The diagnosis and classification of low back-related leg pain

Siobhán Margaret Stynes

A thesis submitted for the degree of

Doctor of Philosophy

June 2017

Arthritis Research UK Primary Care Centre

Keele University

SUBMISSION OF THESIS FOR A RESEARCH DEGREE

Part I. DECLARATION by the candidate for a research degree. To be bound in the thesis

Degree for which thesis being submitted Doctor of Philosophy

Title of thesis The diagnosis and classification of low back-related leg pain

This thesis contains confidential information and is subject to the protocol set down for the submission and examination of such a thesis.

NO

Date of submission 10th February 2017 Original registration date 1st April 2012
(Date of submission must comply with Regulation 2D)

Name of candidate Siobhán Margaret Stynes

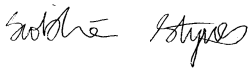
Research Institute Primary Care and Health Sciences

Name of Lead Supervisor Dr Kika Konstantinou

I certify that:

(a) The thesis being submitted for examination is my own account of my own research
(b) My research has been conducted ethically. Where relevant a letter from the approving body confirming that ethical approval has been given has been bound in the thesis as an Annex
(c) The data and results presented are the genuine data and results actually obtained by me during the conduct of the research
(d) Where I have drawn on the work, ideas and results of others this has been appropriately acknowledged in the thesis
(e) Where any collaboration has taken place with one or more other researchers, I have included within an 'Acknowledgments' section in the thesis a clear statement of their contributions, in line with the relevant statement in the Code of Practice (see Note overleaf).
(f) The greater portion of the work described in the thesis has been undertaken subsequent to my registration for the higher degree for which I am submitting for examination
(g) Where part of the work described in the thesis has previously been incorporated in another thesis submitted by me for a higher degree (if any), this has been identified and acknowledged in the thesis
(h) The thesis submitted is within the required word limit as specified in the Regulations

Total words in submitted thesis (including text and footnotes, but excluding references and appendices) …61,802……

Signature of candidate  Date 7th April 2017

Note
Extract from Code of Practice: If the research degree is set within a broader programme of work involving a group of investigators – particularly if this programme of work predates the candidate's registration – the candidate should provide an explicit statement (in an 'Acknowledgments' section) of the respective roles of the candidate and these other individuals in relevant aspects of the work reported in the thesis. For example, it should make clear, where relevant, the candidate's role in designing the study, developing data collection instruments, collecting primary data, analysing such data, and formulating conclusions from the analysis. Others involved in these aspects of the research should be named, and their contributions relative to that of the candidate should be specified (*this does not apply to the ordinary supervision, only if the supervisor or supervisory team has had greater than usual involve*

Declaration

This thesis was nested within a prospective observational cohort study called the ATLAS study, **A**ssessment and **T**reatment of **L**eg pain **A**ssociated with the **S**pine, carried out at the Arthritis Research UK Primary Care Centre, Keele University.

The objectives of the thesis were developed with members of the main ATLAS research team, namely Dr Kika Konstantinou, Professor Kate Dunn, Professor Elaine Hay and Professor Danielle van der Windt. These discussions took place over a 12 month pre-doctoral period of study and training, and the thesis plan formed the basis of a fellowship application to the National Institute of Health Research.

I was a member of the ATLAS study team during its development and design stage and contributed to the ethical approval process and design of the baseline clinic questionnaire. I acted as study coordinator for the ATLAS study during its early recruitment stage from September 2011 to March 2012 to assist with the smooth running of the research clinics and monitor recruitment. I took part in clinician training during the course of the ATLAS study and occasionally had a role as a clinical assessor. I attended clinics during the data collection phase of the reliability study. I assisted in designing the template and coding for the clinical information database and assisted with data entry from the clinical assessment forms.

Specific to this thesis, I planned the reliability study and carried out all data collection. I planned and ran the clinician workshop. Physiotherapy research facilitators based at Keele (Carol Doyle, Lucy Huckfield, Treena Larkin, Tina Hadley-Barrows, Yvonne Rimmer), assisted with identifying potential clinicians to attend the workshop.

I am co-author on the publication of the ATLAS study protocol paper (Konstantinou et al. 2012). This work informed a large part of the data acquisition methods chapter (three) of this thesis. I am an acknowledged study member on the ATLAS baseline results paper (Konstantinou et al. 2015).

The statistical analyses presented within this thesis were planned and undertaken by myself with support and advice from statisticians Dr Reuben Ogollah (chapter six and seven) and Dr Martyn Lewis (chapter five), and my supervisors.

I undertook formal statistical training in advance of carrying out the statistical analyses by attending MSc module statistical courses at Keele University (Quantitative analysis I,II; Reliability analysis; Multivariate analysis) and Erasmus University Rotterdam (Diagnostic Research).

My supervisors Kika Konstantinou, Kate Dunn and Elaine Hay gave advice and feedback on writing of the chapters and on presentation of the data. Joanne Jordan advised on searching strategies for the systematic review and Nadia Corp advised on use of reference management software.

My Doctoral studies are funded by a Clinical Academic Training Fellowship awarded by the National Institute of Health Research and Chief Nursing Officer (2012-2016). My pre-doctoral year of training was funded by a Nursing and Midwifery and Allied Health Professions (NMAHP) Research Support Award from the West Midlands Strategic Health Authority.

Abstract

Low back-related leg pain (LBLP) is clinically diagnosed as referred leg pain or sciatica. The clinical task of differentiating sciatica from referred leg pain can be challenging but is important for the purpose of treatment choices. There is currently no agreement on which clinical criteria best identify sciatica in clinical or research settings and the spectrum of clinical presentation in patients with LBLP is variable. This thesis aimed to identify diagnostic criteria for sciatica and explore and describe clusters of LBLP patients using cross-sectional data from 609 primary care LBLP consulters.

A systematic literature search of LBLP classification systems showed very few systems specifically addressed LBLP classification. Within the systems, there was wide variation in definitions and clinical features of sciatica, with most systems based on clinical opinion.

Reliability was merely fair (kappa = 0.35) amongst clinicians diagnosing sciatica but at higher levels of confidence in diagnosis (≥80%), reliability improved (kappa =0.68). Using high confidence clinical diagnosis as a reference standard, with and without confirmatory MRI findings, diagnostic models for sciatica were developed and compared. A simple scoring tool based on the best performing model was devised showing the probability of having sciatica based on results from five clinical items (subjective sensory changes, below knee pain, leg pain worse than back pain, positive neural tension, neurological deficit). Latent class analysis identified five classes of LBLP patients. One class was clearly a referred leg pain group, the other four classes seemed to represent sciatica with varying clinical profiles.

This thesis provides a diagnostic tool for sciatica with potential application in clinical and research settings. It also reveals clusters of LBLP patients which could represent more homogenous groups amenable to different treatment approaches. This thesis has provided a strong basis for future work to further explore the clinical utility of the findings.

Acknowledgements

There are many people to thank for their support and input during my doctoral studies. Firstly, my fantastic team of supervisors; Dr Kika Konstantinou, Professor Kate Dunn and Professor Elaine Hay. It was a pleasure working with my supervisors who were at all times motivating, encouraging and inspiring. They were always available to give me guidance on my work, but also gave me the confidence and space to develop as a more independent researcher.

I am very grateful to Dr Reuben Ogollah, statistician, who provided patient advice and guidance with the main analysis, and thank-you to Dr Martyn Lewis for his assistance with the reliability study analysis. Thanks to Jo Jordan for her assistance with my systematic search and Nadia Corp for reference manager system advice. Professor Danielle van der Windt was instrumental in the planning of this thesis in the initial stages and I am very grateful for her input.

This doctoral programme was funded by the National Institute for Health Research/Chief Nursing Officer and I am extremely grateful for their support. I'd like to acknowledge Professsor Nadine Foster and Professor George Peat who supported my initial NIHR fellowship application.

My sincere thanks to members of the ATLAS study research team, and all the participating clinicians, patients and the NHS service managers for supporting the ATLAS study. In particular, I thank the clinicians who were involved in the reliability study, and the clinicians who participated in the workshop. Thanks to Lucy Huckfield and Paul Delicata for modelling for the photographs.

To my sisters Katharine and Rachel, my brother Joseph and friend Gill—thank-you for proof-reading at short notice, and Rachel for your much needed help with my formatting.

I extend huge gratitude to my parents, siblings and in-laws for their continued interest in my studies and support. Especially, I thank my husband Ben for his unfailing optimism in me and for becoming a temporary stay at home dad to allow me to fulfil this PhD journey. Finally to my little ladies Roisin and Clodagh Durham, thanks for being the ultimate and best distraction.

Contents

List of Tables

List of Figures

List of Boxes

List of Abbreviations

ACR	American College of Rheumatology
AIC	Akaike Information Criteria
ATLAS	Assessment and Treatment of Leg pain Associated with the Spine
AUC	Area Under the receiver operating characteristic Curve
BAK	Bias adjusted kappa
BIC	Bayesian Information Criterion
BMI	Body mass index
CI	Confidence interval
CT	Computerised Tomography
DF	Dorsiflexion
EMG	Electromyography
GP	General practitioner
HADS	Hospital Anxiety and Depression Scale
IASP	International Association of Pain
ICD	International Statistical Classification of Diseases and Related Health Problems
ICF	International Classification of Functioning Disability and Health
IPQ-R	Revised Illness Perceptions Questionnaire
k	Kappa
LBLP	Low back-related leg pain
LBP	Low back pain
LC	Latent class
LMR	Lo-Mendall - Rubin
LR	Likelihood ratio
LRT	Likelihood ratio test
MeSH	Medical Subject Headings
MRI	Magnetic Resonance Imaging

MSK	Musculoskeletal
NHS	National Health Service
NICE	National Institute for Health Care Excellence
NIHR	National Institute for Health Research
NRI	Nerve root involvement
NRS	Numerical rating scale
OR	Odds ratio
PAK	Prevalence adjusted kappa
PNS	Peripheral nerve sensitisation
PSEQ	Pain self-efficacy questionnaire
QST	Quantitative sensory testing
QTFC	Quebec Task Force Classification
RCT	Randomised Controlled Trial
RMDQ	Roland Morris Disability Questionnaire
ROC	Receiver Operator Characteristic
ROM	Range of movement
SBI	Sciatica Bothersomeness Index
SD	Standard deviation
s-LANSS	Self-report Leeds Assessment of Neuropathic Symptoms and Signs
SLR	Straight leg raise
SPSS	Statistical Package for the Social Sciences
STarT Back	Stratified Targeted Treatment approach for back pain
TRIPOD	Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis
UK	United Kingdom
VIF	Variance inflation factor

Published and presented work associated with this thesis

Research articles from this thesis

Stynes, S., Konstantinou, K., Dunn, K.M. (2016). Classification of patients with low back-related leg pain: a systematic review. *BMC Musculoskeletal Disorders*. 17:226 . DOI: 10.1186/s12891-016-1074-z.

Stynes, S., Konstantinou, K., Dunn, K. M., Lewis, M., Hay. E. (2015). Reliability among clinicians diagnosing low back-related leg pain. *European Spine Journal,* DOI: 10.1007/s00586-015-4359-2.

Research articles associated with this thesis

Konstantinou, K., Dunn, K. M., Ogollah, R., Vogel, S., Hay, E. M. (2015). ATLAS study research team. Characteristics of patients with low back and leg pain seeking treatment in primary care: baseline results from the ATLAS cohort study. *BMC Musculoskeletal Disorders,* 16(1), 332

Konstantinou, K., Beardmore, R., Dunn, K., Lewis, M., Hider, S., Sanders, T., Jowett, S., Somerville, S., Stynes, S., van der Windt, D., Hay, E. (2012). Clinical course, characteristics and prognostic indicators in patients presenting with back and leg pain in primary care The ATLAS study protocol *BMC Musculoskeletal Disorders,* 13, 4.

Published abstracts from this thesis

Stynes, S., Konstantinou, K., Ogollah, R., Hay, E. M., Dunn, K. M. (2016). Identification of nerve root involvement in primary care consulters with low back related leg pain: Development of a multivariable diagnostic model. *Bone & Joint Journal Orthopaedic Proceedings Supplement,* 98-B (Supp 6), 26.

Stynes, S., Konstantinou, K., Dunn, K. M. (2015). Classification of patients with low back-related leg pain: A systematic review. *Bone & Joint Journal Orthopaedic Proceedings Supplement,* 97-B(Supp 2), 1.

Stynes, S., Konstantinou, K., Dunn, K. M., Lewis, M., Hay, E. M. (2014). Agreement among therapists when diagnosing low back-related leg pain. *Bone & Joint Journal Orthopaedic Proceedings Supplement* 96-B: (SUPP 4), 4

Published abstracts associated with this thesis

Ogollah, R. O., Dunn, K. M., Stynes, S., Konstantinou, K. (2016) One-year trajectories of low back related leg pain in primary care patients: Analysis using growth mixture modelling. *Rheumatology*, 55(supp 1), i81.

Work presented from this thesis

<u>Oral presentations:</u>

Stynes, S., Konstantinou, K., Ogollah, R., Hay, E. M., Dunn, K. M. Classification of low back-related leg pain consulters using latent class analysis.
Society of Back Pain Research, Preston, November 2016

Stynes, S., Konstantinou, K., Ogollah, R., Hay, E. M., Dunn, K. M. Diagnosis of nerve root involvement in primary care consulters with low back related leg pain.

International Federation of Orthopaedic Manipulative Physical Therapists (IFOMPT), Glasgow, July 2016.

Stynes, S., Konstantinou, K., Ogollah, R., Hay, E. M., Dunn, K. M. Identification of nerve root involvement in primary care consulters with low back related leg pain: Diagnostic classification using alternative approaches.
International Back and Neck Pain Forum, Buxton, June 2016.

Stynes, S., Konstantinou, K., Ogollah, R., Hay, E. M., Dunn, K. M. Identification of nerve root involvement in primary care consulters with low back related leg pain: Development of a multivariable diagnostic model.
Society of Back Pain Research, Bournemouth, November 2015.

Stynes, S., Konstantinou, K., Dunn, K. M. Classification of patients with low back-related leg pain: A systematic review.
Society of Back Pain Research, Dublin, November 2014.

Stynes, S., Konstantinou, K., Dunn, K. M., Lewis, M., Hay, E. M The agreement among therapists when diagnosing low back-related leg pain.
Society of Back Pain Research, London, November 2013.

Stynes, S. Diagnosis and classification of low back related leg pain in primary care.
Postgraduate Symposium, Research institute for Primary Care and Health Sciences, Keele University , May 2014.

Stynes, S. Development of a multivariable diagnostic model for the identification of nerve root involvement in primary care consulters with low back related leg pain.
Postgraduate Symposium, Research institute for Primary Care and Health Sciences, Keele University, May 2015.

Stynes, S. Classification of low back-related leg pain.
Postgraduate Symposium, Research institute for Primary Care and Health Sciences, Keele University, May 2016.

Poster presentation

Stynes, S., Konstantinou, K., Ogollah, R., Hay, E.M., Dunn, K.M. Identification of nerve root involvement in primary care consulters with low back related leg pain: Diagnostic classification using alternative approaches.
British Society of Rheumatology, Glasgow, April 2016.

Stynes, S., Konstantinou, K., Dunn, K. M., Lewis, M., Hay, E. M The agreement among therapists when diagnosing low back-related leg pain.
British Society of Rheumatology, Glasgow, April 2016.

Stynes, S., Konstantinou, K., Ogollah, R., Hay, E.M., Dunn, K.M. Identification of nerve root involvement in primary care consulters with low back related leg pain: Diagnostic classification using alternative approaches.
Midlands Academy of Medical Sciences Festival, Leicester, April 2016

Awards associated with this thesis
Local Organisers Award: Society of Back Pain Research Conference, Dublin November 2014
Title: Classification of low back-related leg pain: A systematic review.

Chapter One: Introduction

1.1 Background

Despite decades of high quality research, the impact of low back pain (LBP) remains an enormous public health burden at the individual and societal level. The latest Global Burden of Disease Study, published at the end of 2012, ranked LBP as the leading cause of years lived with disability among all non-fatal diseases (Buchbinder et al. 2013). Figures specific to the United Kingdom (UK) show that LBP was the top cause of years lived with disability in both 1990 and 2010, and increased by 12% over this time (Murray et al. 2013).

In 2008 the National Institute for Health Care Excellence (NICE) LBP clinical guidelines estimated the cost of LBP to the UK National Health Service (NHS) was £2.1 billion (NICE Clinical guideline CG88 2009), a rise of over 1/3 from previous estimates by Maniadakis and Gray in 2000 (Maniadakis and Gray 2000) when taking inflation into account. The nation's healthcare spending on spinal pain is considered to exceed spending on most other major medical conditions when direct and indirect costs are taken into account (Lee et al. 2013).

Much of the healthcare burden of LBP is seen in the primary care setting because patients initially present here to seek advice and are signposted to treatment pathways accordingly. To tackle the problem of LBP, an international panel of leading LBP researchers in primary care have highlighted areas of research to prioritise. Ranked number one is the question: "*Can clinically relevant subgroups of LBP be identified"* (Costa et al. 2013). One of the most common subgroups or variations of LBP is back pain radiating to the leg, with about two thirds of LBP patients presenting with back and leg pain in both primary and secondary care settings (Waddell 2004, Hill et al. 2011a, Kongsted et al. 2013).

Low back related leg pain (LBLP), generally considered as any pain or unpleasant sensation below the level of the gluteal fold, is associated with a poorer prognosis in LBP patients (Burton et al. 1995, Cherkin et al. 1996, Shaw et al. 2001, Fransen et al. 2002) and patients with LBLP suffer

more severe pain and disability, take longer to recover and lose more time from work (Andersson 1997, Selim et al. 1998, Miranda et al. 2002, Tubach et al. 2004, Grotle et al. 2005) compared to those with LBP alone.

1.2 Types of Low back-related leg pain (LBLP)

When patients present with LBLP, once serious pathology (such as tumours, cauda equina, fracture, inflammatory causes) is ruled out, the differential diagnosis is between leg pain that is due to spinal nerve root involvement or to non-specific leg pain referred from structures in the back (e.g. disc/muscle/joint) but not involving the nerve root. The majority of published guidelines on LBP (Haswell et al. 2008, Koes et al. 2010) advocate identifying patients with leg pain thought to be due to nerve root involvement, as treatment options for nerve root pain can be different from those for non-nerve root pain (Valat et al. 2010), with options available such as early investigation, pharmacotherapy, injections or surgery (Lee et al. 2013). Appropriate diagnosis may therefore reduce unnecessary tests and interventions and help more timely direct access to appropriate diagnostic and treatment resources (Bogduk and McGuirk 2002).

