectively.

Discussion

Non-specific LBP patients comprise a very heterogeneous group expressing many different diagnoses and functional problems. Assessment of severity or chronicity of LBP can only in part be based on findings in a clinical examination. In addition, when using diagnostic tests, other epidemiological and statistical variables should also be considered[3]. A lack of consensus among participating clinicians regarding duration-based subgroups and a lack of evidence for the validity of sub-grouping form a compelling argument for further research into this clinical practice.

In this study, each group had clearly differing lengths of LBP history. This allowed us to investigate the sensitivity of the clinical tests within the pain-duration-based groupings, a necessity to understanding the true effectiveness of a test[14,15]. Our results provide new and important information about the usefulness of the clinical tests for improving back pain physiotherapy. In our study, the positive test results were modest for general movement, like walking, undressing, and squatting, percentages ranging from 24% to 33% in the CLBP group. For lumbar spine flexion, specific AP mobility, tightness of hip flexors, slump, sensation in the feet, Kibler's, and interspinosus ligament sensitivity, percentages from 42% to 84% occurred in the CLBP group. The low likelihood of assort of these tests supports the notion that not all tests in common clinical use are equally valuable. This seems to be the case when the target group has the diagnosis "non-specific low back pain." These tests could, however, be valuable for differential diagnosis, which emphasizes the urgent need to develop clinical tests in order to choose proper treatment for a particular back problem.

As far as we know, no studies have revealed major differences in the clinical examination between chronic and subacute LBP patients and subjects without LBP patient status. One study has, however, evaluated the relationship between mechanical factors and LBP. Nourbakhsh and Arab[9] came to the conclusion that muscle endurance and weakness are associated with LBP. They also concluded that structural factors such as amount of lumbar lordosis, pelvic tilt, leg length discrepancy, nor length of abdominal, hamstring, and iliopsoas muscles are related to the occurrence of LBP. Conversely, we found that the iliopsoas was significantly tighter in the CLBP and SLBP groups than in the controls, and the size or shape of lumbar lordosis was abnormal in the CLBP and SLBP groups as well. However, the value of inspection and muscle tightness tests should be questioned because of their low reliability.

One commonly tested assessment for patients with chronic LBP is muscle strength, or endurance tests. We elected not to test these variables based on the abundance of earlier research on this subject[16-20]. Instead, we opted to investigate functional tests, which are routinely underexplored for patients with LBP. Ljungquist et al reported functional tests (gait and stair-climbing tests and two lifting tests) as sensitive to change in spinal pain[17]. In our study, difficulties in walking occurred only in subjects in the CLBP group.

Specific patterns of intervertebral motion and intravertebral deformation result in pain in chronic LBP patients, which substantiates their mechanical back pain etiology. The slump test, for instance, assesses the contribution of neural tissue to the referred symptoms associated with spinal pain, especially with discogenic pain[18]. In the present study, the slump was the only one of the neural and neurological tests with a high odds ratio and moderate percentage of positive tests. In addition, our positive findings in the lumbar flexion and extension tests, in specific PA mobility, and the tightness of the iliopsoas test were much greater in LBP patients than in controls. Further, interspinosus ligament sensitivity was a very common finding in our CLBP group. All these findings support a conclusion of mechanical pain etiology[19].

Pain-related factors contribute to muscle inhibition in patients with CLBP, which is especially related to lumbar flexion[21]. Dankaerts et al suggested that maladaptive movement or control impairment and associated faulty strategies result in chronic abnormal tissue loading (associated with either excessive or reduced spinal stability), pain, disability, and distress[22]. Whether impaired flexion is a result of pain or a reason for pain remains an open question.

Limitations

This study has shortcomings. The chronic LBP group may not be comparable to other chronic LBP groups, because we suspected that one subject might need a spinal operation (the general practitioner had wanted to consult an orthopedic surgeon). Second, the clinical findings were not compared with any objective findings. In addition, the non-specific LBP grouping was not determined by physical methods; instead, it was based on pain duration and the orthopedic consultation. Consequently, any comparison between certain objective methods and clinical tests

would have been unlikely to provide any additional information. In addition, the examiners did not know each patient or control's medical history and back pain history, x-ray findings, or treatments, a situation very common in clinical practice on the first visit, when treatment is based only on a clinical examination. In addition, although in the actual study 14 physiotherapists participated and intertester reliability and test-retest reproducibility of all clinical tests were analyzed in all test categories by only 2 physiotherapists specializing in OMT. All physiotherapists were taught by the OMT specialist, with 22 years of experience in this field, to perform the specific clinical tests in the same way.

Conclusions

Overall, a large number of the selected 31 clinical tests were sufficient to sort the CLBP from controls. Tests such as walking, lumbar spine flexion, tightness of hip flexor muscles, slump, Kibler's skin-roll, and the interspinosus ligament sensitivity test seem best in differentiating between these two groups. The best predictor for CLBP was the lumbar spine flexion test. SLBP seemed to differ from the control group in lumbar flexion, in a specific PA mobility test, and in tightness of hip flexor muscles. CLBP differed from SLBP in functional tests, sensation in the feet, and in the different pain provocation tests. Whether these tests are sufficiently sensitive to classify a more specific diagnostic or clinical subgroup remains untested, and further studies with clinical tests to differentiate among pathological conditions are necessary. In future we should also explore the interplay between biological, social, and psychosocial factors.

Acknowledgments

The authors would like to thank Professor Ari Heinonen for his critical comments, and Jukka-Pekka Kesonen, MSc, for his work in the statistical analysis of the data, both from the University of Jyvaskyla, Finland.

REFERENCES

1. Waddell G. The problem, social interactions, treatment: The scientific evidence. In Waddell G, ed. *The Back Pain Revolution*. Edinburgh, UK: Churchill Livingstone, 1998.
2. van Tulder M, Koes B, Bouter L. Conservative treatment of acute and chronic nonspecific low back pain: A systematic review of randomized controlled trials of the most common interventions. *Spine* 1997;22:2128–2156.
3. Nachemson A, Vingård E. Assessment of patients with neck and back pain: A best-evidence synthesis. In Nachemson A & Jonsson E, eds. *Neck and Back Pain: The Scientific Evidence of Causes, Diagnosis, and Treatment*. Philadelphia, PA: Lippincott Williams & Wilkins, 2000.
4. Waddell G, Feder G, Lewis M. Systematic reviews of bed rest and advice to stay active for acute low back pain. *Br J Gen Pract* 2001;47:647–652.
5. Dionne C, Dunn K, Croft P, et al. A consensus approach toward the standardization of back pain definitions for use in prevalence studies. *Spine* 2008;33:95–103.
6. Billis E, McCarthy C, Oldham J. Subclassification of low back pain: A cross-country comparison. *Eur Spine J* 2007;16:865–879.
7. Hides J, Richardson C, Jull G. Multifidus muscle recovery is not automatic after resolution of acute, first-episode low back pain. *Spine* 1996;23:2763–2769.
8. Hides J, Jull G, Richardson C. Long-term effects of specific stabilizing exercises for first-episode low back pain. *Spine* 2001;11:E243–E248.
9. Nourbakhsh M, Arab A. Relationship between mechanical factors and incidence of low back pain. *J Orthop Sports Phys Ther* 2002;32:447–460.
10. Pengel L, Refshauge K, Maher C. Responsiveness of pain, disability, and physical impairment outcome patients with low back pain. *Spine* 2004;29:879–893.
11. Linton S, Andersson T. Can chronic disability be prevented? A randomized trial of a cognitive-behavior intervention and two forms of information for patients with spinal pain. *Spine* 2000;21:2825–2831.
12. Bigos S. Acute low back pain in adults. In US Department of Health and Human Services, *Clinical Practice Guidelines*. AHCPR 95–0643. Philadelphia, PA: Williams & Wilkins, 1994.
13. Potter N, Rothstein J. Intertester reliability for selected clinical tests of the sacroiliac joint. *Phys Ther* 1985;11:1671–1675.
14. Nachemson A, Vingård E. Assessment of patients with neck and back pain: A best–evidence synthesis. In Nachemson A & Jonsson E, eds. *Neck and Back Pain: The Scientific Evidence of Causes, Diagnosis, and Treatment*. Philadelphia: Lippincott, Williams & Wilkins, 2000.
15. van Tulder M, Goossens M, Waddell G. Conservative treatment of chronic low back pain. In Nachemson A & Jonsson E, eds. *Neck and Back Pain: The Scientific Evidence of Causes, Diagnosis, and Treatment*. Philadelphia, PA: Lippincott, Williams & Wilkins, 2000.
16. Andersson G, Deyo R. History and physical examination in patients with herniated lumbar discs. *Spine* 1996;21:10S–18S.
17. Schenk P, Klipstein A, Spillmann S, et al. The role of back muscle endurance, maximum force, balance and trunk rotation control regarding lifting capacity. *Eur J Appl Physiol* 2006;96:146–156.
18. Ljungquist T, Nygren A, Jensen I, et al. Physical performance tests for people with spinal pain-sensitivity to change. *Disabil Rehabil* 2003;25:856–866.
19. Kerr JM, Vijnovich AL, Bradnam L. Changes in alpha-motoneuron excitability with positions that tension neural tissue. *Electromyogr Clin Neurophysiol* 2002;42:459–471.
20. Hicks G, Fritz J, Delitto A, et al. Interrater reliability of clinical examination measures for identification of lumbar segmental instability. *Arch Phys Med Rehabil* 2003;84:1858–1864.
21. Verbunt J, Seelen H, Vlaeyen J, et al. Pain-related factors contributing to muscle inhibition in patients with chronic low back pain: An experimental investigation based on superimposed electrical stimulation. *Clin J Pain* 2005;3:232–240.
22. Dankaerts W, O'Sullivan P, Straker L, et al. The inter-examiner reliability of a classification method for non-specific chronic low back pain patients with motor control impairment. *Man Ther* 2006;11:28–39.

ORIGINAL PAPERS

III

Inter-tester reliability in classifying acute and subacute low back pain patients into clinical subgroups: a comparison of specialists and non-specialists. A pilot study

By

Paatelma, M., Karvonen, E., & Heinonen, A.

The Journal of Manual and Manipulative Therapy 2010, 17, 221-229

Inter-tester Reliability in Classifying Acute and Subacute Low Back Pain Patients into Clinical Subgroups: A Comparison of Specialists and Non-Specialists. A Pilot Study

Markku Paatelma, MSc, PT[1]; **Eira Karvonen,** MSc, PT[2]; **Ari Heinonen,** PhD, PT[3]

ABSTRACT: Many systems have been suggested for classifying low back pain (LBP); the most commonly used among physiotherapists involves a pathoanatomical/pathophysiological tissue classification system. Few studies have examined whether this form of classification of LBP disorders can be performed in a reliable manner between specialists with advanced training, or between specialists with advanced training and non-specialists who lack advanced training. The purpose of this paper was to examine the inter-tester reliability of two specialists, and the ability of a specialist and non-specialist to independently classify patients with LBP, utilizing clinical tests and history-based classification methods after a short educational course on the classification system. Subjects were acute or sub-acute patients with LBP who visited their occupational healthcare or municipal healthcare center. Inter-tester reliability between the specialist and non-specialists was at almost the same level: overall Kappa 0.60 (95%CI; 0.40 to 0.85), overall agreement 70%, as between the two specialists: overall Kappa 0.65 (95%CI; 0.33-0.86), overall agreement 77%. The findings suggest that a short educational course can provide rather reliable examination tools to allow non-specialized physiotherapists to classify patients according to tissue origination.

KEYWORDS: Inter-tester Reliability, Low Back Pain Classification, Orthopedic Manual Therapy

Low back pain (LBP) has a lifetime prevalence of approximately 60 to 80%[1] and is recognized internationally as a major health, social, and economic burden[2]. Despite this fact, the current literature has not conclusively demonstrated a single, specific, most effective treatment method for LBP[3]. Explanations for this failure to identify a single effective treatment likely involve differing etiological factors and variations in the pathoanatomical/pathophysiological/tissue origins of LBP.

Identification of subgroups of LBP has been a focus of major research. Several authors have suggested that because non-specific low back pain (NSLBP) is a benign problem, emphasis should be on clinical tests and assessments; NSLBP should not be viewed as a homogenous condition, and treatment outcomes can be improved when sub-grouping is used to guide treatment decision-making[4-7]. In contrast, others have believed that subgrouping is only one of a number of possible explanations for the manifestations of NSLBP[8].

Although a number of LBP classification systems have been proposed, such as a pathoanatomical/pathophysiological classification system[9], the McKenzie classification[10], treatment-based classification[11], and the movement-impairment classification[12], what is still unclear is which clinical tests between two assessing clinicians are sufficiently reliable to allow subgroup categorization[13]. The reliability and validity of the overall classification systems has been tested[14-21] and has been reported as moderate or good. Nonetheless, reliability has only been shown to be effective in clinicians who receive advanced training. For example, the McKenzie system has been shown to be reliable in LBP sub-grouping classification only by suitably trained examiners and not by minimally trained or untrained assessors[16,17]. In contrast, in the movement-impairment classification studies by Dankaerts et al[20], patients were independently assessed by two "experts" and 13 physicians or physiotherapists participating in workshops or postgraduate training under supervision of the developer of the classification system. Neither length of training nor differences between experts in either comparison were described.

For comprehensive use in a number of clinical settings by multiple healthcare providers, the accuracy and repeatability of subgroup classification should be similar, not only between specialists but also between specialists and non-specialists working with musculoskeletal disorders. With the largest percentage of physiotherapists[22] and physicians[23] using a general pathoanatomical/pathophysiologi-

[1]OMT Instructor, University of Jyvaskyla, Jyvaskyla, Finland
[2]OMT, University of Jyvaskyla, Jyvaskyla, Finland
[3]Professor, University of Jyvaskyla, Department of Health Sciences, PO Box Fin-40014 Finland
Address all correspondence and requests for reprints to: Markku Paatelma, MSc, PT, markku.paatelma@auron.fi

cal classification system for LBP, this area is appropriate for investigation. Although the validity of the many proposed general pathoanatomical/pathophysiological classification systems has not yet been convincingly confirmed by objective methods, clinical experience suggests that even an idea of the origin of the symptoms may aid the therapist as to the best choice of treatment methods[24,25]. Therefore, the aim of this pilot study was to evaluate the inter-tester reliability of pathoanatomical/pathophysiological classification within a group of acute and sub-acute non-specific LBP patients. Reliability was examined between the patient findings of 1) general practitioners in primary care physiotherapy compared to findings classified by a specialist in orthopedic manual therapy (OMT); and 2) the findings of a physiotherapist specialist of OMT with multiple years of training compared to a physiotherapist with short-term experience as a specialist in OMT.

Methods

This pilot study was conducted in two parts. In Part 1, we compared the inter-tester reliability of LBP subgroup classification between a specialist in OMT and non-specialists, and in Part 2 between a physiotherapist specialist of OMT with multiple years of training compared to a physiotherapist with short-term experience as a specialist in OMT (Figure 1).

