From THE DEPARTMENT OF NEUROBIOLOGY, CARE
SCIENCES AND SOCIETY, DIVISION OF PHYSIOTHERAPY
Karolinska Institutet, Stockholm, Sweden

A TREATMENT-STRATEGY-BASED CLASSIFICATION SYSTEM FOR DECISION-MAKING IN PATIENTS WITH LOW-BACK PAIN

A biomedical approach
Development and inter-examiner reliability

Birgitta Widerström

Karolinska
Institutet

Stockholm 2012

Evidence does not make decisions, people do.
(Haynes, R.H., Devereaux, P,J., Guyatt, G.H. 2002)

To all clinical physiotherapists working with spinal pain

ABSTRACT

Background: Low-back pain (LBP) is a common, disabling and costly disorder and its treatment includes a range of interventions. Increased demands are put on effective care and rehabilitation offered to this population. Despite extensive research and sophisticated investigation methods, knowledge of the exact origin of LBP is limited and consequently approximately 80% of cases are classified as non-specific. To identify subgroups among these is therefore a priority research task. Physiotherapists can, through a careful patient interview and physical examination, distinguish different clinical presentations, classify these and then decide on appropriate treatment strategies.

Aim: The main purpose of this thesis was to develop and examine the inter-examiner reliability of a new treatment-strategy-based classification system for clinical decision-making in low-back pain patients in primary health care.

In **Study I** clinical data were collected for the classification system when 16 patients with low-back pain were examined, classified into four different treatment strategies-*pain modulation, stabilisation exercise, mobilisation,* and *training* - and treated according to one of these. The clinical examination and differences in specific clinical signs and symptoms were analysed and resulted in a classification algorithm, in which the classification process can be followed. All treatment was individualised. A progressive treatment flow towards increased physical loading and function as the clinical status improved was reported.

In **Study II** the inter-examiner reliability (agreement) of this classification algorithm was investigated. Two pairs of experienced physiotherapists trained in Orthopaedic manual therapy (OMT), with no previous experience of the classification system, examined and classified 64 adult patients with low- back pain. The agreement in their judgments was compared by calculating raw agreement (%) and the kappa coefficient (κ). Further, inter-examiner reliability was examined for five selected clinical signs and symptoms (examination items), identified as important for classification. Agreement was substantial (80%, $\kappa = 0.72$) when the two pairs classified patients into one of the four classifications. Agreement on the five specific clinical signs and symptoms was diverse. The assessments of *neurological signs and symptoms* had almost perfect agreement (92%, $\kappa = 0.84$), while those for *irritability* and *uni-or bilateral signs* were moderate (82%, $\kappa = 0.41$ and 62%, $\kappa = 0.42$, respectively). For the identification of a *specific movement pattern* and *specific segmental signs* the agreement was fair (68% $\kappa = 0.38$ and 67%, $\kappa = 0.28$, respectively).

Conclusion: The two studies in this thesis have presented and examined the inter-examiner reliability of a new treatment-strategy-based classification system for decision-making in patients with low-back, in primary health care. A classification algorithm where the differences in clinical status are described, and a progressive treatment flowchart, have been presented. The new classification system and three of its examination items can readily and reliably be used by experienced OMT-trained physiotherapists in primary care. The two examination items that had low agreement should be revised or clarified before future use in the classification system.

Key words: agreement, algorithm, classification, clinical decisions, inter-examiner reliability, low-back pain, physiotherapy

SAMMANFATTNING

Bakgrund: Ländryggssmärta är en vanlig, invalidiserande och kostsam sjukdom för individ och samhälle och dess behandling omfattar en rad olika interventioner. Ökade krav ställs på att den vård och rehabilitering som erbjuds denna patientkategori är verksam och effektiv. Trots omfattande forskning och sofistikerade undersökningsmetoder, är kunskapen om ryggsmärtans exakta orsak bristfällig och cirka 80% av fallen klassas som ospecifik ländryggssmärta. Därför är en prioriterad forskningsuppgift att identifiera undergrupper bland dessa. Sjukgymnaster kan genom en noggrann sjukhistoria och klinisk undersökning urskilja olika kliniska symtom och fynd och kan sedan behandla dessa med riktade behandlingsstrategier. Denna process innebär en klassificering av symtom, undersökningsfynd och behandling.

Syfte: Att utveckla och undersöka inter-bedömarreliabiliteten (överensstämmelsen mellan olika undersökare) hos ett nytt klassifikationssystem baserat på behandlingsstrategier, för patienter med ländryggsbesvär i primärvården.

I **Studie I** samlades kliniska data till klassifikationssystemet, när 16 patienter med ländryggsbesvär undersöktes och klassificerades till en av fyra olika behandlingsstrategier; *smärtlindring, stabiliseringsträning, mobilisering,* och *träning*, och därefter behandlades i enlighet med en av dessa. Den kliniska undersökningen och skillnader i specifika undersökningsfynd och symtom analyserades och resulterade i en algoritm, där klassifikationsprocessen kan följas. All behandling var individuellt anpassad. Dessutom redovisades ett progressivt behandlingsflöde mot ökad fysisk belastning och funktion i takt med förbättrat status.

I **Studie II** undersöktes inter-bedömarreliabiliteten för klassifikationsalgoritmen. Två par erfarna sjukgymnaster, vidareutbildade i Ortopedisk medicinsk terapi (OMT), men utan tidigare erfarenhet av klassifikationssystemet, undersökte och klassificerade 64 vuxna patienter med ländryggsbesvär. Deras bedömningar jämfördes genom att beräkna överensstämmelse i procent (%) och i kappa värden (κ). Vidare undersöktes inter-bedömarreliabiliteten för fem utvalda delmoment i den kliniska undersökningen som identifierats som viktiga för klassifikationen. Resultatet visade att överensstämmelsen var hög (80%, κ = 0.72), när de två paren klassificerade patienterna till en av de fyra klassifikationerna. Överensstämmelsen för de fem delmomenten i undersökningen var varierande. Bedömningarna av *neurologiska fynd och symtom* hade nästan perfekt överstämmelse (92%, κ= 0.84). För bedömningarna av *irritabilitet* respektive *uni- eller bilaterala fynd* var överensstämmelsen måttlig (82%, κ= 0.41, respektive 62%, κ= 0.42), medan den var låg för bedömningarna av *specifikt rörelsemönster* respektive *specifika segmentella fynd och symtom* (68% κ= 0.38, respektive 67%, κ= 0.28).

Sammanfattningsvis har de två studierna i denna avhandling presenterat och undersökt inter-bedömarreliabiliteten hos ett nytt klassifikationssystem för patienter med ländryggsbesvär baserat på behandlingsstrategier. En klassifikationsalgoritm med skillnader i kliniskt status, liksom ett flödesschema för behandling har beskrivits. Klassifikationssystemet kan på ett enkelt och tillförlitligt sätt användas av erfarna OMT-utbildade sjukgymnaster i primärvården. Tre delmoment av den kliniska undersökningen hade måttlig till nästan perfekt överensstämmelse, medan de två som hade låg överensstämmelse bör omprövas eller förtydligas innan de används i klassifikationssystemet i framtiden.

LIST OF PUBLICATIONS

I. Widerström B, Olofsson N, Arvidsson I. Manual therapy and a treatment based classification algorithm for patients with low-back pain. A pilot study. *J Back Musculoskelet* 2007;20:61-69

II. Widerström B, Olofsson N, Arvidsson I, Harms-Ringdahl K, Evers Larsson U. Inter-examiner reliability of a proposed decision-making treatment based classification system for low back pain patients. *Man Ther*, In press (2012) DOI. 10.1016/j.math.2011.12.009

CONTENTS

LIST OF ABBREVIATIONS

CI	Confidence interval
CPR	Clinical prediction rule
EMB	Evidence-based medicine
EBP	Evidence-based practice
HRQoL	Health-related quality of life
ICF	International Classification of Functioning, Disability and Health
ICD	International Statistical Classification of Diseases and Related Health 10th Revision
IFOMPT	The International Federation of Manipulative Physical Therapists
LBP	Low-back pain. Pain ache or discomfort, localised below the costal margin and above the gluteal folds with or without referred leg pain
MCID	Minimal important clinical difference
MDT	Mechanical diagnosis and therapy classification system (McKenzie)
MSI	Movement System Impairment classification system
NTPT	Neural tension provocation tests
OMT	Orthopaedic manual therapy
OSW	Oswestry low-back pain disability questionnaire
PCS	SF 36 subscale for physical health
PKB	Prone knee bend
ROM	Range of motion
SLR	Straight leg raise
TBC	Treatment-based classification system
TENS	Transcutaneus electric nerve stimulation
WCPT	World Confederation of Physical Therapy

1 INTRODUCTION

1.1 PREFACE

Patients with low-back pain (LBP) are one of the most common groups of patients I see and treat in the clinic. These patients with pain sometimes radiating to the buttock and/ or legs are a heterogeneous group with variation in symptoms, signs, duration, severity and disability. Diagnostic studies have failed to explain the pathology and/or pathophysiology behind LBP, and therefore a majority of these people are labelled as non-specific low-back-pain patients. All clinical physiotherapists are interested in the outcome of treatment and how the patients are best helped. In my experience the differences in clinical status are crucial for the decision on the treatment likely to be most helpful for the patient. Several randomized clinical trials comparing interventions do not address these differences. Instead patients are randomized into two or more 'treatment-arms' as if they were a homogeneous group. The results from these studies give limited information on how to match treatment with clinical status. I started to question what it was in the patient's clinical status that made me suggest acupuncture and not physical training as initial treatment and whether patients with similar clinical status could be identified. These questions led me into the field of treatment-based classification systems. These systems aim to identify diverse clinical presentations and determine interventions likely to be successful. A system that immediately caught my interest was the Treatment Based Classification System (TBC) first presented by Delitto et al[28], and further developed by Fritz and co-workers[39, 40, 42]. This impairment-based system has classifications that are commonly used in the management of LBP patients, but it also has classifications that are narrow, lacking a necessary flexibility for physiotherapists and patients. I found no existing classification system that acknowledged that physiotherapists commonly use techniques for pain relief and physical exercise as first-line treatments in patients with LBP. The need for a system that includes these frequently used treatment selections and provides a clinical flexibility was the starting point for the work presented in this thesis.

1.2 FRAMEWORK

This thesis concerns patients with LBP seeking physiotherapy treatment in primary health care. In the International Statistical Classification of Diseases and Related Health Problems (ICD-10, 2010) these patients are found in the dorsopathies subclassifications[125]. These include e.g. lumbago, lumbago with sciatica, sciatica, dorsalgia, spinal instabilities, segmental and somatic dysfunction. The main interest has been how these patients may be categorised in order to identify subgroups for which a specific treatment strategy is beneficial. For this purpose a decision-making treatment-based algorithm was developed as part of the present thesis. An algorithm is a description of a stepwise process with set criteria for the pathways in the algorithm which terminate in a result. In this case the result is a classification with a suggested treatment selection.

Theoretically the effect of LBP on the individual can be described by the World Health Organization's International Classification of Functioning, Disability and Health (WHO-ICF) model[126]. This conceptual framework provides a unified and standardised language to describe people's health from the perspective of body, individual and society.
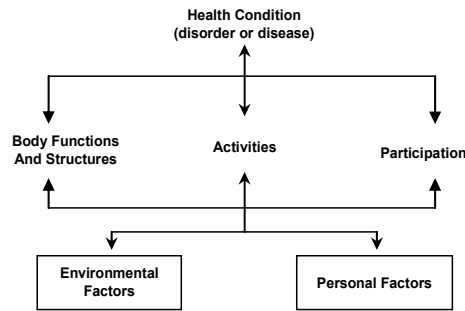
Figure 1. Interaction between the components of the ICF model[126]
(Reproduced with permission from World Health Organization. ID:96222)

As health is related to all components, the ICF has synthesised the components into a bio-psychosocial model; a complex interaction of physiological, psychological, personal and environmental factors (Figure 1).

Although LBP does not necessarily include structural changes by definition, it can cause loss of health due to impairments of body structures and functions, activity limitations and participation restrictions[34, 57, 103]. This loss of health may be caused by physiological events and be affected by personal and/or environmental factors and may have an effect on activity and participation.

In the work presented in this thesis the main concern has been body structure and function (impairments) and activity limitations, from a biomedical approach. These may be identified in the patient's medical history and a physical examination. This examination and the patient's clinical status (signs and symptoms) are fundamental for pre-treatment clinical decision-making[1]. Clinical practice shows that clinical status is not static but fluctuates in response to many factors such as movement, loading and psychological issues[5]. Clinical status will also differ depending on the phase of the clinical course when the patient is examined (Figure 2). Evaluation of the patient's response to physical treatment is essential for how treatment should be selected and adapted accordingly. It is considered in this work that, for full recovery and prevention, improvements in clinical status should lead to a treatment-flow with increased demands on physical function[18]. Similarly, lack of treatment response should lead to reconsideration of treatment selection and, on occasion, to a different classification, or referral for medical consideration. Further, it is considered that the patient's actual pathology – in most cases unknown – is consistent throughout a clinical course while the pathophysiology may differ, and that physical treatments influence this patho-physiology.
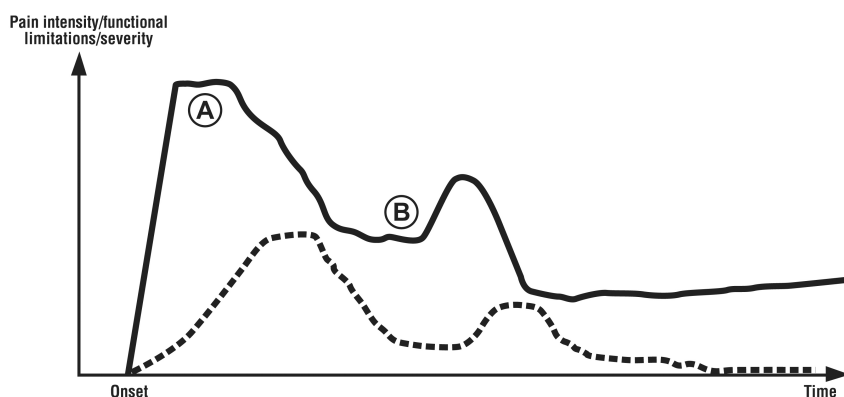
Figure 2. Different clinical courses. The black line illustrates an acute onset of low-back pain, with lingering symptoms. The dashed line illustrates an insidious onset with full recovery. At point A patients will have severe pain and limitations, while at point B pain and limitations have subsided.