1.3 Definitions of LBLP

Despite efforts to standardise the terms used to describe LBLP, varying definitions are used clinically, in the literature and in guidelines, to define populations with LBLP (Konstantinou and Dunn 2008, Genevay et al. 2010, Lewis et al. 2011, Lin et al. 2014). Leg pain due to spinal nerve root involvement is also described as nerve root pain, radicular pain, radiculopathy (in the presence of neurological deficit), neuropathic leg pain secondary to compressive spinal pathology, sciatica, neural pain, nerve root compression or entrapment. Alternative names to describe non-specific leg pain include referred pain, non-neural pain, nociceptive pain or somatic referred leg pain. For consistency in this PhD thesis, the terms “sciatica” and “referred pain” will be used throughout.

The use of the term “referred pain” has been recommended by the International Association of Pain (IASP) (Mersey and Bogduk 1994). Characteristics of referred leg pain include pain described by patients as a deep, dull ache, in a non-specific (non-dermatomal) distribution in the leg, and without any associated findings of pins and needles, numbness, muscle weakness or reflex changes (Bogduk 2009). Referred leg pain generally tends to localise in areas above the knee (Bogduk 2009), but experimental studies stimulating the interspinous ligaments, facet joints (Mooney and Robertson 1976) or the intervertebral discs (O’ Neill et al. 2002) have evoked pain as far as the foot (figure 1.1).

“Sciatica” describes symptoms of leg pain arising from involvement of a spinal nerve root, with or without neurological deficit. The International Association of Pain (IASP) recommended the term sciatica should be abandoned (Mersey and Bogduk 1994) and others refer to the term as archaic (Fairbank 2007) because of its origin from an era when all radiating leg pain was called sciatica. It is recognized that the terms radicular pain and radiculopathy are more accurate (Merskey and Bogduk 1994, Konstantinou and Dunn 2008) to describe leg pain due to nerve root involvement. However “sciatica” is a term understood by clinicians and patients and continues to have widespread use in the literature (Koes et al. 2007, Konstantinou and Dunn 2008, Valat et al. 2010, Lewis et al. 2011, Verwoerd et al. 2014). The new NICE guidelines on LBP, currently in development and under review, due for publication later in 2016, are inclusive of sciatica and are using the term “Low back pain and sciatica”.

As well as a lack of agreement on the use of the term sciatica, the clinical criteria to define sciatica also vary widely (Lin et al. 2014). A call was made to “clarify the definitions of sciatica” in a sciatica review paper by Valat and colleagues in 2010 (Valat et al. 2010). The general consensus from the literature is that sciatica pain is often described as sharp, shooting or of burning quality, usually following a specific distribution (dermatomal) in the leg corresponding to the involved lumbar nerve root (see figure 1.2) and often extending to the foot, with or without pins and needles, numbness or muscle weakness (Deyo and Weinstein 2001, Koes et al. 2007, Valat et al. 2010). In

90% of cases, sciatica is caused by a herniated (prolapsed/ bulging) disc (figure 1.3), but tightening of the central spinal or lateral canal (spinal stenosis) (figure 1.4) compromising the nerve root is also a recognised cause of spinal nerve root compression or irritation (Valat et al. 2010).

Referred leg pain

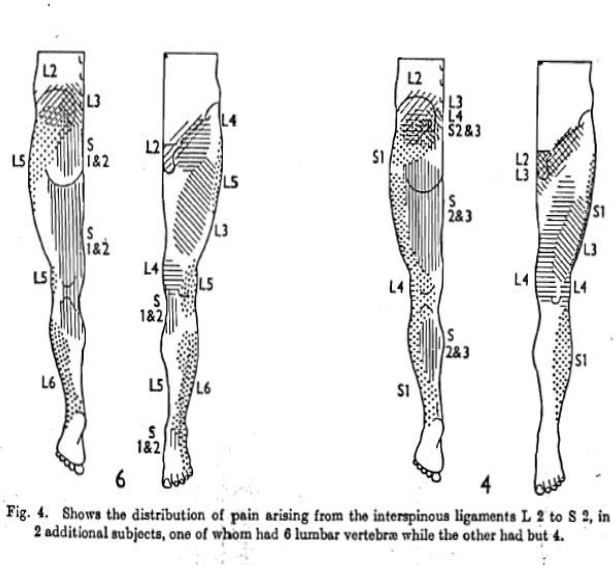


Fig. 4. Shows the distribution of pain arising from the interspinous ligaments L 2 to S 2, in 2 additional subjects, one of whom had 6 lumbar vertebræ while the other had but 4.

Figure 1.1 Patterns of leg pain evoked by noxious stimulation of the interspinous ligaments at the lumbar segments indicated (reproduced from Kellegren 1939 p 39 with permission).

Sciatica leg pain

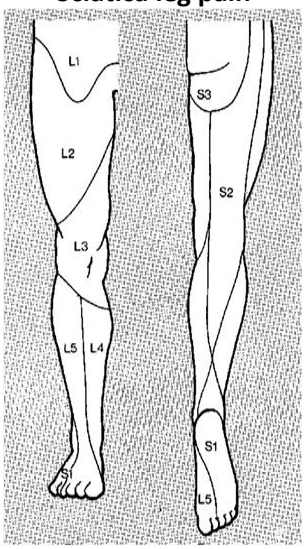


Figure 1.2 Areas of leg pain distribution from corresponding lumbar nerve root involvement at nerve roots L1,L2,L3,L4,L5,S1 (http://www.rcemlearning.co.uk accessed 21/9/2016).

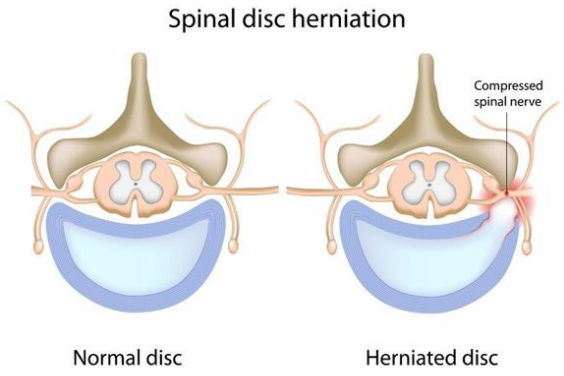


Figure 1.3 Spinal disc herniation (www.spineuniverse.com accessed 07/06/2016)

Spinal Stenosis

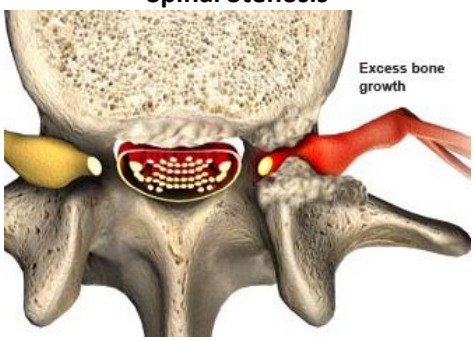


Figure 1.4 Spinal stenosis (www.spineuniverse.com accessed 07/06/2016)

1.4 Prevalence, aetiology and impact of LBLP

Around 6% of the UK general population visit their General Practitioner (GP) with back pain each year (Jordan et al. 2010) and about 61% of patients consulting with LBP report having leg pain (Hill et al. 2011a). This computes to approximately 3.7% of the population reporting LBLP, although not all will be diagnosed as having sciatica. Prevalence estimates of sciatica in the literature vary widely from 1.6% to 43%, as studies have used different populations, definitions and criteria to define or diagnose sciatica (Konstantinou and Dunn 2008). However, recently published results of a large cohort study of primary care LBLP consulters (the ATLAS study), reported 74% of LBLP patients had a clinical diagnosis of sciatica (Konstantinou et al. 2015). Combining this figure with the overall prevalence of LBLP indicates that approximately 2.7% of the general population may have a clinical diagnosis of sciatica each year (Konstantinou et al. 2015). Based on the most recent UK population census figures, with a UK adult population of approximately 50 million, this suggests potentially 1.3 million cases of sciatica present to primary care annually (Office for National Statistics 2016a).

Reviews on the aetiology of sciatica have summarised evidence of intrinsic and extrinsic factors considered to be risk factors associated with the development of sciatica (Stafford et al. 2007, Koes et al. 2007). Personal factors associated with the development of sciatica included: increased height (for men aged 50-64 only), age (incidence peaks in fifth decade), smoking and regular walking. Jogging in those with a previous history of sciatica is a risk factor for recurrence of sciatica. Genetic links with sciatica were established in a study on 9365 pairs of adult twins which estimated a heritability of 20.8% for self-reported episodes of sciatica (Heikkilä et al. 1989). Certain occupations associated with an increased incidence of sciatica are ones that involve exposure to vibration including driving; jobs that involve awkward positions that include flexion/twisting of the trunk; or jobs which require working with arms often above shoulder height. Carpenters and machine operators were more at risk than sedentary workers and full time farmers were more likely to develop sciatica than farmers retired or working part time. Leg pain is

considered to be an obstacle to recovery (Cherkin et al. 1996) or a marker of severity (Hill et al. 2011), and the further the pain radiates, particularly when associated with evidence of positive neurological findings, the greater the likelihood of increased levels of disability and health care use (Selim et al. 1998, BenDebba et al. 2000, Hicks et al. 2008, Kongsted et al. 2013).

It is generally believed that patients with sciatica have a favourable outcome as most cases resolve spontaneously within 12 weeks (Koes et al. 2007, Valat et al. 2010). However, it is reported that that up to 30% of patients will experience persistent symptoms at 1 year (Weber et al. 1993) and an estimated 5-15% of patients with sciatica proceed to disc surgery (Bush et al. 1992, Weber et al. 1993).

1.5 Pathophysiology of LBLP

Referred leg pain: Referred pain is defined as pain perceived as occurring in a body region that is distinct from the region in which the actual source of pain is located (Baron et al. 2016). Arm pain when experiencing a heart attack is a classic example of referred pain. In LBLP, the source of referred leg pain arises in the somatic tissues of the lumbar spine for example disc/ligament/joints/muscle. The theory of convergence provides an explanation for the nature of referred leg pain (Bogduk 2009). Pain fibres from structures in the lumbar spine and from the leg converge on interneurons in the spinal cord and brain. Hence nociceptive input from the painful spinal structures is perceived in areas of the lower limb that share the same innervation as the spinal source (Lewis et al. 2011, Bogduk 2009).

Sciatica: For decades, the leg symptoms associated with sciatica were attributed to mechanical pressure on the spinal nerve root. A landmark paper by Mixter and Barr in 1934, correlated clinical findings in patients with sciatica to pathological specimens of disc material that encroached on the nerve root (Mixter and Barr 1934). This introduced the era of surgery for lumbar disc disorders (Pearce 2007) and any relief of symptoms following removal of the offending disc was attributed to relief of pressure on the nerve root(s). Lumbosacral nerve roots

are considered susceptible to the effects of pressure because they lack a well-developed venous drainage system (Stafford et al. 2007). In a study with 394 patients with sciatica, magnetic resonance imaging (MRI) findings showed a positive correlation between severity of disc disease and pain and disability measures, and those with larger herniations had more leg pain and disability (Porchet et al. 2002). However imaging studies have also added to the evidence base that pressure alone on the nerve root could not be the only mechanism involved in sciatica. Disc pathology and stenosis with apparent nerve root involvement are relatively common in asymptomatic adults (Boden et al. 1990, Jensen et al. 1994, Boos et al. 2000). Conversely, patients with a clinical diagnosis of sciatica can experience symptoms without evidence of nerve root compression on MRI (el Barzouhi et al. 2014, Konstantinou et al. 2015). Outcomes can be favourable despite persistence of disc herniation visible on repeat MRI (Garfin et al. 1991), and removal of disc material causing nerve root compression does not always ameliorate symptoms (Boos et al. 2000). It is now recognised that a complex interaction of inflammatory, immunological and mechanical pressure related processes are involved in the generation and maintenance of nerve root pain (Stafford et al. 2007, Kumar et al. 2011).

Since the late 1940s, hypotheses about alternative mechanisms for generating the pain and symptoms associated with sciatica were developed. Evidence of an inflammatory response on nerve roots during surgical procedures to relieve the pressure on nerve roots (laminectomy) led to the theory of disc material prolapse provoking an inflammatory reaction in the lumbar nerve roots (Lindahl and Rexed 1951). High levels of an enzyme implicated in inflammation, called Phospholipase A2 (PLA_2), was demonstrated in herniated nucleus material from discs in sciatica patients (Saal et al. 1990). Another inflammatory agent, the cytokine TNF-α, appears to be the cytokine most strongly associated with the inflammatory properties of nucleus pulposus (Stafford et al. 2007). It induces synthesis of nitric oxide, a potent mediator of inflammation and several studies have suggested that TNF-α plays an early and prominent role in the pathophysiological events that lead to nerve dysfunction and pain when the nucleus pulposus is approximated to

lumbar nerve roots. Data from trials with anti-TNF medicines have shown some promising results in terms of improvement in sciatic leg pain and reduction in the need for surgery (Genevay et al. 2010, Ohtori et al. 2011). Other pro-inflammatory cytokines and matrix metalloproteinases (MMPs) such as IL-6, IL-8 and MMP-1 are also thought to be associated with sciatica (Karppinen et al. 2008, Genevay et al. 2009). Inflammation in the presence of compression is thought to enhance the inflammatory response.

Following any tissue damage, an inflammatory reaction is normally followed by an immune response (Chandrasoma and Taylor 1998). Studies which analysed circulating autoantibodies against glycosphingolipids, molecules highly expresseed in cells from the nervous system, have shown the presence of these antibodies in patients with sciatica due to disc herniation (Brisby et al. 2002). This suggests that activation of the immune system is possibly involved in the pathophysiology of both acute and chronic sciatica (Brisby et al. 2002, Stafford et al. 2007).

Chronic infection of the lumbar intervertebral discs with the bacteria *Propionibacterium acnes* has also been associated with the inflammation around the nerve root in patients with sciatica (Stirling et al. 2001). Two recent reviews (Ganko et al. 2015, Urquhart et al. 2015) concluded that pathogen micro-organisms are found in about 35% of surgically removed discs, and that *Propionibacterium acnes* is the most frequently identified species and to lesser extent *Staphylococci.* A randomised controlled trial (RCT) investigating the effectiveness of antibiotics for patients with severe chronic LBP and Modic type 1 changes (bone swelling), in the vertebrae next to a previous disc herniation, showed significant improvements compared to the placebo group. A reduction in leg pain was seen in the antibiotic group and although the researchers did not diagnose the leg pain as referred pain or sciatica, they proposed the mechanism for leg pain reduction could be either easing of somatic/referred pain from the disc or reduction in infectious byproducts from the disc that may have been irritating the nerve roots (Albert et al. 2013).

1.6 Diagnosis and classification of LBLP

This thesis concerns both the diagnosis and classification of LBLP. Some argue that there is no distinction between these entities and the phrase "diagnostic classification" is used when describing subgroups of LBP patients (Fairbank et al. 2011). Others define diagnosis as "binary classification" (Austin et al. 2013). Identification of classification criteria has a long tradition in the field of rheumatology because of the complex often multisystem disorders that present to clinicians (Katz et al. 2000). The American College of Rheumatology (ACR) has led numerous international task forces to define clinical criteria for rheumatological conditions including rheumatoid arthritis, osteoarthritis, polymyalgia rheumatic and fibromyalgia. The working definitions that distinguish classification from diagnostic criteria are clear according to the ACR (Aggarwal et al. 2015). Classification criteria are "*standardised definitions that are primarily intended to enable clinical studies to have uniform cohorts for research*" whereas diagnostic criteria are a "*set of signs and symptoms and tests developed for use in routine clinical care to guide the care of individual patients*" (Aggarwal et al. 2015). On the other hand, some sources do not consider this distinction between diagnosis and classification criteria is warranted and suggest they work as a continuum; diagnosis is described as "*nothing different than classification in the individual patient"* (Yazici 2009).

Similar to the issue posed with the lack of agreement on the terminology for LBLP, a similar conundrum arose with use of the terms "diagnosis" and "classification". For the purpose of this thesis, the terms are used separately to reflect the different methodological approaches used to firstly diagnose sciatica in patients with LBLP and secondly to classify LBLP patients into distinct groups with similar characteristics. This distinction between diagnosis and classification is also supported by the descriptions of both terms according to the National Library of Medicine's Medical Subject Headings (MeSH). Diagnosis is described as:

"The determination of the nature of a disease or condition, or the distinguishing of one disease or condition from another".

The MeSH descriptor for classification is:

"The systematic arrangement of entities in any field into categories classes based on common characteristics such as properties, morphology, subject matter, etc.".

The diagnosis process assumes there is an underlying dichotomous disease state in an individual i.e. a person has or does not have the condition/disease (Croft et al. 2015). The process involves interpreting combinations of presenting signs and symptoms and ultimately labelling the set of symptoms/clinical findings using a single phrase (Temple et al. 2001). This process is recognised as often probabilistic and uncertain and this is particularly pertinent in the primary care setting when patients present with signs and symptoms of great variation, severity and duration. The challenge posed is to identify optimal combinations of signs and symptoms to select patients with high probability of having the target condition to guide further management (Croft et al. 2015).

The diagnostic process of assuming there is an underlying 'yes' or 'no' disease state is recognised as a flawed assumption (Moynihan et al. 2012, Croft et al. 2015). The symptoms often represent a broad spectrum of the disease state and the notion of classification addresses this uncertainty. Combining all the important signs and symptoms and searching for patterns/groups or clusters with similar profiles but without a predefined diagnosis, is the basis behind classification. Various methods can be used to decide on classification criteria, namely a judgement based approach ranging from single opinion to consensus from groups of experts, or statistical methods that analyse patient datasets to identify clusters of patients based on similar symptom profiles. These methods will be addressed in more detail in future chapters.

1.6.1 Diagnosis of LBLP

In the UK healthcare system, the majority of LBLP patients are first seen in the primary care setting and are assessed by clinicians such as GPs, physiotherapists, chiropractors or osteopaths, who use a combination of findings from history and physical examination, to evaluate the nature of the leg pain, reach a diagnostic decision and make management plans accordingly.