Participants

For the first part of the study, 21 eligible, consecutive, and voluntary patients with LBP with ages ranging from 18 to 56 years were recruited by an invitation letter from four municipal healthcare centers in central Finland. Additionally, 30 patients were recruited from a private occupational healthcare center (Medivire) in the city of Jyväskylä, Finland (Table 1, Figure 1). All patients who had visited their municipal or occupational healthcare center because of low back pain that had lasted less than 3 months were recruited. The inclusion criteria were 18- to 65-year-old individuals with current low back pain with or without radiating pain to one or both lower legs. The back pain episode could be the first or recurrent with the last episode lasting less than 3 months.

Exclusion criteria were pregnancy, use of psychogenic medications, and diagnosed osteoporosis. The history of present and former LBP episodes and of physical work was asked by a standardized questionnaire. Pain and symptom location were identified with a body chart completed by the patient. All patients showed interest in participating in this study and no one refused. The subjects provided written informed consent before the study and the local ethics committee approved the study protocol.

Procedure

In Part 1, all 21 patients at the healthcare centers were examined by one physiotherapist with 15 years of experience as an OMT and also by four physiotherapists who lacked an OMT specialization, each in his or her own municipal health care center, with the physiotherapists examining six, six, four, and five patients between each. These four non-specialist physiotherapists had a range of 4 to 12 years of clinical experience in physiotherapy and had completed brief post-graduate courses in musculo-skeletal physiotherapy. These physiotherapists were taught by the OMT specialist (EK) to perform the specific clinical tests in order to identify the LBP subgroup classification based on OMT practice. The physiotherapists were blinded to the results of the OMT specialist, who was also blinded to the results of the four non-specialists. A neutral observer supervised the study.

Before initiating the study, all PTs participated in five half-day training sessions, during which they performed the tests used for the classification. This included training on both patients and healthy subjects to standardize the tests by performing the tests according to written instructions and under guidance of the OMT specialist.

In Part 2, the 30 patients at Medivire were examined by two physiotherapists who specialized in OMT, one with 20 years of experience in OMT (MP) and one with 2 years of experience in OMT (JR). MP had five years of clinical experience in the field of musculoskeletal physiotherapy prior to specialization in OMT, and JR had six. Examiners were blinded to each other`s results and the tests were supervised by a neutral observer. OMT specialization in Finland requires 3.5 years of training and includes a post-graduate examination supervised by the Finnish Association of Physiotherapists.

In Parts 1 and 2 of the study, clinical assessments included 50 different tests (Table 2). Each patient assessment session lasted 30 minutes and the data were collected on a lumbar spine-assessment form. This assessment comprised an LBP history, observation of the posture of the low back and lower extremities, function of the lower back and lower extremities, stability tests for lumbar spine and pelvic girdle, specific pain provocation and alleviation tests, mobility tests of the lower back and sacroiliac joints, neurological and neurodynamic tests, and tests for muscle tightness. In all the tests, the decision was binary: the test was either negative (normal finding) or positive (pathological finding). Based on the tests and a defined clinical reasoning process, the patients were classified into one of five mutually exclusive clinical LBP pathoanatomical/pathophysiolocial subgroups:

1. Discogenic pain
2. Clinical lumbar instability
3. Clinical lumbar spinal stenosis
4. Segmental dysfunction/facet pain
5. Sacroiliac joint pain/dysfunction

One patient with a post-operative spine could not be classified by either of two examiners into any of the subgroups and was therefore excluded from statistical analysis. Thus, the number of analyzed patient groups was 20 in Part 1 and 30 in Part 2. The inter- and intra-tester reliability, sensitivity, and specificity of the tests used were evaluated before the testing[26].

Subgroup Classification

The clinical reasoning process by which the physiotherapist reached a classification involved a dedicated deductive

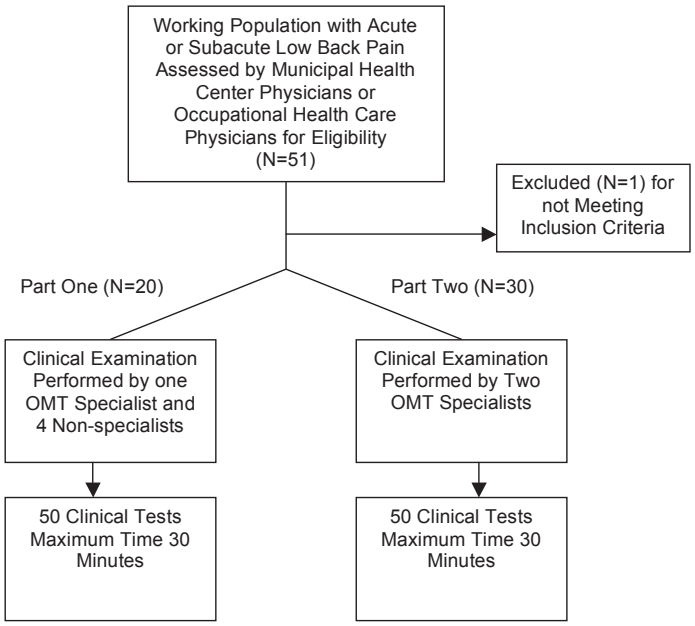


FIGURE 1. Flow chart of subjects during study.

TABLE 1. Characteristics of subjects in subgroup classification between a specialist in OMT and non-specialists, and between two specialists in OMT (an expert vs. a novice).

	OMT specialist vs. non-specialists[1] (N=20)	OMT expert vs. OMT novice[2] (N=30)
Age (years, mean, SD)	40.0 (11.5)	37.9 (4.5)
Gender (female/male)	65/35	74/26
History of present LBP episode		
Acute: < 6 weeks	35	27
Subacute: > 6 weeks <12 weeks	65	73
Chronic: >12 weeks	0	0
On sick-leave because of LBP	0	0
Pain and symptom location		
VAS from 0 to 10 over last 24 hrs (mean, SD):	3 (3.5)	3 (2.9)
Signs of pain drawing in low back	42	45
Signs of pain drawing in lower leg	58	55
Physical work		
Light (%)	65	67
Heavy (%)	35	33

[1]Municipal Health Care (Part 1)
[2]Occupational Health Care (Part 2)
Note: Values are percentages unless stated otherwise.

process. Briefly, *discogenic pain without nerve root irritation* was the diagnosis when a patient's pain (local or referred) could be provoked in modified slump test and when movement into extension was less painful or alleviated the same pain (centralized). *Discogenic pain with nerve root irritation* was recorded when radiating pain was provoked by nerve tension tests (SLR) and by other neurodynamic tests.

The construct for *clinical lumbar instability* involved assessment of three interdependent components: the passive, the active, and the neuromuscular subsystem[27]. Clinical lumbar instability was recorded when the patient reported low back pain or fatigue or both during prolonged sitting/standing/lying down, and when pain during extension was relieved and movement increased with traction. In addition, the classification was made if there were difficulties in a one-leg stance or active straight-leg raise

TABLE 2. Clinical Tests and Findings used in Classification.

Clinical Assessment	Finding
Functions	Normal walking, heel walking, and walking on toes, undressing, squat and rise
Inspection	posture of spine in standing, knees, and feet length difference of lower limbs
Mobility	lumbar spine flexion, extension, lateral flexion (right and left) specific Posterior-Anterior mobility of T12-S1 hip rotation left and right
Pain Provocation	lumbar spine extension with traction from max extension left to max flexion right from max extension right to max flexion left posterior pelvic pain provocation right and left interspinal ligament provocation Kibler`s skin rolling sacroiliac joint provocation right and left L4, L5 rotation provocation right and left
Muscle tightness	hamstrings right and left piriformis right and left gluteus med/min right and left iliopsoas right and left
Stability	one-leg standing right and left active SLR right and left isometric lumbar extension transverse abdominis activity
Neurology and neurodynamics	SLUMP in sitting SLR right and left patella reflex right and left Achilles reflex right and left Ely`s test right and left

(ASLR) or both, or inability to activate either transverse abdominus or lumbar multifidi combined with local interspinal pain, or a combination of these problems.

Clinical central spinal stenosis was recorded when the patient reported a clear pattern of intermittent claudication provoked by extension, which was relieved by sitting or flexed spinal posture. Symptoms and signs could be combined with tightness of hip flexors or positive sciatic or a femoral nerve tension test or several. *Clinical lateral spinal stenosis* was recorded when the patient reported radiating pain with nerve tension tests and during extension/lateral flexion toward the symptomatic side or during transverse process provocation (passive foramina approach), or both.

Segmental dysfunction/facet pain was recorded when pain and movement restrictions were identified during physiological movements in standing and painful hypomobility while lying prone, whereas sacroiliac joint pain/dysfunction was recorded if the patient's pain was provoked while standing on one leg and relieved with a sacroiliac joint belt, or provoked with sacral thrust and/or during posterior pelvic pain provocation (PPPP) or both, or if pain and difficulties occurred during an ASLR.

Examination Techniques

To improve reliability in the examination, all tests were standardized with operational definitions for their use and interpretation, and was taught comprehensively and trained by an OMT specialist prior to the study. In our study, the examination techniques of concern were based on provocation and alleviation techniques[28], sacroiliac joint provocation[29-32], and neurodynamic tests[33,34].

Statistical Analysis

Statistical analyses were performed with SPSS software, version 14.0. Percentage agreement and the kappa statistic served to test inter-tester agreement in the choice of subgroup classification. Clinical agreement was recorded when the two subgroup classifications were exactly the same.

Overall inter-tester agreement and overall kappa coefficient were calculated first using a 2x2 contingency table. Then the prevalence of positive observations and percentages of agreement for all categories and kappa in discogenic pain and clinical instability were calculated.

The kappa statistic estimates the degree of agreement corrected for chance agreement[35]. There is general agreement that for physical therapists, kappa a preferred statistic for estimating the accuracy of nominal and ordinal data in clinical research[36]. The percentage agreement does not take into account agreement due solely to chance[37].

The classification system proposed by Landis and Koch[38] allowed determination of the level of kappa as follows: poor: smaller than zero; slight: zero to 0.20; fair: 0.21 to 0.40; moderate: 0.41 to 0.60; substantial: 0.61 to 0.80; almost perfect: 0.81 to 1.00. Clinical relevance was considered in this study as a kappa of 0.41[38] and the percentage of agreement at 70%[39].

Results

For Part 1, which involved a municipal healthcare center, LBP subgroup prevalence was as follows: clinical instability and discogenic pain were the most common (35% and 30%), followed by segmental dysfunction/facet pain (18%), sacroiliac joint pain/dysfunction (10%), and clinical spinal stenosis (7%). In Part 2, which involved an occupational healthcare setting, clinical instability and discogenic pain were also the most

frequent subgroups (43% and 37%), followed by segmental dysfunction/facet pain (7%), sacroiliac joint pain/dysfunction (7%), and clinical spinal stenosis (6%) (Tables 3 and 4).

Percentage of agreement ranged from 75% to 100% between the experienced physiotherapist with OMT specialization and the four physiotherapists without OMT specialization. Overall inter-tester agreement was 70% and the overall kappa coefficient was 0.60 (95% CI; 0.40 to 0.85) (Table 3). The prevalence of positive observations and percentages of agreement for all categories and the kappa in discogenic pain and clinical instability are presented in Table 5.

Overall agreement between these OMT specialists for Part 2 was 77% with an overall kappa of 0.65 (95% CI; 0.33 to 0.86) (Table 4). Table 6 gives the prevalence of positive observations and percentages of agreement for all the categories and the kappa in discogenic pain and clinical instability. Percentages of agreement ranged from 75% to 100% (Table 6).

Discussion

This pilot study demonstrated that inter-tester reliability of categorization of LBP subgroups between an experienced physiotherapist in OMT and four physiotherapists without OMT specialization and then between two clinicians with OMT training, one experienced and one inexperienced, was acceptable. Comparison with other inter-tester reliability studies is difficult because in the majority of comparative studies, the study group consisted of chronic LBP patients[13,14,19] or the length of LBP was not reported[16-18,20]. However, at least two studies support our findings that a short training period in the workplace is effective. Fritz et al[40]found acceptable overall agreement on classification of a decision-making algorithm using physical therapists with varying levels of experience, identifying no significant differences based on level of experience. When clinicians were newly trained in a classification system for acute LBP during a one-day course, that classification showed moderate reliability[41].

Worth noting is that the reliability of discogenic and sacroiliac joint pain/dysfunction clinical tests varies from fair to good, but the reliability of tests for segmental dysfunction/facet pain is poor[42]. This may also explain, at least partly, the low prevalence of this subgroup. The tests for clinical lumbar instability also vary from poor to good[43,44]. At present, there are no reliability studies that exist for detecting clinical central or lateral stenosis. Only a self-reported history questionnaire has proven successful as a diagnostic tool for lumbar spinal stenosis[45].

In the present pilot study, the number of patients was small, which thus reduces the generalizability of the findings. Prevalence of some of the subgroups was also low, for example, sacroiliac joint pain/dysfunction and clinical spinal stenosis, and thus calculation of kappa was not possible. Despite this limitation, a notable strength was that the patients were recruited from routine referrals, and because the care was provided for free, cost containment did not limit physiotherapy examination. Although the time-limited examination was 30 minutes, some LBP patients may require longer assessment or even several assessments, with their responses to specific interventions and applications of home-exercises contributing to the final subgroup categorization[46].

Concepts of diagnosis and classification of LBP have a long history in medicine, whereas formal schemes of diagnostic classification in physical therapy are relatively new[47]. Because of this short history, some questions concerning the validity of the LBP classification are still prevalent. In addition, the few high-quality studies that exist demonstrate either conflicting evidence or

TABLE 3. Cross-tabulation of diagnostic groups between an experienced physiotherapist (15 years) in OMT shown vertically, and horizontally figures by four physiotherapists without OMT specialization. Cases (%). Number of subjects 20.

	Four physiotherapists without OMT specialization					
Experienced OMT-physiotherapist diagnosis	Discogenic pain	Clinical instability	Clinical spinal/ lateral stenosis	Segmental dysfunction/ facet pain	Sacroiliac & pelvic pain	Total
Discogenic pain	**5 (25)**	1 (5)	0 (0)	1 (5)	0 (0.0)	7 (35)
Clinical instability	0 (0)	**5 (24)**	0 (0)	1 (5)	0 (0.0)	6 (30)
Clinical spinal/lateral stenosis	0 (0)	0 (0)	**1 (5)**	1 (5)	0 (0.0)	2 (10)
Segmental dysfunction/facet pain	0 (0)	2 (10)	0 (0)	**1 (5)**	0 (0.0)	3 (15)
Sacroiliac & pelvic pain	0 (0)	0 (0)	0 (0)	0 (0)	**2 (10)**	2 (7)
Total	5 (25)	8 (40)	1 (5)	4 (20)	2 (10)	20 (100)
Agreement %	70					
Kappa (95% CI)	0.60 (0.40 to 0.85)					

Bold represents counts for agreement.