There are many physiotherapy treatment methods and techniques with similar purpose. Instead of restricting treatment selections in each classification to one specific method or technique as most classification systems do, treatment strategies allow more than one single technique to be possible in each classification. It is here proposed that clinical practice in physiotherapy for LBP has four main treatment strategies, each with a specific purpose. They are to reduce pain and tension *(pain modulation)*, to provide dynamic stability and control to the lumbar spine (*stabilisation exercise),* to normalise or increase mobility *(mobilisation)*, and to increase motor timing, coordination and tolerance of spinal loading (*training*).

Evidence based medicine (EBM) was initially focused on applying the best research evidence to a clinical problem[35]. The evidence is determined in systematic reviews. These reviews select high-quality research on individual interventions and analyse the results to determine the effectiveness and subsequent evidence regarding the different interventions for a specific population. Upgraded versions of the practice of EBM, i.e. evidence-based practice (EBP), have emphasised that scientific evidence and an evidence hierarchy alone are not a sufficient and adequate guide to action[45, 54, 98]. In the updated model by Haynes et al[54], clinical expertise is a key element, a fourth element, that overlays the other three components showing the importance of the clinician's knowledge, skill and experience for the overall clinical decision-making (Figure 3).
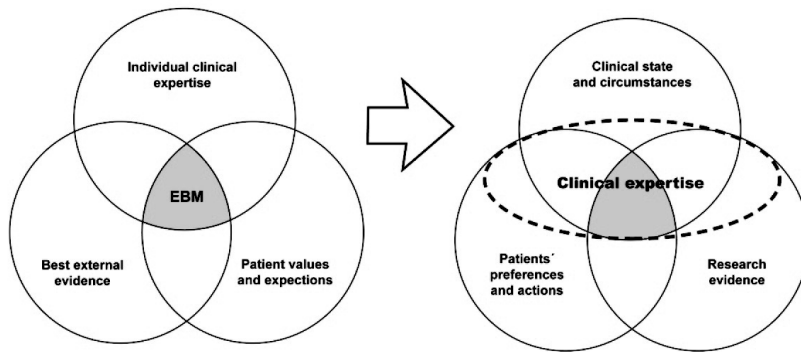
Figure 3. The early and the updated models for evidence-based clinical decision-making[54]

(Reproduced from Haynes RB, Devereaux PJ, Guyatt GH. Clinical expertise in the era of evidence-based medicine and patient choice. *Evid Based Med* 2002;7:36-38. With kind permission from BMJ Publishing Group Ltd. License number 2862600927472)

2 BACKGROUND

Low-back pain is as world-wide health problem and one of the most common reasons for patients in the Western countries to seek medical treatment[122]. LBP may be defined as "pain, ache or discomfort, localised below the costal margin and above the gluteal folds, with or without referred leg pain"[85]. Although often benign in nature, LBP stands for individual suffering and extensive cost to society. An investigation of the health-related quality of life (HRQoL) of patients with different diseases showed that LBP scored lower than angina pectoris, diabetes, asthma and neck and shoulder pain[17]. The expenditure and loss of productivity is substantial: in the United States health-care costs among back-pain patients increased by 65% from 1997 to 2005, more rapidly than overall health-care costs[82, 83]. In Sweden, statistics from 2009 show that back pain was the second most common reason for sick leave and that the expense for the Swedish social security system were 4,144 billion SEK[75]. Current research has found limited or conflicting evidence for improved outcomes with common physiotherapy interventions and optimal physical treatment for LBP remains unknown[6, 53, 77, 119, 120]. This indicates an urgent need for investigations on how these patients may best be helped.

The LBP patient group is not uniform, but includes patients with a cluster of signs and symptoms from the back, in different stages of impairment and disability. Individual interventions must be equally diverse. The outlook for the majority of LBP patients is generally recovery within three months, though recurrence, episodes with intermittent flares, is so high that is seems to be a part of its natural history and some sufferers do not fully recover[6, 56, 57, 66, 71, 103].

For most spinal disorders the major symptom is pain. Pain may arise from soft tissue (muscles, tendons and connective tissue), nerves, joints or bones[110]. Neither the considerable research aiming to establish the exact aetiology, nor sophisticated imaging techniques, have been able to determine an exact pathology in patients with LBP[7, 130]. This has resulted in a wide variation of diagnostic labels and nomenclature denoting spinal disorders (ICD-10)[125]. The interpretation and usage of these terms differ extensively depending on whether the diagnosis is made by a physician, a rheumatologist or an orthopaedic surgeon.

For the purpose of physiotherapy intervention, diagnostic labelling is even more diverse. Clinicians agree that LBP is a heterogeneous condition[64], but disagree on how to label disorders and on the most appropriate methods for classifying these patients. Classification systems that use anatomic site or pathologic process as the basis for differentiation result in a large group of non-specific LBP patients, without subsequent guidance on management. The limited high-quality-research evidence for the effectiveness of conservative management of LBP has resulted in a plenitude of practice patterns[64, 65, 93]. Therefore, a top research priority is to find reliable and valid classification methods for the non-specific LBP population, to identify specific sub-groups and consequently their specific physiotherapy management.

2.1   LOW-BACK PAIN: PATHOLOGY AND PAIN MECHANISMS

In most cases LBP is not a sign of severe pathology[20]. Nevertheless, screening for red flags, i.e. severe medical pathology such as infection, tumour, inflammatory process, fracture or radicular syndrome, is pertinent so that appropriate medical investigations and treatments can be undertaken. Imaging studies have indicated that LBP can occur although lumbar anatomy is normal[7]. The development of non-specific LBP is therefore believed to be multi-factorial, potentially related to combinations of physical characteristics, genetic, behavioural, psychological, anatomical and societal factors[44]. The factors of social, psychological and cognitive origin that influence the patients' pain, i.e. yellow flags, may be addressed by a screening process using specific questions during the patient interview. When yellow flags are considered a dominant factor for the LBP, the patient should be advised to seek the appropriate treatment in addition to physiotherapy.

Pain is often the major symptom and of the greatest concern for the patient. Pain is also one of the most sensitive measures when treatment effects are assessed in LBP. Symptom relief, daily functioning and work status are more associated to outcome than are range of movement (ROM) and back strength, and are therefore important to address and monitor in treatment[6, 81]. One can expect that, for many patients, a mixture of anatomical structures such as the intervertebral disc, the zygapophysial joints, ligaments and muscles are involved in their pain[110]. In most patients seen by physiotherapists in primary care the local back pain experienced is nociceptive[109]. Nociceptive pain is a response to noxious (painful) stimuli of sensory receptors capable of transducing noxious stimuli (nociceptors) as a result of inflammation, oedema, or ischemia, caused by trauma or repetitive or excessive mechanical loading (pressure or tension)[89, 134]. This noxious stimulus is modified in the spinal cord and brain by peripheral and central mechanisms. Peripheral sensitisation refers to an increased responsiveness, reduced threshold of nociceptors and an increase of receptive field size, mediated by several pain- and inflammatory substances[44, 110]. Central sensitisation occurs in response to the peripheral neural events described above, giving increased excitation and/or decreased inhibition of central neurons[106, 110]. These sensitisation mechanisms lead to that stimuli of neighbouring uninjured areas may be experienced as hurting, and may also cause innocuous (non-painful) stimuli to be experienced as painful. In the clinic this may be seen as pain and tenderness over a large area, and/or distant to the site of injury and increased response to painful stimuli (hyperalgesia), and tenderness to gentle touch (allodynia).

A subgroup of LBP patients, approximately 10%, have peripheral neurological signs and symptoms, e.g. leg pain, motor and/or sensory disturbances, indicating nerve root symptoms, indicative of nerve tissue damage[51, 101, 116] These symptoms have been associated with disorder severity and prediction of chronicity, work absence and higher health-care costs[101]. The patients often have a prolonged healing process, and therefore need longer treatment and more carefully- dosed and-progressed interventions than patients without these signs and symptoms[5, 116].

2.2 PHYSIOTHERAPY AND ORTHOPAEDIC MANUAL THERAPY

Human movement is the central concept in physiotherapy. There are subspecialisation areas in physiotherapy among which Orthopaedic manual therapy (OMT) is one. The International Federation of Manipulative Physical Therapists (IFOMPT) defines OMT as "a specialized area of physiotherapy/physical therapy for the management of neuro-musculo-skeletal conditions, based on clinical reasoning, using highly specific treatment approaches including manual techniques and therapeutic exercises. OMT also encompasses … "the available scientific and clinical evidence and the bio-psychosocial framework of each individual patient" (www.ifompt.org). In general, manual therapy is a term referring to thrust and non-thrust techniques, but sometimes also to other hands-on treatment procedures such as soft-tissue techniques and massage. OMT is a postgraduate specialisation in physiotherapy while basic OMT techniques are part of undergraduate education.

2.2.1 Physiotherapy examination

The physiotherapy examination procedure for LBP include four equally important parts; patient interview, active movement examination, peripheral neurological examination and passive movement examination. The interview will yield information on; how the patient experiences the disorder; the area and nature of the pain; the progression of the disorder; earlier treatment and treatment response; other medical problems possibly associated with the LBP and activity limitations.

The active movement examination will identify posture and how the patient can move in daily life. It is focused on impairments (deficits in mobility, balance and/or coordination) and associated pain. Active stability tests recognise the active control of the spine during specific movements where spinal control is pertinent (e.g. single-leg stand, active straight leg raise).

The passive movement examination will yield information on mobility, including segmental movement; range, quality and associated pain. The range may be denoted as normal, hypo- or hyper mobile. Quality refers to the characteristic end-feel of each joint and depends on the anatomy of the joint and the direction of the movement tested[62]. In the spine, segmental signs may either be unilateral, bilateral, or bilateral but predominantly unilateral.

Neurological examination includes active and passive tests and will identify altered reflexes and/or sensation, motor disturbances (e.g. muscle weakness) and/or altered neuro-dynamic function. Neuro-dynamic tests comprise tension tests; the slump test; straight leg raise (SLR); prone knee bend (PKB); and palpation of neural tissue (the sciatic and femoral nerves)[46].

2.2.2 Physiotherapy interventions

The scientific evidence for most physiotherapy interventions is yet limited due to small effect sizes and short-term benefits[76, 114, 119]. However, European and American clinical guidelines for the management of LBP recommend to, "..stay active, self-care options (advice) and use medication with proven benefits". For those who do not improve,"..consider spinal manipulation for acute LBP" and for sub-acute or chronic LBP, ".. consider exercise, acupuncture, spinal manipulation, yoga, relaxation, cognitive-behavioural therapy and multidisciplinary rehabilitation for non-specific LBP[22, 70]. Of these guideline-endorsed interventions for sub-acute or chronic LBP some

are cost-effective, exercise, acupuncture, spinal manipulation/mobilisation and multidisciplinary rehabilitation, while staying active, advice, medication, massage, yoga or relaxation are not[76]. Regardless of the magnitude of research support physiotherapists use interventions to alleviate pain and normalise function such as transcutaneous electrical nerve stimulation (TENS), hot or cold packs, ultrasound, low-intensity laser, taping, acupuncture, massage, trigger-point techniques, joint mobilisation or manipulation, traction, neuro-dynamic techniques, and active exercises[47, 93]. The rationale for their use is probably multidimensional. The observation of patient improvements, individualised treatment and the use of concomitant interventions may be parts of this rationale.