Differentiating sciatica from referred leg pain is considered of primary importance because the clinical course and therapeutic interventions can be different for these two conditions (Laslett et al. 2005, Valat et al. 2010). However, the diagnosis of sciatica in clinical practice can be difficult, and clinicians may disagree as to its presence or absence in a patient with LBLP (Vroomen et al. 2000, Fairbank 2007). Various methodological issues have hampered research into accurate diagnosis of sciatica; particularly (i) the lack of agreement on which clinical items best identify sciatica and (ii) selecting a reference standard to use for modelling diagnostic studies on LBLP patients.

On balance, evidence for the diagnostic accuracy of physical examination tests to identify sciatica due to disc prolapse suggests that although some tests perform better than others in confirming or excluding the condition, most of them have poor individual performance (Vroomen et al. 1999, van der Windt et al. 2010, Al Nezari et al. 2013). It has been suggested that accuracy may improve with combinations of tests from the clinical assessment (van der Windt et al. 2010), and this has been evaluated to some extent, but the majority of studies have been conducted in secondary care settings (Verwoerd et al. 2014) with patients that exhibited severe and therefore easier to recognise symptoms (Vroomen et al. 1999), and/or in patients that were surgical candidates (Vucetic et al. 1999). These are not typical of the majority of patients assessed and managed in primary care.

Additionally, studies show that reliability of the tests performed during physical examination of LBP patients is variable and often poor which may in turn explain the poor diagnostic performance of most individual clinical assessment items (van der Windt et al. 2010, McCarthy et al. 2007). The reliability of the overall clinical diagnosis of LBLP has received less attention (Vroomen et al. 2000) and will be addressed in this thesis.

There is no accepted “gold standard” for diagnosing sciatica. The reference standard in many secondary care studies has been findings from MRI or from surgery, for example, the presence of

disc herniation material (van der Windt et al. 2010). The issue of a reference standard (MRI, surgical findings) in these studies has been debated, as surgery tends to be reserved for those with the most obvious presentations of the condition (van der Windt et al. 2010). Also, as mentioned previously, positive MRI findings can be found in asymptomatic individuals (Jensen et al. 1994) and about 20% of patients in secondary care settings, clinically diagnosed with radiculopathy, have negative MRIs (Peul et al. 2007). A diagnostic systematic review on MRI for diagnosing lumbar pathology (Wassenaar et al. 2012) showed that a considerable proportion of patients may be incorrectly classified by MRI for herniated disc and spinal stenosis.

Development of diagnostic criteria is problematic across all fields of medicine because of a lack of relevant gold standards. A provisional diagnosis in medicine is frequently based on a cluster of signs and symptoms, for example in conditions such as congestive heart failure, asthma or myocardial infarction. Studies in other musculoskeletal conditions (e.g. carpal tunnel) suggest that expert clinicians can accurately identify cases more so than history, physical examination or laboratory findings (Katz et al. 2000). For some conditions, expert clinician opinion is considered the best proxy for a gold standard and subsequently determining which clinical assessment findings match the clinical impression or diagnosis of the expert clinician (Katz et al. 2000).

1.6.2 Classification of LBLP

The challenge in choosing an optimal reference standard in diagnostic modelling is a considerable one and a way of overcoming this is to employ statistical modelling in order to identify clusters of patients with similar clinical characteristics, as this method circumvents the need for a reference standard. Latent class (LC) modelling has been used in different health conditions, including LBP, to identify groups with similar characteristics, but this may be at the cost of clinical relevance. Such analyses however may help in identifying homogenous subgroups of LBLP with distinct patterns of presentation (Nazareth et al. 2006) that reflect the full spectrum of the condition. The usefulness of this approach has not yet been evaluated in the clinical syndrome of LBLP. A

potential explanation for the small effect sizes seen in trials evaluating treatment in LBP is the heterogeneity of the patients. This has been responsible for the drive to identify relevant subgroups of back pain patients that may respond more favourably to certain interventions or management approaches (Fritz et al. 2007, Schafer et al. 2009a).

1.7 Rationale for thesis

There is a clear gap in the evidence regarding how combination of tests may perform in diagnosing sciatica in the primary care setting, and whether positive or negative tests help identify subgroups of patients with LBLP with distinct presentation of signs and symptoms. Whilst it is perhaps not necessary to make very specific diagnoses in primary care (e.g. disc herniation, spinal stenosis), this does not obviate the need for the early identification and differentiation of symptoms of LBLP which are important for informing prognosis, formulating treatment plans and guiding the need for referrals to specialist services in a timely fashion. This is particularly important for those patients that continue to have significant symptoms of sciatica, but may also reduce unnecessary investigations or interventions in those that do not have such symptoms.

1.8 Conclusion

In this PhD thesis the contribution of history and examination items to the diagnosis and classification of LBLP will be explored in an unselected primary care population presenting with LBLP. Firstly the agreement among clinicians when diagnosing LBLP will be examined. Following this, the optimal combination of items will be investigated for (i) the diagnosis of sciatica using clinical judgement with and without confirmatory MRI as the reference standards, and (ii) the classification of LBLP using statistical modelling to identify subgroups of LBLP patients. The clinical usefulness of the findings from the thesis will be explored with clinicians. This research aims to provide diagnostic information to practitioners to assist with timely identification of patients with leg pain due to sciatica and provide an empirically based method of classifying patients with leg pain that could be feasible and useful in primary care.

Chapter Two: Aims, objectives and study design

2.1 Introduction

This chapter presents the overall aim of the thesis. To address this main aim, five stages of research are planned, each requiring different methods of analysis. An overview of the aims, objectives and study design for these five stages are given.

Overall aim

The overall aim of this doctoral thesis is to investigate the contribution of findings from self-report, history and physical examination in the diagnosis and classification of patients who consult in primary care with LBLP, using empirical and statistical methods.

Overall objectives

(i) To identify the combination of clinical assessment items that best identify sciatica in the primary care setting in unselected populations with LBLP.

(ii) To explore whether combinations of items from clinical assessment lead to the formation of clinically relevant subgroups of LBLP patients with distinct presentations and characteristics.

2.2 Stage one

Aim: To provide a systematic review of the literature on proposed classification systems for LBLP.

Specific objectives

(i) Describe the various ways LBLP is classified.

(ii) Describe the methods used to derive the classification systems that include patients with LBLP.

(iii) Appraise the classification systems using a specific tool.

(iv) Identify how sciatica is described and diagnosed in the various systems.

(v) Explore the applicability of using the classification systems in primary care.

2.3 Stage two

Aim: To investigate the agreement and inter-rater reliability amongst clinicians diagnosing patients with LBLP

Specific objectives

(i) To describe the agreement and reliability amongst experienced physiotherapists, taking part in an observational cohort study of LBLP patients (the ATLAS study), when diagnosing patients presenting in primary care with symptoms of LBLP.

(ii) To gain a broader insight into current agreement on the clinical diagnosis of LBLP amongst health care professionals by investigating the agreement and reliability between the ATLAS study clinicians and other health care professionals, not involved in the ATLAS study, when diagnosing LBLP.

(iii) To explore the relationship between different levels of confidence in diagnosis with agreement and reliability indices.

(iv) To identify elements of the assessment that led clinicians to their diagnosis, using a standardised proforma, and use this information to gain insight into reasons for potential disagreement.

2.4 Stage three

Aim: To develop a clinical diagnostic model that identifies the combination of history items and physical examination items from the clinical examination that best discriminates between patients with and without a clinical diagnosis of sciatica.

Specific objectives

(i) Describe the characteristics of the LBLP sample used for the diagnostic model and compare to consulters not used in the analysis.

(ii) Identify predictors to enter into the modelling process and calculate their individual diagnostic accuracy.

(iii) Enter chosen predictors into a multivariable logistic regression model to obtain a regression function, evaluate performance of the model and assess the model's internal validity using bootstrapping techniques.

(iv) Devise a simple scoring tool suitable for clinical use.

(v) Repeat all these procedures using different reference standards and compare the final models and their performance.

2.5 Stage four

Aim: To investigate whether distinct subgroups of patients with LBLP can be identified using statistical analysis of history and clinical assessment findings.

Specific objectives

Using cases with high levels of diagnostic confidence:

(i) Compare the agreement between two statistically derived groups identified by Latent class (LC) modelling and the clinically defined groups with and without a diagnosis of sciatica.

(ii) Explore if additional statistically derived classes within this sample can be identified.

(iii) Compare the characteristics of the statistically derived classes identified by LC modelling to the clinically defined groups with and without a diagnosis of sciatica.

Using cases with any level of diagnostic confidence:

(iv) Identify classes of LBLP patients, with statistically distinct characteristics, using LC modelling.

(v) Describe these classes in terms of demographics, pain and physical function, psychosocial and work features, risk of persistent LBLP related disability, and MRI findings, and then compare the classes to the clinically defined subgroups with and without a diagnosis of sciatica.

2.6 Stage five

Aim: To explore the clinical relevance of the main thesis findings on diagnosis and classification of LBLP with clinicians.

Specific objectives

(i) To gauge clinicians opinions on the clinical relevance of the diagnostic tool for sciatica

(ii) To investigate the validity of the LBLP classification system in terms of face and content validity and perceived usefulness in clinical practice.

A narrative synthesis will be given based on the opinions and feedback from clinicians involved in the management of spinal pain patients. The main themes identified from the discussions will be used to inform the interpretation and clinical relevance of the research findings.

2.7 Methodological overview of thesis

The analysis performed in the thesis is quantitative in nature using cross-sectional data. For the two rater inter-rater reliability study (stage two), clinical data from video-recording of the clinical assessment will be used.

The PRISMA guidelines (Moher et al. 2015) will be followed to write the systematic review (stage one). Guidelines for Reporting Reliability and Agreement Studies (GRRAS) (Kottner and Streiner 2011) will be followed to write the reliability study (stage two). The Tripod statement (Guidelines for Transparent Reporting of a multivariable model for Individual Prognosis or Diagnosis) (Moons et al. 2015) will be used to guide the writing of the diagnostic model chapter (stage three).

2.8 Summary

This chapter presented a brief overview of the five stages of research planned to address the principal aim of this research. The next chapter details the design and methods of data acquisition for this thesis.

Chapter Three: Design and methods of data acquisition

3.1 Introduction

This chapter outlines the designs, methods and details of all data collection for the analyses carried out in the thesis. Methods for each stage of the thesis are described in full detail in the corresponding chapters. The cohort taking part in the ATLAS study provided the data for all analyses in this thesis. The ATLAS study was part of the "leg pain workstream" within a five year programme of research on the primary care management of spinal pain, funded by National Institute for Health Research (NIHR) Programme Grants for Applied Research (PGfAR) (2009-2013). The main aim of the ATLAS study was to investigate the clinical course, characteristics and prognostic indicators in a cohort of patients presenting in primary care with LBLP. Full details of the ATLAS study are available in the published protocol (Konstantinou et al. 2012a).

3.2 Study design and setting

The ATLAS study was a multicentre, prospective observational cohort study of adults consulting in primary care with symptoms of back and leg pain with or without sciatica. Nested within the ATLAS study was the reliability study which was part of this PhD thesis. The analysis carried out for this thesis is based on cross-sectional baseline data from the ATLAS study. The thesis researcher did not design or lead the main ATLAS cohort study but was a member of the ATLAS study team (see Declaration page i). The reliability study was designed and conducted by the thesis researcher.

3.2.1 Ethical approval

Ethical approval for the ATLAS study and the nested reliability study was granted by the South Birmingham Research Ethics Committee, reference number 10/H1207/82.

3.2.2 Participants

Adults aged 18 years and over with LBLP of any duration and severity, who consulted their GP at one of 17 GP practices in North Staffordshire and Stoke-on-Trent, UK, were eligible to take part in the ATLAS study. Leg pain was defined as any pain or unpleasant/abnormal sensation such as pins and needles or numbness, spreading from the back beyond the gluteal fold into the leg.

3.2.3 Exclusion criteria

The exclusion criteria for patients in the ATLAS study are listed below (box 3.1). The same criteria applied to patients recruited to the nested reliability study.

Box 3.1 Exclusion criteria for the ATLAS study

- 'Red flags' indicative of possible serious spinal pathology
- Previous lumbar spinal surgery
- Serious co-morbidity or mental health problems
- Pregnancy
- Currently receiving physiotherapy, osteopathy or chiropractic treatment for the same problem
- Under a secondary care doctor for the same problem
- Unable to communicate in English

3.2.4 Recruitment procedure

Patients were recruited between April 2011 and March 2013. When a patient attended their GP and reported LBLP, the GP received a "pop up" prompt screen on their computer when a relevant Read Code (Hassey et al. 2001) was entered. This prompt asked two questions:

"*Do you think this patient has leg pain associated with a back problem?*

"Is this patient suitable to be invited to the Community Low Back and Leg Pain Clinic?

The prompt also reminded GPs of the exclusion criteria for the study. If the GP answered 'yes' to both questions, the patient was identifiable as suitable for the research study. On a weekly basis,

staff from the informatics team of the West Midlands North Primary Care Research Network (WMN PCRN) downloaded contact details for patients who were flagged by the GPs as suitable. Letters were sent to these patients inviting them to telephone the research centre at Keele to make an appointment at the Community Low Back and Leg Pain Clinic (the ATLAS research clinic). In the letter, information was included about the study (Appendix A) and a baseline questionnaire (Appendix B). The letter also outlined that attendance at the clinic did not oblige patients to take part in the study and that they would still have an appropriate assessment and follow up care. Patients who telephoned the clinic administrator at the research centre were offered a clinic appointment within 10 working days. If the GP felt a patient needed to be seen urgently at the clinic they could telephone the clinic administrator to book an urgent appointment at the next clinic.

Potential participants for the reliability study were recruited consecutively at the ATLAS research clinics by the thesis researcher between August 2011 and July 2012. Full details are given in chapter five.

ATLAS research clinic

The ATLAS research clinics ran three times per week and were hosted at two NHS sites, the Haywood hospital in Stoke-on-Trent, and the Midway Primary Care Centre in Newcastle-under-Lyme. Personnel at the clinics included an administrator to meet and greet patients and oversee paperwork; a research nurse and two study physiotherapists. All patients attending the clinic were assessed by a study physiotherapist and managed according to best practice irrespective of whether or not they were eligible or agreed to take part in the study.

At the clinic, a research nurse screened the patients for potential eligibility and obtained signed informed consent if the patient was potentially eligible and wished to take part in the study. For potentially eligible participants who gave preliminary consent, their baseline study questionnaire which had been mailed to them with their appointment letter, was checked for any missing

information. They were also asked to complete a further brief clinic baseline questionnaire (Appendix C) before they proceeded to see the study physiotherapist. At this point the patient had the option to agree or decline to have their assessment video-recorded for the reliability study part of this PhD thesis (see chapter five). Patients were then seen by one of the two study physiotherapists based at the clinic. A standardised clinical assessment was carried out and the physiotherapist established if the patient had the condition of interest and was fully eligible to take part in the study. Figure one shows the flow of patients through the ATLAS study. Reasons for exclusion after clinical assessment have been reported elsewhere (Konstantinou et al. 2015). The most common reason was absence of pain in the leg.

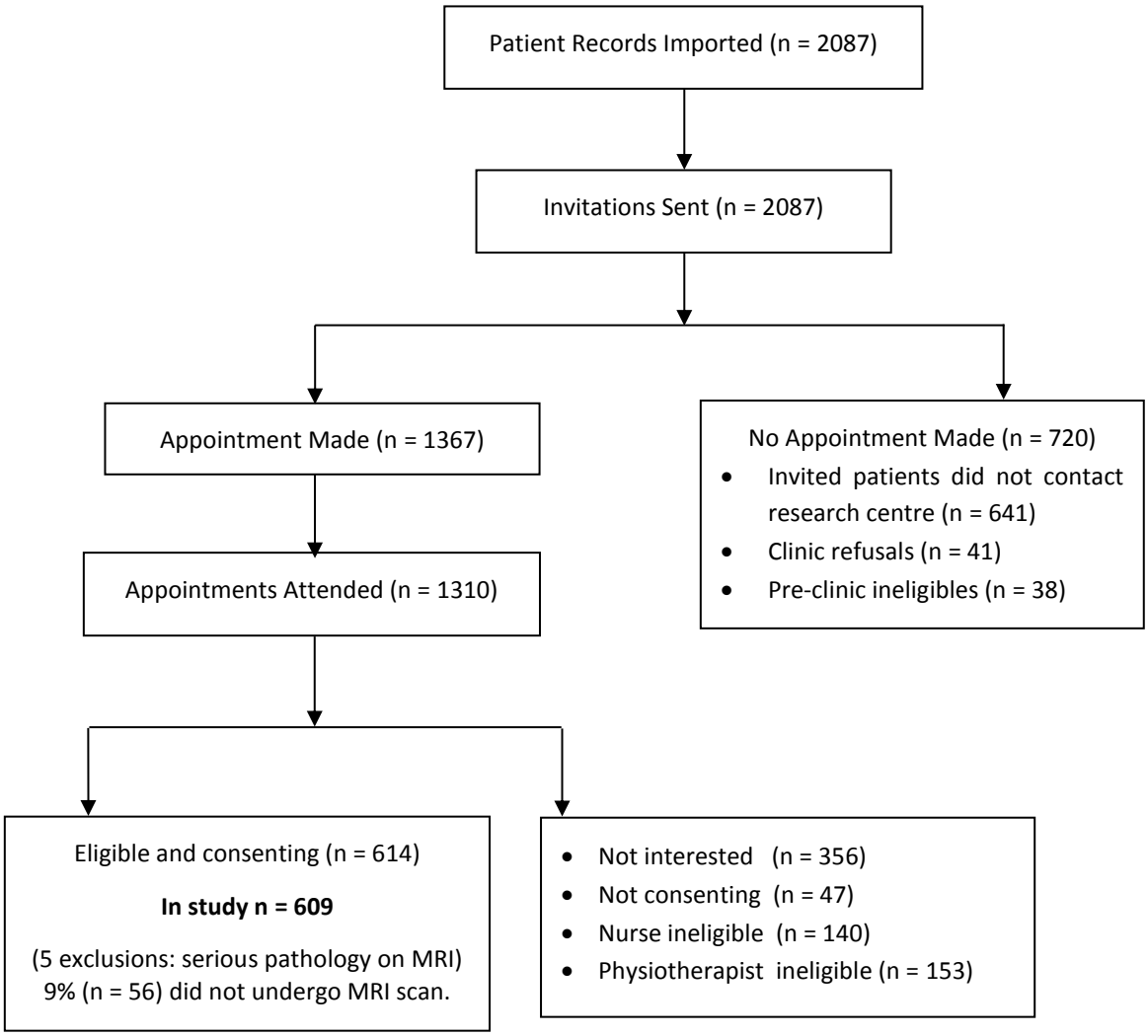


Figure 3.1 ATLAS study flow diagram (from Konstantinou et al. 2015).