TABLE 4. Cross-tabulation of diagnostic groups between two physiotherapists specialized in OMT: one experienced (20 years) in OMT shown vertically, and a physiotherapist with less experience (shown horizontally) (2 years). Cases (%). Number of subjects 30.

	Novice physiotherapist in OMT					
Experienced OMT-physiotherapist	Discogenic pain	Clinical instability	Clinical spinal/ lateral stenosis	Segmental/ dysfunction facet pain	Sacroiliac & pelvic pain	Total
Discogenic pain	**9 (30)**	0 (0)	0 (0)	1 (3)	0 (0)	10 (33)
Clinical instability	3 (10)	**10 (33)**	0 (0)	0 (0)	0 (0)	13 (43)
Clinical spinal/lateral stenosis	0 (0)	0 (0)	**2 (7)**	0 (0)	0 (0)	2 (7)
Segmental dysfunction /facet pain	0 (0)	3 (10)	0 (0)	**0 (0)**	0 (0)	3 (10)
Sacroiliac & pelvic pain	0 (0)	0 (0)	0 (0)	0 (0)	**2 (7)**	2 (7)
Total	12 (40)	13 (43)	2 (7)	1 (3)	2 (7)	30 (100)
Agreement %	77					
Kappa (95 % CI)	0.65 (0.33 to 0.86)					

Bold represents counts for agreement.

TABLE 5. Inter-examiner reliability in assessment of low back pain subgroup classification between a physiotherapist specialized in OMT and by four physiotherapists without OMT specialization. Number of subjects 20.

	Number of positive observations by examiner					
LBP classification	Experienced Examiner	Inexperienced Examiners	Agree[1]	Disagree[2]	Agreement %	Kappa (95 % CI)
Discogenic pain	7	5	18	2	90	0.76 (0.35 to 1.00)
Clinical instability	6	8	16	4	80	0.57 (0.14 to 0.90)
Clinical spinal/ lateral stenosis	2	1	19	1	95	Φ
Segmental dysfunction/ facet pain	3	4	15	5	75	Φ
Sacroiliac & pelvic pain	2	2	20	0	100	Φ

[1]Examiners agreed that a given number of patients had the classification.
[2]Examiners disagreed whether patients had the classification.
Φ=calculation of Kappa impossible, due to low number of subjects in subgroups.

TABLE 6. Inter-examiner reliability in assessment of low back pain subgroup classification between two physiotherapists specialized in OMT (expert vs. novice). Number of subjects 30.

	Numbser of positive observations by examiner					
LBP classification	Expert	Novice	Agree[1]	Disagree[2]	Agreement %	Kappa (95 % CI)
Discogenic pain	10	12	26	4	87	0.71 (0.40 to 0.93)
Clinical instability	13	13	24	6	80	0.59 (0.28 to 0.86)
Clinical spinal/ lateral stenosis	2	2	30	0	100	Φ
Segmental dysfunction/ facet pain	3	1	26	4	75	Φ
Sacroiliac & pelvic pain	2	2	30	0	100	Φ

[1]Examiners agreed that a given number of patients had the classification.
[2]Examiners disagreed whether patients had the classification.
Φ=calculation of Kappa impossible, due to low number of subjects in subgroups.

only moderate evidence of LBP classification reliability[48].

In the present study, the inter-tester reliability was good, particularly in the clinical instability and discogenic pain subgroups in which most of our patients were classified. Although instability is fraught with conceptual and terminological difficulties without widely used accepted clinical diagnostic criteria, it is the most common subgroup in clinical physiotherapy praxis[49]. Discogenic pain was also common, in accordance with many findings[50-52].

In addition to the low prevalence of given subgroups, the low mean age of the patients may explain why only 7% were classified into a clinical spinal stenosis subgroup, a finding in line with the study of Leinonen et al[53]. Furthermore, although research has shown that pain arising from the facet joint is difficult to characterize by clinical examination variables[9,54], we concluded that 11% of patients were classified into this subgroup. This agrees with a primary subgroup classification, that is, with sacroiliac joint pain/dysfunction appearing in only 8%. In the literature, prevalence of sacroiliac joint pain varies but is routinely consider to be around 9% to 20%[27,30,55].

The type of pathoanatomical/pathophysiological/tissue origin classification might be very useful in cases of acute and subacute pain, because the examination is non-invasive and widely available. If no major trauma or suspected malignancy exists, invasive diagnostic techniques are not recommended in acute and subacute cases, techniques that may differ greatly among chronic cases. Classification in chronic LBP is often even more complicated, with its high levels of distress and disability. Psychological factors consistent with fear-avoidance models are associated with the development of chronic LBP. In addition, supplementing behavioral treatment options by treatment-based classification (TBC) physical therapy intervention for acute and sub-acute LBP patients was shown to be ineffective for improving important outcomes related to development of chronic LBP[56].

European clinical guidelines for LBP recommend early referral of appropriate patients to health services such as physiotherapy. Casserley-Feeny et al showed significantly higher percentages of acute LBP patients in the private setting than in the public setting; they also found longer wait times and a higher number and longer duration of physiotherapy treatments in the public setting, suggesting the need to develop primary healthcare in the aim of preventing acute LBP from becoming chronic[57]. Our classification was used previously by Paatelma et al in a randomized controlled trial that compared OMT and the McKenzie Method with advice only for LBP treatment[58], and included patients from a full spectrum of chronicity. Whether this kind of classification used in our study would improve the efficacy of LBP pain treatment compared to a classification based on time-duration of pain[59-61] or diagnostic triage[62] is unknown. Still, although empirical clinical evidence has shown the efficacy of physiotherapy for acute first-episode low back disorders, it has not shown the same efficacy in patients with chronic low back disorders[63]. Whether classification might enhance the efficacy in all subgroups is also unknown.

Conclusions

The results of this study suggest that inter-tester reliability in classifying patients with LBP into the clinical subgroups seems to be moderate or high regardless of experience level of the OMT-specialist physiotherapist or whether the physiotherapist had a background in OMT principles. This clinically relevant and clearly defined pain pattern system uses key elements of the history and examination to classify patients with low back pain. However, larger trials using those tests in every subgroup, including those which have high odd ratios, are necessary before we can make general statements about the reliability of subgroup classification in the early stage of low back pain.

REFERENCES

1. 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;23:2616–2624.
2. Hemmilä HM. Quality of life and cost of care of back pain patients in Finnish general practice. *Spine* 2002;27:647–653.
3. Cohen SP, Argoff CE, Carragee EJ. Management of low back pain. *BMJ* 2008;337:a2718 DOI: 10.1136/bmj.a2718.
4. McKenzie R, May S. *Mechanical Diagnosis and Therapy: The Lumbar Spine.* Waikanae, New Zealand: Spinal Publications, 2003.
5. O'Sullivan P. Diagnosis and classification of chronic low back pain disorders: Maladaptive movement and motor control impairments as underlying mechanism. *Man Ther* 2005;10:242–255.
6. Fritz JM, Lindsay W, Matheson JW, et al. Is there a subgroup of patients with low back pain likely to benefit from mechanical traction? Results of a randomized clinical trial and subgrouping analysis. *Spine* 2007;32:E793–E800.
7. 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; 31:623–631.
8. Wand BM, O'Connell NE. Chronic non-specific low back pain: Sub-groups or a single mechanism? *BMC Musculoskelet Disord* 2008;9:11.
9. Coste J, Paolaggi JB, Spira A. Classification of nonspecific low back pain. II. Clinical diversity of organic forms. *Spine*. 1992;17: 1038–1042.
10. McKenzie R. *The Lumbar Spine: Mechanical Diagnosis and Therapy*. Waikanae, New Zealand: Spinal Publications, 1981.
11. Delitto A, Erhard RE, Bowling RW. A treatment-based classification approach to low-back syndrome: Identifying and staging patients for conservative treatment. *Phys Ther* 1995;75:470–485.
12. Sahrmann S. *Diagnosis and Treatment of Movement Impairment Syndromes*. St. Louis, MO: Mosby, 2002.
13. McCarthy CJ, Gittins M, Roberts C, Oldham JA. The reliability of the clinical tests and questions recommended in international guidelines for low back pain. *Spine* 2007;32:921–926.
14. Petersen T, Olsen S, Laslett M, et al. Inter-tester reliability of a new diagnostic classifi-

cation system for patients with non-specific low back pain. *Aus J Physiother* 2004;50:85–94.

15. Laslett M, McDonald B, Tropp H, Aprill C, Oberg B. Agreement between diagnoses reached by clinical examination and available reference standards: A prospective study of 216 patients with lumbopelvic pain. *BMC Musculoskelet Disord* 2005;6:28.
16. George S, Delitto **A.** Clinical examination variables discriminate among treatment-based classification groups: A study of construct validity in patients with acute low back pain. *Phys Ther* 2005;85:306–314.
17. Kilpikoski S, Airaksinen O, Kankaanpää M, Leminen P, Videman T, Alen M. **Interexam**iner reliability of low back pain assessment using the McKenzie method. *Spine* 2002;8:E207–E214.
18. Clare HA, Adams R, Maher CG. **Reliabil**ity of McKenzie classification of patients with cervical or lumbar pain. *J Manipulative Physiol Ther* 2005;28:122–127.
19. Van Dillen L, Sahrmann S, Norton B, Caldwell C, McDonnell M, Bloom N. Movement system impairment-based categories for low back pain: Stage 1 validation. *J Orthop Sports Phys Ther* 2003;33:126–142.
20. Dankaerts W, O'Sullivan P, Straker L, Burnett A, Skouen J. **The inter-examiner reliab**ility of a classification method for non-specific chronic low back pain patients with motor control impairment. *Man Ther* 2006;11:28–39.
21. Luomajoki H, Kool J, de Bruin ED, Airaksinen O. Reliability of movement control tests in the lumbar spine. *BMC Musculoskelet Disord* 2007;8:90.
22. Spoto MM, Collins J. **Physiotherapy diag**nosis in clinical practice: A survey of orthopaedic certified specialists in the USA. *Physiother Res Int* 2008;13:31–41.
23. Billis EV, McCarthy CJ, Oldham JA. Subclassification of low back pain: A cross-country comparison. *Eur Spine J* 2007;16:865–879.
24. 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;31:623–631.
25. Rothwell PM. Treating individuals 2. Subgroup analysis in randomised controlled trials: Importance, indications, and interpretation. *Lancet* 2005;365:176–186.
26. Paatelma M, Karvonen E, Heiskanen J. How do clinical test results differentiate chronic and subacute low back pain patients from "non-patients"? *J Man Manip Ther* 2009;17:11–19.
27. Demoulin C, Distrée V, Tomasella M, Crielaard JM, Vanderthommen M. Lumbar functional instability: A critical appraisal of the literature. *Ann Readapt Med Phys* 2007;50:677–684, 669–676.
28. Kaltenborn F, Evjenth O, Kaltenborn T, Morgan D, Vollowitz E. *The Spine. Manual Mobilization of the Joints.* Vol II. Oslo, Norway: Norli, 2003.
29. Laslett M, Williams M. The reliability of selected pain provocation tests for sacroiliac joint pathology. *Spine* 1994;19:1243–1249.
30. Laslett M. Evidence-based diagnosis and treatment of the painful sacroiliac joint. *J Man Manip Ther* 2008;16:142–152.
31. Mens J, Vleeming A, Snijders C, Koes B, Stam H. Validity of active straight leg raise test for measuring disease severity in patients with posterior pelvic pain after pregnancy. *Spine* 2002;27:196–200.
32. van der Wurff P, Buijs E, Groen G. A multitest regimen of pain provocation tests as an aid to reduce unnecessary minimally invasive sacroiliac joint procedures. *Arch Phys Med Rehabil* 2006;87:10–14.
33. Butler D. *Mobilisation of the Nervous System*. 1st ed. Edinburgh, UK: Churchill Livingstone, 1991.
34. Shacklock M. *Clinical Neurodynamics: A New System of Musculoskeletal Treatment*. 1st ed. Edinburgh, UK: Elsevier, 2005.
35. Kramer M, Feinstein A. Clinical biostatistics LIV: The biostatistics of concordance. *Clin Pharmcol Ther* 1981;29:111–123.
36. Haley S, Osberg J. Kappa coefficient calculation using multiple ratings per subject: A special communication. *Phys Ther* 1989; 69:970–974.
37. Hunt R. Percentage agreement, Pearson`s correlation, and Kappa as measures of inter-examiner reliability. *J Dent Res* 1986;65:128–130.
38. Landis J, Koch G. The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159–174.
39. Potter N, Rothstein J. Intertester reliability for selected clinical tests of the sacroiliac joint. *Phys Ther* 1985;65:1671–1675.
40. Fritz J, Brennan G, Clifford S, Hunter S, Thackeray A. An examination of the reliability of a classification algorithm for subgrouping patients with low back pain. *Spine* 2006;31:77–82.
41. Heiss D, Fitch D, Fritz J, Sanchez W, Roberts K, Buford J. The interrater reliability among physical therapists newly trained in a classification system for acute low back pain. *J Orthop Sports Phys Ther* 2004;34:430–439.
42. Hancock MJ, Maher CG, Latimer J, et al. Systematic review of tests to identify the disc, SIJ or facet joint as the source of low back pain. *Eur Spine J* 2007;16:1539–1550.
43. Abbott JH, McCane B, Herbison P, Moginie G, Chapple C, Hogarty T. Lumbar segmental instability: A criterion-related validity study of manual therapy assessment. *BMC Musculoskelet Disord* 2005;6:56.
44. Kasai Y, Morishita K, Kawakita E, Kondo T, Uchida A. A new evaluation method for lumbar spinal instability: Passive lumbar extension test. *Phys Ther* 2006;86:1661–1667.
45. Konno S, Kikuchi S, Tanaka Y, et al. **A diag**nostic support tool for lumbar spinal stenosis: A self-administered, self-reported history questionnaire. *BMC Musculoskelet Disord* 2007;30;8:102.
46. Werneke M, Hart D. Categorizing patients with occupational low back pain by use of the Quebec Task Force Classification system versus pain pattern classification procedures: Discriminant and predictive validity. *Phys Ther* 2004;84:243–254.
47. Zimny N. Diagnostic classification and orthopaedic physical therapy practice: What we can learn from medicine. *J Orthop Sports Phys Ther* 2004;34:105–109.
48. May S, Littlewood C, Bishop A. Reliability of procedures used in the physical examination of non-spesific low back pain: A systematic review. *Aust J Physiot* 2006;52:91–102.
49. Barr K, Griggs M, Cadby T. Lumbar stabilization: Core concepts and current literature, Part 1. *Am J Phys Med Rehabil* 2005; 84:473–480.
50. Derby R, Lee S, Kim B, Chen Y, Aprill C, Bogduk N. Pressure-controlled lumbar discography in volunteers without low back symptoms. *Pain Med* 2005;6:213–221.
51. Schwarzer A, Aprill C, Derby R, Fortin J, Kine G, Bogduk N. The prevalence and clinical features of internal disc disruption in patients with chronic low back pain. *Spine* 1995;20:1878–1883.
52. Freemont AJ, Peacock TE, Goupille P, Hoyland JA, O`Brien J, Jayson MI. Nerve ingrowth into diseased intervertebral disc in chronic back pain. *Lancet* 1997;350:178–181.
53. Leinonen V, Maatta S, Taimela S, et al. **Im-**

paired lumbar movement perception in association with postural stability and motor- and somatosensory-evoked potentials in lumbar spinal stenosis. *Spine* 2002;27:975–983.