In the present work, physiotherapy interventions for musculoskeletal disorders follow a four-step process; pain alleviation, movement normalisation; movement control and tolerance of loading (Figure 4). These steps sometimes overlap. For patient expectations, confidence and reassurance at the start of treatment, the physiotherapist's control and activity are pertinent. As treatment proceeds the patient's activity and responsibility will increase. This is to meet higher physical demands that will be put on the patient as he/she improves, but also for future self-management of his/her condition and the prevention of recurrence[18, 19]. As pain is often the major symptom, of most concern for the patient and the main reason to seek physical treatment, it must be monitored throughout the whole process. Normalisation or restoration of normal function to a joint may either include active mobility exercises and /or passive mobilisation techniques. Both active and passive techniques may be specific or general. Many different techniques are described and used[62, 80, 86]. Movement control is achieved through guided and graded active exercises specifically addressing motor timing and coordination and has to be achieved before loading the spine with more weight or complex movements[97]. Tolerance of loading is considered as the last step in the rehabilitation process and will include a mixture of loaded, complex and combined active exercises[32].
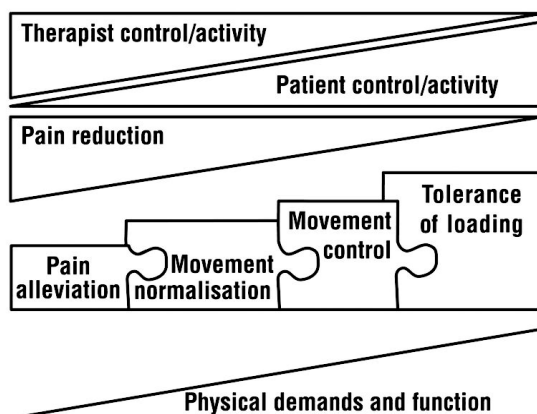

Figure 4. Physiotherapy intervention process for muscoluskeletal disorders considered in this thesis

2.3 CLINICAL DECISION-MAKING

Clinical decision-making or clinical reasoning – the two terms are used interchangeably – refers to the cognitive process used by medical professionals in the evaluation and management of a patient[61]. This process has important consequences for patients, as it will guide intervention, and is a challenge for all clinicians[123]. Applied to physiotherapy this includes; collecting and analysing information and generating hypotheses concerning the cause or nature of the patient's problem (patient interview); testing these hypotheses through further data collection (physical examination) and, determining optimal diagnostic and treatment selections and prognosis (clinical and scientific evidence). Elstein et al[33] concluded that clinical reasoning is specific to one's area of work and depends on the clinician's organization of knowledge in a particular area. Relevant to physiotherapy this includes; facts (anatomy, pathology and pathophysiology, sources of pain and dysfunction); procedures (examination and treatment strategies); concepts (e.g. instability, positive neurological signs, sensitisation mechanisms); principles (treatment selections, extent of treatment, precautions and contraindications); and patterns of presentations (clusters of symptoms and signs)[33]. Further, full competence in physiotherapy includes experience, intuition and social, manual, communication, and clinical skills.

A clinical reasoning model for physiotherapists has been described by Tyni-Lenné[117]. This model has five parts; examination, diagnosis, goals/planning, intervention and evaluation: each part can be related to the ICF terminology. The patient's problem can be examined, analysed, and diagnosed in terms of body function and structure, activities, participation, environmental and personal factors. Goals, interventions and evaluation can be determined in terms of changes/ improvements in functioning and disabilities as well as in contextual factors.

An illustration of the physiotherapy clinical reasoning with reference to the different concepts presented above is presented in Figure 5.
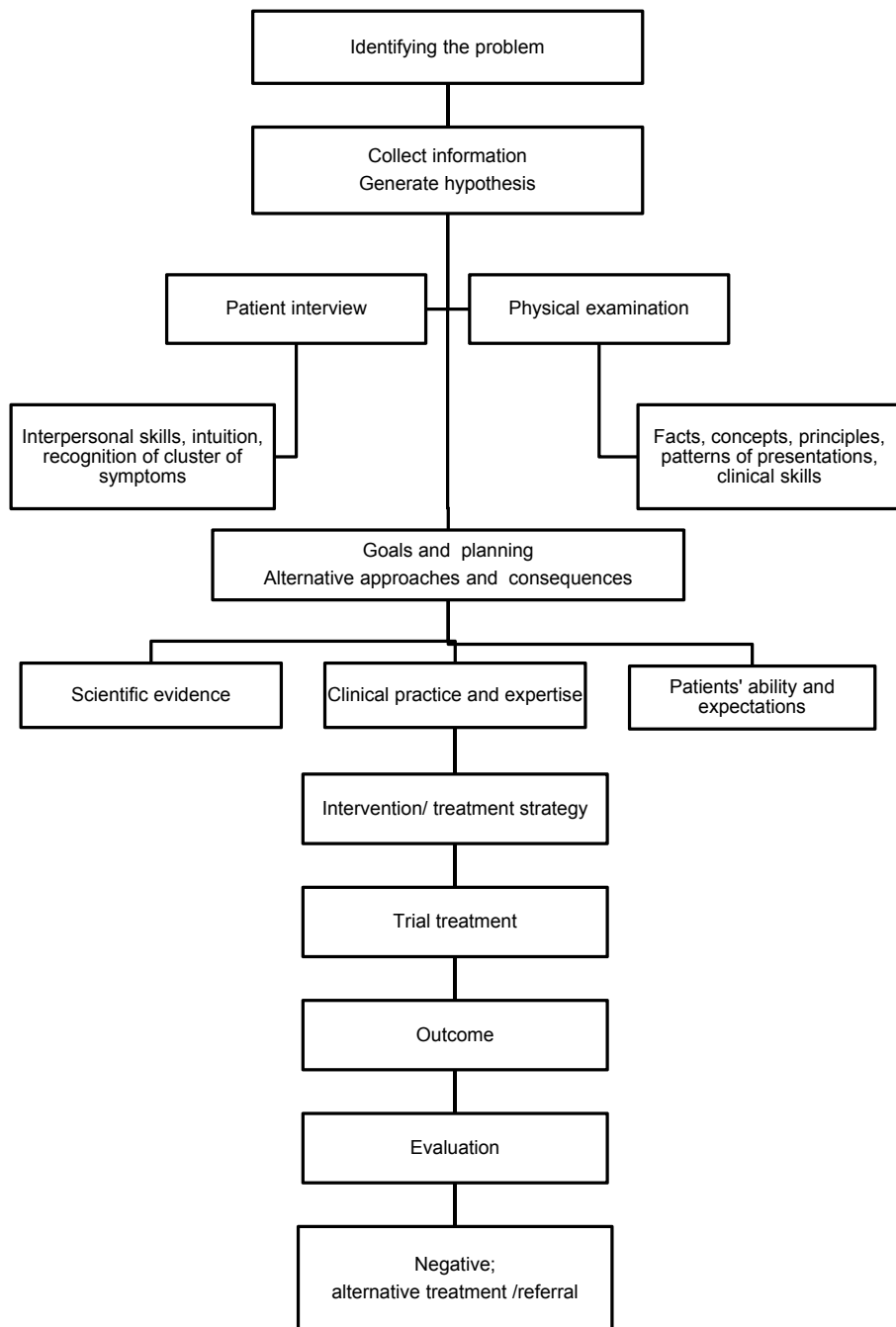
Figure 5. The physiotherapy clinical decision-making

2.4 CURRENT LOW-BACK PAIN CLASSIFICATION SYSTEMS

Classification systems refer to theoretical and clinical models in which patients can be categorised into classifications[85]. These have specific attributes to which patients may be associated. The systems are often based on an algorithm; a description of a step-by-step procedure which terminates with a result. Criteria for the pathways in the algorithm and the resulting classification are set. These criteria may derive from hypotheses, theories, clinical experience, expert opinion, and/or study results. The terms; 'classification model' or 'classification system' are often used interchangeably, and so are 'subgroup', 'category' and 'classification'. The terms classification systems and classifications are used consistently throughout this thesis.

In 2007 Billis et al[10] identified 39 different diagnostic and treatment-based classification systems. Three classification paradigms were identified; biomedical, psychosocial and bio-psychosocial (Figure 6). Psychosocial models are designed for use in medical or multi-disciplinary settings, while in physiotherapy settings most models follow the biomedical paradigm and just a few have a mixed bio-psychosocial approach. The majority of systems are based on a judgmental approach, relying on clinical experience and intuition. The biomedical paradigm consists of two main systems with a pathoanatomical or a clinical features/ impairment orientation[10]. Pathoanatomical systems focus on diagnosis and classify into syndromes, each assumed to refer to a specific pathological condition without guidelines for treatment[69, 91]. By contrast, impairment systems classify patients based on clusters of signs and symptoms to guide treatment, without assumptions about pathoanatomical causes. Several impairment-based systems have been presented [28, 84, 100], and some have been found valid with good inter-examiner reliability[40, 48, 67, 118, 131]. However, some may be considered incomplete; others complicated and time-consuming; some include clinical features and nomenclature not commonly known by physiotherapists; or they require specifically-trained physiotherapists, limiting their utility and generalizability. One impairment-based system, the TBC system[28, 42] has been of special interest and inspiration, and now forms part of the present new classification system. The TBC system has been investigated in several studies and has shown preliminary evidence of the effectiveness of the decision-making classification[41] and, further, evolving support for classification and matched physiotherapy treatment may result in better clinical outcomes[13]. It has also shown moderate-to-good inter-examiner reliability[40, 42]. However, the TBC system was developed and preliminarily validated in patients with acute exacerbation of LBP and, further, has classifications that are narrow, lacking a necessary clinical flexibility for physiotherapists and patients. Two classifications; traction and specific exercises, are specific treatment selections for patients with signs of nerve-root involvement, and for those who will respond favourably to repeated end-range movements as described by McKenzie[84], respectively. The single use of these treatments has not been fully supported in systematic reviews[23, 77]. As there may be other treatment selections from which these two subgroups of LBP patients might benefit, this restricts the clinical utility of TBC system.
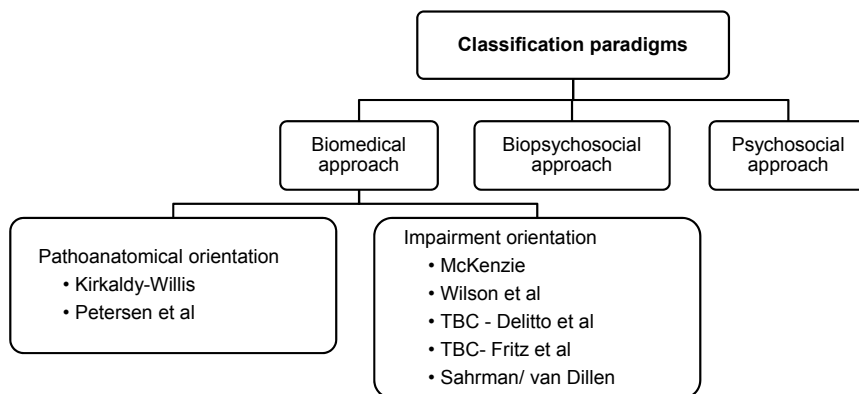
```
                    ┌─────────────────────────┐
                    │  Classification paradigms │
                    └─────────────────────────┘
        ┌──────────────────┬──────────────────┬──────────────────┐
  ┌───────────────┐ ┌────────────────┐ ┌───────────────┐
  │  Biomedical   │ │ Biopsychosocial│ │  Psychosocial │
  │   approach    │ │    approach    │ │   approach    │
  └───────────────┘ └────────────────┘ └───────────────┘
```

Pathoanatomical orientation
 • Kirkaldy-Willis
 • Petersen et al

Impairment orientation
 • McKenzie
 • Wilson et al
 • TBC - Delitto et al
 • TBC- Fritz et al
 • Sahrman/ van Dillen

Figure 6. The three current classification paradigms according to Billis et al[10]

2.5 RELIABILITY AND VALIDITY

Reliability refers to how far a test, method or instrument measures the same attribute each time it is used. It has to do with consistency, reproducibility and repeatability and can be defined as the degree to which a test or measure is free from error. There are different aspects of reliability in clinical testing; test-retest reliability, intra-examiner reliability and inter-examiner reliability. Test-retest reliability concerns the consistency of repeated measurements over time, when subjects are believed to be stable concerning the measured attribute. Intra-examiner reliability refers to how consistent repeated measures made by the same examiner on two or more occasions are, while inter-examiner reliability refers to agreement between two or more examiners[102]. For a classification system to be clinically useful, good inter-examiner reliability is crucial as it shows that the system can be applied consistently by different clinicians. The simple approach to assessing inter-examiner agreement is to calculate how many exact agreements were observed, denoted as raw agreement measured in percentage. Raw agreement does not account for agreement just by chance; hence, a chance-corrected measurement is needed. This chance-corrected measure of agreement is called kappa (κ). It has a maximum of 1.00 when agreement is perfect. A value of zero indicates no agreement better than chance[4]. Though examining inter-examiner reliability is pertinent, good inter-examiner reliability is not sufficient for a method to be considered valid. Validity refers to the degree to which an instrument or test measures what it intends to measure[27]. The different types of validity are: face, content, construct and criterion validity. The different types have to be established prior to generalisation of an instrument or test in clinical work.

2.6 RATIONALE FOR THE THESIS

LBP is a common disorder with suffering for the individual and high costs for society. Many of these patients are treated by physiotherapists and there is a need for improved management for this patient group. Although several classification systems have been presented in the literature and some are reliable and valid, all have limitations and are not necessarily readily applied in clinical practice. The literature revealed no classification system that included a warranted flexibility in treatment selections suitable for clinical practice. Further, the existing classification systems have not

clearly acknowledged the growing support for the insight that individualised pain treatment and physical training are beneficial for LBP patients[3, 114, 119, 120, 133].

The work reported in this thesis aimed for a classification system that is; as inclusive as possible for LBP patients seeking physiotherapy in primary health care; is easy to understand; does not require extensive familiarisation or specific equipment; considers examination time limits; includes known clinical features and common treatment selections, and provides clinical flexibility for patients and physiotherapists. The present work has initiated the development of a system that includes these criteria, and provides evolving evidence for its future utility in clinical practice.

2.7   SUMMARY OF PROBLEM AREAS

Problems relevant in this area are:

- LBP is one of the most common reasons for patients to seek medical treatment, indicating an urgent need to find out how these patients best may be helped

- LBP is a heterogeneous condition that needs individualised and varied interventions

- Diagnosing LBP is difficult and may be viable only in approximately 10 % of cases

- Classification systems may be one way to identify subgroups and the optimal physical treatment for each of these subgroups

- There is a need for a classification system that is feasible and dynamic for patients and physiotherapists

3 AIMS

The overall aim of the work presented in this thesis was to develop and describe a treatment-strategy-based classification system for decision-making in patients with non-specific low-back pain, and examine its inter-examiner reliability.