3.3 Standardised clinical assessment

This standardised assessment was developed by the Principal Investigator of the ATLAS study (Kika Konstantinou) and the ATLAS study team, based on results from a national Delphi study aimed at developing consensus on the content of the clinical assessment for adults (≥18 years old) consulting with LBLP in primary care (Konstantinou et al. 2012b). The aim of the assessment schedule was to provide the optimal clinical assessment to guide diagnosis of the LBLP (i.e. determine if the leg pain was due to nerve root involvement or not) and inform appropriate management. Members of the Delphi study included over 40 local and national experts from all disciplines with relevant clinical and research experience in LBP (Konstantinou et al. 2012b). A list of items from history and clinical examination were rated on a scale of 1 to 9 of importance in (i) the assessment of a LBLP patients and (ii) the diagnosis of leg pain due to nerve root involvement (figure 3.2).

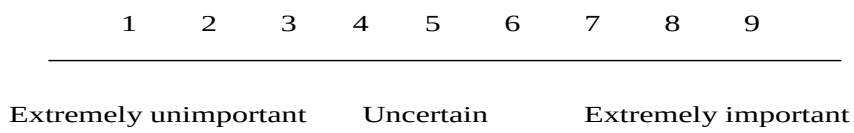


Figure 3.2 Likert scale for Delphi questions

Items rated between '7 to 9' by >70% of the Delphi participants were classed as important and included in the assessment schedule.

An overview of the history and physical examination from the clinical assessment is shown in table 3.1. The full clinical assessment form can be found in Appendix D.

Table 3.1 Clinical examination overview

History

- List of patient's complaints.
- Establish changes in daily functioning in terms of impairments, restrictions in activity, participation problems.
- Establish nature, severity of complaint, area of symptoms (mark on body chart).
- How it started-how it progressed, functioning levels before onset of complaint.
- Natural course of symptoms (better, worse, the same), effects on function.
- Aggravating and easing factors.
- Previous LBP history, tests/treatments and results.
- Rule out any red flags (suggesting serious pathology).
- Assessment of psychosocial (yellow flags) and work issues (blue/black flags).
- Coping with problem, perception of problem.
- Previous and current medical history.
- Drug History (frequency of use, effect).
- Social History (work status, financial status, effect of problem on emotions and mood, active leisure activities, family support).

Physical Examination

- Observation.
- Movements (range) and symptoms on movement.
- Neurological examination.
- Neural tension tests (straight leg raise, femoral stretch test, slump test).
- Specific, manual joint/muscle tests (as necessary).
- Specific differentiation tests - hip/back/sacroiliac joint (as necessary).

Following the standardised assessment schedule, the assessing physiotherapist classified the LBLP according to the presence or absence of sciatica (nerve root involvement) based on clinical findings, and rated confidence in their diagnosis on a 0-100% scale, where 100% means absolutely certain/confident. In patients diagnosed with nerve root involvement, the assessing physiotherapist listed up to 4 reasons for their diagnosis (box 3.2).

Box 3.2 ***Questions clinicians documented for eligible patients at the end of the assessment***

1. Is this low back pain with nerve root involvement? Yes / No
2. How confident are you in your clinical impression (rate on a 0-100% scale where 100% means absolutely certain/confident)?
3. List up to 4 most relevant items that led you to your clinical impression/diagnosis

The following section describes the items from the clinical assessment that were used for data analysis for this thesis.

3.3.1 History items

Positive cough or sneeze: A positive response was recorded if a patient's leg pain was reproduced or increased on coughing, sneezing or straining.

Below knee pain: A positive response was recorded if the assessing physiotherapist marked any areas of pain, pins and needles or numbness below the knee on the body chart manikin.

Leg pain worse than back pain: A positive response was recorded if the patient reported that their leg pain is worse/ or more bothersome than their back pain.

Subjective sensory changes: A positive response was recorded if the patient reported that they had noticed symptoms such as numbness, pins and needles or tingling in their leg.

3.3.2 Physical examination items

Aspects of the physical examination used in analyses for this thesis are described and displayed overleaf. To ensure optimum standardisation of procedures among the assessing clinicians, techniques for performing the clinical examination were demonstrated and agreed upon at training sessions.

Pain on movement: Reproduction of leg pain on either forward bending or backward bending from a neutral standing position was recorded (figure 3.3).

Forward bending **Backward bending**

Figure 3.3 Pain on movement

Neurological examination of the lower limbs:

Methods of neurological examination of the lower limb have gradually evolved and there is no one specific way of conducting a neurological examination. In the clinical setting it is often based on methods taught during clinical training, preference within a department/service or habit. Variations across sources exist on positioning, recording or interpretations of responses to neurological testing (myotomes, reflexes and dermatomes). The neurological examination tests described below represent methods generally taught to physiotherapists at undergraduate training and techniques described in text books.

Neural tension tests were performed as described by David Butler in his text-book "The Sensitive Nervous System" (Butler 2000).

For consistency neurological tests were performed in this order: myotomes, reflexes and sensation. Where applicable, the unaffected side was tested first.

Myotome deficit: Plantar flexion (walk on tip-toes; testing S1, S2 spinal nerve roots), dorsi flexion (heel walking; L4, L5 spinal nerve roots) and knee extension (squat on one leg; L3, L4 spinal nerve roots) were examined in weight bearing position (figure 3.4). If the patient was unable to weight bear, the tests were performed in lying. Hip flexion (L1, L2 spinal nerve roots), ankle inversion (L4

spinal nerve root (not shown in figure 3.4), big toe extension and ankle eversion (L5 spinal nerve root) were performed in lying (figure 3.4). Assessors documented myotome muscle strength on a six point Oxford manual muscle testing scale from 0 (no contraction/total paralysis) to 5 (normal movement/full strength against full resistance) (Kendall et al. 2005).

Figure 3.4 Myotome testing lower limb

Reflex deficit: Knee reflex (L3-L4) was tested initially in sitting (figure 3.5), and could be tried in supine if unable to elicit in sitting. Ankle reflex (S1-S2) was tested supine initially with tested leg crossed over the opposite lower leg. If unable to elicit the reflex, other positions could be tried; supine tapping fingers placed over ball of the foot, prone with knee slightly bent (figure 3.5), or in sitting. In cases when the reflex was not elicited, reinforcement techniques via muscular exertion away from the site tested were used e.g. asking patient to clench fists or teeth (Delwaide and Toulouse 1981). Testing was done using a Queens Clinical reflex hammer (38cm) tapping the tendon up to six times. Reflexes were documented on a 5 point scale as either normal, absent; slightly reduced; significantly reduced or brisk.

Knee reflex **Ankle reflex**

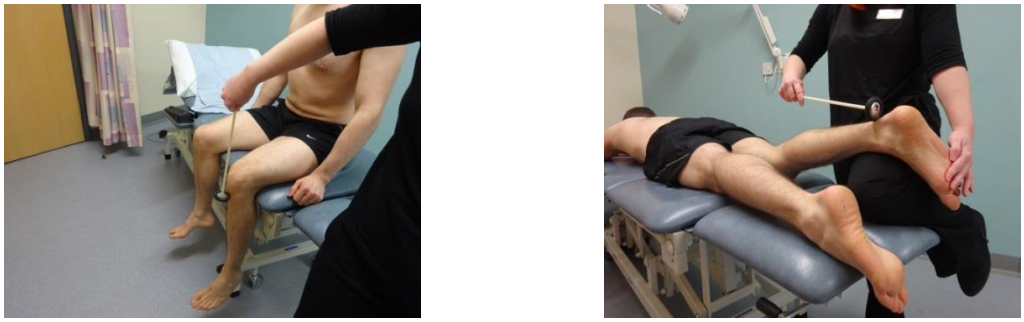


Figure 3.5 Reflex testing lower limb

Sensory deficit: A Neurotip (a sterile single use neurological examination pin) was used to test sensation. Starting on the unaffected leg with patient lying on their back, response to pin prick stimulus over "signature zones" (areas of skin in the lower limb where there is the least overlap of sensory innervation from the lumbo-sacral nerve roots (Nitta et al. 1993) (see diagram in Appendix E p8)) and/or over the area in the leg where the patient reported any altered skin sensation was recorded. Patients were asked to clarify if they felt a scratch from a pin and how it compared to the unaffected leg. Responses were recorded on a 4 point scale as either: normal; anaesthesia (no feeling); loss of pin prick (not scratchy but felt blunt) or reduced pin prick (scratchy but reduced compared to unaffected side).

Sensory testing lower limb **Neurotip for sensation testing**

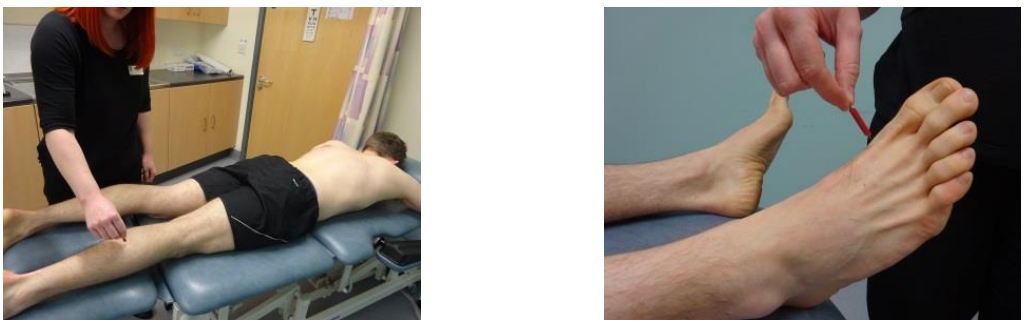


Figure 3.6 Sensory testing lower limb

Neural tension tests: An option of performing four different neural tension tests was given, depending on symptom distribution. All assessors performed the straight leg raise (SLR) and

crossed straight leg raise test. If these tests were negative or inconclusive, further exploration of neural tension was performed by carrying out a slump test. If distribution of leg pain was on the anterior aspect of the leg, the femoral nerve stretch (also known as the prone knee bend) test was performed. All tests were performed firstly on the unaffected leg.

SLR: Assessing the lower nerve roots L5, S1. With the patient lying on their back the assessor passively elevated the patient's extended leg (figure 3.7). The test was recorded as positive for neural tension if the patient's own leg symptoms were reproduced. Dorsiflexion (DF) of the foot could be added to confirm a positive test. Symptoms would be expected to increase with DF then ease with release of DF.

Crossed SLR: This was recorded as positive if a patient's leg symptoms were reproduced on passive elevation of the contralateral unaffected leg.

Slump test: This test also assesses lower spinal nerve roots. The patient sits with hands behind their back, in a slumped position with flexion of the thoracic and lumbar spines. Flexion of the neck is added and the slump position is passively maintained by the assessor. The patient then extends the unaffected leg followed by dorsi flexion of the foot (figure 3.7). Any pain response at each addition of trunk or limb movement is noted. The test is repeated on the opposite side. A positive test is based on reproduction of the patient's typical leg pain/symptoms in the affected leg, as opposed to muscle tightness restricting knee extension or ankle dorsiflexion. Raising the head from flexion back to neutral is performed and if the leg pain/symptoms ease, this confirms a positive slump test.

Femoral nerve stretch test: Assesses the upper and middle spinal lumbar nerve roots L2, L3, L4. The test is performed with the patient lying prone and the assessor passively bending the patient's knee (figure 3.7). A positive test is recorded on reproduction of the patient's leg pain/symptoms.

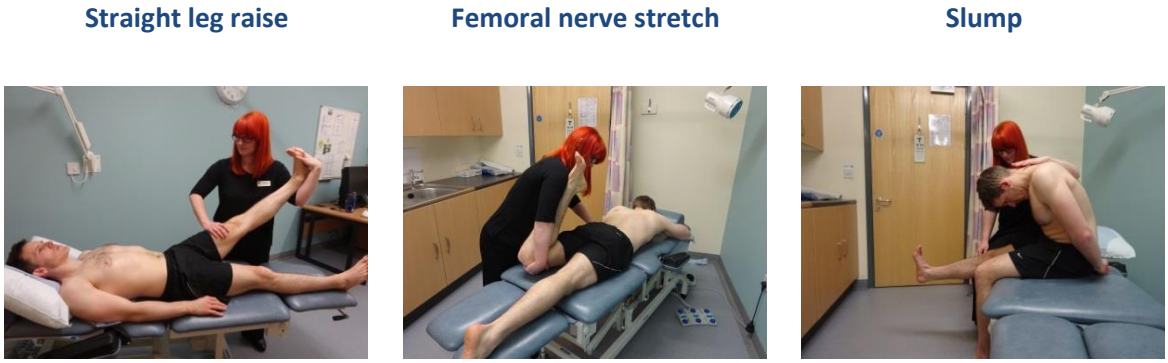


Figure 3.7 Neural tension testing lower limb

3.4 MRI scans

Within 14 days of the clinical assessment, patients eligible and consenting to take part in the ATLAS study received an MRI scan of the lumbar spine as part of the research study, providing there were no contraindications to the procedure or they did not have an MRI scan in the previous 6 months and their clinical presentation remained unchanged. Absolute contraindications included presence of: ferromagnetic aneurysm clips; cardiac pacemaker; orbital metallic foreign body; cochlear implant. Relative contraindications were: presence of a transcutaneous nerve stimulator; epilepsy or severe claustrophobia.

MRI is considered the gold standard diagnostic imaging modality for LBLP leg pain (Li et al. 2015). It provides excellent resolution of spinal nerve roots allowing for assessment of nerve root compression for any reason. MRI was performed using a mobile 0.2T magnetic resonance unit and a body spine surface coil to image the lumbar region. Patients underwent a standard lumbar spine MRI (including sagittal T1 and T2 weighted spin echo sequences), similar to that undertaken in routine clinical practice.

3.4.1 Scoring of MRI

The MRI scans were scored by a single assessor: a senior consultant musculoskeletal radiologist, who was blind to any clinical information about the patient's symptoms other than that the patient had LBLP (not specifying which leg). The radiologist provided a clinical report indicating:

- definite, possible or absence of nerve root compression
- the lumbar level(s) involved (3 lower lumbar levels)
- side (right or left) of the nerve root involvement
- Reason for the compression if present (e.g. disc prolapse, stenotic features)

The primary purpose of the MRI scan was for research purposes. Clinicians involved in the patients' care had access to the MRI report and films and any relevant scan findings were addressed as appropriate by the treating clinician.

3.5 Clinical assessors

Seven experienced physiotherapists participated in the assessment of the patients at the ATLAS research clinics. Of the physiotherapists who performed the clinical assessments, three were spinal specialist physiotherapists and four were senior musculoskeletal (MSK) physiotherapists (table 3.2). The average years in practice since qualification was 19.3 years (range 7-41 years) and they had an average of 14.9 years' experience (range 6-27 years) in predominately treating a musculoskeletal caseload. The thesis researcher was a clinician assessor who provided clinic cover in the event of illness or holiday leave of the other assessors. This was a clinical role and not carried out as part of the thesis.

Table 3.2 Clinical assessors' role and years qualified

Assessor	Years qualified	Years as MSK Physiotherapist
Senior MSK physio	7	6
Senior MSK physio	17	15
Senior MSK physio	41	27
Senior MSK physio *	18	14
Spinal specialist physio	17	15
Spinal specialist physio	12	10
Spinal specialist physio	23	17

MSK, Musculoskeletal; *Thesis researcher

3.6 Training

Prior to commencement of the study, the assessors took part in 2.5 days training for the ATLAS study.[1] The training was led by the study's principal investigator. The initial two days of training (11/12 January 2011) focused on the standardised evidence-based clinical assessment according to agreed protocols, and on the management of LBLP. An overview of the research study, the clinics and planned flow of patients through the study was given. The background to the development of the evidence-based study assessment schedule was presented. Common LBLP presentations were discussed and emphasis was given on the differentiation between leg pain due to sciatica and referred leg pain. For example table 3.3, used in the training manual, gives suggested differentiating signs and symptoms between sciatica and referred pain.

Table 3.3 Differentiating between sciatica and referred leg pain

	Sciatica leg pain	**Referred (non nerve root) pain**
Pain descriptors	Sharp, toothache-like, cramping, tingling, burning	deep-dull ache
Pain distribution	Dermatomal distribution Leg often worse than back	Non dermatomal distribution Not often below the knee
Cough/sneeze/strain	Often worse with coughing/sneezing/straining	Not effected by coughing/sneezing
Neurodynamic testing	Often positive neurodynamic tests (e.g. straight leg raise)	Normal neurodynamic tests
Neurological testing	Variable neurological findings	Normal neurological findings

A demonstration of the assessment was given using role play and practical sessions and all assessors practised the physical examination techniques to ensure standardisation of procedures.

[1] One assessor (thesis researcher) was not available to partake in the initial two day training. Subsequently she had individual mentoring with the study team principal investigator and observed several assessments beforing carrying out an assessment at the research clinic. She participated in the refresher half day training session.

The training also included sessions with a GP with special interest in musculoskeletal conditions who discussed signs and symptoms of serious spinal pathology or red flags and a Consultant in Pain Management who discussed medication use in LBP and sciatica management.

A follow-up half-day refresher two months later (31 March 2011) allowed time to answer any questions from the assessors arising since the previous training days. It also included a review of the clinical assessment procedures and discussed general procedures and paperwork for the study clinic.

A comprehensive Examiner Manual (Appendix E) supplemented the training.

3.6.1 Adherence to clinic assessment schedule

A Quality Control system was used to audit adherence to the clinic assessment schedule to ensure that the necessary information was collected in a standardised way. The principal investigator of the study observed a random selection of six to eight assessments for each study physiotherapist over the course of the recruitment period. Feedback as needed was given to the assessor to ensure the required standards of a research assessment were met.

3.7 Measurements (from questionnaires)

Measurements used in this thesis collected from the questionnaires are detailed below. The coding and categorisation for variables within the questionnaire were prepared by the ATLAS statistician as part of the main study. Variables that were prepared specifically for analysis in this thesis are stated in the thesis text.

3.7.1 Socio demographic variables

Age at time of baseline questionnaire completion, gender of each participant and current smoking status were obtained from the baseline questionnaire. BMI was calculated from height and weight measured in clinic, using the formula *mass (kg)/height (m^2).* BMI categories were defined according to BMI score ranges as: normal/underweight (<25), overweight (25 to <30) or obese/morbidly obese (30 to > 40).