54. Manchikanti L, Manchukonda R, Pampati V, Damron K, McManus C. Prevalence of facet joint pain in chronic low back pain in postsurgical patients by controlled comparative local anesthetic blocks. *Arch Phys Med Rehabil* 2007;88:449–455.
55. Riddle D, Freburger J. Evaluation of the presence of sacroiliac joint region dysfunction using a combination tests: A multicenter intertester reliability study. *Phys Ther* 2002;82:772–781.
56. George SZ, Zeppieri G Jr, Cere AL, et al. A randomized trial of behavioral physical therapy interventions for acute and subacute low back pain. *Pain* 2008;140:145–157.
57. Casserley-Feeny S, Bury G, Daly L, Hurley D. Physiotherapy for low back pain: Differences between public and private healthcare sectors in Ireland. A retrospective survey. *Man Ther* 2008;13:441–449.
58. Paatelma M, Kilpikoski S, Simonen R, Heinonen A, Alen M, Videman T. Orthopaedic manual therapy, McKenzie method or advice only for low back pain in working adults: A randomized controlled trial with one year follow-up. *J Rehabil Med* 2008;40:858–863.
59. Spitzer WO. Low back pain in the workplace: Attainable benefits not attained. *Br J Ind Med* 1993;50:385–388.
60. Merskey H, Bogduk N. *Classification of Chronic Pain.* Seattle, WA: IASP Press, 1994.
61. Bogduk N, McGuirk B. *Medical Management of Acute and Chronic Low Back Pain: An Evidence-based Approach.* Edinburgh, UK: Elsevier, 2002.
62. Waddell G. Subgroups within "nonspecific" low back pain. *J Rheumatol* 2005;32:395–396.
63. Hides J, Jull G, Richardson C. Long-term effects of specific stabilizing exercises for first-episode low back pain. *Spine* 2001;26:E243–248.

ORIGINAL PAPERS

IV

Orthopaedic manual therapy, McKenzie-method or advice only for low back pain in working adults: a randomized controlled trial with one year follow-up

By

Paatelma, M., Kilpikoski, S., Simonen, R, Heinonen, A., Alèn, M. & Videman, T.

Journal of Rehabilitation Medicine 2008, 40: 858-863

J Rehabil Med 2008; 40: 858–863

ORIGINAL REPORT

ORTHOPAEDIC MANUAL THERAPY, MCKENZIE METHOD OR ADVICE ONLY FOR LOW BACK PAIN IN WORKING ADULTS: A RANDOMIZED CONTROLLED TRIAL WITH ONE YEAR FOLLOW-UP

Markku Paatelma, MSc, PT[1], Sinikka Kilpikoski, MSc, PT[1], Riitta Simonen PhD, PT[2], Ari Heinonen, PhD, PT[1], Markku Alen, MD, PhD[3,4] and Tapio Videman, MD, PhD[5]

From the [1]Department of Health Sciences, University of Jyväskylä, Jyväskylä, [2]Finnish Institute of Occupational Health, Helsinki, [3]Department of Medical Rehabilitation, Oulu University Hospital, [4]Institute of Health Sciences, University of Oulu, Oulu, Finland and [5]Faculty of Rehabilitatin Medicine, University of Alberta, Edmonton, Canada

Objective: **To examine the effects of 2 manual therapy methods compared with one counselling session with a physiotherapist with "advice-only to stay active" for treating low back pain/leg pain and disability.**

Design: **A randomized, controlled trial with a 1-year follow-up.**

Subjects: **A total of 134 subjects with low back disorders.**

Methods: **Participants with acute to chronic first or recurrent low back pain, excluding those with "red flag" criteria, were assigned randomly to one of the 3 intervention groups: an orthopaedic manual therapy group ($n=45$), a McKenzie method group ($n=52$), and an "advice only to be active" group (advice-only) ($n=37$). Data on leg and low back pain intensity and disability (Roland-Morris Disability questionnaire) were collected at baseline, and at 3-, 6-, and 12-month follow-up points.**

Results: **At the 3-month follow-up point, significant improvements had occurred in all groups in leg and low back pain and in the disability index, but with no significant differences between the groups. At the 6-month follow-up, leg pain (–15 mm; 95% confidence interval (CI) –30 to –1), back pain (effect: –15 mm; –27 to –4), and disability index (–4 points; –7 to –1) improved ($p<0.05$) more in the McKenzie method group than in the advice-only group. At the 1-year follow-up, the McKenzie method group had ($p=0.028$) a better disability index (–3 points; –6 to 0) than did the advice-only group. In the orthopaedic manual therapy group at the 6-month and 1-year follow-up visits, improvements in the pain and disability index were somewhat better than in the advice-only group ($p=0.067$ and 0.068, respectively). No differences emerged between the orthopaedic manual therapy and McKenzie method groups in pain- and disability-score changes at any follow-up.**

Conclusion: **The orthopaedic manual therapy and McKenzie methods seemed to be only marginally more effective than was one session of assessment and advice-only.**

Key words: educational booklet, McKenzie method, orthopaedic manual therapy, low back pain, back pain.

J Rehabil Med 2008; 40: 858–863

Correspondence address: Markku Paatelma, Härkälahdenkuja 9, FI-00850 Helsinki, Finland. E-mail: markku.paatelma@tcr.fi

Submitted May 31, 2007; accepted June 12, 2008

INTRODUCTION

Back pain is extremely common all over the industrial world, with its high prevalence leading to personal and socioeconomic consequences. For example, in Finland, 33% of women and 29% of men reported having had low back pain (LBP) in the previous month, and 11% of women and 10% of men reported being diagnosed or treated by a physician for LBP in the previous year. Moreover, the prevalence of LBP has not changed over the last 30 years (1).

The European guidelines for management of LBP recommend the treatment of acute non-specific LBP (less than 6 weeks' duration) with advice to "stay active and continue normal daily activities including work if possible". They also recommend "considering spinal manipulation for those who are failing to return to normal activities" (2). For non-specific chronic (more than 12 weeks' duration) LBP in conjunction with supervised exercise therapy "short courses of manipulative therapy can also be considered" (3).

One systematic review of the best synthesis of efficacy of manipulative therapy for LBP showed manual therapy providing either similar or better pain outcomes in the short- and long-term than did placebo or other treatments (4). Another systematic review of the McKenzie method showed that the approach resulted in a greater decrease in pain and disability in the short term than did standard therapies (5). One assessment by a physician and a physiotherapist compared with patients` continuation of daily activity as tolerated proved equally effective in recovery from LBP symptoms (6). In addition, intensive physiotherapy in combination with a neurophysiology education component was effective among patients with chronic LBP (7). One subgroup of patients with LBP with severe disability not responding to conservative treatment were recommended to be treated with a multidisciplinary approach (8). Moreover, the UK Back pain and Exercise And Manipulation (BEAM) Trial study for the effectiveness of physical treatments for back pain in primary care concluded that regarding "best care" in general practice, manipulation followed by exercise achieved a moderate benefit at 3 months and a small benefit at 12 months (9).

A number of conservative treatment methods have been studied for LBP, but controversy remains as to the most ef-

fective. Despite those promising studies, no data currently have compared orthopaedic manual therapy (OMT) or the McKenzie method with advice to stay active in subjects with acute to chronic LBP. The aim of this study was to compare the effectiveness of OMT, the McKenzie method, and advice-only to stay active (advice-only) for low back and leg pain intensity and disability among a working-age population.

MATERIAL AND METHODS

Procedure

Participants were recruited from 4 occupational health care centres in the city of Jyväskylä, Finland. Occupational physicians were instructed to identify eligible subjects. Everyone who visited the occupational health care centre because of low back trouble and fulfilled the inclusion criteria was recruited. Those patients who visited their occupational physicians 0–7 days after their last episode of LBP had started commenced treatment on day 8, and latest on day 14. Participants completed the questionnaires, were assessed physically, and randomized into the study groups. Outcome measures, which included a battery of self-reported measures (use of healthcare services due to other problems and other back pain treatments) were assessed at 3-, 6-, and 12-month visits (Fig. 1). The measurements were made by one research assistant and coded by another who was blinded to the patient's group assignment. All the subjects provided written informed consent before the study, and the study protocol was approved by the local ethics committee.

Participants and eligibility

Participants were selected according to the following inclusion criteria: 18–65-year-old employed people with current non-specific LBP with or without radiating pain to one or both lower legs. The back pain episode could be acute to chronic, the first or recurrent. Exclusion criteria were: pregnancy, low back surgery less than 2 months previously, and "red flags" that indicate serious spinal pathology (10).

Randomization

Randomization of the participants into the treatment groups was by a stack of sealed envelopes, numbered in an order prepared from a random number table. The aim was to investigate 180 patients during 3 years, but the final number of participants available was 136, which explains the imbalance in the number of subjects between groups: 45 had been allocated to the OMT group, 52 to the McKenzie, and 37 to the advice-only group. In addition, 2 subjects were excluded as not fulfilling the inclusion criteria (Fig. 1). No significant differences existed between groups at baseline in age, gender, or clinical characteristics (Table I).

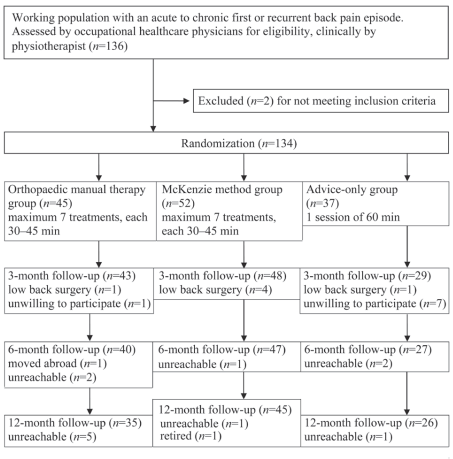


Fig. 1. Flow of patients through the trial.

Clinical examination

An LBP history was taken, and a structured examination lasting 45–60 min was carried out by the research assistant before randomization.

Clinical trial

Orthopaedic manual therapy. The OMT group underwent spinal manipulation if indicated (11), specific mobilization, and muscle-stretching techniques (12, 13). In addition, the following mobilization or high velocity, low-force manipulation techniques were performed: (*i*) translatoric thrust manipulation or mobilization of the thoracic-lumbar junction with the patient supine or lying on their side; (*ii*) translatoric thrust manipulation or mobilization of L1 to L5 with the patient prone or lying on their side; (*iii*) the sacroiliac manipulation/mobilization technique used in this study was a ventral or dorsal gliding of the ileum on the sacrum with the patient prone. Furthermore, these patients were taught to perform self-mobilization and stretching exercises at home once a day. Usually 3–5 individually selected home-exercises were prescribed to actively mobilize the low back, with 2–3 sets of 15–20 repetitions for each exercise, and lumbar stabilization exercises with 10 repetitions of 10 sec, and stretching exercises to be performed once a day for 45–60 sec.

McKenzie method of mechanical diagnosis and therapy. In the McKenzie method group, the participants were clinically assessed and classified into the mechanical syndromes. If a non-mechanical syndrome was present, the subjects were transferred from conservative care for further investigations. If a syndrome was present, then one of the treatment principles of mechanical therapy was selected as the management strategy. This consisted of an educational component, supported with the book *Treat Your Own Back* (14, 15), and an active therapy component provided instructions in exercises repeated several times a day according to the principles of the approach (10–15 repetitions every 1–2 h with or without a sustained end-range position on a regular basis

Table I. *Baseline demographics of 134 patients*

	Groups			
	OMT ($n=45$)	McKenzie ($n=52$)	Advice-only ($n=37$)	*p*-value
Age, years, mean (SD)	44 (10)	44 (9)	44 (15)	0.93
Gender, female/male	19/26	15/37	13/24	
History of previous LBP				0.64
First episode, %	9	11	6	
1–5 episodes, %	39	45	44	
≥6 episodes, %	52	44	50	
On sick-leave because of LBP, %	16	17	8	
Symptom location				0.23
Low back pain only, %	29	19	31	
Radiating pain to buttock above knee, %	30	54	37	
Radiating pain below knee, %	41	27	32	
Physical work				0.43
Sedentary, %	38	33	54	
Light, %	51	61	32	
Heavy, %	11	6	14	

OMT: orthopaedic manual therapy; LBP: low back pain; SD: standard deviation.

according to symptom response). On occasions, if improvements were not sustained or were too slow developing, patient-generated forces were supplemented by clinician-generated forces: therapist's over-pressure or mobilization or both within same treatment direction principle of management. "High velocity, low-force" manipulation techniques were avoided in this group during this trial (16).

Advice only. Subjects in the advice-only group received 45–60 min counselling from a physiotherapist concerning the good prognosis for LBP and concerning pain tolerance, medication, and early return to work. The patients in this group were told to avoid bed rest and advised to continue their routine as actively as possible, including exercise activities, within the limits permitted by their back pain. They were also instructed to contact their physicians if their symptoms worsened. For support, a 2-page educational back booklet (translated into Finnish from Burton et al. (18)) was also supplied (17). Other treatments during follow-up were minimal, with no differences between groups.

Number of visits

The number of visits was one for subjects in the advice-only group, and ranged from 3 to 7 in the OMT and McKenzie groups (mean 6 treatments in each group).

Therapists

In both treatment groups, the physical therapists treated their subjects independently by the method in which they were certified. All treatments were provided to each individual participant by the same therapist. The OMT was carried out by a physiotherapist (MP) with 20 years of clinical experience in this field. The McKenzie method was carried out by a physiotherapist (SK) with 10 years of experience in this therapy method. The physiotherapist (RS) who advised the subjects to stay active and continue normal daily living had 5 years of clinical experience in treating patients with LBP.

Outcome measures

Intensity of leg and low back pain. A visual analogue scale (VAS) allowed the subject to rate his or her current intensity of leg and LBP from 0 (no pain/symptoms) to 100 (worst imaginable pain/symptoms) (19).

Disability. A 0–24-point scale Roland-Morris Disability Questionnaire allowed measurement of disability in daily activities in relation to low back trouble in the previous 3 months (20).