Specific aims

Specific aims were

- to describe differences in clinical status for each classification (Study I)
- to describe the classification process so it can be used by physiotherapists in clinical practice (Study I)
- to present a classification system that allows for a progressive treatment-flow with adaptation to change and improvements in clinical status (Study I)
- to examine the inter-examiner reliability of the new treatment-strategy-based classification system and five of its specific examination items (Study II)

4 METHODS

4.1 DESIGNS AND ETHICAL APPROVALS

This thesis is based on two studies. The first part of Study I is descriptive, resulting in an individualised clinical decision-making algorithm. The second part is a multiple case study using a pre-post-test design. Study II investigates inter-examiner reliability, employing a mixed and simultaneous examiner design. For both studies no data could be linked to any individual, and patients could withdraw at any time without giving any reason. Participation or non-participation would not influence future physical treatment. The studies were approved by the Regional Ethics Committee in Umeå (Study I) and the Regional Medical Research Committee in Stockholm (Study II).

4.2 STUDY SAMPLES AND SETTINGS

In both studies the patients were a convenience sample of adult, consecutive, consenting patients with LBP, who sought physiotherapy treatment at outpatient clinics in primary health-care. The clinic in Study I (n=16) is situated in Östersund in the northern part of Sweden, while the clinics in Study II are located in two different parts of Greater Stockholm, Sweden, one suburban (n=34) and one urban (n=30). For both studies patients were given written and oral information about the study and gave their informed consent to participate. Included were those with LBP regardless of duration, with or without radiating pain to the lower extremities and with no difficulty understanding the Swedish language. Exclusion criteria were previous back surgery, pregnancy, and known neurological or rheumatic disease.

4.3 EXAMINERS

The single examiner in Study I was a physiotherapist with 27 years of clinical experience (the author), specialising in OMT, with a master's degree in physiotherapy and OMT. The four volunteer examiners in Study II were all experienced (8-25 years), but had various levels of OMT training. Two of the examiners (pair A) had master's degrees in OMT, while in pair B one had a university postgraduate certificate, and the other a clinical postgraduate certificate, in OMT. To ensure examiner autonomy, crucial for reliability studies, none was involved in the formation of the algorithm and all worked geographically far from the developer of the classification system.

4.4 CLINICAL EXAMINATION PROCEDURE

The clinical examination procedure used in both studies followed the process outlined in section 2.2.1 (Physiotherapy examination). The patient interview focused on area and course of symptoms, history of injury and changes over time, general health and level of irritability[80, 134]. This level was determined to be mild, moderate or high, using two questions; how easily symptoms were aggravated by activity, and the estimated time for symptoms to subside after aggravating activity. The physical examination had two parts. In the first the examiner observed the patient´s posture, malalignments and signs of muscle hypotrophy, after which he/she instructed, observed and judged the patient´s active movements. Active movement examination was performed in all anatomical movement planes and focused on identifying movement patterns, denoted as present or

not (Box 1). In patients with aberrant movement patterns, active stability tests were performed. These tests, where active control of the lumbar spine is tested, were at the examiner's discretion and could include single-leg balance, single active straight leg raise, static and/or dynamic lunges and single-leg-hip flexion in sitting. These were judged positive or negative when performed with poor or good control of the spine, respectively.

Box 1. The different movement patterns used in the new classification system

Aberrant	Specific	Non-specific	Multidirectional
• Deviation during movements and/or • Painful arc and/or • Reversed lumbar-pelvic rhythm and/or • Thigh-climbing	• Pain and limitation in a flexion/opening/tension pattern (flexion and lateral- flexion to the opposite side from the pain) or • Pain and limitation in an extension/closing/compression pattern (extension and lateral-flexion to the same side as the pain)	• A mixture of flexion and extension patterns	• Pain and limitations in all movement directions

In the second part of the physical examination the examiner performed passive movement- and neurological examinations. The passive and accessory movement testing sought to evaluate spinal segmental mobility and pain response to the testing. Mobility was denoted as hypomobile, normal or hypermobile. The signs and associated pain were denoted as 1) unilateral, 2) bilateral or 3) bilateral but predominantly unilateral. In patients with radiating pain to the lower extremities, a peripheral neurological examination was performed. It included nerve conduction tests; muscle strength, reflexes and sensation, denoted as positive or negative (normal). In patients with radiating pain but normal neurological tests, neuro-dynamic tests were performed. These tests were; the slump position; straight leg raise (SLR); prone knee bend (PKB) and palpation of nerve structures[46]. All these tests were denoted positive or negative.

In the inter-examiner reliability study (Study II) the examination procedure had to consider systematic bias. Therefore, the examiners in each pair were assigned number 1 or 2, changing for every other patient (Study II, Figure 2). To minimise patient variability and ensure that the examiners were given the same information, both examiners were present during the patient interviews and active movement testing, but only examiner number 1 questioned the patient and instructed on active movements. As active movements may change with repeated examination, these were carried out once. This single-active-movement examination enabled the examiners´ judgments to be based on the same information, but still be independently interpreted. Each examiner separately performed the second part of the examination.

4.5   MEASUREMENTS

Baseline data, age and symptom duration, were obtained orally during the patient interview. It has been proposed that a battery of instruments should be used for establishing pain intensity, well-being and level of disability in the LBP population as well as for measuring treatment outcome [31]. Three self-reported instruments were used to meet these requirements. The Borg CR 10 scale[12] was used to assess pain intensity (Studies I and II). The Swedish version of the Oswestry low-back pain questionnaire (OSW) [36] was used to measure functional disability (Studies I and II) and the Swedish version of the SF 36[113] was used to measure well-being (Study I). All three self-reported instruments were also used for outcome measurements in Study I.

4.6   THE ALGORITHM AND THE NEW CLASSIFICATION SYSTEM

The algorithm and the new classification system are based on the clinical decision-making described in section 2.3 (Clinical decision-making) and in Figure 5. Further, the new classification system is partly based on the TBC system[28], itself based on patient interviews and clinical examinations to categorise patients with into one of four treatment classifications mobilisation, stabilisation, specific exercise or traction. The new system uses two classifications similar to the TBC system's *stabilisation exercise* and *mobilisation*, plus two new ones *pain modulation* and *training*. The latter two were formed empirically from clinical practice based on the observations that individual pain treatment and physical exercise are commonly used by clinical physiotherapists and that patients seem to benefit from these treatments. The framework for the development of the new treatment-strategy-based classification system is presented in Figure 7.
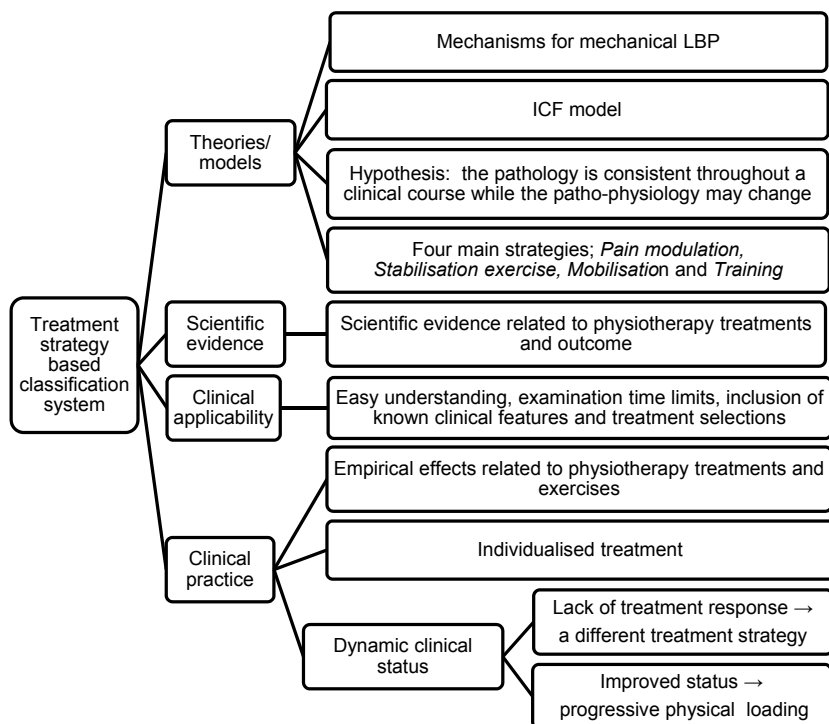
```
                    ┌─────────────────────────────────────────┐
                    │        Mechanisms for mechanical LBP     │
                    └─────────────────────────────────────────┘
                    ┌─────────────────────────────────────────┐
                    │               ICF model                  │
        Theories/   └─────────────────────────────────────────┘
        models      ┌─────────────────────────────────────────┐
                    │ Hypothesis: the pathology is consistent  │
                    │ throughout a clinical course while the   │
                    │ patho-physiology may change              │
                    └─────────────────────────────────────────┘
```

Framework figure:

- Treatment strategy based classification system
  - Theories/models
    - Mechanisms for mechanical LBP
    - ICF model
    - Hypothesis: the pathology is consistent throughout a clinical course while the patho-physiology may change
    - Four main strategies; *Pain modulation, Stabilisation exercise, Mobilisation* and *Training*
  - Scientific evidence
    - Scientific evidence related to physiotherapy treatments and outcome
  - Clinical applicability
    - Easy understanding, examination time limits, inclusion of known clinical features and treatment selections
  - Clinical practice
    - Empirical effects related to physiotherapy treatments and exercises
    - Individualised treatment
    - Dynamic clinical status
      - Lack of treatment response → a different treatment strategy
      - Improved status → progressive physical loading

Figure 7. Framework for the algorithm and the new classification system

4.6.1   The classifications

The examination procedure and the combination of examination signs and symptoms that each classification embraces are presented in Study I and Study II (Table 2). A description of the patient characteristics and treatment selections; aims, possible effects and evidence, for each classification follows.

Pain modulation
The *pain modulation* classification was formed empirically to cover patients with the most severe symptoms and difficulties to perform daily activities. These patients may have pain at rest and in several active movement directions. Spinal passive movement evaluation may be inconclusive due to perceived pain with movements and testing. Due to the severity of signs and symptoms, patients with signs of nerve-root involvement (radiating pain to the leg and altered reflexes, sensation and power), and patients with positive neuro-dynamic tests (radiating pain to the leg but no motor and or/sensory disturbances)[108] were classified to *pain modulation.*

The suggested treatment selections in this classification aim to reduce pain intensity and enhance relaxation. These may be acupuncture, TENS, soft-tissue techniques including trigger- point techniques, traction techniques in pain-free positions and low-grade mobilisations I-II, i.e. in pain-free position, with large amplitude, slow, smooth and gentle, off the resistance of the joint[62]. Treatments for patients with nerve-root

involvement may be traction[72], acupuncture[133] or specific extension-oriented exercises as described by McKenzie[84]. For patients with mechanical sensitisation of neural tissue, neuro-dynamic treatment techniques could be considered as a treatment selection[46].

There is moderate evidence that acupuncture, and soft-tissue techniques, reduce LBP[43, 76, 133]. The pain-modulating effects of manual techniques are yet not fully understood. However, it is expected to include mechanisms such as mechanoreceptor stimuli resulting in neurophysiological responses[9].There is basic scientific evidence that TENS has an analgesic effect but, due to poor study design and small sample sizes in clinical trials, TENS is not fully proved to relieve LBP in patients[30]. The efficacy of traction is unclear because of generally poor study design and because those patients most likely to benefit have not been specifically studied[3]. However, it is suggested that traction benefits patients with LBP and radicular pain and concomitant neurological deficit[49, 72]. Extension-oriented exercises may be effective in patients with LBP and distal symptoms[15].

Stabilisation exercise

The *stabilisation exercise* classification was adapted from the TBC system and covers a sub-group of LBP patients who have decreased capacity controlling segmental movements[90, 95, 112]. These patients are found to be young, have excessive ROM and possibly increased segmental mobility (hypermobility), aberrant active movements and positive active stability tests[13, 21, 59, 112]. These clinical findings and a history of recurrent symptoms and major limitations caused by minimal provocations have been included as key features in this classification[68].

The treatment selection under *stabilisation exercise*, specific retraining and co-activation of the deep abdominal and spinal muscles aims to provide dynamic stability to the lumbar spine and reduce associated pain[88, 94, 95]. These exercises differ from general exercises being more specific and require more attention and precision from the patient. They should be carefully and individually dosed and, most importantly, slowly graded into loaded positions[94]. For load and grade progression a multitude of exercises are described[94, 97]. Many different tools can be used; Swiss balls, balance plates, weights and pulling machines. The selection of exercises will be guided by the experience and skill of the treating physiotherapist and by the patient's ability to perform the exercises accurately.

The loss of a normal pattern of spinal motion and control is considered to cause pain and/or neuromuscular dysfunction[2, 88, 89]. This has found some support in studies using imaging techniques showing a correlation between segmental hypermobility and high incidence and slow recovery from LBP[2, 63, 73]. Management using stabilising exercises reduces disability, pain and the recurrence of LBP[60, 88, 94].

Mobilisation

The *mobilisation* classification is adapted from the TBC system and covers LBP patients with lumbar hypomobility but without distal neurological signs and symptoms[24, 38]. In the new system, it also covers patients with a specific movement pattern (Box 1). It has not yet been established whether thrust techniques are suitable alternatives to non-thrust mobilisations: the terminology is inconsistent and does not distinguish between thrust and non-thrust techniques. In clinical practice one meets many patients with LBP with hypomobility and non-radiating pain with long-term-fluctuating symptoms. For these patients, mobilisation techniques may be a better treatment selection than manipulations, which reportedly benefit patients with short duration of symptoms[38].