Socio-economic status was determined based on the participant's current or most recent paid job (Appendix B, page 16, question 3). The Standard Occupational Classification system was used to classify job titles into four levels: Managerial and professional occupations (higher); Intermediate occupations (intermediate); Routine and manual occupations (routine); Never worked and long-term unemployed (Office of National Statistics 2010). Conditions of employment range from higher managerial and professional occupations with salary scales, promotional prospects, sick pay, and discretion over planning work, through to routine occupations with hourly pay or piece work with no promotional prospects and minimal benefits (Office of National Statistics 2016b).

Work: Interference of pain with work performance was measured on a 0 to 10 scale where 0 is "not at all" and 10 is "the pain is so bad I am unable to do my job". Proportion of patients with certified time off work was gathered from two questions "have you self-certified time off work because of your current bout/episode of back or leg pain" and "have you been given any "sick notes" or "fit notes" from your doctor because of your current bout/episode of back or leg pain? (Appendix B page 18 question 12 & 13).

3.7.2 Low back and leg pain characteristics

Duration of symptoms (weeks) was established from the questions (Appendix B page 3-4, question 1 & 8). "Have you had this current bout /episode of (back/leg) pain for 2 to 6 weeks; 6 to 12 weeks; 3 to 6 months; 7 to 12 months; more than 12 months".

Pain trajectory over the previous year was assessed from the question "how has your back and/or leg pain been over the past 12 months" (Appendix C, page 5, question 1) with seven pictorial and written options ranging from "first ever episode" to "severe pain all the time" (based on Dunn and Croft 2006). The seven responses were dichotomised as either mild (first ever episode of pain; a few episodes mainly pain free periods in between; some pain all the time/ few episodes of severe pain; pain that has gradually improved) or moderate severe pain trajectory (pain that goes up and down all the time; severe pain nearly all the time; pain that has gradually got worse).

Back pain intensity (0-10) was calculated from the mean of three 0 (no pain) to 10 (pain as bad as could be) numerical rating scales for least, current and usual pain over the last two weeks (Dunn et al. 2010). Leg pain intensity (0-10) was calculated in the same manner. Both questions were sourced from the baseline self-report questionnaire [2] (Appendix B, page 3, questions 2 & 3; page 4, questions 6 & 7; page 10, questions 29 & 30). Disability was measured using the leg pain version of the Roland Morris Disability Questionnaire (RMDQ) (Roland and Morris 1983, Patrick et al. 1995). 23 items are scored from 1 to 23 (Appendix B, page 8 & 9). Higher scores indicate greater disability. Most of the RMDQ items reflect physical activity, one item asks about sleep, one is about psychological functioning and one is on social functioning. Measurement of how patients perceive their LBLP symptoms was measured using the Sciatica Bothersomeness Index (SBI) (Patrick et al. 1995). The scale includes self-reported ratings (0-6) of bothersomeness of symptom intensity of: (i) leg pain, (ii) numbness or tingling in the leg, foot or groin, (iii) weakness in the leg/foot, and (iv) back or leg pain while sitting (Appendix B, page 5, question 1-4). A composite score (0-24) was calculated by summing the ratings of the four symptoms. Higher scores indicate worse symptoms. The SBI has been used in several studies to evaluate symptom severity of leg pain (Atlas et al. 2005, Peul et al. 2007, Grovle et al. 2010).

Neuropathic pain (0-24) was measured using the seven item self-report Leeds Assessment of Neuropathic Symptoms and Signs (s-LANSS) pain scale questionnaire (Appendix C, page 3 & 4, question 1-7). The s-LANSS aims to distinguish pain of predominantly neuropathic origin from nociceptive pain, without the need for clinical examination. It has been validated in chronic pain patients (Bennett et al. 2005) and used in describing LBP (El Sissi et al. 2010) and LBLP (Schafer et

[2] A similar question was asked in the clinical assessment about current leg pain at worst, at best and average. This was not coded for entry into the database due to a large volume of missing data. Hence the leg pain intensity question from the self-report questionnaire was used instead.

al. 2009a) populations. Scores range from 0-24 and an overall score of 12 or more indicates possible neuropathic pain (Bennett et al. 2005).

Risk of poor outcome (low, medium, high): Each patient completed the STarT Back Tool, a reliable and valid nine item tool with cut off scores to predict poor prognosis, in terms of back pain related disability, in LBP patients (Hill et al. 2008). A dichotomised response format ('agree' or 'disagree') is used for eight of the questions and a Likert scale for the bothersomeness item. (Appendix B, page 7, questions 1-9). All positive items are summed and a psychosocial subscale score ranging from 0 to 5 is produced by summing the bothersomeness, fear, catastrophising, anxiety, and depression questions. Scores of 0-3 are considered low risk, a score of 4 or more with <3 on the psychosocial subscale is medium risk, and a score of ≥4 on the psychosocial score is high risk of future persistent LBP related disability.

3.7.3 Psychological measure

Anxiety and depression were measured using the Hospital Anxiety and Depression Scale (HADS), scored from 0 (no anxiety or depression) to 21 (high level of anxiety or depression) (Zigmond and Snaith 1983). A score between 0 to 7 is considered normal range, a score of 8 to 10 indicates mild/possible depression and a score of ≥ 11 is considered indicative of probable/moderate/severe depression (Appendix C, page 7-9, question 1-14).

The Revised Musculoskeletal Illness Perception Questionnaire (IPQ-R) Short form was used to assess patient's illness beliefs about their LBLP (adapted from Moss-Morris et al. 2002). Three subscales of the Revised Illness Perception Questionnaire (IPQ-R) were used (illnesss identity, timeline and personal control) which have been identified as strong independent predictors of outcome in primary care (Foster et al. 2010). Illness identity, measured on a scale of 0 to 7, evaluated the perceived symptoms and their possible relation to the illness. Higher scores indicate a stronger belief that a greater sum of symptoms is attributable to the respondent's LBLP.

The timeline subscale was evaluated with the question "*my back pain will last for a long time"* and the question evaluating personal control was "*what I do can determine whether back/leg pain gets better".* For both these domains, the patient rated their level of agreement on a five point Likert scale from strongly disagree to strongly agree. (Appendix B: illness identity page 13 question 1; timeline page 13, question 2; personal control, page 14, question 5).

The pain self-efficacy questionnaire (PSEQ) was used to measure patients' expectations and confidence that they can perform certain tasks despite their pain (Nicholas 2007). Ten questions reflect daily activities, work, socialising and coping without medication (Appendix B, page 12, question 1-10). Patients rate their confidence on a 7-point Likert scale ranging from 0 (not confident at all) to 7 (completely confident). Scores are summed to give a total score out of 60 with higher scores representing greater pain self-efficacy beliefs. This questionnaire has shown high internal consistency and test-retest reliability (Nicholas 2007).

3.7.4 Health measures

Patients recorded how many comorbidities they had from a list of five conditions (chest problems, heart problems, hypertension, diabetes, circulation problems in legs). (Appendix C page 6 question 1).

The assessing clinician asked about sleep disturbances due to LBLP during the clinical assessment. Answers were recorded as either yes or no.

General health was measured using Short Form Health Questionnaire (SF 1) (Ware 2000) with patients' health ranked as either good/very good/excellent or fair/poor (Appendix B, page 11, question 6). A single index value for health status was calculated from the EQ-5D-3L (EuroQoL Group 1990) (Appendix B, page 11, questions 1-5). The EQ-5D-3L asks respondents to indicate the most appropriate statement (no, some or extreme problems) corresponding to the five domains of: mobility, self-care, usual activities, pain/discomfort and anxiety/depression (http://www.euroqol.org/ accessed 08.02.2016). An index score between zero and one is calculated from the responses, with values closer to one indicating better quality of health.

3.8 Sample size

The sample size for the ATLAS cohort study was calculated a priori, to detect, with 80% power, a difference of 15% in the proportion of people with poor outcome in relation to LBLP related disability, between the subgroups of leg pain patients with and without sciatica, assuming 20% loss to follow up and 5% two-tailed significance level (Konstantinou et al. 2012a). A final figure of 609 eligible patients were included in the ATLAS study and this number was sufficient to address the three stages of the thesis that involved quantitative analysis (sample size justifications for stage two—reliability study; stage three—diagnostic model and stage five—latent class modelling are addressed within chapters four, five and six respectively). A formal sample size calculation was done for the reliability study which required at least 30 subjects for analysis, at 90% power in order to detect a kappa of 0.6 (from a null hypothesis value of 0 (α=0.05)) with a 95% confidence interval (see chapter five, section 5.2.6, page 88).

3.9 Minimisation of missing data

At the research clinic, the clinic administrator checked the baseline questionnaire and addressed any missing or ambiguous information with the patient before they left the clinic. Clinicians were required to check through their clinical assessment form at the end of their assessment, to ensure all relevant sections were completed.

3.10 Data management

All data collected from the self-report questionnaires were transformed into custom made Microsoft Access databases by study administrators or using a teleform system. One in 10 checks was performed during all stages of data entry to identify and correct any errors or omissions. The clinical assessment forms completed by the physiotherapists at the research clinics were photocopied and the original copy remained at the NHS site as part of the patients' physiotherapy treatment records. A copy was returned to the Research Institute for Primary Care and Health Sciences at Keele University and data from the assessment forms was entered into a Microsoft

Access database by members of the ATLAS study team, at least one of which was one of the assessing clinicians. In the event of ambiguous findings or missing data, the clinic notes were cross checked with the original files at the clinical sites. The thesis researcher contributed with data entry for the clinical assessment forms.

Data cleaning of all the ATLAS study variables was completed by the ATLAS study statistician (Reuben Ogollah). Data for the reliability study was handled by the thesis researcher. All data used within this thesis (baseline questionnaire, clinical assessment form, MRI findings) was transferred into an SPSS file by the study's statistician (RO). Prior to performing analysis for this thesis, the PhD researcher checked data needed for the thesis analysis, for unlikely or ambiguous values and any unaccounted for missing data.

3.11 Statistical analyses

Quantitative analyses within this thesis were performed by the thesis researcher primarily using SPSS version 21. For some aspects of the analysis, SPSS was insufficient to perform the required data analysis and in such cases the analysis was performed in conjunction with the statistician (RO) in STATA (version 13) (chapter six diagnostic model) and Mplus (version 5) for latent class modelling (chapter seven). Some computations for evaluating diagnostic tests (sensitivity, specificity, predictive values, likelihood ratios (chapter five) were carried out by the thesis researcher in an on-line statistical programme (http://vassarstats.net/clin1.html).

Variables requiring categorisation from continuous scores (e.g. HADS scores, the s-LANSS score), were computed and coded in advance by the statistician (RO). Any additional variables required specifically for this thesis were generated by the thesis researcher and stored in SPSS. Specific details of analyses are described in the analysis and result sections of chapters five, six and seven.

3.12 Summary

The ATLAS study was a prospective observational cohort study of primary care consulters with LBLP. This chapter outlined the ATLAS study design, methods, data collection and quality

assurance procedures. All analyses undertaken for this thesis were based on baseline self-report, clinical and imaging data collected from patients taking part in the ATLAS study. Subsequent chapters will present in detail the analysis undertaken to address the objectives of this thesis

Chapter Four: Classification of low back-related leg pain: a systematic review

4.1 Introduction

Reassessment of the list of research priorities for LBP in primary care, initially established in the 1990s, saw the classification of LBP remain top of the research agenda (Costa et al. 2013). Extensive work has been published on classification systems where researchers and clinicians have attempted to subgroup LBP patients into homogeneous populations with similar characteristics, with the aim of optimising management and improving patient outcomes. Despite the implications for the patient in terms of pain and disability, and the implications for the wider community of having LBLP (including sciatica), the classification of this presentation has received limited attention in the literature and guidelines, compared to LBP alone. A systematic review of the scientific literature and practice guidelines was conducted to compile a summary of proposed classification systems for LBLP. A systematic search was developed and carried out, studies were included on the basis of defined criteria, and standardised quality assessment and data extraction were completed. A narrative synthesis of the results was produced in which the strengths and weaknesses of the various current approaches to classifying LBLP are discussed.

4.1.1 Study aim and objectives

Aim: To provide a systematic review of the literature on proposed classification systems for LBLP.

Objectives:

(i) Describe the various ways LBLP is classified.
(ii) Describe the methods used to derive the classification systems that include patients with LBLP.
(iii) Appraise the classification systems using a specific tool.
(iv) Identify how sciatica is described and diagnosed in the various systems.
(v) Explore the applicability of using the classification systems in primary care.

4.2 Methods

4.2.1 Inclusion criteria

Types of Studies: The following research designs were assessed for inclusion in the review: original articles describing a classification system, cohort studies, systematic reviews, diagnostic studies, treatment based studies and clinical trials evaluating effectiveness of treatments for patient subgroups within LBLP classifications. Case studies and case series designs were excluded.

Types of participants: Adults over the age of 18 years with low back pain and related leg pain of any duration. Leg pain was defined as pain below the gluteal fold. Studies looking at specific spinal "red flag" conditions such as cauda equina syndrome, tumours or spinal fractures or a specific disease cohort such as diabetes, were excluded.

Types of publications: Published studies were included if they fulfilled any of the criteria below:

- Developed and described an original classification system for back pain that included patients with LBLP.
- Adapted an existing classification system that was designed for or included LBLP patients.
- Provided approaches to appraising and validating an existing classification system for LBLP.

Studies that only used expensive or advanced investigations or technology, more likely to be feasible for secondary care settings (e.g. electromyography, surgical findings, imaging or expensive kinematic equipment) for classification of patients, were excluded.

4.2.2 Search strategy for identification of studies

An electronic search of the databases MEDLINE, EMBASE, CINAHL, AMED, PEDro, Web of Science, Cochrane library, DARE and HTA was performed during July 2013 to September 2013. All databases were searched from their inception and no date or language restriction was applied. An updated search was performed in August 2015. The Medline search strategy is detailed in table 4.1 with the number of 'hits' per search in brackets. Search strategies for other databases used are listed in Appendix F.

Table 4.1 Search terms for Medline

Database: Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations and Ovid MEDLINE(R) <1946 to Present> Date of Search: 26 July 2013		
Back	1	Back Pain/ (14558)
	2	Spine/ (20674)
	3	Back/ (3711)
	4	lumbo$.ti,ab. (11604)
	5	backache.ti,ab. (1956)
	6	back ache.ti,ab. (52)
	7	(spinal or spine).ti,ab. (247442)
	8	lumbar.ab,ti. (73557)
	9	"back pain".ab,ti. (28820)
	10	Low Back Pain/ (14064)
	11	1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 (321052)
Leg	12	(leg adj3 pain).ti,ab. (3559)
	13	(nerve adj3 pain).ti,ab. (2193)
	14	(radi$ adj3 pain).ti,ab. (6371)
	15	neuropathic.ti,ab. (16915)
	16	(referr$ adj3 pain).ti,ab. (2479)
	17	"nerve root$".ti,ab. (8152)
	18	Polyradiculopathy/ (2224)
	19	Nerve Compression Syndromes/ (8997)
	20	radicul$.ti,ab. (10340)
	21	12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 (52960)
Back and Leg	22	11 and 21 (20495)
Sciatica	23	Sciatica/ (4167)
	24	sciatic$.ti,ab. (21982)
	25	23 or 24 (23651)
Stenosis	26	Spinal Stenosis/ (4152)
	27	spinal stenosis.ti,ab. (2985)
	28	26 or 27 (5250)
Disc	29	Intervertebral Disc Displacement/ (15413)
	30	((disc or discs) adj1 (displacement$ or hernia$ or protru$ or avulsion$)).ti,ab. (6448)
	31	((disk or disks) adj1 (displacement$ or hernia$ or protru$ or avulsion$)).ti,ab. (2500)
	32	29 or 30 or 31 (18525)
Non specific LBP	33	"non specific low back pain".ti,ab. (417)
	34	"nonspecific low back pain".ti,ab. (323)
	35	"low back-related leg pain".ti,ab. (16)
	36	33 or 34 or 35 (744)
All back and leg pain	37	22 or 25 or 28 or 32 or 36 (59153)
Classification	38	Diagnosis/ (16563)
	39	Diagnosis, Differential/ (372889)
	40	(clinical adj1 predict$).ti,ab. (8801)
	41	(clinical adj1 rule$).ti,ab. (186)
	42	(predict$ adj3 (model$ or rule$)).ti,ab. (65237)
	43	(diagnos$ adj3 (model$ or rule$)).ti,ab. (4638)
	44	(classification or classified).ti,ab. (340737)
	45	identification.ti,ab. (427912)
	46	"subgroup$".ti,ab. (131764)
	47	"sub-group$".ti,ab. (6528)
	48	38 or 39 or 40 or 41 or 42 or 43 or 44 or 45 or 46 or 47 (1312408)
Classification of LBLP	49	37 and 48 (5248)

This search strategy was supplemented by 'hand searching' the reference lists of all the included full-text papers and selected review papers. Authors were contacted when clarification on aspects of their study was needed or to request if additional supporting information was available on their classification system. A search in Pubmed of first named authors of the included classification systems was also performed to identify any additional relevant published work.

4.2.3 Study selection

All citations identified from the electronic databases searches were directly imported into online reference management software (Refworks 2.0). Duplicates were removed. Titles were screened and citations were excluded when it was immediately apparent that they did not meet the inclusion criteria. When eligibility could not be determined on the basis of the title, abstracts were reviewed and citations were selected if they were potentially relevant for inclusion in the review. Screening up to this point was done by one reviewer (SS).

To select potential full text papers, two reviewers independently screened the titles and abstracts of the remaining citations (SS and KK). When a citation was deemed relevant or when the title and abstract were insufficient for including or excluding from the review, the full text article was retrieved. Any disagreements were resolved through discussion and consensus. Selected full text papers were screened independently by the same two reviewers and final agreement was reached on which papers to include for (i) quality appraisal and (ii) supporting data on the identified classification systems such as validity, reliability and generalisability. Additional papers identified by 'hand searching' of reference lists of the included full text papers were also reviewed independently by two reviewers, and selected if appropriate for inclusion.

4.2.4 Data extraction and quality appraisal

Standardised forms were used for (i) describing and (ii) critically appraising each identified classification system. This comprehensive descriptive and critical appraisal framework was originally used by Buchbinder et al. (1996) to evaluate classification systems for soft tissue

disorders of the neck and upper limb. The authors adapted the approach from work done on constructing health status measures and the psychological literature. It has since been used for classifying LBP populations in other reviews (Riddle 1998, Petersen et al. 1999, McCarthy et al. 2004, Billis et al. 2007).