Data analysis

The data was analysed by the intention-to-treat principle with *post hoc* tests using analysis of variance (ANOVA). *Post hoc* between-group comparisons were performed using Sheffe's adjustment for multiple comparison. An alternative analysis was conducted that accounted for drop-outs at follow-up, whereby missing values were replaced with imputed values generated by a series of estimated marginal means of measuring 2-tailed equations; subjects' previous scores were used to determine a predicted value that reduced the variance of the value for each variable. Baseline characteristics were summarized for descriptive purposes with medians and quartiles used for continuous measures and percentages for categorical measures. For all comparisons, a probability of < 0.05 was considered statistically significant (2-tailed) (21, 22).

RESULTS

The absolute values of pain and disability indices at the 3-, 6-, or 12-month follow-up are shown in Table II.

At the 3-month follow-up, improvements occurred in all groups in leg pain (from 11 to 19 mm), LBP (from 14 to 21 mm), and in the Roland-Morris index (from 5 to 6 index-points), but with no significant differences between groups (Fig. 2).

Table II. *Outcome measures at baseline, 3-, 6-, and 12-month follow-up visits, median (25th and 75th quartiles)*

	Groups		
	OMT ($n=45$)	McKenzie ($n=52$)	Advice only ($n=37$)
Baseline values			
Leg pain, VAS, mm*	20 (0, 54)	16 (0, 30)	16 (0, 30)
Low back pain, VAS, mm*	35 (20, 50)	32 (20, 42)	37 (21, 50)
Roland-Morris, 0–24†	9 (5, 8)	9 (4, 6)	8 (4, 1)
Outcome measures at 3 months			
Leg pain, VAS, mm*	6 (0, 14)	1 (0, 3)	4 (0, 10)
Low back pain, VAS, mm*	18 (11, 28)	10 (2, 22)	17 (10, 28)
Roland-Morris, 0–24†	2 (0, 5)	1 (0, 6)	0 (0, 3)
Outcome measures at 6 months			
Leg pain, VAS, mm*	4 (0, 11)	1 (0, 4)	8 (5, 24)
Low back pain, VAS, mm*	14 (10, 21)	10 (5, 15)	22 (15, 39)
Roland-Morris, 0–24†	1 (0, 4)	0 (0, 4)	1 (0, 7)
Outcome measures at 12 months			
Leg pain, VAS, mm*	2 (0, 10)	0 (0, 8)	8 (0, 21)
Low back pain, VAS, mm*	11 (3, 22)	8 (0, 23)	16 (7, 33)
Roland-Morris, 0–24†	0 (0, 2)	1 (0, 2)	0 (0, 3)

*Self-reported measures included a VAS.
†0–24-point scale on Roland-Morris Disability questionnaire.
OMT: orthopaedic manual therapy; VAS: visual analogue scale.

After the 6-month follow-up, the improvement in back pain was better in the McKenzie group (effect: –15 mm; 95% CI: –27 to –4; $p=0.009$) and in the disability index (–4 index-points; –7 to –1; $p=0.003$), than in the advice-only group. In addition, leg pain improved significantly or almost significantly more in the McKenzie method group (–15 mm; –30 to –1; $p=0.036$) and in the OMT group (–14 mm; –28 to 1; $p=0.075$) than in the advice-only group (Table III). Leg pain decreased 18 mm in the OMT group, 19 mm in the McKenzie group, and 4 mm in the advice-only group (Fig. 2). The corresponding reductions in back pain was 17 mm in the OMT group, 21 mm in the McKenzie method group, and 8 mm in the advice-only group; and the Roland-Morris index changes were 7, 8, and 4 points, respectively (Fig. 2).

After the 12-month follow-up, improvements on the disability index were 3 index-points larger in the McKenzie method group (95% CI: –6 to 0; $p=0.028$) and in the OMT group (95%CI; –6 to 0; $p=0.068$) than in the advice-only group (Table III). Leg pain decreased in all groups from 7 mm to 17 mm (Fig. 2), back pain from 15 mm to 20 mm (Fig. 2), and Roland-Morris index from 4 to 8 points (Fig. 2).

No significant differences emerged between the OMT and McKenzie method groups in pain and disability scores at any follow-up point. In addition, no inter-group differences emerged during follow-up in visits to physicians or other healthcare professionals or in the use of pain-killers.

Drop-outs

The drop-out rate during the follow-up year ranged from 14% in the McKenzie method group, to 22% in the OMT group, to 30% in the advice-only group. Fig. 1 reveals drop-outs at different follow-up points and shows the reasons for withdrawal. The baseline background values and outcome measures for those subjects who had withdrawn did not differ from

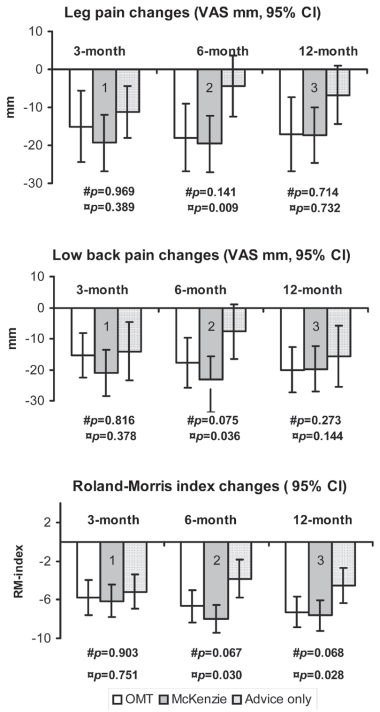


Fig. 2. Mean changes from baseline at 3-, 6- and 12-month follow-up points in leg pain, low back pain, and Roland-Morris disability index among 26 participants in the advice-only, 35 in the OMT, and 45 in the McKenzie method group who completed the 12-month follow-up. Error bars represents 95% confidence intervals and *p*-values indicate treatment effects in the OMT(#) and the McKenzie method(¤) groups compared with the advice-only group. CI: confidence interval; VAS: visual analogue scale.

those of participants who completed the study except with regards to leg pain. Six participants who had low back surgery during the follow-up period were excluded from this analysis: one each from the OMT and advice-only groups and 4 from the McKenzie method group.

DISCUSSION

The short-term outcomes of this study are in accordance with those of other recent studies (23, 24) showing that the majority of acute LBP disorders are resolved within a 4-week period. This may indicate spontaneous LBP recovery in the short term, but recurrences of LBP are frequent (25). Our results seemed to be somewhat inconsistent. We found no treatment effect immediately after the treatment period at the 3-month check-up. Our results showed small treatment effects only at the 6-month follow-up point in back pain and disability index in favour of the McKenzie method group; the treatment effect was almost significant in leg pain in favour of the OMT and McKenzie method groups and, in addition, at the 12-month follow-up point, the treatment effect was detected in the disability index in the McKenzie method group and also a trend in the OMT group. Our results are in line with those of Cherkin et al. (27), who found that differences in extent of dysfunction among physical therapy, chiropractic manipulation, and educational booklet groups were small and non-significant and approached significance only at one year, with greater dysfunction in the booklet group than in the other 2 groups (27). In addition, the UK BEAM Trial study showed a moderate benefit at 3 months and a small benefit at 12 months after manipulation followed by exercise (9).

In the present study, the drop-out rate was rather high, and was highest in the advice-only group. This is an obvious weakness of this study. Reasons for the rather high drop-out rate in the advice-only group included disappointment that having only one treatment possibility, for many subjects being unreachable because of working in a paper mill with much travelling due to required work responsibilities. It is also possible that those patients who suffered no more LBP during

Table III. *Therapy-group differences compared with the advice-only group (mean, 95% confidence interval (CI)) at 3-, 6-, and 12-month visits in pain and disability variables by intention-to-treat analysis*

	OMT group		McKenzie group	
	Difference (95% CI)	*p*-value[1]	Difference (95% CI)	*p*-value[1]
3-month follow-up				
Leg pain, (VAS, mm)	–4 (–18 to 11)	0.810	–8 (–22 to 6)	0.378
LBP, (VAS, mm)	–1 (–14 to 12)	0.396	–7 (–20 to 6)	0.389
Roland-Morris, (0–24)	–1 (–4 to 3)	0.903	–1 (–4 to 2)	0.751
6-month follow-up				
Leg pain, (VAS, mm)	–14 (–28 to 1)	0.075	–15 (–30 to –1)	0.036
LBP, (VAS, mm)	–10 (–22 to 2)	0.141	–15 (–27 to –4)	0.009
Roland-Morris, (0–24)	–3 (–6 to 0)	0.067	–4 (–7 to –1)	0.003
12-month follow-up				
Leg pain, (VAS, mm)	–10 (–25 to 5)	0.273	–10 (–23 to 2)	0.144
LBP, (VAS, mm)	–4 (–17 to 9)	0.714	–4 (–17 to 9)	0.732
Roland-Morris, (0–24)	–3 (–6 to 0)	0.068	–3 (–6 to 0)	0.028

[1]*p*-values are for the between-group differences in the ANOVA (analysis of variance).
OMT: Orthopaedic manual therapy; LBP: low back pain; VAS: visual analogue scale; CI: confidence interval.

the follow-up were less willing to participate in follow-up measurements. However, the fact that drop-outs did not differ in outcome measures from those who completed the follow-ups strengthens validity. Furthermore, sample size was quite small for a 3-arm trial, making type II error possible. For example, the confidence intervals in Fig. 2 seem to indicate differences, especially between the therapy groups and the advice-only group. A small number of subjects usually makes confidence intervals larger and thus weakens results. Unfortunately, we did not calculate sample size beforehand, but the power analysis shows that the F tests will detect differences between groups equal to those implied by the sample difference.

This study had several strengths, e.g. its randomized controlled design and the fact that therapists in the OMT and McKenzie method therapy were very experienced, with over 20 and 10 years experience in the field as therapists as well as teachers. Furthermore, the validity of the VAS and Roland-Morris has been proven to be good (28, 29). Since our participants were recruited by routine referrals from occupational healthcare services, and interventions included commonly delivered treatments, our results can be generalized. According to subjects' characteristics at baseline, our subjects were similar and corresponded to those of similar studies (30, 31).

The decrease in back pain in the treatment groups was in line with earlier findings (31–33). These studies showed changes similar (56–63% at 3–12 months follow-up on LBP) to those in this study. The Roland-Morris index improved in the treatment groups and the advice-only group much the same as in the studies of Frost et al. (34) and Wand et al. (31). They found from 18% to 69% changes in an up-to-12-month follow-up in their physiotherapy group or in their advice and therapy group. Their advice group showed changes from 10% to 57%, which was of the same magnitude as we observed.

In the present study, we did not classify patients into acute, subacute or chronic LBP, although it would have been useful to examine each of these subgroups separately. Because of our small sample size, however, such sub-group analyses based on duration of LBP could not be conducted. Earlier history, clinical signs of back problems, and other physical, psychosocial, and individual factors are associated with back pain-related outcomes among a working population (35) and should be taken into consideration in interpreting any results.

Clinical practice guidelines for management of sub-acute LBP recommend advice, mobilization, manipulation, exercise or analgesics or a combination of these. OMT and the McKenzie method include advice, exercise, and mobilization, and, if needed, manipulation. Ideally, there would be some way of identifying the subgroups most likely to benefit from one or all of these therapies, using for instance a prediction rule depending on the patient's present status to achieve better results (36–38).

Previous studies (18, 30) and the recent study of Hancock et al. (39) suggest that advice supported by a booklet is a useful intervention compared with the usual care given by a general practitioner, but only if the information is reinforced by all involved in the patient's care (18, 29). Counselling is cheap and quite easy to implement and probably well accepted, and thus should be an alternative treatment in non-specific LBP. Although the difference between OMT and McKenzie methods compared with advice-only favoured the therapy groups to some extent, this finding was not clinically meaningful at any stage of follow-up. However, in all groups, the reduction in pain and disability was clinically significant at one year. For example, a reduction of approximately 2 points (20 mm) or of approximately 30% (30 mm) on a VAS in LBP and 5 points or 50%, respectively, on the Roland-Morris Disability Index, represent a clinically important difference (40, 41).

In conclusion, some improvements appeared in all groups in leg and LBP and in disability. The OMT and McKenzie groups showed no consistent treatment effect at different follow-up points compared with the advice-only group in these quite heterogenic non-specified LBP patients. However, a slightly trend emerged that the OMT and McKenzie method groups showed some small treatment effect compared with the advice-only group.

REFERENCES

1. Helakorpi S, Uutela A, Prättälä R, Puska P, editors. Health behaviour among Finnish adult population., Helsinki: Publications of the National Public Health Institute B 10/1997; Spring 1997.
2. van Tulder M, Becker A, Bekkering T, Breen A,Carter T, Real M, et al. European guidelines for the management of acute nonspecific low back pain in primary care. 2004. Eur Spine J 2006; 15 Suppl 2: S169-S191.
3. Airaksinen O, Brox JI, Cedraschi C, Hildebrant J, Klaber-Moffet J, Kovacs F, et al. European guidelines for the management of chronic non-specific low back pain. 2004. Eur Spine J 2006 Mar; 15 Suppl 2: S192-S300.
4. Haas M, Evans RL, Bouter LM. Efficacy of spinal manipulation and mobilization for low back pain and neck pain: a systematic review and best evidence synthesis. Spine J 2004; 4: 335–356.
5. Clare H, Adams R, Maher C. A systematic review of efficacy of McKenzie therapy for spinal pain. Aust J Physiother 2004; 50: 209–216.
6. Karjalainen K, Malmivaara A, Pohjolainen T, Hurri H, Mutanen P, Rissanen P, et al. Mini-interventions for subacute low back pain: a randomized controlled trial. Spine 2003; 28: 533–540.
7. Moseley GL, Nicholas MK, Hodges PW. A randomized controlled trial of intensive neurophysiology education in chronic low back pain. Clin J Pain 2004; 20: 324–330.
8. Bogduk N. Management of chronic low back pain. Med J Aust 2004; 180: 79–83.
9. UK BEAM Trial 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.
10. Bigos S, editor. Acute low back problems in adults. Clinical practice guideline. Agency for Health Care Policy and Research, AHCPR Publication No. 95-0643: Rockville, USA; 1994.
11. Krauss J, Evjenth O, Creighton D, editors. Translatoric spinal manipulation. A Lakeview Media LLC Publication: Minneapolis, USA; 2006.
12. Kaltenborn FM, Evjenth O, Kaltenborn TB, Morgan D, editors. Manual mobilization of the joints. The Kaltenborn method of joint examination and treatment, valid. The spine. 4th edn. Norli Bokhandel: Oslo, Norway; 2003.
13. Evjenth O, Hamberg J, editors. Auto-stretching. Alfta Rehab Förlag: Alfta, Sverige; 1989.
14. McKenzie R, editor. Treat your own back. Waikanae, New Zealand: Spinal Publication Ltd; 1985.
15. Udermann B, Spratt KF, Donelson R, Mayer J, Graves J. Can a