The treatment selections under *mobilisation* may be active mobility exercises, passive manual mobilisation techniques[62, 80] and/or a combination[86]. They aim to normalise or increase lumbar mobility. Passive manual mobilising techniques, traction, compression or gliding[62, 80] may be used when patients are unable to perform active exercises due to stiffness and/or pain. These mobilisations (grades III-IV) are carried out near end-range of the joint, more firmly, at higher speed, smaller amplitude and longer duration, than grades I-II. Manipulation (grade V) refers to thrust techniques with low force and high velocity. None of the techniques should provoke pain, although brief discomfort may be accepted, as long as the patient is informed and consents. Several mobilisation-methods are described, as well as manipulation techniques[62, 80, 86]. No specific method has been proved superior to another, so no restrictions are made under *mobilisation*.

The use of spinal manual mobilisation/manipulation is guideline-endorsed and reportedly cost-effective for sub-acute and chronic LBP[3, 14,76]. The exact mechanism of the mobilising effect of passive mobilisation and manipulation is not clear[11]. Early concepts of pain-modulating effects of manual techniques have been predominantly mechanistic in nature, such as moving joint inclusions or disc fragments, dividing adhesions or repositioning sub-luxed vertebral segments[25, 80]. Later theories have proposed that manual therapy is a stimulus that might affect the nervous system[105, 107], such as inhibition of nociceptive afferent input to the spinal cord (gate control theory) or inhibition of muscle spasm due to a decline in neural discharge with repeated movements[134]. Recent research suggests that manual mobilisation techniques are likely to have multiple effects yet not fully understood[9, 105, 128, 129]

Training

The *training* classification was formed empirically to cover patients with symptoms in remission who seek physiotherapy to increase function and prevent recurrence[19]. It also includes patients who have been in one of the other classifications and improved so that training can further improve their function.

The treatment selection in this classification, physical exercises, aims to improve function and increase tolerance of loading. These include a warm-up before an individualised progressive exercise programme including exercises for; spinal mobility; balance; fitness, lower-extremity strength; coordination between extremities and trunk, and control of the trunk during complex whole-body movements. In contrast to the *stabilisation exercise* classification, initial exercises in the *training* classification have higher loading and demands on function, and a more rapid progression. There are extensive exercises and tools to be used and to date there is no evidence that one

specific mix of exercises is more efficient than another, leaving the selection to the treating physiotherapist and the ability and preference of the patient.

Scientific evidence for short-time benefit of physical training and exercises in patients with LBP is increasing[76, 114, 119, 120]. This benefit and the necessity of muscle training are supported by research findings of muscle alterations in LBP patients, leading to muscle fatigue [29] and/or deficits in normal timing and recruitment (motor function) of the back muscles, not always spontaneously resolved when symptoms alleviate[79]. Further, patients with recurrent LBP have altered and rigid postural control strategies[16]. These findings are suggested as factors for recurrence, making a mix of exercises addressing these functions important for prevention. While some studies have shown that physical exercises have positive effects on pain and disability[114, 119, 120], it is neither clear how patients are best selected for exercise therapy, nor what exercises or dosages are most beneficial.

4.6.2 The specific examination items

Information from all parts of the examination procedure (section 2.2.1) was used in the algorithm. From the patient interview, radiating pain, pain in rest, clinical instability symptoms, neurological symptoms and level of irritability were stressed. The active movement examination stressed the identification of different movement patterns. This identification has been described and used in several classifications systems[84, 87, 91, 99] and is considered in some as a key feature for intervention[28, 84]. The passive movement testing stressed; the passive segmental movement range and quality (normal-hypo-or hyper mobility); the identification of symptomatic segmental level by reproducing perceived pain; whether the signs were specific or multilevel, and whether the signs were uni-or bilateral. The peripheral neurological examination stressed conduction deficits (altered reflexes and/or sensation, motor disturbances)[132] and/or altered neuro-dynamic function[46].

Among these signs and symptoms five key characteristics were selected by the developer as specific examination items. These items give information on the severity of the disorder, direct the selection of classification and guide how treatment within the classification may be performed account taken of extent, manner and dosage The five specific examination items included in Study II were labelled; 1) level of irritability, 2) specific movement pattern, 3) specific segmental signs, 4) uni-or bilateral signs and 5) neurological signs and symptoms.

4.6.3 Familiarisation

The two pairs of examiners in Study II were familiarised with the algorithm during a single three-hour session at each clinic. The procedure was outlined and clinical decisions, main characteristics and possible treatment selections for each classification were explained. The studies were performed in clinical practice and intended to mirror everyday clinical work, therefore the examiners were instructed to maintain their ordinary examination procedure. The specific examination items, outlined in a checklist, were presented and discussed (Study II, Box 1).

4.7   DATA ANALYSES

All statistical methods applied in the studies are presented in Table 1. Descriptive data for both studies were given as means for continuous data and as medians for ordinal data, and min-max values.

In Study I, all outcome data analysed derived from self-reported instruments for pain, disability and generic well-being, and were all ordinal data. Ratings, baseline and on discharge, were compared individually and no comparisons between patients were made. For pain intensity, minimum clinical important change (MCID) was set at at least 30% difference in the patients´ ratings, as recommended for assessing individual patients[37]. For disability (OSW) improvements were set to at least six points or a 50% improvement in patients´ ratings[58]. The median values for pain intensity (CR 10) at baseline were calculated and compared with a point value on discharge. For each subject, changes in points in the OSW scores (initial OSW score – discharge OSW score) and for percentage change (initial OSW score – discharge OSW score/initial OSW score x 100%) were calculated. The scores on the SF 36 were presented as point values at baseline and on discharge and compared to the Swedish population mean[113].

In Study II the differences in patient characteristics and distribution at the two different clinics were analysed using Student's t-test, the Mann-Whitney U test and χ^2. Agreement between the examiners in each pair was calculated as a percentage (%) of observed agreement (raw agreement) and as the chance-corrected agreement (kappa coefficient) and corresponding 95% confidence intervals (CI). The un-weighted kappa coefficient (κ) was calculated for categorical variables (classification, specific movement pattern, specific segmental-, neurological- and uni- or bilateral symptoms and signs). The linear weighted kappa coefficient (κ_w) was calculated for the irritability variable. The answers to the two questions on irritability; 1) how easily symptoms were aggravated by activities with three category answers (hard, moderate, easy) and 2) the time for symptoms to subside after aggravation also with three category answers (rapid, moderate, slow); were transferred to one ordinal variable scored 1-5 (Table 2). This was to obtain an aggregated result of the two questions, for all four examiners. Kappa values were interpreted as; $\leq$ 0.20 poor, 0.21-0.40 fair, 0.41-0.60 moderate, 0.61-0.80 substantial, and 0.81-1.00 almost perfect agreement[74]. Differences in distribution of patients to classifications were calculated using Fischer´s exact test.

Table 1. All methods applied in the data analyses

Statistics	Study I	Study II
Descriptive statistics	●	●
Student's *t*- test		●
Mann-Whitney *U* test		●
Chi-square test		●
Fischer´s exact test		●
Kappa statistics		
-unweighted		●
-linear weighted		●

Table 2. Ordinal scale for scoring irritability

		Question 2. Time for aggravation of symptoms to subside?		
		rapid	moderate	slow
Question 1. How easily are your symptoms aggravated by activities?	hard	1	2	3
	moderate	2	3	4
	easy	3	4	5

Answers to question 1 (hard, moderate, easy) and 2 (rapid, moderate, slow) were combined to an ordinal scale, illustrating level of irritability.

5 RESULTS

5.1 STUDY SAMPLES

The baseline characteristics of the patients in Studies I and II presented in Table 3 show that a majority had subacute or chronic LBP. In numbers, there were more females than males, a wide range in age, and on average moderate pain or disability ratings (Table 3).

Table 3. Characteristics of the patients included in Study I and II

Variable	Study I (n=16)	Study II (n= 64)
Age (years)		
mean (min-max)	48.4 (21-81)	46.5 (17-77)
Gender		
Male, n (%)	7 (44)	27 (42)
Female, n (%)	9 (56)	37 (58)
Symptom duration in weeks		
median (min-max)	77(10 - >1000)	12 (1-572)
Pain intensity-CR 10[1]		
median (min-max)	3 (1-6)	3.5 (0-9)
Oswestry score-OSW[2]		
median (min-max)	24 (2-62)	30 (2-60)
SF 36		
mean (min-max)	32.57 (15.54-52.49)	

[1] Borg's pain scale[12]
[2] Oswestry low-back pain disability questionnaire[36]

5.2 STUDY I

The main result was a treatment-strategy-based classification algorithm for decision-making, that illustrates the physiotherapist's clinical reasoning process and classifies patients with LBP into one of four classifications; *pain modulation, stabilisation exercise, mobilisation* and *training* (Study I, Figure 1). All the patients included (n=16) were classified into one of the four classifications (Figure 8). The *stabilisation exercise* classification was the least used while *mobilisation* was the most frequent. The examination, including classification, was conducted within the normal scheduled time frame for a first visit (45-60 min).
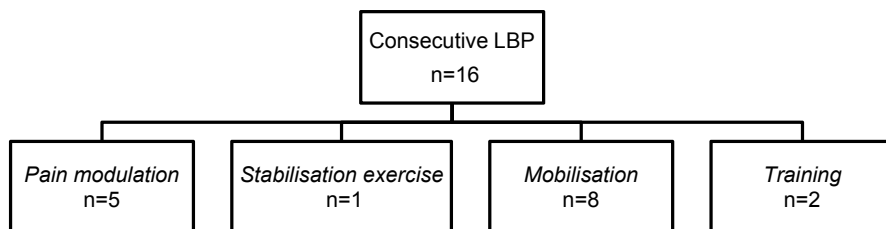

Figure 8. Distribution of patients to classifications in Study I

All the patients were treated according their assigned classification, but treatments were individualised, appropriate to clinical practice. Two patients were excluded during the treatment period; one due to pregnancy unknown at the start of the study and one due to progressive symptoms and subsequently referred for medical investigation. Short-term treatment outcome for the remaining 14 patients showed that all but one had improved pain intensity scores (Study I, Table 2). Eight patients had improved disability scores (OSW), two were unchanged and four patients considered themselves as worse (Study I, Figure 3). For the subscale for physical health (PCS) in the SF 36,12 patients considered their health improved, while two scored a decline (Study I, Figure 4). The treatment flow chart demonstrated that two of the five patients initially classified to *pain modulation* and seven of the eight patients classified to *mobilisation* were transferred to the *training* classification when their clinical status improved. The two patients initially classified to *training* and the single patient assigned to *stabilisation exercise* remained in their initial classifications throughout the study (Study I, Figure 2).

5.3 STUDY II

The main result of Study II demonstrated that the new classification system had substantial inter-examiner reliability, when experienced OMT physiotherapists, newly introduced to the system, independently classified LBP patients (n=64) into one of the four classifications (80%, κ=0.72, CI=0.59-0.85). For each pair (A and B) the agreement was also substantial (A=76%, κ=0.66, CI=0.45-0.86 and B= 83%, κ=0.75, CI=0.52-0.98) (Study II, Table 3). There were no differences in distribution of patients made by the two pairs to the classifications *stabilisation exercise*, *mobilisation* or *training*. For *pain modulation* the distribution differed (p=0.008) so that pair B classified more patients to this classification than pair A did. (Study II, Table 4).

Agreement within each classification showed that 19 of 21 patients (90%, κ=0.77, CI=0.46-1.07) were classified to *pain modulation* by both examiners. Corresponding figures for *stabilisation exercise* were 5 of 6 patients (83%, κ=0.67, CI=0.07-1.26), for *mobilisation* 11 of 19 patients (58%, κ=0.11, CI= -0.37-0.58), and for *training* 16 of 18 patients (89%, κ=0.75, CI=0.43-1.08). The interpretation of these values (κ) was that inter-examiner reliability was substantial for three of the classifications; *pain modulation, stabilisation exercise* and *training*. For the classification *mobilisation* inter-examiner reliability was poor (Study II, Table 4).

Agreements on the five specific examination items were diverse. One item had almost perfect agreement (presence of neurological signs and symptoms, 92%, (κ=0.84, CI=0.71-0.97) , two had moderate agreement; level of irritability (82%, κ=0.4, CI=0.25-0.56) and presence of uni-or bilateral signs (62%, κ=0.42, CI=0.23-0.60), and two had fair agreement presence of a specific movement pattern (68%, κ=0.38, CI=0.15-0.61) and specific segmental signs (67%, κ= 0.28,CI=0,03-0.53) (Study II, Table 3).

There were no differences in characteristics in the patients included at the two clinics, except for pain intensity, for which patients at clinic B reported higher scores than those at clinic A did (p=0.007) (Study II, Table 1). All classifications were used by the two pairs of examiners (Figure 9). Stabilisation exercises was the least used classification.
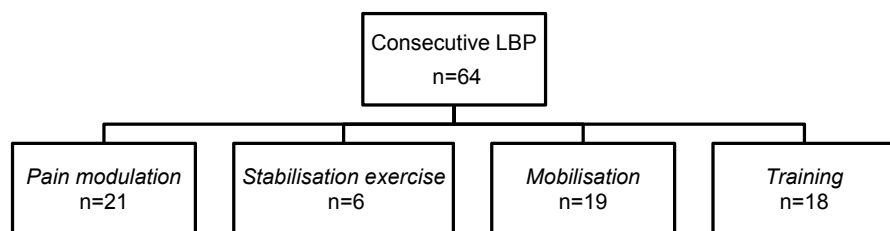

Figure 9. Distribution of patients to classifications in Study II

6 DISCUSSION

The research reported in this thesis focused on how patients with LBP may be categorised according to their clinical status and consequently what treatment would be most beneficial. In Study I a framework for a new treatment-strategy-based classification system and a clinical decision algorithm were described. Further, a progressive treatment-flow-chart with adaptation to improvements in clinical status was presented. Study II showed that, although the physiotherapists were newly introduced to the algorithm, the agreement was substantial when they independently classified patients with LBP to one of the four treatment-strategy-based classifications. Moreover, three of five specific examination items had at least moderate agreement (irritability, uni-bilateral signs, and neurological signs symptoms) while two had fair agreement (specific movement pattern and specific segmental signs). These results suggest that the new classification system may be reliably and readily applied by OMT-trained physiotherapists in out-patient settings and that the two specific items with fair agreement should be clarified or reconsidered.