The descriptive items for the classification systems identified in this review include:

- *purpose* of the study
- *method of development* (judgement based e.g. author's opinion; or statistical approach) and the professional background of the developer(s)/author(s)
- *domain of interest* referring to the patient population and setting
- *specific exclusions* for patients
- *categories* within the system and whether other dimensions *(additional axis)* to the condition were considered (e.g. severity or chronicity of symptoms)
- *criteria used* to assign patients to categories (e.g. clinical examination items)
- *training and personnel needed* to perform the classification.

If a classification system was described in more than one publication, data was extracted initially from the original source and complemented with information from subsequent papers that reported on its use or adaptation. Seven criteria were addressed to appraise the methodological quality of the classification systems; these are described in table 4.2. They include appropriateness of purpose; content validity; face validity; feasibility; construct validity; reliability and generalisability. A score of 1 was awarded for meeting a criterion, 0.5 for partially meeting a criterion and 0 for not meeting a criterion or being unable to score it due to lack of information/evidence. Summating the scores from each criterion gave the overall total score out of a maximum score of 7.

Table 4.2 Criteria used to appraise classification systems (Buchbinder et al. 1996)

Criteria	Description
Purpose	• Is the purpose, population and setting clearly specified?
Content validity	• Is the domain and all specific exclusions from the domain clearly specified? • Are all relevant categories included? • Is the breakdown of categories appropriate, considering the purpose? • Are the categories mutually exclusive? • Was the method of development appropriate? • If multiaxial, are criteria of content validity satisfied for each additional axis?
Face validity	• Is the nomenclature used to label the categories satisfactory? • Are the terms used based upon empirical (directly observable) evidence? • Are the criteria for determining inclusion into each category clearly specified? • If yes do these criteria appear reasonable? • Have the criteria been demonstrated to have reliability or validity? • Are the definitions of criteria clearly specified? • If multiaxial are criteria of face validity satisfied for each additional axis?
Feasibility	• Is the classification simple to understand? • Is the classification easy to perform? • Does it rely on clinical examination alone? • Are special skills, tools and/or training required? • How long does it take to perform?
Construct validity	• Does it discriminate between entities that are thought to be different in a way appropriate for the purpose? • Does it perform satisfactorily when compared to other classification systems which classify the same domain?
Reliability	• Does the classification system provide consistent results when classifying the same conditions? • Is the intraobserver and interobserver reliability satisfactory?
Generalisability	• Has it been used in other studies and/or settings?

4.3 Results

4.3.1 Search results

A flow diagram summarising the systematic search and study selection process is given in figure 4.1. The initial database search yielded 16,891 references, and an additional 21 papers were identified through hand searches of reference lists. After removal of duplicates, 13,358 records remained. Following initial screening of titles and abstracts to exclude papers that were clearly irrelevant, 417 remaining titles and abstracts were selected and subsequently screened. 121 articles were identified for full text review. 72 papers were excluded, reasons for exclusion included:

- the paper did not describe the development or use of a classification system for LBLP
- the classification system did not include patients with back-related leg pain
- the classification system did include leg pain patients, but the identified subgroups did not allow the reader to distinguish between who did and did not have leg pain in each subgroup.

49 papers were selected for inclusion which reported on 21 classification systems (figure 4.1).

4.3.2 Data extraction and appraisal of selected studies

Based on approaches used in previous LBP classification reviews (McCarthy et al. 2004, Billis et al. 2007, Fairbank et al. 2011), the 21 classification systems were organised into five themes reflecting the purpose and criteria of the classification system: (i) clinical features (ii) pathoanatomical source of pain (iii) treatment based approach (iv) screening tools and clinical prediction rules and (v) pain mechanisms.

Data extraction from the papers is presented in tables 4.3 to 4.7. Each table presents one of the classification system themes and gives a descriptive summary of the individual papers within each theme.

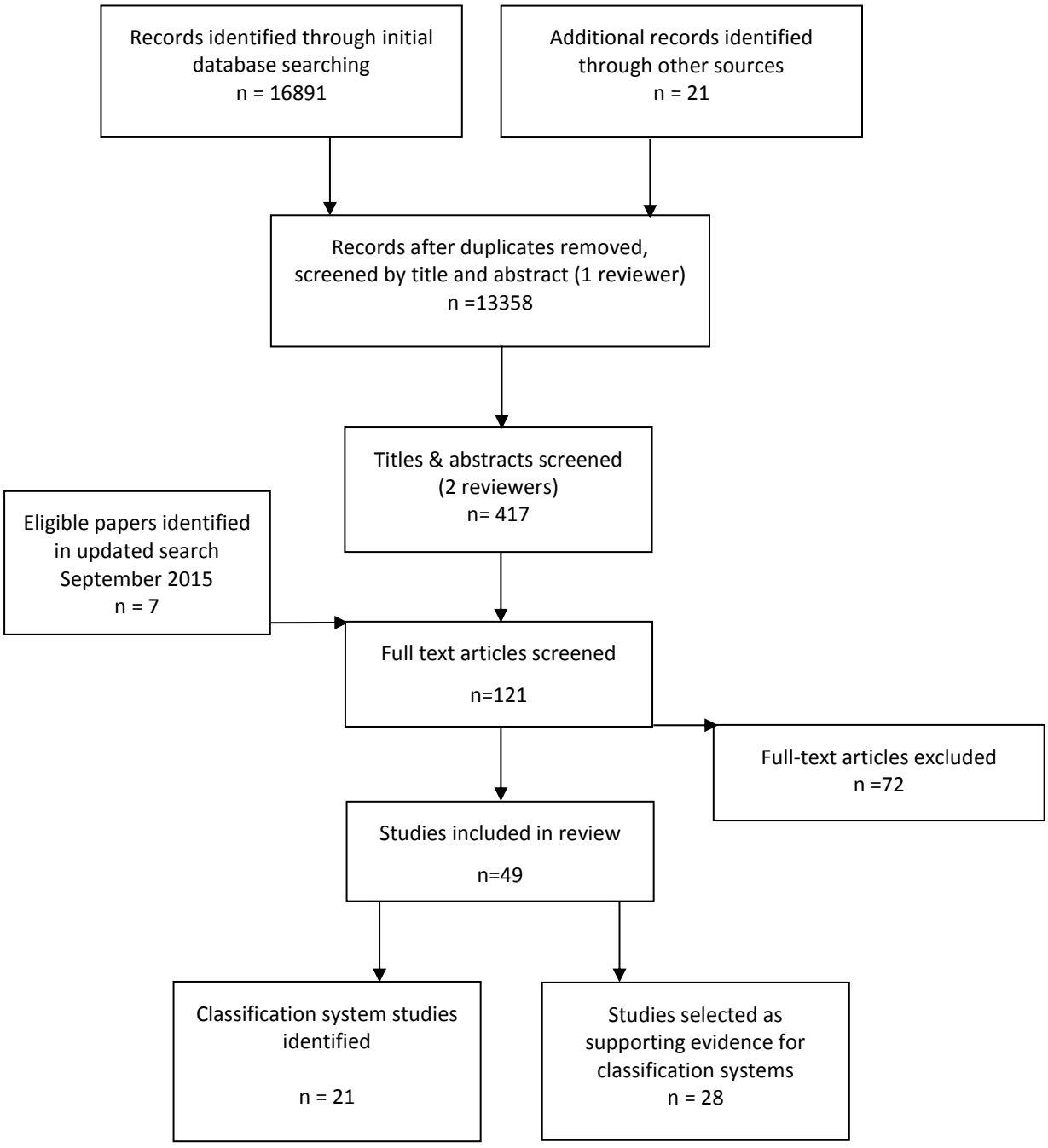


Figure 4.1 Study selection flow diagram

Table 4.3 – 4.7 Data extraction for classification systems

Table 4.3 Systems classifying by Clinical Features

Primary Author	Purpose	Method of Development	Domain of Interest	Specific Exclusions	Categories	Criteria used	Training / Personnel needed
Barker (1990)	Devise classification meaningful to General Practitioner (GP).	Judgement approach. Profession of developer: GP.	Low Back Pain (LBP). 486 patients attending authors' GP practice.	Febrile illness, backache accompanied by many other complaints.	1: Acute lumbago 2: Acute mechanical derangement 3: Acute sciatica 4: Sacro-iliac joint (SIJ) 5: Mild sciatica	Patient history, pain location drawings, clinical examination.	None.
Ben Debba et al. (2000)	Assign LBP patients into one of four modified Quebec Task Force Classification categories.	Judgement and statistical approach. Profession of developer: Neurosurgeon.	Persistent LBP. 1,997 patients from tertiary care.	Age under 25, ≥1 prior surgical or interdiscal procedure, no pain in the small of the back.	1: Back pain only 2: Back and above knee pain 3: Back and below knee pain 4: Back and below knee pain with positive straight leg raise (SLR)	Spatial distribution of patient's pain (from questionnaire). Results of SLR test.	Standardisation of SLR performed by clinician or technician.
Nachem-son and Andersson (1982)	Introduce a simple classification system suitable for use in epidemiological screening.	Judgement approach. Profession of developer: Orthopaedic spine surgeon.	LBP.	None.	1: Insufficienta dorsi 2: Lumbago 3: Sciatica 4: Rhizopathy 5: Lumbago sciatica Additional axis: Yes-Duration and recurrence	Patient history and clinical examination. Radiographic results can be used.	Authors report it is simple to use.
Spitzer et al. (1987)	Compile a diagnostic classification system for:	Judgement approach. Profession of	LBP.	None.	1: Pain without radiation 2: Pain + radiation proximal extremity 3: Pain + radiation distal	Patient history. Clinical examination and paraclinical test results (laboratory tests,	Able to interpret investigative tests.

Table 4.3 Systems classifying by Clinical Features

Primary Author	Purpose	Method of Development	Domain of Interest	Specific Exclusions	Categories	Criteria used	Training / Personnel needed
	clinical decision making; establishing prognosis; evaluating quality of care; Conducting scientific research.	developer: Multidisciplinary task force representing wide range of disciplines.			extremity 4: Pain + radiation to upper limb/lower limb with neurological signs 5: Presumptive root compression, +ve image 6: Root compression, +ve image 7: Spinal stenosis 8: Post surgical < 6 months 9: Post surgical > 6 months 10: Chronic pain syndrome 11: Other diagnoses Additional Axis: Yes Work and duration	radiography, imaging methods, Electromyography (EMG) nerve blocks).	
Sweetman et al. (1992)	Describe common patterns of LBP and identify clinical tests to help recognize the patterns.	Statistical approach. Profession of developer: Rheumatologist.	LBP. 301 patients referred from GP to rheumatology clinic.	Less than 15 or over 75 years old.	1: Persistent unilateral back pain and sciatica 2: Back pain or sciatic switching sides (sacroiliitis) 3: Central/ bilateral back pain 4: Lateral flexion or rotation cause pain on the opposite side(facet joint) 5: Back pain at rest on one side but pain on opposite side with several tests (unstable L4/5 syndrome) 6: Dorso lumbar junction	Questionnaire and clinical examination and x-ray.	Uses a computer algorithm for pattern recognition.

Table 4.3 Systems classifying by Clinical Features

Primary Author	Purpose	Method of Development	Domain of Interest	Specific Exclusions	Categories	Criteria used	Training / Personnel needed
					conditions 7. Persistent unilateral back pain and sciatica with loss of lower limb reflex (Disc with nerve root compression)		

Table 4.4 Systems classifying by Pathoanatomy

Primary Author	Purpose	Method of Development	Domain of Interest	Specific Exclusions	Categories	Criteria used	Training/ Personnel needed
Bernard and Kirkaldy Willis (1987)	Determine pathology causing LBP.	Judgement approach. Profession of developer: Orthopaedic surgeon.	LBP. Medical record review of 1293 patients, majority of whom had failed initial treatment by primary care physicians.	None.	Group A: well recognised syndromes 1. Herniated nucleus pulposus 2. Lateral spinal stenosis 3. Central spinal stenosis 4. Spondylolisthesis 5. Segmental instability Group B: less well recognised syndromes 6. Sacroiliac joint 7. Posterior joint 8. Maigne's syndrome 9. Gluteus maximus 10. Gluteus medius 11. Quadratus lumborum 12. Piriformis 13. Hamstring origin 14. Tensor fascia latae Group C: remaining syndromes	Medical records and response to treatment which included: manipulation/stretching; injections; radiofrequency denervation; palpation; joint motion tests, neural tension tests and neurological testing, response to surgery, pain provocation palpation, xray and computed tomography (CT) scans.	None.

Table 4.4 Systems classifying by Pathoanatomy

Primary Author	Purpose	Method of Development	Domain of Interest	Specific Exclusions	Categories	Criteria used	Training/ Personnel needed
					15. Pseudarthrosis 16. Non specific 17. Post fusion stenosis 18. Anklyosing spondylitis 19. Disc space infection 20. Tumour 21. Arachnoiditis 22. Lateral femoral nerve entrapment		
Cassisi et al. (1993)	Explore differences between two groups of chronic LBP patients.	Judgement approach. Profession of developer: Neurosurgeon.	Chronic LBP. 151 patients in tertiary care.	Neoplasm, mechanical, toxic-metabolic, inflammatory-infectious, vascular and psycho-physiological conditions.	Myofascial pain. Disc herniation.	Patient history and clinical examination.	None.
Hahne et al. (2011)	Identify patho-anatomical subgroups with subacute LBP. For use in a randomised controlled trial (RCT): the STOPS trial.	Judgement approach including an expert panel of physiotherapists. Profession of developer: Physiotherapist.	LBP +/- leg pain. Subacute pain lasting between 6 weeks and 6 months.	Red flags, recent spinal injections, previous spinal surgery, recent regular physiotherapy treatment.	1: Reducible discogenic pain 2: Non reducible discogenic pain (not responsive to mechanical loading strategies) 3: Disc herniation with associated radiculopathy 4: Facet joint dysfunction 5: Multi-factorial persistent pain	Patient history and clinical examination.	Unclear what specific training is needed for classification.

Table 4.4 Systems classifying by Pathoanatomy

Primary Author	Purpose	Method of Development	Domain of Interest	Specific Exclusions	Categories	Criteria used	Training/ Personnel needed
Paatelma et al. (2009)	Evaluate the reliability of a patho-anatomical classification system.	Judgement approach. Profession of developer: Physiotherapist.	LBP +/- leg pain. 21 patients.	Age over 56, LBP > 3 months.	1: Discogenic pain 2: Lumbar instability 3: Spinal Stenosis 4: Segmental dysfunction/facet pain 5: SIJ dysfunction/pain	Patient history and clinical examination.	5 ½ day training sessions to standardise tests. 30 minute assessment.
Petersen et al. (2003)	Develop a classification system with pathoanatomic orientation for use in primary care.	Judgemental approach. Profession of developer: Physiotherapist. Slightly modified version of Laslett and van Wijmen (1999) classification system.	Non-specific LBP.	Red flag symptoms, hip disorders, suspected referred pain from viscera.	1: Disc syndrome (reducible;irreducable and non-mechanical) 2: Adherent nerve root 3: Nerve root entrapment 4: Nerve root compression 5: Spinal stenosis 6: Zygapophysial joint 7: Postural 8: Sacro-iliac joint 9: Myofascial pain 10: Adverse neural tension 11: Abnormal pain 12: Inconclusive	Patient history and clinical examination.	Some training required and experience of the McKenzie assessment. Takes 1 hour to complete.
Vining et al. (2013)	Create a classification system based on available evidence for use in research and clinical setting	Judgement approach. Based on Petersen et al. (2003) model. Profession of developer: Chiropractor.	LBP.	none	1. Screening 2. Nociceptive - Discogenic - SIJ - Zygapophyseal joint -Myofascial 3. Neuropathic - Compressive radiculopathy - Non compressive	Patient history and clinical examination. Questions and physical component of the Leeds Assessment for Neuropathic Symptoms and Signs (LANSS). Arterial brachial index test for neurogenic	None.

Table 4.4 Systems classifying by Pathoanatomy

Primary Author	Purpose	Method of Development	Domain of Interest	Specific Exclusions	Categories	Criteria used	Training/ Personnel needed
					radiculopathy - Neurogenic claudication - Central pain 4. Functional instability 5. Other diagnoses	claudication if indicated	

Table 4.5 Systems classifying by Treatment based approach

Primary Author	Purpose	Method of Development	Domain of Interest	Specific Exclusions	Categories	Criteria used	Training/ Personnel needed
Delitto et al. (2012)	Classify, define and assign treatment approaches to musculoskeletal conditions using the World Health Organisation terminology related to International Classification of Functioning, Disability and Health.	Judgement approach. Profession of developer: Content experts appointed by Orthopaedic section of the American Physical Therapy Association.	LBP.	Serious medical conditions.	1: Lumbosacral segmental/somatic dysfunction with mobility deficits 2: Spinal instabilities with movement coordination impairments 3: Flatback syndrome or lumbago due to displacement of disc 4: Of acute low back pain with related (referred) lower extremity pain 5: Lumbago with sciatica 6: Low back pain/ strain/lumbago -with related cognitive or affective tendencies 7: Of chronic LBP with related generalized pain	Patient history and clinical examination. Questionnaires for category with related cognitive or affective tendencies.	None.

Table 4.5 Systems classifying by Treatment based approach

Primary Author	Purpose	Method of Development	Domain of Interest	Specific Exclusions	Categories	Criteria used	Training/ Personnel needed
					Additional axis-Yes- acute, subacute, chronic		
Hall et al. (1994)	Identify typical patterns of pain and determine treatment direction.	Judgement approach. Profession of developer: Spinal surgeon and physical therapist.	LBP.	None.	I: LBP +/- referred pain aggravated by flexion, slow onset lasting weeks 2: LBP +/- referred pain aggravated by extension, sudden onset lasts 1-2 weeks 3: Leg dominant pain due to nerve involvement, aggravated by flexion, slow onset, lasts weeks 4: Leg dominant pain due to nerve involvement aggravated by activity and extreme sustained extension, relieved by rest. Rapid onset 5: Abnormal pain behaviour, chronic pattern associated work/sleep/psycho/social issues	Patient history and clinical presentation.	None.
McKenzie (1981)	Develop a classification to determine choice of treatment.	Judgement approach. Profession of developer: Physiotherapist.	LBP.	Constant pain, serious pathology, neurological deficit.	1: Postural 2: Dysfunction 3: Derangement 1-7	Patient history and clinical examination.	Training in McKenzie assessment desired.