patient educational book change behaviour and reduce pain in chronic low back pain patients? Spine 2004; 4: 425–435.
16. McKenzie RA, May S, editors. The lumbar spine mechanical diagnosis and therapy. Waikanae, New Zealand: Spinal Publication Ltd; 2003.
17. Takala J-P, editor. Älä suotta kärsi selkäkivusta. [Don't suffer needlessly]. Muskeli Project, 1995. University of Huddersfield, Spinal Research Unit: Back Pain; 1995 (in Finnish).
18. Burton A, Waddell G, Tillotson K, Summerton N. Information and advice to patients with back pain can have a positive effect. A randomized controlled trial of a novel educational booklet in primary care. Spine 1999; 24: 2484–2491.
19. Scott J, Huskisson E. Vertical and horizontal visual analogue scales. Ann Rheum Dis 1978; 38: 560.
20. Roland M, Fairbank J. The Roland – Morris Disability Questionnaire and the Ostwestry Disability Questionnaire. Spine 2000; 25: 3115–3124.
21. Hicks CM, editor. Research methods for clinical therapists. Applied project design and analysis (3rd edn) Churchill Livingstone, Edinburgh, UK: Harcourt Publisher Ltd; 2000.
22. Sim J, Wright C. Research in health care: concepts, design and methods. Cheltenham: Stanley Thornes; 2000.
23. Burton AK, McClune T, Clarke R, Main C. Long-term follow-up of patients with low back pain attending for manipulative care: outcomes and predictors. Manual Therapy 2004; 9: 30–35.
24. Gudavalli M, Cambron J, McGregor M, Jedlicka J, Keenum M, Ghanayem A, et al. A randomized clinical trial and subgroup analysis to compare flexion-distraction with active exercise for chronic low back pain. Eur Spine J 2005; 8: 1–13.
25. Croft PR, Macfarlane GJ, Papageorgiou AC, Thomas E, Silman AJ. Outcome of low back pain in general practice: a prospective study. BMJ 1998; 316: 1356–1359.
26. Hides JA, Jull GA, Richardson CA. Long-term effects of specific stabilizing exercises for first-episode low back pain. Spine 2001; 26: E243–E248.
27. Cherkin D, Deyo R, Battié M, Street J, Barlow W. A comparison of physical therapy, chiropractic manipulation, and provision of an educational booklet for the treatment of patients with low back pain. N Engl J Med 1998; 339: 1021–1029.
28. Price D, McGrath P, Rafii A, Buchingham B. The validation of visual analogue scales as ratio scale measures for chronic and experimental pain. Pain 1983; 17: 45–56.
29. Rocchi M, Sisti D, Benedetti P, Valentini M, Bellagamba SA. Critical comparison of nine different self-administered questionnaires for the evaluation of disability caused by low back pain. Eura Medicophys 2005; 41: 275–281.
30. Little P, Roberts L, Blowers H, Garwood J, Cantrell T, Langridge J, et al. Should we give detailed advice and information booklets to patients with back pain? A randomised controlled factorial trial of self management booklet and doctor advice to take exercise for back pain. Spine 2001; 26: 2065–2072.
31. Wand B, Bird C, McAtrhur J, McDowell M, De Souza L. Early intervention of acute low back pain. A single-blind randomized trial of biopsychosocial education, manual therapy, and exercise. Spine 2004; 21: 2350–2356.
32. Niemisto L, Rissanen P, Sarna S, Lahtinen-Suopanki T, Lindgren KA, Hurri H. Cost-effectiveness of combined manipulation, stabilizing exercises, and physician compared to physician consultation alone for chronic low back pain: a prospective randomized trial with 2-year follow-up. Spine 2005; 30: 1109–1115.
33. Pengel L, Refshauge K, Maher C. Responsiveness of pain, disability, and physical impairment outcome patients with low back pain. Spine 2004; 29: 879–883.
34. Frost H, Lamb S, Doll H, Carver T, Stewart-Brown S. Randomized controlled trial physiotherapy compared with advice for low back pain. BMJ 2004; 239: 708–714.
35. Videman T, Ojajärvi A, Riihimäki H, Troup J. Low back pain among nurses. A follow-up beginning at entry to the nursing school. Spine 2005; 30: 2334–2341.
36. Childs J, Fritz J, Flynn T, Irrgang J, Johnson K, Majkowski G, 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: 920–928.
37. Laslett M, McDonald B, Tropp H, Aprill CN, Oberg B. Agreement between diagnoses reached by clinical examination and available reference standards: a prospective study of 216 patients with lumbopelvic pain. BMC Musculoskelet Disord 2005; 6: 28.
38. Laslett M, McDonald B, Aprill CN, Tropp H, Oberg B. Clinical predictors of screening lumbar zygapophyseal joint blocks: development of clinical prediction rules. Spine J 2006; 6: 370–379.
39. Hancock MJ, Maher CG, Latimer J, McLachlan AJ, Cooper CW, Day RO, et al. Assessment of diclofenac or spinal manipulative therapy, or both, in addition to recommended first-line treatment for acute low back pain: a randomised controlled trial. Lancet 2007; 370: 1638–1643.
40. Lauridsen HH, Hartvigsen J, Manniche C, Korsholm L, Grunnet-Nilsson N. Responsiveness and minimal clinically important difference for pain and disability instruments in low back pain patients. BMC Musculoskelet Disord 2006; 25: 82.
41. Farrar JT, Young JP Jr, LaMoreaux L, Werth JL, Poole RM. Clinical importance of changes in chronic pain intensity measured on an 11-point numerical pain rating scale. Pain 2001; 94: 149–158.

STUDIES IN SPORT, PHYSICAL EDUCATION AND HEALTH

1 KIRJONEN, JUHANI, On the description of a human movement and its psychophysical correlates under psychomotor loads. 48 p. 1971.
2 KIRJONEN, JUHANI JA RUSKO, HEIKKI, Liikkeen kinemaattisista ominaispiirteistä, niiden psykofyysisistä selitysyhteyksistä ja näiden muutoksista psykomotorisen kuormituksen ja kestävyysharjoittelun vaikutuksesta. - On the kinematic charac-teristics and psychophysical correlates of a human movement and their changes during psychomotor loading and endurace conditioning. 156 p. 1971.
3 SARVIHARJU, PEKKA J., Effects of psycho-physical loading and progressive endurance conditioning on selected biochemical correlates of adaptive responses in man. 95 p. 1973.
4 KIVIAHO, PEKKA, Sport organizations and the structure of society. 54 p. 1973.
5 KOMI, PAAVO V., NELSON, RICHARD C. AND PULLI, MATTI, Biomechanics of skijumping. 53 p. 1974.
6 METELI, Työolot, terveys ja liikuntakäyttäytyminen metallitehtaissa. Kartoittavan kyselyn aineistot ja toteuttaminen. 178 p. 1974.
7 TIAINEN, JORMA M., Increasing physical education students' creative thinking. 53 p. 1976.
8 RUSKO, HEIKKI, Physical performance characteristics in Finnish athletes. 40 p. 1976.
9 KIISKINEN, ANJA, Adaptation of connective tissues to physical training in young mice. 43 p. 1976.
10 VUOLLE, PAULI, Urheilu elämänsisältönä. Menestyneiden urheilijoiden elämänura kilpailuvuosina - Top sport as content of life. 227 p. 1977.
11 SUOMINEN, HARRI, Effects of physical training in middle-aged and elderly people with special regard to skeletal muscle, connective tissue, and functional aging. 40 p. 1978.
12 VIITASALO, JUKKA, Neuromuscular performance in voluntary and reflex contraction with special reference to muscle structure and fatigue. 59 p. 1980.
13 LUHTANEN, PEKKA, On the mechanics of human movement with special reference to walking, running and jumping. 58 p. 1980.
14 LAAKSO, LAURI, Lapsuuden ja nuoruuden kasvuympäristö aikuisiän liikuntaharrastusten selittäjänä: retrospektiivinen tutkimus. - Socialization environment in childhood and youth as determinant of adult-age sport involvement: a retrospective study. 295 p. 1981.
15 BOSCO, CARMELO, Stretch-schortening cycle inskeletal muscle function with special reference to elastic energy and potentiation of myoelectrical activity. 64 p. 1982.
16 OLIN, KALEVI, Päätöksentekijöiden viiteryhmät kaupunkien liikuntapolitiikassa. - Reference groups of decision-makers in the sport politics of cities. 155 p. 1982.
17 KANNAS, LASSE, Tupakointia koskeva terveyskasvatus peruskoulussa. - Health education on smoking in the Finnish comprehensive school. 251 p. 1983.
18 Contribution of sociology to the study of sport. Festschrift Book in Honour of Professor Kalevi Heinilä. Ed. by OLIN, K. 243 p. 1984.
19 ALÉN, MARKKU, Effects of self-administered, high-dose testosterone and anabolic steroids on serum hormones, lipids, enzymes and on spermatogenesis in power athletes. 75 p. 1985.
20 HÄKKINEN, KEIJO, Training and detraining adaptations in electromyographic, muscle fibre and force production characteristics of human leg extensor muscles with special reference to prolonged heavy resistance and explosive type strength training. 106 p. 1986.
21 LAHTINEN, ULLA, Begåvningshandikappad ungdom i utveckling. En uppföljningstudie av funktionsförmåga och fysisk aktivitet hos begåvningshandikappade ungdomar i olika livsmiljöer. 300 p. 1986.
22 SILVENNOINEN, MARTTI, Koululainen liikunnanharrastajana: liikuntaharrastusten ja liikuntamotiivien sekä näiden yhteyksien muuttuminen iän mukana peruskoululaisilla ja lukiolaisilla. - Schoolchildren and physically active interests: The changes in interests in and motives for physical exercise related to age in Finnish comprehensive and upper secondary schools. 226 p. 1987.
23 POHJOLAINEN, PERTTI, Toimintakykyisyys, terveydentila ja elämäntyyli 71-75-vuotiailla miehillä. - Functional capacity, health status and life-style among 71-75 year-old men. 249 p. Summary 13 p. 1987.
24 MERO, ANTTI, Electromyographic acticity, force and anaerobic energy production in sprint running; with special reference to different constant speeds ranging from submaximal to supramaximal. 112 p. Tiivistelmä 5 p. 1987.
25 PARKATTI, TERTTU, Self-rated and clinically measured functional capacity among women and men in two age groups in metal industry. 131 p. Tiivistelmä 2 p. 1990.
26 HOLOPAINEN, SINIKKA, Koululaisten liikuntataidot. - The motor skills of schoolboys and girls. 217 p. Summary 6 p. 1990.
27 NUMMINEN, PIRKKO, The role of imagery in physical education. 131 p. Tiivistelmä 10 p. 1991.
28 TALVITIE, ULLA, Aktiivisuuden ja omatoimivuuden kehittäminen fysioterapian tavoitteena. Kehittävän työntutkimuksen sovellus lääkintävoimistelijan työhön. - The development of activity and self-motivation as the aim of physiotherapy. The application of developmental work research in physiotherapy. 212 p. Summary 8 p. 1991.
29 KAHILA, SINIKKA, Opetusmenetelmän merkitys prososiaalisessa oppimisessa - auttamis-

käyttäytymisen edistäminen yhteistyöskentelyn avulla koululiikunnassa. - The role of teaching method in prosocial learning - developing helping behavior by means of the cooperative teaching method in physical education. 132 p. Summary 2 p. 1993.

30 LIIMATAINEN-LAMBERG, ANNA-ESTER, Changes in student smoking habits at the vocational institutions and senior secondary schools and health education. 195 p. Yhteenveto 5 p. 1993.

31 KESKINEN, KARI LASSE, Stroking characteristics of front crawl swimming. 77 p. Yhteenveto 2 p. 1993.

32 RANTANEN, TAINA, Maximal isometric strength in older adults. Cross-national comparisons, background factors and association with Mobility. 87 p. Yhteenveto 4 p. 1994.

33 LUSA, SIRPA, Job demands and assessment of the physical work capacity of fire fighters. 91 p. Yhteenveto 4 p. 1994.

34 CHENG, SULIN, Bone mineral density and quality in older people. A study in relation to exercise and fracture occurrence, and the assessment of mechanical properties. 81 p. Tiivistelmä 1 p. 1994.

35 KOSKI, PASI, Liikuntaseura toimintaympäristössään. - Sports club in its organizational environment. 220 p. Summary 6 p. 1994.

36 JUPPI, JOEL, Suomen julkinen liikuntapolitiikka valtionhallinnon näkökulmasta vuosina 1917-1994. - Public sport policy in Finland from the viewpoint of state administration in 1917-1994. 358 p. Summary 7 p. 1995.

37 KYRÖLÄINEN, HEIKKI, Neuromuscular performance among power- and endurance-trained athletes. 82 p. Tiivistelmä 3 p. 1995.

38 NYANDINDI, URSULINE S., Evaluation of a school oral health education programme in Tanzania: An ecological perspective. 88 p. Tiivistelmä 2 p. 1995.

39 HEIKINARO-JOHANSSON, PILVIKKI, Including students with special needs in physical education. 81 p. Yhteenveto 4 p. 1995.

40 SARLIN, EEVA-LIISA, Minäkokemuksen merkitys liikuntamotivaatiotekijänä. - The significance of self perception in the motivational orientation of physical education. 157 p. Summary 4 p. 1995.

41 LINTUNEN, TARU, Self-perceptions, fitness, and exercise in early adolescence: a four-year follow-up study. 87 p. Yhteenveto 5 p.1995.

42 SIPILÄ, SARIANNA, Physical training and skeletal muscle in elderly women. A study of muscle mass, composition, fiber characteristics and isometric strength. 62 p. Tiivistelmä 3 p. 1996.

43 ILMANEN, KALERVO, Kunnat liikkeellä. Kunnallinen liikuntahallinto suomalaisen yhteiskunnan muutoksessa 1919-1994. - Municipalities in motion. Municipal sport administration in the changing Finnish society 1919-1994. 285 p. Summary 3 p. 1996.

44 NUMMELA, ARI, A new laboratory test method for estimating anaerobic performance characteristics with special reference to sprint running. 80 p. Yhteenveto 4 p. 1996.

45 VARSTALA, VÄINÖ, Opettajan toiminta ja oppilaiden liikunta-aktiivisuus koulun liikuntatunnilla. - Teacher behaviour and students' motor engagement time in school physical education classes. 138 p. Summary 4 p. 1996.

46 POSKIPARTA, MARITA, Terveysneuvonta, oppimaan oppimista. Videotallenteet hoitajien terveysneuvonnan ilmentäjinä ja vuorovaikutustaitojen kehittämismenetelmänä. - Health counselling, learning to learn. Videotapes expressing and developing nurses´ communication skills. 159 p. Summary 6 p. 1997.

47 SIMONEN, RIITTA, Determinants of adult psychomotor speed. A study of monozygotic twins. - Psykomotorisen nopeuden determinantit identtisillä kaksosilla. 49 p. Yhteenveto 2 p. 1997.