LBP is a common, disabling and costly disorder and its management includes a range of different physiotherapy interventions[47]. It is not clear how patients are appropriately selected to these interventions. Various tools to facilitate decision-making in assessment and treatment in individual patients, Clinical predictions rules (CPR), have been found not confidently applicable, due to poor study design[50]. The new classification system may have clinical advantages over these CPRs. In conjunction with preliminarily validated clinical characteristics from two CPRs presented in the TBC system, (stabilisation and mobilisation)[21, 38], the new classification system also includes new combinations of signs and symptoms for clinical decisions (Study II, Table 2). Further, the treatments presented in many CPRs are such that only one single treatment is considered appropriate, while the new classification system uses wider concepts for treatment, i.e. treatment strategies. These strategies provide greater flexibility in treatment selection, which may benefit the response to treatment in the individual patient. Lastly, it enables the treatment to be adapted and adjusted continuously according to the clinical status, so that in the same patient initial pain-alleviating treatment may be followed by physical training, when status improves. However, that the clinical phase or these combinations of signs and symptoms is appropriate for guiding treatment is still to be investigated.

The new system has a similar number of categories to other reported impairment classification systems. The TBC system[42] has four (manipulation, specific exercise, stabilisation and traction), the McKenzie (MDT) system has three primary categories (derangement, dysfunction and posture)[84] and the Movement System Impairment classification system (MSI) has five (rotation-extension, extension, rotation, rotation-flexion and flexion)[100]. To be efficient and of clinical utility a system must discriminate patient characteristics and use all classifications. The four classifications in the new system were all used by examiners in the two studies. This is similar to the TBC system[42], while the MDT system places a majority of the patients in the derangement classification (90%)[67], and MSI places a majority (84%) in two classifications (rotation-extension and rotation) and two classifications were never selected (extension and flexion)[115]. In the new system, the classifications are comprehensible and include commonly used treatments for LBP, of which some are guidelines endorsed for the management and prevention of sub-acute and chronic LBP (mobilisations

/manipulations, physical exercises, and acupuncture). However, whether classifying patients in this way will improve outcomes remains unkown.

With respect to the multi-factorial causes of LBP[44], the biomedical approach and treatment selections mainly directed towards impairments may be seen as a limitation of the new system. However, it does not exclude additional treatments such as individualized treatment regimen, simple home-exercises, ergonomic advice and/or cognitive-behavioural interventions. A therapeutic approach including guidance and support during a treatment period is appropriate in all clinical physiotherapy. In the chronic LBP population there will be many patients with various degrees of depression, psychosocial distress, insomnia and/ or movement fear-avoidance[52, 92]. These signs and symptoms need to be carefully considered and their management may need the consultation and treatment provided by other professionals.

Inter-examiner reliability is pertinent for a classification system as it shows how the system may be used consistently by different clinicians. However, the present study offers no evidence for the accuracy of the examiners' classification, as no investigation on treatment response was carried out. The result of Study II corresponds to recent inter-examiner reliability studies on other impairment-based classification systems (Study II, Box 2). It is difficult to compare kappa values from different studies as the interpretation of the magnitude of the kappa coefficient can be influenced by prevalence, numbers of categories, bias and independent ratings[4, 102]. Further, several studies on agreement have used various levels of training time and experience in the systems investigated, which also influences agreement[26, 48, 111, 127]. The guidelines for the interpretation of kappa, among which Landis and Koch[74] have provided one set, are all arbitrary. Clinically acceptable agreement depends on circumstances [4, 121]. The inter-examiner reliability of the TBC system has shown kappa values from poor to moderate for classification. A study where the examiners were unfamiliar with the TBC system[55] showed a value of κ =0.15 for overall classification, whereas the examiners in the present Study II, newly introduced to the new classification system, had a much higher agreement value (κ=0.72). This is promising for future studies and generalisability.

In contrast to the substantial agreement on classification, the agreement on specific examination items was diverse. This suggests that classification was made not on the five examination items only, but on a compilation of subjective and physical examination findings, and that the algorithm and included specific examination items were used as an aid together with other clinical judgments. Of the five examination items included, three were moderate-to-almost perfect, while two were fair (Study II, Table 3). This is in line with other studies, showing that agreement on clinical tests is difficult to achieve and requires strict protocols and sufficient training time[40, 59, 111, 118, 121]. Various studies have concluded that agreement increases with familiarity[26, 111, 127]. However, the major strengths of Study II are that the examiners were un-involved in the formation, spent just a few hours becoming familiar with the system and carried out examinations without a strict protocol. These may all be reasons for the diverse results on specific items. Moreover, some items were fairly new to the examiners (irritability and specific movement pattern) and another (specific segmental signs) has inherently un-reliabile components such as small segmental mobility situated deeply below the palpation surface, all of which certainly influenced agreement. A reliable classification system must contain examination items that can be measured in a consistent manner and must use a decision-making algorithm that can be applied consistently by different examiners[40]. Therefore, one may argue that the items that had only fair agreement

could potentially undermine the validity of the classification system. These items should therefore be clarified or reconsidered before use in future studies.

The algorithm and the new classification system have unique components and possible clinical advantages that other classification systems lack. These components; the use of treatment strategies; the adaptation to the patient's clinical status and reclassification during a progressive treatment flow, need to be established in clinical studies.

6.1 METHODOLOGICAL CONSIDERATIONS AND LIMITATIONS

6.1.1 Study samples and settings

It has been advocated that clinical studies of LBP should be conducted in patients seeking care, as these patients are thought to best represent the LBP population[51]. Both the present studies used consecutive samples of adult consenting patients seeking physiotherapy intervention in primary-health-care out-patient clinics. All three clinics were connected to the Swedish social security system and included a mix of referred and self-referred patients, normal for Swedish conditions. The wide inclusion criteria in both studies, including patients with radiating pain to the lower extremities, cover most patients with LBP seen by physiotherapists in primary care. Regardless of where patients were sampled (Östersund or Stockholm), their baseline characteristics were similar, and also comparable to those reported in patients seen in primary health care in other studies[34, 104]. The natural history of LBP has often a recurrent course and therefore the estimated durations in these patients were longer than would be expected in the general population[20, 51]

6.1.2 Examiners

The four examiners in Study II, volunteered to participate in the study due to interest in the new classification system. All four were experienced and trained in the OMT method by different education providers and to a diverse extent. As the algorithm included specific examination items that require manual experience and skills, it was considered that the physiotherapists in Study II should have OMT training, although the level of training was not stipulated. Interestingly the differences in training, between pairs and within pair B in Study II, did not influence the inter-examiner reliability values for classification (Study II, Table 3). The extensive clinical experience of all examiners probably influenced the agreement positively.

It may be argued that the examiners included were not representative of most professionals in primary care due their experience and extensive post-graduate training. However, OMT is part of undergraduate training and many physiotherapists in primary health care use OMT and attend post-graduate courses in OMT, though not always to a certificate or a Master's. In hindsight it would have been interesting to have included one inexperienced pair with undergraduate training only. This would have provided more information on how the differences in experience and OMT training may influence the agreement and how readily the algorithm may be understood and correctly applied.

6.1.3 Study I

In Study I, one examiner only classified all the patients and performed all the treatment, indicating bias. However, this was to maintain consistency of examination and treatment approach, as this was a pilot study that aimed to collect data for the development of an algorithm. The formation of the algorithm was based on a mixture of theories, scientific evidence, clinical practice in Sweden, parts of the TBC system and the developer's experience. The included concepts, examination items and treatment selections are commonly used within the field nationally and internationally, and some are seen in other classification systems. The two new classifications, *pain modulation* and *training* were empirically formed and to date, no examination of the validity of the new classifications system or the two new classifications has been conducted. Validity has to be established before generalised clinical use. However, the algorithm has been presented to experts in OMT, clinicians in primary health care and senior physiotherapy students, in Sweden. These completed a questionnaire with questions on comprehension and clinical relevance, applicability, contents and concordance to national clinical patterns as they knew them. Preliminary compilation of data suggested that face and content validity were sufficient for further investigation of the new system. In addition, patients responding to mobilisation and stabilisation in the TBC system have been identified [21, 38].

6.1.4 Study II

There was no measure of whether the patients remained stable between the two passive examinations. Such a measure could have decreased the risk of disagreement due to changes in examination responses caused by repeated clinical tests[121]. However, it would have been difficult to establish the degree of fluctuation that would influence the passive and neurological examinations so that disagreements would occur.

The mixed simultaneous and independent examiner design could potentially overestimate the kappa values, as inter-examiner reliability studies require independent examiners who fully repeat the examination[102]. It was therefore surprising that inter-examiner reliability was not higher than fair for the item, presence of specific movement pattern, showing that the interpretation of active movements may differ between examiners despite concurrent observations. The other item collected from the part of the examination where both physiotherapists were present, level of irritability, had a moderate weighted kappa value. Feedback from the examiners upon completion of the study showed that the irritability concept was fairly new to them and not used routinely prior to the study. The moderate kappa value was influenced by this novelty rather than the simultaneously given information and shows that the information was independently interpreted. Further, the answers from this item were put in a table with five categories, where not all were used. Since raw agreement was high (82%), the explanation of the moderate agreement might therefore be a situation of limited variation resulting in incorrectly low kappa values[121].

There are several methods for examining agreement on judgments from physical examinations. These include repeated examinations on the same day, on separate days, concurrent examinations or using videotaped examinations[8, 26, 40, 59, 67, 78, 96]. Study II used examinations on the same day. This was for practical reasons, but also to avoid fluctuations in status from day to day, which could deflate agreement. A use of videotape examinations would definitely decrease patient variability, but may only be used for one part of the examination procedure, the active movement testing. Further, the external validity of such studies is limited, as judgments from videos are not used

under ordinary clinical conditions. The inclusion of more examiners than four and randomly assigning them to pairs would have been the ideal method. This would have shown exhaustively whether the new classification system could be reliably used by different examiners. However, this method has obvious logistic difficulties. The number of patients included in the study was based on a power calculation and is higher than in most inter-reliability studies on impairment-based classification systems (Study II, Box2).

6.1.5 External validity

External validity refers to whether research findings obtained from a small sample can be extrapolated to a population as a whole. For this, subject sampling and setting are of great importance. For this reason the present two studies included examiners who normally would perform the examination procedure under study and patients who normally would go through the same. Both studies were performed in an out-patient clinic using ordinary examination procedures, time limits and an appropriate clinical flexibility for physiotherapists and patients. However, as all examiners had OMT training the results can only be extrapolated to physiotherapists with similar training. Examiner autonomy is of concern for the external validity of inter-examiner reliability studies. For this, Study II did not include the developer among the examiners, as several studies of classification systems have done [26, 48, 118, 127]. In these studies the developers´ judgements are used as "gold standard" and require extensive training time to ensure all examiners will examine and judge accordingly [26, 118, 127].

6.1.6 Internal validity

Internal validity refers to the confidence that one can place in the cause-effect relationship in a study. This is especially important in outcome studies where conclusions on effectiveness of interventions are drawn from study results. Study I used a consecutive sample without randomisation, a small sample size and a pre-post-test design, all of which that no conclusion on treatment outcome could be drawn, nor could evidence be provided that classification in this way improves outcome. However, the aim of this part of the study was not to investigate the treatment outcome as such, but to follow up on individual response to intervention in order to guide progression and treatment-flow.

The examiners in both studies maintained their ordinary examination procedure without strict protocols, since it is unrealistic to expect physiotherapists to use an unanimous examination procedure in clinical practice. This makes it possible to measure the normal variability in examinations and judgments, which increases the applicability and generalisability of the results. However, OMT training includes a specific examination procedure, therefore it could be expected that all examinations were performed in a similar manner. The examination procedure was outlined with account taken of examiner bias as well as patient convenience and variability. The availability of clinical information from patients to examiners prior to the physical examination increases sensitivity in studies of diagnostic accuracy[124]. As physiotherapy examinations include patient history, research on examination must be performed likewise, although this type of clinical review bias is likely to occur. As active movements may change with repeated examination, these were carried out once. This single-active-movement examination enabled the judgments to be based on the same information, but still to be independently interpreted. Each examiner separately performed the passive movement examination and the peripheral neurological examination. The response to these tests may also change with repeated examination, but for independent interpretation these

hands-on tests must be performed individually. The examiners were blinded to each other's judgments.

6.2 FUTURE RESEARCH

Pre-treatment clinical decision-making is fundamental to the physiotherapy management of patients with LBP. Further studies are needed to identify clusters of signs and symptoms that may target groups for specific physiotherapy interventions. The cause-effect of classification to treatment outcome and different aspects of validity of the new classification system have to be investigated before the system can be generalised to clinical practice in primary health care. Studies of validity and causality are currently being planned.

6.3 CLINICAL IMPLICATIONS

The two studies presented in this thesis are a first step in the multistep process that the development of a new classification system employs. The system aims to be an aid in the decision-making and in the identification of sub-groups in the LPB population and by extension to find optimal physiotherapy treatments for each sub-group. The results show that the new system can be reliably used by experienced OMT-trained physiotherapists. Although, single items showed less inter-reliability, there is good reason to believe that the new system is reliable, easy to understand and readily applied. It may be interesting to, and be used by, clinical physiotherapists working with spinal pain. Use of the new system does not require expensive equipment or specific tools. It is based on ordinary physiotherapy examination procedure and includes known clinical features and interventions. Its use may lead to improved physiotherapy management for the LBP population in primary health care. However, it has to be further examined, and therefore the clinical implications are limited to date.