Table 4.5 Systems classifying by Treatment based approach

Primary Author	Purpose	Method of Development	Domain of Interest	Specific Exclusions	Categories	Criteria used	Training/ Personnel needed
Albert et al. (2012)	Examine the association between treatment outcome and baseline type of disc lesion.	Judgement approach. Profession of developer: Physiotherapist.	Radicular pain with dermatomal distribution to knee or below. 176 patients with sciatica involved in large RCT.	Over 65 years old, leg pain < 3 on 1-10 scale, duration < 2 weeks or > 1 year, red flags, previous back surgery, serious comorbidities.	5 groups based on their pain response: 1: Abolition centralisation 2: Reduction centralisation 3: Unstable centralisation 4: Peripheralisation 5: No change	Response to repeated moving testing. Lumbar magnetic resonance imaging (MRI).	Training from McKenzie accredited physiotherapist.

Table 4.6 Systems classifying by Screening Tools/Clinical Prediction Rules

Primary Author	Purpose	Method of Development	Domain of Interest	Specific Exclusions	Categories	Criteria used	Training/ Personnel needed
Fritz et al. (2007)	To identify if there is a subgroup of patients likely to respond to traction	Judgement and statistical approach. Profession of developer: Physiotherapist.	LBP with signs of nerve root compression Primary care.	Over 60 years old, red flags, previous spinal surgery in past 6 months, pregnancy, absence of symptoms when sitting.	Patients likely to benefit from traction have: leg symptoms; signs of nerve root compression; symptom peripheralisation on extension movement; positive crossed SLR	Patient history and clinical examination	None.
Roach et al. (1997)	To develop screening tests to place patients into a predetermined	Judgemental and statistical approach. Profession of	LBP. 106 tertiary care patients.	Back pain treatment within last year, history of back surgery,	1: Disc 2: Spinal stenosis 3: Disc disease with spinal stenosis 4: Benign low back pain	Questionnaire (Pain response to activity and position questionnaire).	None.

Table 4.6 Systems classifying by Screening Tools/Clinical Prediction Rules

Primary Author	Purpose	Method of Development	Domain of Interest	Specific Exclusions	Categories	Criteria used	Training/ Personnel needed
	structure-based diagnostic classification system.	developer: Physiotherapist.		unconfirmed diagnosis at end of study.		Additional advanced diagnostic tools such as CT/MRI and lab work.	
Scholz et al. (2009)	Test the utility of a tool (Standardised Evaluation of Pain (StePs)) to differentiate between radicular and axial pain.	Statistical approach. Profession of developer: Anesthesiologist and Pharmacologist.	Chronic LBP.	Over 18 years old, Pain < 3 months, global pain intensity in week prior to recruitment <6 severe psychiatric or medical illness, another painful or neurological disease or local infection.	Axial low back pain. Radicular low back pain. Most discriminatory items for radicular pain: positive SLR, deficit in detection of cold and reduced response to pinprick Also identified subtypes of radicular and axial LBP based on clusters of signs and symptoms.	Brief structured interview of 6 questions and 10 standardized physical tests.	Training in administering the tests in physical examination to assess cutaneous changes, pressure; pinprick; vibration; thermal sensitivity and proprioception.

Table 4.7 Systems classifying by Pain Mechanisms

Primary Author	Purpose	Method of Development	Domain of Interest	Specific Exclusions	Categories	Criteria used	Training/ Personnel needed
Schafer et al. (2009a)	Identify the predominant pain mechanisms responsible for patient's back and leg pain, to guide treatment	Judgement approach. Profession of developer: Physiotherapist.	Low back related leg pain.	Recent surgery or nerve root block, diabetes vascular disease in lower extremities, systematic	1. Central sensitisation 2. Denervation 3. Peripheral nerve sensitisation 4. Musculoskeletal	Patient history and clinical examination. Leeds Assessment for Neuropathic Symptoms and Signs (LANSS) questionnaire.	None.

Table 4.7 Systems classifying by Pain Mechanisms

Primary Author	Purpose	Method of Development	Domain of Interest	Specific Exclusions	Categories	Criteria used	Training/ Personnel needed
	decisions.			disease. Inflammatory arthropathies.			
Smart et al. (2011)	Identify signs and symptoms of patients categorised according to mechanism-based classification of pain.	Judgement and statistical approach. Profession of developer: Expert consensus panel to develop clinical criteria list.	LBP +/- leg pain. 464 patients.	History of diabetes, central nervous system injury, pregnancy, non musculo-skeletal LBP.	1. Centralisation pain 2. Peripheral neuropathic 3. Nociceptive	Patient history and clinical examination.	Practical training with an assessment manual provided.
Nijs et al. (2015)	Apply a pain classification system to LBP patients	Judgement approach Profession of developer: Expert opinion of 18 international pain experts	LBP	none	1. Nociceptive pain 2. Neuropathic pain 3. Central sensitisation	Patient history, clinical examination, diagnostic investigations	None.

Quality appraisal of each methodological criterion (according to the tool used and described in section 4.2.4, page 43) for the 21 classification systems was done and the overall score for each system was calculated (out of a maximum score of 7). These results are presented in Appendix G. To derive this score, any supporting studies (reporting on reliability, construct validity, generalisability) for the classification systems were also included. A list of the papers organised in themes and the overall appraisal quality score for each paper is presented in table 4.8.

Table 4.8 Overview of classification systems organised by themes and accompanying quality scores

Clinical Features		**Pathoanatomy**		**Treatment Based Approach**		**Screening Tools/ Clinical Prediction Rules**		**Pain Mechanisms**	
Barker 1990	2	Bernard, Kirkaldy Willis 1987	2	Albert et al. 2012	4	Fritz et al. 2007	3	Schafer et al. 2009	5
Bendebba et al. 2000	3.5	Cassisi et al. 1993	3	Hall et al. 1994	5	Roach et al. 1997	3	Smart et al. 2011	5
Nachemson, Andersson 1982	3.5	Hahne et al. 2011	3	Mckenzie 1981	5.5	Scholz et al. 2009	4	Nijs et al. 2015	2.5
Spitzer et al. 1987	4	Paatelma et al. 2009	3.5	Delitto et al. 2012	3.5				
Sweetman et al. 1982	2.5	Petersen et al. 2003	4						
		Vining et al. 2013	3.5						

4.3.3 General summary of classification systems organised by themes

(i) Clinical Features

Five papers described classification systems according to clinical presentation of signs and symptoms (table 4.3). Overall, on the appraisal tool these systems scored low (median score=3.5, interquartile range (IQR) =1.5) and with the exception of the Quebec Task Force Classification

(QTFC) system (Spitzer et al. 1987), there was no supporting work on the systems' validity and generalisability.

Purpose

The common purpose among these systems was to assign LBP patients into categories based on similar clinical characteristics which would be useful in both clinical and research settings.

Method of development

Development methods were mostly "judgement based" but varied in approach. They included the opinion of the author(s) who developed the system (Nachemson and Andersson 1982, Barker et al. 1990) and use of expert international panels (Spitzer et al. 1987). Statistical methods were used by Sweetman et al. (1992) to develop their system. The authors used cluster analysis to statistically derive seven patterns of LBP using clinical information from a cohort of 301 patients. The authors subsequently used their clinical judgement to propose pathoanatomic labels to match the statistically derived patterns.

Categories

The QTFC (Spitzer et al. 1987) system has eleven categories, based on location of pain and/or identification of pathology from clinical examination and confirmed by imaging tests. It also includes two additional axes of classification: duration of symptoms and work status. Ben Debba et al. (2000) proposed a simple classification system based on the first four categories of the QTFC with the addition of the SLR (described in chapter three, section 3.3.2, page 27) to determine the presence or absence of neurological signs. The authors used the SLR because they state it is the clinical test most commonly used to determine "presence of neurological signs" in LBP.

The systems proposed by Barker (1990) and Nachemson and Andersson (1982) have five similar categories. Barker's (1990) categories were based on observations of 486 LBP patients attending his general practice over five years. No validation or reliability work has been done on this system

whereas Nachemson and Andresson (1982) refer to several studies that have used their system in Sweden.

Validity

Despite the robust methodological approach to classification development by Sweetman et al. (1992), subsequent testing of the system's validity was not performed and allocation to the categories required a full computerised assessment, hence it was not feasible for use in a primary care setting. Validation work was done by BenDebba et al. (2000) which showed that the four classes of the QTFC, with the inclusion of response to SLR, differed from each other on pain and function measures. The QTFC is by far the most extensively investigated and adapted classification system in this theme of clinical features. In particular the first four categories, which are based on pain location and clinical signs and do not involve advanced imaging techniques such as MRI. These four categories have shown good discriminative and predictive ability, with category 3 (LBP and below knee pain) and category 4 (LBP and below knee pain with neurological signs) considered the most severe categories associated with poorer function and inability to return to work (Loisel et al. 2002), poorer movement quality (Marras et al. 1995), higher presence of neuropathic pain (Attal et al. 2011), and showing less favourable response to treatment (DeRosa and Peterfield 1992). Atlas et al. (1996) showed that in LBP patients with radiating leg pain, severity of their symptoms and probability of surgical treatment increased from category 2 (LBP and above knee pain) to category 6 (spinal nerve root compression confirmed by imaging). Frank et al. (2000) used the QTFC system to group 657 LBP patients referred to a rheumatology service. They found that patients with leg pain were significantly more disabled and depressed than those without leg pain. A modified QTFC system, which also included symptoms duration and insurance status, classified 263 LBP patients and reported outcomes following intensive physiotherapy treatment intervention (Hearne 1997). Poorer outcomes seemed to be associated with patients that had predominantly leg pain. In a Danish study, 2673 patients were classified into one of the first four categories of the QTFC (Kongsted et al. 2012). Patients with leg pain were more severely

affected than those with local back pain, and those with signs of nerve root involvement were the ones most severely affected in terms of pain, disability, work participation and psychosocial profile. In a prospective study using the same dataset (Kongsted et al. 2013), leg pain with or without neurological signs predicted activity limitation and time off work but was not influenced by whether the pain is above or below the knee.

One study has compared the discriminative and predictive properties of the QTFC system (categories 1 to 4) with a classification system based on whether the leg pain 'centralised' towards the back or 'peripheralised' towards the feet in response to movement (Werneke and Hart 2004). Both systems could differentiate between groups' baseline pain intensity and disability but the classification method by leg pain "centralisation" or "peripheralisation" was superior in predicting treatment outcomes and long term work status.

Reliability

No additional studies were identified that reported on the reliability of the classification systems within the "clinical features" theme. Within their original paper, Sweetman et al. (1992) reported 70% reproducibility when their classification algorithm was used on a smaller test sample of 80 LBP patients.

(ii) Pathoanatomy

Six systems were identified in the pathoanatomical theme (table 4.4). Overall on the appraisal tool the systems scored low (median score=3.25, IQR =0.875) mainly due to lack of supporting work on the systems' validity and generalisability.

Purpose

The purpose of the five pathoanatomical classification systems was to identify a pathology or anatomical structure responsible for a person's LBP +/- leg pain. In two of the six systems (Petersen et al. 2003, Hahne et al. 2011) specific treatments were suggested for the identified categories.

Method of Development

All systems used a judgement approach for development. A group of Canadian chiropractors created a diagnostic classification system based on available evidence (Vining et al. 2013). They began with Petersen et al.'s (2003) pathoanatomical system and modified some of the categories, but also added or updated diagnostic criteria based on available evidence. No validation or reliability work has been subsequently published to support this system.

Categories

The number of categories in all the systems ranged from two (Cassisi et al. 1993) to twenty-two (Bernard and Kirkaldy-Willis 1987). Advanced diagnostics such as injections and imaging tests are included in some of Bernard and Kirkaldy-Willis (1987) classification categories. All studies recognised the lumbar disc as a source of pain, with two studies describing the type of disc herniation as reducible or non-reducible (Petersen et al. 2003, Hahne et al. 2011). Patients classified as having a "reducible disc" are considered to present with a directional spinal movement preference and likely to have pain which moves ("centralises") from the leg to the low back because of a theoretical "posterior or posterolaterally migrated nucleus pulposus that can be 'reduced' into a more central and non-pain provoking position" (Hahne et al. 2011). The non-reducible disc pain category is assigned to patients who do not demonstrate a directional preference in response to spinal movement. Hahne et al. (2011) also describe disc herniation with associated radiculopathy[3]. Facet joint dysfunction as a source of pain was included in five of the systems, and four systems included stenosis. Leg pain due to sciatica was considered in all six groups but under varying nomenclature and criteria for diagnosis (Table 4.9). Cassisi et al. (1993) used the term "myofascial pain ", but their criteria for diagnosing myofascial pain could also have fitted the diagnostic criteria for facet joint pain as defined by Hahne et al. (2011). The system

[3] Radiculopathy was defined by Hahne et al. (2011) as leg pain with at least one clinical sign of: reflex deficit, myotome deficit, sensory deficit or positive neural tension, and evidence of disc herniation consistent with clinical findings on imaging.

proposed by Vining et al. (2013) bases many of their diagnostic criteria on available statistically derived diagnostic models. For example, compressive radiculopathy is a subcategory within a neuropathic pain category. They used clinical assessment items identified from a statistically derived diagnostic tool (Vroomen et al. 2002) and added the score from questions and the physical component of the LANSS neuropathic pain tool questionnaire, to assign patients to this category.

Validity

There was some evidence of supporting work that reported on validity of the systems**.** Petersen et al. (2003) detail content validity for several of their categories, citing studies that have shown the ability to discriminate between different categories. A follow up study tested how six of the most common pathoanatomical categories in their system, diagnosed by physiotherapists in chronic LBP patients, agreed with selected reference standards (advanced imaging tests, injections or discography) (Laslett et al. 2005). Agreement was low with a kappa of 0.31 (Laslett et al. 2005). Nerve root pain was the second most common diagnosis after discogenic pain and had poor agreement with the reference standard (Laslett et al. 2005). Hahne et al. (2011) designed their classification system for use in a clinical trial to compare specific physiotherapy treatment to advice for the five subgroups of LBP. Recently published results show a reduction in activity limitation and back and leg pain intensity across a 52-week follow-up for patients who received 10 individual sessions relative to two sessions of guideline-recommended advice (Ford et al. 2016).

Reliability

Inter-rater reliability was reported on two systems. Paatelma et al. (2009) demonstrated good reliability (kappa=0.6) between specialists and non-specialists and slightly higher reliability (kappa=0.65) among the specialists. Petersen et al. (2004) demonstrated good reliability (kappa= 0.62) among therapists trained to use their classification system.

Feasibility

Petersen et al. (2003) acknowledged that their system was difficult to perform and would take up to one hour. Paatelma et al. (2009) reported their system took 30 minutes in order to subdivide patients into one of its five categories. These were the only two out of all 22 classification systems that gave information on how long the assessments would take to perform. Both these systems and the one by Hahne et al. (2011) required training for the clinicians using them.

(iii) Treatment Based Approach

Four systems were considered to be treatment based approach classifications (table 4.5). They scored a median of 4.5 (IQR 1.5). This higher score reflects the subsequent supporting published work on the Mckenzie (1981) and Hall et al. (1994) studies.

Purpose

The common purpose among the treatment based classification approach is to guide or match specific treatment allocation to LBP patients that have been grouped according to clinical features, functional status or response to spinal movement. Development of all systems was considered "judgement based" Delitto et al.'s classification (2012) also involved an expert panel.

Categories

In the McKenzie system (Mckenzie 1981), also known as Mechanical Diagnosis and Therapy (MDT), patients are classified, according to how their symptoms respond to repeated spinal movements and sustained positions, into one of three syndromes (Postural, Dysfunction or Derangement) with several subsyndromes (Hefford 2008). Nested within the Derangement syndrome are subsyndromes that include patients with leg pain with or without sciatica symptoms. Identification of a specific syndrome guides the therapist to select an appropriate treatment approach. The system was developed in 1981 based on observations from the clinical experience of its founder, Robin Mckenzie. Since its development, several studies have been published supporting its validity, reliability and generalisability.

Hall et al. (1994) endeavoured to offer a "non-threatening" diagnosis to patients, as they believed attributing a patient's pain to a definitive pathological diagnosis was rarely possible. They recognised five distinct patterns of pain, with the dominant pattern determining the appropriate treatment. Patterns I and II can include referred leg pain, patterns III and IV have leg dominant pain due to involvement of spinal nerves. The fifth pattern involves abnormal pain behaviour with associated work, sleep and psychosocial issues. Despite basing their classification on the clinical picture as opposed to anatomy or pathology, the authors did offer explanations for patients of 'painful disc', 'worn spinal joints', 'pinched nerve' and 'bony spurs within the spine' for categories I to IV respectively.

The system described by Albert et al. (2012) classified sciatica patients according to whether their leg pain "centralised" (relocated towards the low back), "peripheralised" (leg pain moved distal to the back towards the feet) or did not change. Other sources refer to this classification as pain pattern classification (Werneke and Hart 2004). The absence of centralisation predicts a poorer outcome, purported to be based on the McKenzie theoretical disc model where the interdiscal pressure is no longer intact due to interruption of the annulus hence centralisation cannot occur.

Clinical guidelines published by the Orthopaedic group of the American Physical Therapy Association (Delitto et al. 2012) proposed a function/ impairment based LBP classification system. LBP presentations without related serious medical or psychological issues were divided into three categories associated with clinical findings: (a) mobility impairment of the spine; (b) associated leg pain and (c) generalised pain. These three categories were further subdivided using International Statistical Classification of Diseases and Related Health Problems (ICD) categories and the International Classification of Functioning Disability and Health (ICF) categories. Experts in LBP, who developed the system, categorised LBP patients with and without leg pain into eleven different mutually exclusive impairment patterns upon which to base intervention strategies. It is considered a treatment based classification as the authors' state that it parallels the treatment based classification system first described by Delitto et al. (1995) but with the addition of ICF

impairments of body function terminology; a mental and sensory impairments category and inclusion of time since onset of symptoms and relationship between pain and movement. Each category is given a recommended treatment approach.

Validity

A systematic review and meta-analysis (Machado et al. 2006) evaluated the effectiveness of the McKenzie treatment approach and concluded that in patients with acute LBP, the McKenzie method produces similar improvements in pain or disability as passive therapy and advice to stay active. In a later RCT where all patients received information, advice and either manipulation or McKenzie based treatment, the authors reported a favourable outcome with the McKenzie approach at 2 month follow up (Petersen et al. 2011).

In Albert et al.'s study (2012), all patients received exercise and advice. Similar improvements in activity limitation and leg pain were seen in those who were categorised as 'centralisers' and in those whose symptoms 'peripheralised', which refutes the proposed McKenzie model theory. Additionally, symptoms centralised in over 90% of patients with MRI confirmed sequestrated or extruded discs.