48 NEVALA-PURANEN, NINA, Physical work and ergonomics in dairy farming. Effects of occupationally oriented medical rehabilitaton and environmental measures. 80 p. (132 p.) 1997.

49 HEINONEN, ARI, Exercise as an Osteogenic Stimulus. 69 p. (160 p.) Tiivistelmä 1 p. 1997.

50 VUOLLE, PAULI (Ed.) Sport in social context by Kalevi Heinilä. Commemorative book in Honour of Professor Kalevi Heinilä. 200 p. 1997.

51 TUOMI, JOUNI, Suomalainen hoitotiedekeskustelu. - The genesis of nursing and caring science in Finland. 218 p. Summary 7 p. 1997.

52 TOLVANEN, KAIJA, Terveyttä edistävän organisaation kehittäminen oppivaksi organisaatioksi. Kehitysnäytökset ja kehittämistehtävät terveyskeskuksen muutoksen virittäjänä. - Application of a learning organisation model to improve services in a community health centre. Development examples and development tasks are the key to converting a health care. 197 p. Summary 3 p. 1998.

53 OKSA, JUHA, Cooling and neuromuscular performance in man. 61 p. (121 p.) Yhteenveto 2 p. 1998.

54 GIBBONS, LAURA, Back function testing and paraspinal muscle magnetic resonance image parameters: their associations and determinants. A study on male, monozygotic twins. 67 p (128 p.) Yhteenveto 1p. 1998.

55 NIEMINEN, PIPSA, Four dances subcultures. A study of non-professional dancers´ socialization, participation motives, attitudes and stereotypes. - Neljä tanssin alakulttuuria. Tutkimus tanssinharrastajien tanssiin sosiaalistumisesta, osallistumismotiiveista, asenteista ja stereotypioista. 165 p. Yhteenveto 4 p. 1998.

56 LAUKKANEN, PIA, Iäkkäiden henkilöiden selviytyminen päivittäisistä toiminnoista. - Carrying

out the activities of daily living among elderly people. 130 p. (189 p.). Summary 3 p. 1998.

57 AVELA, JANNE, Stretch-reflex adaptation in man. Interaction between load, fatigue and muscle stiffness. 87 p. Yhteenveto 3 p. 1998.

58 SUOMI, KIMMO, Liikunnan yhteissuunnittelumetodi. Metodin toimivuuden arviointi Jyväskylän Huhtasuon lähiössä. - Collaborative planning method of sports culture. Evaluation of the method in the Huhtasuo suburb of the city of Jyväskylä. 190 p. Summary 8 p. 1998.

59 PÖTSÖNEN, RIIKKA, Naiseksi, mieheksi, tietoiseksi. Koululaisten seksuaalinen kokeneisuus, HIV/AIDS-tiedot, -asenteet ja tiedonlähteet. - Growing as a woman, growing as a man, growing as a conscious citizen. 93 p. (171 p.). Summary 3 p. 1998.

60 HÄKKINEN, ARJA, Resistance training in patients with early inflammatory rheumatic diseases. Special reference to neuromuscular function, bone mineral density and disease activity. - Dynaamisen voimaharjoittelun vaikutukset nivelreumaa sairastavien potilaiden lihasvoimaan, luutiheyteen ja taudin aktiivisuuteen. 62 p. (119 p.) Yhteenveto 1 p. 1999.

61 TYNJÄLÄ, JORMA, Sleep habits, perceived sleep quality and tiredness among adolescents. A health behavioural approach. - Nuorten nukkumistottumukset, koettu unen laatu ja väsyneisyys. 104 p. (167 p.) Yhteenveto 3 p. 1999.

62 PÖNKKÖ, ANNELI, Vanhemmat ja lastentarhanopettajat päiväkotilasten minäkäsityksen tukena. - Parents´ and teachers´ role in self-perception of children in kindergartens. 138 p. Summary 4 p. 1999.

63 PAAVOLAINEN, LEENA, Neuromuscular characteristics and muscle power as determinants of running performance in endurance athletes with special reference to explosive-strength training. - Hermolihasjärjestelmän toimintakapasiteetti kestävyyssuorituskykyä rajoittavana tekijänä. 88 p. (138 p.) Yhteenveto 4 p. 1999.

64 VIRTANEN, PAULA, Effects of physical activity and experimental diabetes on carbonic anhydrace III and markers of collagen synthesis in skeletal muscle and serum. 77 p. (123 p.) Yhteenveto 2 p. 1999.

65 KEPLER, KAILI, Nuorten koettu terveys, terveyskäyttäytyminen ja sosiaalistumisympäristö Virossa. - Adolescents' perceived health, health behaviour and socialisation enviroment in Estonia. - Eesti noorte tervis, tervisekäitumine ja sotsiaalne keskkond. 203 p. Summary 4p. Kokkuvõte 4 p. 1999.

66 SUNI, JAANA, Health-related fitness test battery for middle-aged adults with emphasis on musculoskeletal and motor tests. 96 p. (165 p.) Yhteenveto 2 p. 2000.

67 SYRJÄ, PASI, Performance-related emotions in highly skilled soccer players. A longitudinal study based on the IZOF model. 158 p. Summary 3 p. 2000.

68 VÄLIMAA, RAILI, Nuorten koettu terveys kyselyaineistojen ja ryhmähaastattelujen valossa. - Adolescents' perceived health based on surveys and focus group discussions. 208 p. Summary 4 p. 2000.

69 KETTUNEN, JYRKI, Physical loading and later lower-limb function and findings . A study among male former elite athletes. - Fyysisen kuormituksen yhteydet alaraajojen toimintaan ja löydöksiin entisillä huippu-urhelijamiehillä. 68 p. (108 p.) Yhteenveto 2 p. 2000.

70 HORITA, TOMOKI, Stiffness regulation during stretch-shortening cycle exercise. 82 p. (170 p.) 2000.

71 HELIN, SATU, Iäkkäiden henkilöiden toimintakyvyn heikkeneminen ja sen kompensaatioprosessi. - Functional decline and the process of compensation in elderly people. 226 p. Summary 10 p. 2000.

72 KUUKKANEN, TIINA, Therapeutic exercise programs and subjects with low back pain. A controlled study of changes in function, activity and participation. 92 p. (154 p.) Tiivistelmä 2 p. 2000.

73 VIRMAVIRTA, MIKKO, Limiting factors in ski jumping take-off. 64 p. (124 p.) Yhteenveto 2 p. 2000.

74 PELTOKALLIO, LIISA, Nyt olisi pysähtymisen paikka. Fysioterapian opettajien työhön liittyviä kokemuksia terveysalan ammatillisessa koulutuksessa. - Now it's time to stop. Physiotherapy teachers' work experiences in vocational health care education. 162 p. Summary 5 p. 2001.

75 KETTUNEN, TARJA, Neuvontakeskustelu. Tutkimus potilaan osallistumisesta ja sen tukemisesta sairaalan terveysneuvonnassa. - Health counseling conversation. A study of patient participation and its support by nurses during hospital counseling. 123 p. (222 p.) Summary 6 p. 2001.

76 PULLINEN, TEEMU, Sympathoadrenal response to resistance exercise in men, women and pubescent boys. With special reference to interaction with other hormones and neuromuscular performance. 76 p. (141 p.) Yhteenveto 2 p. 2001.

77 BLOMQVIST, MINNA, Game understanding and game performance in badminton. Development and validation of assessment instruments and their application to games teaching and coaching. 83 p. Yhteenveto 5 p. 2001.

78 FINNI, TAIJA, Muscle mechanics during human movement revealed by *in vivo* measurements of tendon force and muscle length. 83 p. (161 p.) Yhteenveto 3 p. 2001.

79 KARIMÄKI, ARI, Sosiaalisten vaikutusten arviointi liikuntarakentamisessa. Esimerkkinä Äänekosken uimahalli. - Social impact

100 RANTA, SARI, Vanhenemismuutosten eteneminen. 75-vuotiaiden henkilöiden antropometristen ominaisuuksien, fyysisen toimintakyvyn ja kognitiivisen kyvykkyyden muutokset viiden ja kymmenen vuoden seuranta-aikana. - The progress of aging processes. A 5- and 10-year follow-up study of the changes in anthropometrical characteristics and physical and cognitive capacities among 75-year-old persons. 186 p. Summary 2 p. 2004.

101 SIHVONEN, SANNA, Postural balance and aging. Cross-sectional comparative studies and a balance training intervention. - Ikääntyminen ja tasapaino. Eri ikäisten tasapaino ja tasapainoharjoittelun vaikuttavuus ikääntyneillä palvelukodissa asuvilla naisilla. 65 p. (106 p.) Yhteenveto 2 p. 2004.

102 RISSANEN, AARO, Back muscles and intensive rehabilitation of patients with chronic low back pain. Effects on back muscle structure and function and patient disability. - Selkälihakset ja pitkäaikaista selkäkipua sairastavien potilaiden intensiivinen kuntoutus. Vaikutukset selkälihasten rakenteeseen ja toimintaan sekä potilaiden vajaakuntoisuuteen. 90 p. (124 p.) Yhteenveto 2 p. 2004.

103 KALLINEN, MAURI, Cardiovascular benefits and potential hazards of physical exercise in elderly people. - Liikunnan hyödylliset ja mahdolliset haitalliset vaikutukset ikääntyneiden verenkiertoelimistöön. 97 p. (135 p). Yhteenveto 2 p. 2004.

104 SÄÄKSLAHTI, ARJA, Liikuntaintervention vaikutus 3–7-vuotiaiden lasten fyysiseen aktiivisuuteen ja motorisiin taitoihin sekä fyysisen aktiivisuuden yhteys sydän- ja verisuonitautien riskitekijöihin. - Effects of physical activity Intervention on physical activity and motor skills and relationships between physical activity and coronary heart disease risk factors in 3–7-year-old children. 153 p. Summary 3 p. 2005.

105 HÄMÄLÄINEN, PIIA, Oral health status as a predictor of changes in general health among elderly people. 76 p. (120 p.) Summary 2 p. 2005.

106 LIINAMO, ARJA, Suomalaisnuorten seksuaalikasvatus ja seksuaaliterveystiedot oppilaan ja koulun näkökulmasta. Arviointia terveyden edistämisen viitekehyksessä. - Sexual education and sexual health knowledge among Finnish adolescents at pupil and school level. Evaluation from the point of view of health promotion. 111 p. (176 p.) Summary 5 p. 2005.

107 ISHIKAWA, MASAKI, *In vivo* muscle mechanics during human locomotion. Fascicle-tendinous tissue interaction during stretch-shortening cycle exercises. - Venytysrefleksin muutokset liikkeessä ja väsymyksessä. 89 p. (228 p.) Yhteenveto 1 p. 2005.

108 KÄRKI, ANNE, Physiotherapy for the functioning of breast cancer patients. Studies of the effectiveness of physiotherapy methods and exercise, of the content and timing of post-operative education and of the experienced functioning and disability . - Rintasyöpäleikattujen toimintakyky ja siihen vaikuttaminen fysioterapiassa ja harjoittelussa. 70 p. (138 p.) Yhteenveto 3 p. 2005.

109 RAJANIEMI, VESA, Liikuntapaikkarakentaminen ja maankäytön suunnittelu. Tutkimus eri väestöryhmät tasapuolisesti huomioon ottavasta liikuntapaikkasuunnittelusta ja sen kytkemisestä maankäyttö- ja rakennuslain mukaiseen kaavoitukseen. - Sports area construction and land use planning – Study of sports area planning that considers all the population groups even-handedly and integrates sports area planning with land use planning under the land use and building act. 171 p. Summary 6 p. 2005.

110 WANG, QINGJU, Bone growth in pubertal girls. Cross-sectional and lingitudinal investigation of the association of sex hormones, physical activity, body composition and muscle strength with bone mass and geometry. 75 p. (117 p.) Tiivistelmä 1 p. 2005.

111 ROPPONEN, ANNINA, The role of heredity, other constitutional structural and behavioral factors in back function tests.- Perimä, muut synnynnäiset rakenteelliset tekijät ja käyttäytymistekijät selän toimintakykytesteissä. 78 P. (125 p.) Tiivistelmä 1 p. 2006.

112 ARKELA-KAUTIAINEN, MARJA, Functioning and quality of life as perspectives of health in patients with juvenile idiopathic arthritis in early adulthood. Measurement and long-term outcome. - Toimintakyky ja elämänlaatu terveyden näkökulmina lastenreumaa sairastaneilla nuorilla aikuisilla. Mittaaminen ja pitkäaikaistulokset. 95 p. (134 p.) Tiivistelmä 2 p. 2006.

113 RAUTIO, NINA, Seuruu- ja vertailututkimus sosioekonomisen aseman yhteydestä toimintakykyyn iäkkäillä henkilöillä. - A follow-up and cross-country comparison study on socio-economic position and its relationship to functional capacity in elderly people. 114 p. (187 p.) Summary 3 p. 2006.

114 TIIKKAINEN, PIRJO, Vanhuusiän yksinäisyys. Seuruutukimus emotionaalista ja sosiaalista yksinäisyyttä määrittävistä tekijöistä. - Loneliness in old age – a follow-up study of determinants of emotional and social loneliness. 76 p. (128 p.) Summary 2 p. 2006.

115 AHTIAINEN, JUHA, Neuromuscular, hormonal and molecular responses to heavy resistance training in strength trained men; with special reference to various resistance exercise protocols, serum hormones and gene expression of androgen receptor and insulin-like growth factor-I. - Neuromuskulaariset,

hormonaaliset ja molekulaariset vasteet voimaharjoittelussa voimaurheilijoilla. 119 p. (204 p.) Yhteenveto 2 p. 2006.

116 Pajala, Satu, Postural balance and susceptibility to falls in older women. Genetic and environmental influences in single and dual task situations. - Iäkkäiden naisten tasapainokyky yksinkertaisissa sekä huomion jakamista vaativissa tilanteissa ja kaatumisriskiperimän merkitys yksilöiden välisten erojen selittäjinä. 78 p. (120 p.) Yhteenveto 3 p. 2006.

117 Tiainen, Kristina, Genetics of skeletal muscle characteristics and maximal walking speed among older female twins. - Lihasvoiman ja kävelynopeuden periytyvyys iäkkäillä naiskaksosilla. 77 p. (123 p.) Yhteenveto 2 p. 2006.

118 Sjögren, Tuulikki, Effectiveness of a workplace physical exercise intervention on the functioning, work ability, and subjective well-being of office workers – a cluster randomised controlled cross-over trial with one-year follow-up. - Työpaikalla tapahtuvan fyysisen harjoitteluintervention vaikuttavuus toimistotyöntekijöiden toimintakykyyn, työkykyyn ja yleiseen subjektiiviseen elämänlaatuun – ryhmätasolla satunnaistettu vaihtovuorokoe ja vuoden seuranta. 100 p. (139 p.) Tiivistelmä 3 p. 2006.