7 CONCLUSION

This thesis has presented

- a new treatment-strategy-based classification system which describes the classification process and differences in clinical status

- a system that includes treatment strategies

- a classification process that includes adaptation to clinical status and a progressive treatment flow

- the knowledge that this new system and three of its five examination items can be reliably and readily applied by experienced and OMT-trained physiotherapists in out-patient settings in primary health care

- two specific examination items that need to be clarified, reconsidered or replaced to improve the reliability and validity of the new system

8   ACKNOWLEDGEMENTS

I wish to express my sincere gratitude to all those who have supported me and helped to make this work possible. In particular I thank:

Ulla Evers Larsson, Physiotherapist and PhD, my supervisor. For taking me on and giving me your time, guidance, engagement, and for your eagle eye correcting my thousands of spelling mistakes. Thank you for bearing with me.

Karin Harms-Ringdahl, Physiotherapist and Professor and my co-supervisor. For sharing your profound knowledge of research, your time, valuable comments and ideas.

Inga Arvidsson, Physiotherapist and PhD, and previous supervisor. Without your belief in my topic this work would never have been done. Thank you for taking me on and helping me through the first struggling steps towards these studies.

Niclas Olofsson, Statistician and co-author. For explaining the mysteries of statistics and always being on-side throughout this work. Your encouragement and support, especially in times of despair has been crucial for the completion of this work.

Eva Olsson, Physiotherapist and friend. For long friendship and interest in my work, and for carrying out the patient examinations, together with clinical physiotherapists, *Ulrika Jederlund, Beatrice Bakalis* and *Fredrik Swedberg*. I am indebted to you all: without your help this work would not have been possible.

Annette Heijne and Cecilia Fridén, Heads of the Division of Physiotherapy at the Department of Neurobiology, Care Sciences and Society, Karolinska Institutet, for the opportunity for me to accomplish this thesis at the Division.

Tim Crosfield, for valuable language revision.

Lisbet Broman, for kindly and patiently helping me to format this thesis.

The Karin Harms-Ringdahl research group at the Division of Physiotherapy, for valuable input and revision of my thesis.

Bertil Nordström, Illustrator. For transferring my pen drawings into clean computerized pictures. Thank you for your helpfulness and invaluable help.

Agneta Sandström, Psychologist and PhD, my mentor and friend. For friendship, encouragement and support, and interesting conversations over long lunches.

All former colleagues at Fysioterapi-och Idrottsskade Metropolen, for interesting conversations on clinical matters over the years. An especially warm thank you to *Mia Tharander* for sampling patients for Study I.

Marie-Louise Bromée, Physiotherapist and friend. For clinical discussions, good travel companion on OMT conferences and good talks about life.

Mia Ruthman, Medical translator and friend. For valuable conversations and nice dinners during my trips to Stockholm over the years.

Marie Sjödin, Nurse and close friend. For friendship and supportive conversations during long dog walks.

My parents, *Ingrid* and *Carl-Yngve Bohlin,* for loving support and encouragement.

Micael Widerström, my husband and friend. For love, support, encouragement and help. And, lastly to our astonishing daughters *Rebecka, Lovisa* and *Matilda,* for always being on-side and supportive. You are the pride and love of my life.

9 REFERENCES

1.   American Physical Therapy Association. *Guide to Pysical Therapist Practice* 2nd ed. Alexandria, VA: American Physical Therapy Association; 2001.
2.   Abbott JH, Fritz JM, McCane B, et al. Lumbar segmental mobility disorders: comparison of two methods of defining abnormal displacement kinematics in a cohort of patients with non-specific mechanical low back pain. *BMC Musculoskelet Disord*. 2006;7:45.
3.   Airaksinen O, Brox JI, Cedraschi C, et al. Chapter 4. European guidelines for the management of chronic non-specific low back pain. *Eur Spine J*. 2006;15 Suppl 2:S192-300.
4.   Altman DG. *Practical Statistics for Medical Research*. London: Chapman-Hall; 1991.
5.   Axen I, Jones JJ, Rosenbaum A, et al. The Nordic Back Pain Subpopulation Program: validation and improvement of a predictive model for treatment outcome in patients with low back pain receiving chiropractic treatment. *J Manipulative Physiol Ther*. 2005;28:381-385.
6.   Balague F, Mannion AF, Pellise F, Cedraschi C. Clinical update: low back pain. *Lancet*. 2007;369:726-728.
7.   Beattie P. The relationship between symptoms and abnormal magnetic resonance images of lumbar intervertebral discs. *Phys Ther*. 1996;76:601-608.
8.   Bertilson BC, Bring J, Sjoblom A, Sundell K, Strender LE. Inter-examiner reliability in the assessment of low back pain (LBP) using the Kirkaldy-Willis classification (KWC). *Eur Spine J*. 2006;15:1695-1703.
9.   Bialosky JE, Bishop MD, Price DD, Robinson ME, George SZ. The mechanisms of manual therapy in the treatment of musculoskeletal pain: a comprehensive model. *Man Ther*. 2009;14:531-538.
10.   Billis EV, McCarthy CJ, Oldham JA. Subclassification of low back pain: a cross-country comparison. *Eur Spine J*. 2007;16:865-879.
11.   Bolton PS, Budgell BS. Spinal manipulation and spinal mobilization influence different axial sensory beds. *Med Hypotheses*. 2006;66:258-262.
12.   Borg G. *Borg's Perceived exertion and pain scales*. Champaign, Ill.: Human Kinetics; 1998.
13.   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.
14.   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;4:335-356.
15.   Browder DA, Childs JD, Cleland JA, Fritz JM. Effectiveness of an extension-oriented treatment approach in a subgroup of subjects with low back pain: a randomized clinical trial. *Phys Ther*. 2007;87:1608-1618; discussion 1577-1609.
16.   Brumagne S, Janssens L, Knapen S, Claeys K, Suuden-Johanson E. Persons with recurrent low back pain exhibit a rigid postural control strategy. *Eur Spine J*. 2008;17:1177-1184.
17.   Burstrom K, Johannesson M, Diderichsen F. Swedish population health-related quality of life results using the EQ-5D. *Qual Life Res*. 2001;10:621-635.

18. Burton AK, Balague F, Cardon G, et al. Chapter 2. European guidelines for prevention in low back pain. November 2004. *Eur Spine J*. 2006;15 Suppl 2:S136-168.

19. Burton AK, Balague F, Cardon G, et al. How to prevent low back pain. *Best Pract Res Clin Rheumatol*. 2005;19:541-555.

20. Cassidy JD, Cote P, Carroll LJ, Kristman V. Incidence and course of low back pain episodes in the general population. *Spine (Phila Pa 1976)*. 2005;30:2817-2823.

21. Childs JD, Fritz JM, Flynn TW, et al. 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.

22. Chou R, Qaseem A, Snow V, et al. Diagnosis and treatment of low back pain: a joint clinical practice guideline from the American College of Physicians and the American Pain Society. *Ann Intern Med*. 2007;147:478-491.

23. Clarke J, van Tulder M, Blomberg S, de Vet H, van der Heijden G, Bronfort G. Traction for low back pain with or without sciatica: an updated systematic review within the framework of the Cochrane collaboration. *Spine (Phila Pa 1976)*. 2006;31:1591-1599.

24. Cleland JA, Fritz JM, Brennan GP. Predictive validity of initial fear avoidance beliefs in patients with low back pain receiving physical therapy: is the FABQ a useful screening tool for identifying patients at risk for a poor recovery? *Eur Spine J*. 2008;17:70-79.

25. Cyriax J. *Textbook of Orthopaedic Medicine; Diagnosis of soft tissue lesions* 7th ed. London: Bailliere-Tindall; 1978.

26. 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.

27. Dawson B, Trapp R. *Basic and Clinical Biostatistics. Chapter13. Reading the medical literature* 3rd ed. United States of America Lange. Medical Book/ McGraw-Hill; 2001.

28. 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; discussion 485-489.

29. Demoulin C, Crielaard JM, Vanderthommen M. Spinal muscle evaluation in healthy individuals and low-back-pain patients: a literature review. *Joint Bone Spine*. 2007;74:9-13.

30. DeSantana JM, Walsh DM, Vance C, Rakel BA, Sluka KA. Effectiveness of transcutaneous electrical nerve stimulation for treatment of hyperalgesia and pain. *Curr Rheumatol Rep*. 2008;10:492-499.

31. Deyo RA, Battie M, Beurskens AJ, et al. Outcome measures for low back pain research. A proposal for standardized use. *Spine*. 1998;23:2003-2013.

32. Elphinston J, Pook P. *Bålstabilitet : fakta och övningar med balansboll*. Farsta: SISU; 2003.

33. Elstein S. Cognitive processes in clinical inference and decision making. In: Turk D, Salovey P, eds. *Reasoning, Inference and Judgment in Clinical Psychology*. New York, NY: The Free Press; 1988:17-50.

34. Enthoven P, Skargren E, Öberg B. Clinical course in patients seeking primary care for back or neck pain: a prospective 5-year follow-up of outcome and health care consumption with subgroup analysis. *Spine (Phila Pa 1976)*. 2004;29:2458-2465.

35. Evidence-based Medicine Working Group. Evidence-based medicine. A new approach to teaching the practice of medicine. *JAMA*. 1992;268:2420-2425.

36. Fairbank JC, Couper J, Davies JB, O'Brien JP. The Oswestry low back pain disability questionnaire. *Physiotherapy*. 1980;66:271-273.

37. 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.

38. Flynn T, Fritz J, Whitman J, et al. A clinical prediction rule for classifying patients with low back pain who demonstrate short-term improvement with spinal manipulation. *Spine*. 2002;27:2835-2843.

39. Fritz JM. Use of a classification approach to the treatment of 3 patients with low back syndrome. *Phys Ther*. 1998;78:766-777.

40. Fritz JM, Brennan GP, Clifford SN, Hunter SJ, Thackeray A. An examination of the reliability of a classification algorithm for subgrouping patients with low back pain. *Spine*. 2006;31:77-82.

41. Fritz JM, Delitto A, Erhard RE. Comparison of classification-based physical therapy with therapy based on clinical practice guidelines for patients with acute low back pain: a randomized clinical trial. *Spine*. 2003;28:1363-1371; discussion 1372.

42. 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*. 2000;25:106-114.

43. Furlan AD, Imamura M, Dryden T, Irvin E. Massage for low-back pain. *Cochrane Database Syst Rev*. 2008;CD001929.

44. George S. Spinal Pain. In: Sluka KA, eds. *Mechanisms and Management of Pain for the Physical Therapist*. Seattle, WA, USA: IASP Press; 2009:317-332.

45. Guyatt G, Cook D, Haynes B. Evidence based medicine has come a long way. *BMJ*. 2004;329:990-991.

46. Hall TM, Elvey RL. Nerve trunk pain: physical diagnosis and treatment. *Man Ther*. 1999;4:63-73.

47. Harman K, Fenety A, Hoens A, Crouse J, Padfield B. Physiotherapy and low back pain in the injured worker: an examination of current practice during the subacute phase of healing. *Physiother Can*. 2009;61:88-106.

48. Harris-Hayes M, Van Dillen LR. The inter-tester reliability of physical therapists classifying low back pain problems based on the movement system impairment classification system. *PM R*. 2009;1:117-126.

49. Harte AA, Gracey JH, Baxter GD. Current use of lumbar traction in the management of low back pain: results of a survey of physiotherapists in the United Kingdom. *Arch Phys Med Rehabil*. 2005;86:1164-1169.

50. Haskins R, Rivett DA, Osmotherly PG. Clinical prediction rules in the physiotherapy management of low back pain: A systematic review. *Man Ther*. 2012;17:9-21.

51. Hayden JA, Chou R, Hogg-Johnson S, Bombardier C. Systematic reviews of low back pain prognosis had variable methods and results: guidance for future prognosis reviews. *J Clin Epidemiol*. 2009;62:781-796 e781.

52. Hayden JA, Dunn KM, van der Windt DA, Shaw WS. What is the prognosis of back pain? *Best Pract Res Clin Rheumatol*. 2010;24:167-179.

53. Hayden JA, van Tulder MW, Malmivaara A, Koes BW. Exercise therapy for treatment of non-specific low back pain. *Cochrane Database Syst Rev*. 2005;CD000335.

54. Haynes RB, Devereaux PJ, Guyatt GH. Clinical expertise in the era of evidence-based medicine and patient choice. *Evid Based Med* 2002;7:36-38.

55. Heiss DG, Fitch DS, Fritz JM, Sanchez WJ, Roberts KE, Buford JA. 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.

56. Hestbaek L, Leboeuf-Yde C, Engberg M, Lauritzen T, Bruun NH, Manniche C. The course of low back pain in a general population. Results from a 5-year prospective study. *J Manipulative Physiol Ther*. 2003;26:213-219.

57. Hestbaek L, Leboeuf-Yde C, Manniche C. Low back pain: what is the long-term course? A review of studies of general patient populations. *Eur Spine J*. 2003;12:149-165.

58. Hicks GE, Fritz JM, Delitto A, McGill SM. Preliminary development of a clinical prediction rule for determining which patients with low back pain will respond to a stabilization exercise program. *Arch Phys Med Rehabil*. 2005;86:1753-1762.

59. 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;84:1858-1864.

60. Hides JA, Jull GA, Richardson CA. Long-term effects of specific stabilizing exercises for first-episode low back pain. *Spine (Phila Pa 1976)*. 2001;26:E243-248.

61. Jones MA. Clinical reasoning in manual therapy. *Phys Ther*. 1992;72:875-884.

62. Kaltenborn FM. *Manual Mobilizations of the Joints, Volume II. The Spine*. Oslo: Olaf Nordlis Bokhandel;2009.

63. Kanemura A, Doita M, Kasahara K, Sumi M, Kurosaka M, Iguchi T. The influence of sagittal instability factors on clinical lumbar spinal symptoms. *J Spinal Disord Tech*. 2009;22:479-485.

64. Kent P, Keating J. Do primary-care clinicians think that nonspecific low back pain is one condition? *Spine (Phila Pa 1976)*. 2004;29:1022-1031.

65. Kent P, Marks D, Pearson W, Keating J. Does clinician treatment choice improve the outcomes of manual therapy for nonspecific low back pain? A metaanalysis. *J Manipulative Physiol Ther*. 2005;28:312-322.

66. Kent PM, Keating JL. The epidemiology of low back pain in primary care. *Chiropr Osteopat*. 2005;13:13.

67. Kilpikoski S, Airaksinen O, Kankaanpaa M, Leminen P, Videman T, Alen M. Interexaminer reliability of low back pain assessment using the McKenzie method. *Spine*. 2002;27:E207-214.

68. Kirkaldy-Willis WH, Farfan HF. Instability of the lumbar spine. *Clin Orthop Relat Res* 1982;110-123.

69. Kirkaldy-Willis WH, Hill RJ. A more precise diagnosis for low-back pain. *Spine*. 1979;4:102-109.

70. Koes BW, van Tulder M, Lin CW, Macedo LG, McAuley J, Maher C. An updated overview of clinical guidelines for the management of non-specific low back pain in primary care. *Eur Spine J*. 2010;19:2075-2094.

71. Kongsted A, Leboeuf-Yde C. The Nordic back pain subpopulation program-individual patterns of low back pain established by means of text messaging: a longitudinal pilot study. *Chiropr Osteopat*. 2009;17:11.

72. Krause M, Refshauge KM, Dessen M, Boland R. Lumbar spine traction: evaluation of effects and recommended application for treatment. *Man Ther*. 2000;5:72-81.

73. Kulig K, Powers CM, Landel RF, et al. Segmental lumbar mobility in individuals with low back pain: in vivo assessment during manual and self-imposed motion using dynamic MRI. *BMC Musculoskelet Disord*. 2007;8:8.

74. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33:159-174.

75. Lidwall U. Socialförsäkringsrapport 2011:4. Available at: http://www.forsakringskassan.se.

76. Lin CW, Haas M, Maher CG, Machado LA, van Tulder MW. Cost-effectiveness of guideline-endorsed treatments for low back pain: a systematic review. *Eur Spine J*. 2011;20:1024-1038.

77. Luijsterburg PA, Verhagen AP, Ostelo RW, van Os TA, Peul WC, Koes BW. Effectiveness of conservative treatments for the lumbosacral radicular syndrome: a systematic review. *Eur Spine J*. 2007;16:881-899.

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

79. MacDonald D, Moseley GL, Hodges PW. People with recurrent low back pain respond differently to trunk loading despite remission from symptoms. *Spine (Phila Pa 1976)*. 2010;35:818-824.

80. Maitland J. *Vertebral manipulation*. 5th London: Butterworths; 1986.

81. Mannion AF, Balague F, Pellise F, Cedraschi C. Pain measurement in patients with low back pain. *Nat Clin Pract Rheumatol*. 2007;3:610-618.

82. Martin BI, Deyo RA, Mirza SK, et al. Expenditures and health status among adults with back and neck problems. *JAMA*. 2008;299:656-664.

83. Martin BI, Turner JA, Mirza SK, Lee MJ, Comstock BA, Deyo RA. Trends in health care expenditures, utilization, and health status among US adults with spine problems, 1997-2006. *Spine (Phila Pa 1976)*. 2009;34:2077-2084.

84. McKenzie RA. *The Lumbar Spine: Mechanical Diagnosis and Therapy*. Waikanae, New Zealand: Spinal Publications Ltd; 1981.

85. Merskey H, Bogduk N. *Classification of chronic pain: descriptions of chronic pain syndroms and definitions of pain terms/ prepared by the Task Force on Taxonomy of The International Study of Pain*. 2nd ed. Seattle: IASP Press; 1994.

86. Mulligan BR. *Manual Therapy, NAGS, SNAGS MWMS etc*. 4th Edition. Wellington, New Zealand: Plane View Sevices Ltd; 1999.

87. 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.

88. O'Sullivan PB, Twomey LT, Allison GT. Evaluation of specific stabilizing exercise in the treatment of chronic low back pain with radiologic diagnosis of spondylolysis or spondylolisthesis. *Spine*. 1997;22:2959-2967.

89. Panjabi MM. Clinical spinal instability and low back pain. *J Electromyogr Kinesiol*. 2003;13:371-379.

90. Panjabi MM. The stabilizing system of the spine. Part II. Neutral zone and instability hypothesis. *J Spinal Disord Tech*. 1992;5:390-396; discussion 397.

91. Petersen T, Laslett M, Thorsen H, Manniche C, Ekdahl C, Jacobsen S. Diagnostic classification of non-specific low back pain. A new system integrating patho-anatomic and clinical categories. *Physiother Theory Pract*. 2003;19:213-237.

92. Pincus T, Burton AK, Vogel S, Field AP. A systematic review of psychological factors as predictors of chronicity/disability in prospective cohorts of low back pain. *Spine (Phila Pa 1976)*. 2002;27:E109-120.

93. Poitras S, Blais R, Swaine B, Rossignol M. Practice patterns of physiotherapists in the treatment of work-related back pain. *J Eval Clin Pract*. 2007;13:412-421.

94.   Rasmussen-Barr E, Äng B, Arvidsson I, Nilsson-Wikmar L. Graded exercise for recurrent low-back pain: a randomized, controlled trial with 6-, 12-, and 36-month follow-ups. *Spine (Phila Pa 1976)*. 2009;34:221-228.

95.   Rasmussen-Barr E, Nilsson-Wikmar L, Arvidsson I. Stabilizing training compared with manual treatment in sub-acute and chronic low-back pain. *Man Ther*. 2003;8:233-241.

96.   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;30:368-383; discussion 384-389.

97.   Richardson CA, Jull GA. Muscle control-pain control. What exercises would you prescribe? *Man Ther*. 1995;1:2-10.

98.   Sackett DL, Rosenberg WM, Gray JA, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. *BMJ*. 1996;312:71-72.

99.   Sahrmann SA. Diagnosis by the physical therapist-a prerequisite for treatment. A special communication. *Phys Ther*. 1988;68:1703-1706.

100.   Sahrmann SA. *Movement impairment syndromes of the lumbar spine* St. Louis: Mosby Inc; 2002.

101.   Selim AJ, Ren XS, Fincke G, et al. The importance of radiating leg pain in assessing health outcomes among patients with low back pain. Results from the Veterans Health Study. *Spine (Phila Pa 1976)*. 1998;23:470-474.

102.   Sim J, Wright CC. The kappa statistic in reliability studies: use, interpretation, and sample size requirements. *Phys Ther*. 2005;85:257-268.

103.   Skargren EI, Carlsson PG, Öberg BE. One-year follow-up comparison of the cost and effectiveness of chiropractic and physiotherapy as primary management for back pain. Subgroup analysis, recurrence, and additional health care utilization. *Spine*. 1998;23:1875-1883; discussion 1884.

104.   Skargren EI, Öberg BE. Predictive factors for 1-year outcome of low-back and neck pain in patients treated in primary care: comparison between the treatment strategies chiropractic and physiotherapy. *Pain*. 1998;77:201-207.

105.   Skyba DA, Radhakrishnan R, Rohlwing JJ, Wright A, Sluka KA. Joint manipulation reduces hyperalgesia by activation of monoamine receptors but not opioid or GABA receptors in the spinal cord. *Pain*. 2003;106:159-168.

106.   Sluka KA. Central pain mechanisms involved in pain processing. In: Sluka KA, eds. *Mechanisms and Mangement of Pain for the Physical Therapist*. Seattle: WA,USA: IASP Press; 2009:41-64.

107.   Sluka KA, Skyba DA, Radhakrishnan R, Leeper BJ, Wright A. Joint mobilization reduces hyperalgesia associated with chronic muscle and joint inflammation in rats. *J Pain*. 2006;7:602-607.

108.   Smart KM, Blake C, Staines A, Doody C. The discriminative validity of "nociceptive," "peripheral neuropathic," and "central sensitization" as mechanisms-based classifications of musculoskeletal pain. *Clin J Pain*. 2011;27:655-663.

109.   Smart KM, Blake C, Staines A, Doody C. Self-reported pain severity, quality of life, disability, anxiety and depression in patients classified with 'nociceptive', 'peripheral neuropathic' and 'central sensitisation' pain. The discriminant validity of mechanisms-based classifications of low back (+/-leg) pain. *Man Ther*. 2012;17:119-125.

110.   Staud R. Evidence for shared pain mechanisms in osteoarthritis, low back pain, and fibromyalgia. *Curr Rheumatol Rep*. 2011;13:513-520.

111.   Strender LE, Sjoblom A, Sundell K, Ludwig R, Taube A. Interexaminer reliability in physical examination of patients with low back pain. *Spine (Phila Pa 1976)*. 1997;22:814-820.

112. Stuge B, Veierod MB, Laerum E, Vollestad N. The efficacy of a treatment program focusing on specific stabilizing exercises for pelvic girdle pain after pregnancy: a two-year follow-up of a randomized clinical trial. *Spine*. 2004;29:E197-203.

113. Sullivan M, Karlsson J, Ware JE, Jr. The Swedish SF-36 Health Survey-I. Evaluation of data quality, scaling assumptions, reliability and construct validity across general populations in Sweden. *Soc Sci Med*. 1995;41:1349-1358.

114. Taylor NF, Dodd KJ, Shields N, Bruder A. Therapeutic exercise in physiotherapy practice is beneficial: a summary of systematic reviews 2002-2005. *Aust J Physiother*. 2007;53:7-16.

115. Trudelle-Jackson E, Sarvaiya-Shah SA, Wang SS. Interrater reliability of a movement impairment-based classification system for lumbar spine syndromes in patients with chronic low back pain. *J Orthop Sports Phys Ther*. 2008;38:371-376.

116. Tubach F, Beaute J, Leclerc A. Natural history and prognostic indicators of sciatica. *J Clin Epidemiol*. 2004;57:174-179.

117. Tyni-Lenné R. Sjukgymnastik-fysioterapi processen (Physiotherapy process). *Sjukgymnasten* 1983;14:17-20.

118. Van Dillen LR, Sahrmann SA, Norton BJ, et al. Reliability of physical examination items used for classification of patients with low back pain. *Phys Ther*. 1998;78:979-988.

119. van Middelkoop M, Rubinstein SM, Kuijpers T, et al. A systematic review on the effectiveness of physical and rehabilitation interventions for chronic non-specific low back pain. *Eur Spine J*. 2011;20:19-39.

120. van Middelkoop M, Rubinstein SM, Verhagen AP, Ostelo RW, Koes BW, van Tulder MW. Exercise therapy for chronic nonspecific low-back pain. *Best Pract Res Clin Rheumatol*. 2010;24:193-204.

121. 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;10:256-269.

122. van Tulder M, Koes B, Bombardier C. Low back pain. *Best Pract Res Clin Rheumatol*. 2002;16:761-775.

123. Watts NT. Clinical decision analysis. *Phys Ther*. 1989;69:569-576.

124. Whiting P, Rutjes AW, Reitsma JB, Glas AS, Bossuyt PM, Kleijnen J. Sources of variation and bias in studies of diagnostic accuracy: a systematic review. *Ann Intern Med*. 2004;140:189-202.

125. WHO. *World Health Organization. ICD-10. International Statistical Classification of Diseases and Related Health Problems,* 10 th Revision. Geneva: WHO; 2010.

126. WHO. *World Health Organization. International classification of functioning, disability and health. ICF*. Geneva: WHO; 2001.

127. Vibe Fersum K, O'Sullivan PB, Kvale A, Skouen JS. Inter-examiner reliability of a classification system for patients with non-specific low back pain. *Man Ther*. 2009;14:555-561.

128. Vicenzino B, Collins D, Benson H, Wright A. An investigation of the interrelationship between manipulative therapy-induced hypoalgesia and sympathoexcitation. *J Manipulative Physiol Ther*. 1998;21:448-453.

129. Vicenzino B, Paungmali A, Buratowski S, Wright A. Specific manipulative therapy treatment for chronic lateral epicondylalgia produces uniquely characteristic hypoalgesia. *Man Ther*. 2001;6:205-212.

130. Videman T, Battie MC, Gibbons LE, Maravilla K, Manninen H, Kaprio J. Associations between back pain history and lumbar MRI findings. *Spine*. 2003;28:582-588.
131. Wilson L, Hall H, McIntosh G, Melles T. Intertester reliability of a low back pain classification system. *Spine*. 1999;24:248-254.
132. Vroomen PC, de Krom MC, Knottnerus JA. Consistency of history taking and physical examination in patients with suspected lumbar nerve root involvement. *Spine (Phila Pa 1976)*. 2000;25:91-96; discussion 97.
133. Yuan J, Purepong N, Kerr DP, Park J, Bradbury I, McDonough S. Effectiveness of acupuncture for low back pain: a systematic review. *Spine (Phila Pa 1976)*. 2008;33:E887-900.
134. Zusman M. Irritability. *Man Ther*. 1998;3:195-202.