119 Lyyra, Tiina-Mari, Predictors of mortality in old age. Contribution of self-rated health, physical functions, life satisfaction and social support on survival among older people. - Kuolleisuuden ennustetekijät iäkkäässä väestössä. Itsearvioidun terveyden, fyysisten toimintojen, elämään tyytyväisyyden ja sosiaalisen tuen yhteys iäkkäiden ihmisten eloonjäämiseen. 72 p. (106 p.) Tiivistelmä 2 p. 2006.

120 Soini, Markus, Motivaatioilmaston yhteys yhdeksäsluokkalaisten fyysiseen aktiivisuuteen ja viihtymiseen koulun liikuntatunneilla. - The relationship of motivational climate to physical activity intensity and enjoyment within ninth grade pupils in school physical education lessons. 91 p. 2006.

121 Vuorimaa, Timo, Neuromuscular, hormonal and oxidative stress responses to endurance running exercises in well trained runners. - Neuromuskulaariset, hormonaaliset ja hapettumisstressiin liittyvät vasteet kestävyysjuoksuharjoituksiin hyvin harjoitelleilla juoksijoilla. 93 p. (152 p.) Yhteenveto 3 p. 2007.

122 Mononen, Kaisu, The effects of augmented feedback on motor skill learning in shooting. A feedback training intervention among inexperienced rifle shooters. - Ulkoisen palautteen vaikutus motoriseen oppimiseen ammunnassa: Harjoittelututkimus kokemattomilla kivääriampujilla. 63 p. Yhteenveto 4 p. 2007.

123 Sallinen, Janne, Dietary Intake and Strength Training Adaptation in 50–70 -year old Men and Women. With special reference to muscle mass, strength, serum anabolic hormone concentrations, blood pressure, blood lipids and lipoproteins and glycemic control. - Ravinnon merkitys voimaharjoittelussa 50–70 -vuotiailla miehillä ja naisilla 103 p. (204 p.) Yhteenveto 3 p. 2007.

124 Kasila Kirsti, Schoolchildren's oral health counselling within the organisational context of public oral health care. Applying and developing theoretical and empirical perspectives. 96 p. (139 p.) Tiivistelmä 3 p. 2007.

125 Pyöriä, Outi, Reliable clinical assessment of stroke patients' postural control and development of physiotherapy in stroke rehabilitation. - Aivoverenkiertohäiriöpotilaiden toimintakyvyn luotettava kliininen mittaaminen ja fysioterapian kehittäminen Itä-Savon sairaanhoitopiirin alueella. 94 p. (143 p.) Yhteenveto 6 p. 2007.

126 Valkeinen, Heli, Physical fitness, pain and fatigue in postmenopausal women with fibromyalgia. Effects of strength training. - Fyysinen kunto, kipu- ja väsymysoireet ja säännöllisen voimaharjoittelun vaikutukset menopaussi-iän ohittaneilla fibromyalgiaa sairastavilla naisilla. 101 p. (132 p.) Yhteenveto 2 p. 2007.

127 Hämäläinen, Kirsi, Urheilija ja valmentaja urheilun maailmassa. Eetokset, ihanteet ja kasvatus urheilijoiden tarinoissa. - An athlete and a coach in the world of sports. Ethos, ideals and education in athletes' narratives. 176 p. Tiivistelmä 2 p. 2008.

128 Aittasalo, Minna, Promoting physical activity of working aged adults with selected personal approaches in primary health care. Feasibility, effectiveness and an example of nationwide dissemination. - Työikäisten liikunnan edistäminen avoterveydenhuollossa – työtapojen toteuttamiskelpoisuus ja vaikuttavuus sekä esimerkki yhden työtavan levittämisestä käytäntöön. 105 p. (161 p.) Yhteenveto 3 p. 2008.

129 Portegijs, Erja, Asymmetrical lower-limb muscle strength deficit in older people. - Alaraajojen lihasvoiman puoliero iäkkäillä ihmisillä. 105 p. (155 p.) Yhteenveto 3 p. 2008.

130 Laitinen-Väänänen, Sirpa, The construction of supervision and physiotherapy expertise: A qualitative study of physiotherapy students' learning sessions in clinical education. - Opiskelijan ohjauksen ja fysioterapian asiantuntijuuden rakentuminen: Laadullinen tutkimus fysioterapiaopiskelijan oppimistilanteista työharjoittelussa. 69 p. (118 p.) Yhteenveto 3 p. 2008.

131 Iivonen, Susanna, Early Steps -liikunta-ohjelman yhteydet 4–5-vuotiaiden päiväkotilasten motoristen perustaitojen kehitykseen. - The associations between an Early Steps physical education curriculum and the fundamental motor skills development of 4–5-year-old preschool children. 157 p. Summary 4 p. 2008.

132 Ortega-Alonso, Alfredo, Genetic effects on mobility, obesity and their association in older female twins. 87 p. 2009.

133 Hulmi, Juha, Molecular and hormonal responses and adaptation to resistance exercise and protein nutrition in young and older men. - Voimaharjoittelun fysiologiset ja molekyylibiologiset vaikutukset lihaskasvunsäätelyssä lisäproteiinia nautittaessa tai ilman. 109 p. (214 p.) Yhteenveto 2 p. 2009.

134 Martinmäki, Kaisu, Transient changes in heart rate variability in response to orthostatic task, endurance exercise and training. With special reference to autonomic blockades and time-frequency analysis. - Sykevaihtelun muutokset ortostaattisessa testissä, kestävyysliikunnassa ja kestävyysharjoittelussa käyttäen hyväksi autonomisen säätelyn salpauskokeita ja aika-taajuusanalyysiä . 99 p. (151 p.) Yhteenveto 2 p. 2009.

135 Sedliak, Milan, Neuromuscular and hormonal adaptations to resistance training. Special effects of time of day of training. 84 p. (175 p.) 2009.

136 Nikander, Riku, Exercise loading and bone structure. 97 p. (141 p.) Yhteenveto 1 p. 2009.

137 Korhonen, Marko T., Effects of aging and training on sprint performance, muscle structure and contractile function in athletes. - Ikääntymisen ja harjoittelun vaikutukset nopeussuorituskykyyn, lihasten rakenteeseen ja voimantuotto-ominaisuuksiin urheilijoilla. 123 p. (211 p.) Tiivistelmä 5 p. 2009.

138 Javanainen-Levonen, Tarja, Terveydenhoitajat liikunnanedistäjinä lastenneuvolatyössä. - Public Health Nurses as Physical Activity Promoters in Finnish Child Health Clinics. 104 p. (148 p.) Summary 6 p. 2009.

139 Klemola, Ulla, Opettajaksi opiskelevien vuorovaikutustaitojen kehittäminen liikunnan aineenopettajakoulutuksessa.- Developing student teachers´ social interaction skills in physical education teacher education. 92 p. (138 p.) Summary 4 p. 2009.

140 Niemi, Reetta, Onks tavallinen koe vai sellanen, missä pitää miettii? Ympäristölähtöisen terveyskasvatuspedagogiikan kehittäminen narratiivisena toimintatutkimuksena. - Is this a normal test or do we have to think? Developing environmentally oriented health education pedagogy through narrative action research . 215 p. 2009.

141 von Bonsdorff, Mikaela, Physical activity as a predictor of disability and social and health service use in older people. - Fyysinen aktiivisuus toiminnanvajauden ja sosiaali- ja terveyspalvelujen käytön ennustajana iäkkäillä henkilöillä 101 p. (134 p.) Yhteenveto 2 p. 2009.

142 Palomäki, Sanna, Opettajaksi opiskelevien pedagoginen ajattelu ja ammatillinen kehittyminen liikunnanopettajakoulutuksessa. - Pre-service teachers' pedagogical thinking and professional development in physical education teacher education. 118 p. (163 p.) Summary 3 p. 2009.

143 Vehmas, Hanna, Liikuntamatkalla Suomessa. Vapaa-ajan valintoja jälkimodernissa yhteiskunnassa. - Sport tourism in Finland – leisure choices in the post-modern society. 205 p. Summary 10 p. 2010.

144 Kokko, Sami, Health promoting sports club. Youth sports clubs' health promotion profiles, guidance, and associated coaching practice, in Finland. 147 p. (230 p.) Yhteenveto 5 p. 2010.

145 Kääriä, Sanna, Low back disorders in the long term among employees in the engineering industry. A study with 5-, 10- and 28-year follow-ups. - Metalliteollisuuden työntekijöiden alaselän sairaudet ikääntyessä: METELI–tutkimuksen 5–, 10– ja 28–vuotisseurantatutkimus. 76 p. (102 p.) Yhteenveto 2 p. 2010.

146 Santtila, Matti, Effects of added endurance or strength training on cardiovascular and neuromuscular performance of conscripts during the 8-week basic training period. - Lisätyn voima- ja kestävyysharjoittelun vaikutukset varusmiesten hengitys- ja verenkiertoelimistön sekä hermo-lihas järjestelmän suorituskykyyn kahdeksan viikon peruskoulutuskauden aikana. 85 p. (129 p.) Yhteenveto 2 p. 2010.

147 Mänty, Minna, Early signs of mobility decline and physical activity counseling as a preventive intervention in older people. - Liikkumiskyvyn heikkenemistä ennakoivat merkit ja liikuntaneuvonta liikkumisvaikeuksien ehkäisyssä iäkkäillä henkilöillä. 103 p. (149 p.) Yhteenveto 2 p. 2010.

148 Rantalainen, Timo, Neuromuscular function and bone geometry and strength in aging. - Neuromuskulaarinen suorituskyky luun geometrian ja voiman selittäjänä ikääntymisen yhteydessä. 87 p. (120 p.) Yhteenveto 1 p. 2010.

149 Kuitunen, Sami, Muscle and joint stiffness regulation during normal and fatiguing stretch-shortening cycle exercise . - Lihas- ja niveljäykkyyden säätely normaalin sekä väsyttävän venymis-lyhenemissykli – tyyppisen harjoituksen aikana . 76 p. (142 p.) Yhteenveto 1 p. 2010.

150 Piitulainen, Harri, Functional adaptation of sarcolemma to physical stress. - Lihassolukalvon toiminnallinen mukautu minen fyysiseen kuormitukseen. 103 p. (178 p.) Yhteenveto 2 p. 2010.

151 Viljanen, Anne, Genetic and environmental effects on hearing acuity and the association between hearing acuity, mobility and falls in older women. - Kuulon tarkkuuden periy-tyvyys ja yhteys liikkumiskykyyn sekä kaatumisiin iäkkäillä naisilla. 85 p. (116 p.) Yhteenveto 2 p. 2010.

152 Kulmala, Jenni, Visual acuity in relation to functional performance, falls and mortality in old age. - Heikentyneen näöntarkkuuden vaikutus toimintakykyyn, kaatumisiin ja kuolleisuuteen iäkkäillä henkilöillä. 98 p. (140 p.) Yhteenveto 3 p. 2010.

153 Nivala, Sirkka, Kokemuksellinen vanheneminen sotainvalideilla. Suomalaisten sotainvalidien kokemus elämänkulustaan ja ikääntymisestään. - Disabled war veterans and experiential ageing. Finnish disabled war veterans and their experience of the course of their lives and growing older. 178 p. Summary 4 p. 2010.

154 Rinne, Marjo, Effects of physical activity, specific exercise and traumatic brain injury on motor abilities. Theoretical and pragmatic assessment. 86 p. (134 p.) Tiivistelmä 2 p. 2010.

155 Mikkola, Tuija, Genetic and environmental contributions to bone structural strength in postmenopausal women. - Perimän ja ympäristötekijöiden vaikutus luun lujuuteen vaihdevuosi-iän ohittaneilla naisilla. 77 p. (130 p.) Yhteenveto 2 p. 2010.

156 Salo, Petri, Assessing physical capacity, disability, and health-related quality of life in neck pain. 93 p. (132 p.) Yhteenveto 2 p. 2010.

157 Ronkainen, Paula, Towards powerful old age. Association between hormone replacement therapy and skeletal muscle. - Vaihdevuosioireisiin käytettävän HRT:n yhteys luurankolihaksiston rakenteeseen ja toimintaan. 118 p. (170 p.) Yhteenveto 2 p. 2010.

158 Kilpikoski, Sinikka, The McKenzie method in assessing, classifying and treating non-specific low back pain in adults with special reference to the centralization phenomenon. - McKenzien mekaaninen diagnostisointi- ja terapiamenetelmä tutkittaessa, luokitel-taessa ja hoidettaessa aikuisten epäspesifiä alaselkäkipua. 90 p. (130 p.) Yhteenveto 2 p. 2010.

159 Mutikainen, Sara, Genetic and environmental effects on resting electrocardiography and the association between electrocardiography and physical activity, walking endurance and mortality in older people. - Lepo-EKG –muuttujien periy-tyvyys sekä yhteydet fyysiseen aktiivisuuteen, kävelykestävyyteen ja kuolleisuuteen iäkkäillä henkilöillä. 84 p. (131 p.) Yhteenveto 3 p. 2010.

160 Völgyi, Eszter, Bone, fat and muscle gain in pubertal girls. Effects of physical activity. 76 p. (138 p.) Tiivistelmä 1 p. 2010.

161 Sillanpää, Elina, Adaptations in body composition, metabolic health and physical fitness during strength or endurance training or their combination in healthy middle-aged and older adults. 113 p. (179 p.) Yhteenveto 3 p. 2011.

162 Karavirta, Laura, Cardiorespiratory, neuromuscular and cardiac autonomic adaptations to combined endurance and strength training in ageing men and women. - Yhdistetyn kestävyys- ja voima-harjoittelun vaikutukset hengitys- ja veren-kiertoelimistön sekä hermo-lihasjärjestel-män toimintaan ja sydämen autonomiseen säätelyyn ikääntyvillä miehillä ja naisilla. 108 p. (178 p.) Yhteenveto 2 p. 2011.

163 Hynynen, Esa, Heart rate variability in chronic and acute stress with special reference to nocturnal sleep and acute challenges after awakening. - Sykevariaatio-mittaukset kroonisen ja akuutin stressin seurannassa käyttäen hyväksi yöunen ja akuuttien tehtävien aikaisia vasteita. 74 p. (109 p.) Yhteenveto 3 p. 2011.

164 Pavelka, Béla, Open Water as a Sportscape. Analysis of canoeing in Finland for developing sport infrastructure and services. 116 p. 2011.

165 Pesonen, Jyri, Opettajat oppijoina. Toiminta tutkimus liikunnanopettajien pätevöittämis-koulutuksen käynnistämisestä ja kehittämisesestä. - Teachers as learners – An action research on starting the development of qualifying training for teachers of physical education. 204 p. Summary 4 p. 2011.

166 Borremans, Erwin, Asperger syndrome and physical exercise. A study about senso-motor profiles, physical fitness, and the effectiveness of an exercise training program in a group of adolescents with Asperger syndrome. 111 p. (181 p.). Yhteenveto 3 p. 2011.