Hall et al. (2009) published results of their primary care based study which compared outcomes in patients classified according to their system (n=1356), to patients managed without a classification system (n=754). They concluded that classification had a positive effect on pain relief post treatment, resulted in less treatment days and patients were less likely to use pain medication. Weaknesses of this validation study included use of a double - cohort study design, i.e. comparison of two cohorts and not an RCT, which meant potential differences at baseline between the usual care groups and the classified group. The intervention for the non-classified patients was also poorly described.

Although no supporting work has been done to examine reliability, validity or generalisability of the system proposed by Delitto et al. (2012), the authors designed the treatment system based on

the validation work done by others including: (i) the original four category treatment based classification approach of Delitto et al. (1995) which uses specific exercise[4], stabilisation[5], traction[6] and manipulation[7]; (ii) Mckenzie centralisation and directional preference exercises, (iii) flexion exercise for stenosis, (iv) lower quadrant nerve mobilisation procedures, (v) patients education and counselling and (vi) progressive endurance exercise and fitness activities. The guidelines, in which the Delitto system is described, are to be reviewed in 2017 or sooner if new evidence is available.

Reliability

The reliability of the Hall and McIntosh (1994) classification was examined by Wilson et al. (1999) using 59 examiners and 204 patients. Among the experienced raters agreement was 80% with good reliability (kappa=0.6), and similar results were seen among the novice raters (agreement; 77%, kappa = 0.6).

A review concluded that there is high strength of evidence of substantial agreement among clinicians certified (formally trained and successfully completing an examination) in the McKenzie approach for classifying patients, and evidence of less agreement for subsyndromes and among non-certified clinicians (Fairbank et al. 2011). However a more recent reliability study involving 1662 patients and 47 raters indicated that inter-rater reliability was not acceptable for therapists at any level of McKenzie training (Werneke et al. 2014).

(iv) Screening Tools/ Clinical Prediction Rules

Three papers were grouped under the theme of screening tools and prediction rules (table 4.6). On the appraisal tool, they had a median score of 3, (IQR 1). All papers combined judgement and

[4] Extension or flexion based exercises, or exercise to correct a lateral shift of the pelvis.
[5] Trunk strengthening and stability exercises or wearing a back support brace.
[6] Mechanical traction force to purportedly stretch or decompress the spine.
[7] Thrust mobilisation technique applied to the low back area.

statistical methodological approaches to develop their systems but had limited follow up supporting studies.

Purpose

The common purpose of these classification approaches was to identify clinical features that either guide diagnosis (Roach et al. 1997, Scholz et al. 2009) or assist with treatment selection (Fritz et al. 2007). They are grouped together because they have similar concepts and methodology.

Method of Development and Categories

Scholz et al. (2009) used statistical analysis to identify the most discriminatory items from a neuropathic pain assessment tool (Standardized Evaluation of Pain (StEP)) to differentiate between LBP patients with and without radicular leg pain (sciatica). Using cluster analysis they also identified four subtypes with similar pain patterns.

Roach et al. (1997) developed screening test algorithms, based on patients' answers to a Pain Response to Activity and Position questionnaire, to place patients into four predetermined "structure-based" diagnostic classifications. A judgement approach method was used to preselect the four LBP categories of disc, spinal stenosis, disc disease with spinal stenosis and benign LBP. Fritz et al. (2007) identified a subgroup of LBLP patients with signs of nerve root compression, likely to respond to mechanical traction and found baseline variables associated with greater improvements with traction were peripheralisation of leg symptoms with extension movement, and a crossed SLR.

Validity

Performance of the tools was assessed by both Roach et al. (1997) and Scholz et al. (2009). Subsequent validation of the algorithms (using sensitivity/specificity/positive and negative predictive values) in Roach et al.'s study identified misclassification of a substantial number of patients. The authors concluded that patients' responses to position and activity were not

sufficient for diagnosis, but they may be useful in 'ruling in' or 'ruling out' a particular diagnosis, hence the algorithms could be a first step in directing and focusing the clinical assessment. Scholz et al.'s tool for the distinction between radicular and axial[8] LBP, identified patients with radicular pain with high sensitivity (92%; CI 83%, 97%) and specificity (97%; CI 89%, 100%). Compared with MRI, their tool had substantially higher diagnostic accuracy than MRI (MRI; 96% sensitivity, 18% specificity). However, not all patients in their study received an MRI scan. No published work was available on application of the rules to another population sample for external validation.

Reliability

The reliability of the reference standard or diagnostic categories was not tested in any of the studies. Test-retest reliability of the screening algorithms was carried out by Roach et al. (1997). The kappa values ranged from 0.57 to 0.91, for the four diagnostic categories, suggesting good to almost perfect reliability. Scholz et al. (2009) recommended that future studies were needed to address test-retest and inter and intra-rater reliability of their questionnaire. No reliability data was available for Fritz et al.'s (2007) treatment prediction rule.

(v) Pain Mechanisms

Three pain mechanism classification system studies were identified (table 4.7) scoring a median 5 points (IQR 2.5) on the appraisal tool. Schafer et al. (2009a) specifically designed their system for LBLP patients. There is considerable overlap in the categories proposed by the three systems.

Purpose

The purpose of the pain classification systems was to subgroup patients into categories that reflect the underlying neurophysiological mechanism responsible for causing and maintaining

[8] The authors describe "axial" LBP as "patients without clinical signs of nerve root involvement", but it is not clear if these patients had leg pain. Attempts to contact the authors were made to clarify "axial pain" but no response was received.

patients' pain. A specific aim of Schafer et al.'s (2009a) system was to improve outcome for patients identified as most likely to respond to 'neural mobilisation'.[9]

Method of Development

All systems were initially developed using a judgement approach. Smart et al. (2011) subsequently used statistical analysis to identify discriminatory clusters of signs and symptoms associated with each of their pain mechanism categories.

Schafer et al. (2009a) based their categories on the clinical experience of the authors as well as a review of relevant observational and experimental research literature. Smart et al. (2011) initially described their classification system as applicable to all musculoskeletal pain, based on previously published theories of pain mechanisms by Woolf (2004), Yunus (2005) and Lidbeck (2002). In Smart et al.'s (2011) study, clinicians completed a 38 item checklist (derived from expert consensus) of clinical criteria suggesting a dominance of various pain mechanisms. The authors subsequently used multivariable analysis to identify discriminatory clusters of signs and symptoms associated with a clinically determined dominance of each of the categories, in 464 LBP patients with and without leg pain (Smart et al. 2012a). Nijs et al. (2015) base their categories on a consensus approach of pain experts.

Categories

Schafer et al. (2009a) described a four category system: (i) central sensitisation (renamed in a later paper as "neuropathic sensitisation" (Schafer et al. (2011)); (ii) denervation; (iii) peripheral nerve sensitisation, and (iv) musculoskeletal pain. Mechanisms and signs and symptoms associated with each category are proposed. The LANSS questionnaire score was used for all four categories.

[9]Passive treatment techniques aimed at mobilising and positively influencing peripheral neural structures (Schafer et al 2011).

Smart et al.'s (2011) system has three categories: (i) central sensitisation pain (ii) peripheral neuropathic pain and (iii) nociceptive pain. Standardised clinical interview and examination are used to categorise patients plus a number of additional pain response symptoms to touch tests. These responses include spontaneous paroxysmal pain, hyperalgesia[10] and allodynia.[11] Response to nerve palpation was also evaluated. Smart et al. (2011) classified 464 LBP patients using their three pain mechanism categories. The authors evaluated the discriminant validity of their system by looking at the extent to which patients in each classification differed from one another in terms of health measures and pain (Smart et al. 2012b). They showed that the central sensitisation pain group had poorer outcomes, followed by the peripheral neuropathic pain group then the nociceptive pain group. Nijs et al. (2015) proposed a three category pain mechanism system of (i) nociceptive, (ii) neuropathic and (iii) central sensitisation pain. The authors state that chronic lumbar radicular pain is the most common neuropathic pain syndrome and apply classification criteria for neuropathic pain to LBP. These screening criteria include evidence of nerve root compromise from diagnostic investigations such as MRI, pain extending below the knee and quantitative sensory testing (QST). Tools such as tuning forks to test vibration, soft brush and sharp pin for touch and cold/warm objects to assess response to temperature are used for QST to assess the relationship between the stimulus and the patient's perceived sensation. It is difficult to interpret where patients fit in the classification system if they have some of the neuropathic symptoms i.e. below knee pain, dermatomal pain distribution, burning/ shooting/prickling pain but do not score positively on the sensory testing. The paper focused predominantly on identification and treatment options for the central sensitisation pain group using criteria from Smart et al.'s (2011) work.

Validity

[10] Increased pain from a stimulus that does usually provoke pain (Jensen and Finnerup 2014)
[11] Pain due to a stimulus that does not normally produce pain (Jensen and Finnerup 2014)

Work has been published by Schafer et al.'s group to support the validity of their system. The authors showed predictive ability of one of the subgroups of their system, the peripheral nerve sensitisation group (PNS) (Schafer et al. 2011). 77 LBLP patients were classified according to their system and all had seven sessions of neural mobilisation. As hypothesised, improvement in outcomes was greatest for the PNS group.

Another study showed that the PNS group had the greatest disability of all four groups and more fear avoidance beliefs compared to central sensitisation and denervation groups (Walsh and Hall 2009). This was considered a surprising outcome as these results would have been expected more from the central sensitisation group and suggestive that the criteria for the described classification schemes do not clearly differentiate between these three subgroups (O' Hearne et al. 2009). Further work to demonstrate the construct validity of their system showed differences in pain hypersensitivity between the neuropathic (central) sensitisation group and denervation group compared to controls (Schafer et al. 2014). However, no significant differences were found between the four pain groups which the authors recognised as weakening the construct validity of the classification system.

Reliability

Inter-rater reliability has been reported for two of the three systems. Smart et al. (2010) additionally assessed intra-rater reliability. No reliability data was found for Nijs et al.'s (2015) system. Schafer et al.'s (2009b) reliability study design involved five pairs of examiners and the main author was always one of the examining pair. In 40 patients with LBLP, agreement was 80%, reliability was substantial (kappa=0.72; CI 0.57, 0.86). Smart et al. (2010) used two examiners and the developer of the system was the first examiner on all cases. Agreement was 87.5%, kappa= 0.77 (CI 0.56-0.96). Intra-rater reliability was almost perfect (kappa = 0.96; CI 0.92-1.0) with the developer of the system re-examining the LBP within 6-56 days of their initial assessment. The authors acknowledged that the simultaneous examiner design may have introduced bias towards inflating the kappa value in both studies.

4.3.4 Applicability to primary care

Evaluation of the generalisability of the systems is assessed by their use in other settings. Generally, there was some evidence of use of some of the systems in different settings but mainly to test issues of validity and reliability. There was evidence of one system (Hall et al. 1994) currently being implemented in the Canadian province of Saskatchewan as part of the Saskatchewan clinical spine pathway (SSP) to manage patients with LBP in primary care (Fourney et al. 2011). Based on a sample of 87 LBP patients, implementation of the pathway showed reduction in MRI utilisation and referrals seen by surgeons for nonoperative care, suggesting a potential for cost savings (Kindrachuk and Fourney 2014). The process also suggests reduction in waiting times and costs by timely direction of suitable patients for surgical review (Wilgenbusch et al. 2014). Further studies are underway to assess the efficacy of the SSP.

Numerous studies have used the QTFC system (Spitzer et al. 1987), especially the first four categories, which support the external validity of the system. There is no current evidence that the QTFC is being implemented in primary care.

4.3.5 Summary of classification system appraisal

A summary of the strengths and weaknesses of the 21 classification systems based on the quality appraisal scoring tool is presented in figure 4.2. The majority of the systems were clear on the purpose of their classification system. Validity of the systems scored poorly, in particular content and construct validity. Reliability data was available on a small number of the systems. Only two systems commented on the time needed to train clinicians and how long it took to carry out the classification system with patients. There was evidence of some of the systems being used in different settings but mainly to test issues of validity and reliability. Only one system is currently being implemented in primary care.

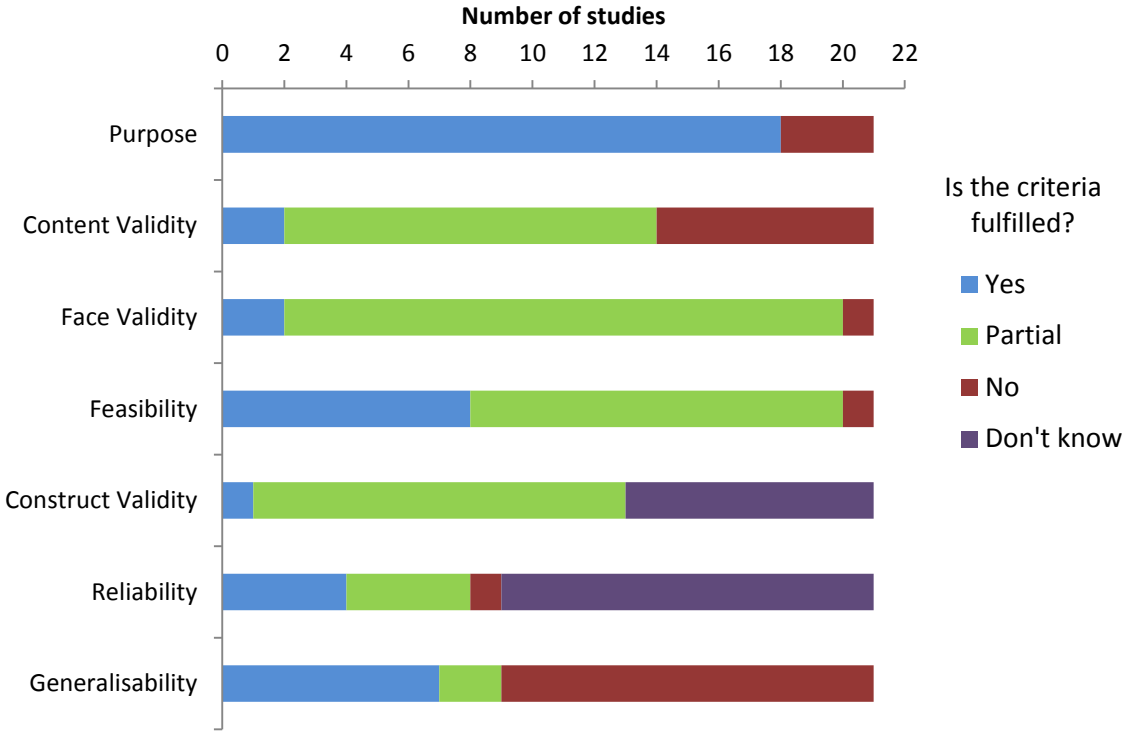


Figure 4.2 Methodological quality summary of the 21 classification systems based on the appraisal tool

4.3.6 Description and diagnosis of sciatica within the systems

It is recognised that variation exists in clinical practice and in the literature over the nomenclature and clinical criteria used to define and diagnose sciatica (Lin et al. 2014). The definition of a subgroup within a classification system is a group with similar characteristics. Hence one of the aims of this review was to consider if groups of patients with sciatica would present with similar clinical characteristics. To explore this issue, the terminology used for categories with sciatica within the classification systems was listed to assess consistency of terms and the clinical criteria within these categories were explored. This is presented in table 4.9. Up to eleven different terms were used to describe sciatica leg pain presentations. The most frequently used terms were sciatica, nerve root, disc and spinal stenosis. But within these terms there was variation, for example nerve root was described as involvement, adherent, compression or irritation. Table 4.9 quantifies how often features from history and physical examination were used in total by the 21 classification systems.

Among items from history taking, ten of the 21 systems mentioned “pain below the knee”; eight systems mentioned “dermatomal distribution of pain” and four out of 21 used “patient’s leg pain was greater than the back pain”. Findings from clinical examination also showed considerable variation. Just over half of the systems used all three neurological deficits (sensory, strength and reflex deficit) and the item most consistently used among all systems was “positive neural tension” (17 systems). A combination of neurological deficits and positive neural tension tests were mentioned in 14 of the 21 systems, but criteria varied from being quite prescriptive, specifying at least one of reflex, sensory or muscle strength deficit, to being quite vague with phrases such as “may have” neurological deficits.

Table 4.9 Descriptors and clinical criteria for sciatica subgroups within the classification systems

	Terms to describe sciatica	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Clinical Features																
Baker 1990	• Sciatica	x					x		x	x	x	x				
Bendebba et al. 2000	• Back and below knee pain with positive straight leg raise	x										x				
Nachemson 1982	• Sciatica; Rhizopathy		x	x	x				x	x	x	x				
Spitzer et al. 1987	• Pain &lower limb radiation with neurological signs; Spinal stenosis	x							x	x	x	x				
Sweetman et al. 1992	• Sciatica	x	x							x	x	x				
Pathoanatomy																
Bernard 1987	• Herniated nucleus pulposis; Spinal stenosis	x	x	x		x	x		x	x	x	x	x			
Cassisi et al. 1993	• Disc herniation	x	x			x			x	x	x	x				
Hahne et al. 2011	• Disc herniation with radiculopathy	x							x	x	x	x				x
Paatelma et al. 2009	• Discogenic pain with nerve root irritation; Spinal stenosis							x				x				
Petersen et al. 2003	• Disc syndrome: reducible/irreducible							x	x	x	x	x		x		
Vining et al. 2013	• Radiculopathy: non/compressive; Neurogenic claudication			x		x		x		x	x			x	x	
Treatment Approach																
Delitto et al. 2012	• Lumbago with sciatica				x				x	x	x	x		x		
Hall et al. 1994	• Leg dominant pain due to nerve root involvement					x		x	x	x	x	x		x		
McKenzie 1981	• Derangement; Adherent nerve root				x									x		
Albert et al. 2012	• Sciatica	x	x						x	x	x	x				
Screening tools /Prediction rule									x	x	x	x				
Fritz et al. 2007	• Low back pain with signs of nerve root involvement															x
Roach et al. 1997	• Disc; Spinal stenosis									x		x			x	
Scholz et al. 2009	• Radicular pain															
Pain Mechanisms																
Smart et al. 2011	• Peripheral neuropathic	x	x									x				
Schafer et al. 2009	• Denervation; Peripheral nerve sensitization		x		x				x	x	x	x			x	
Nijs et al. 2015	• Neuropathic/ radicular pain	x	x		x		x		x						x	x

Key for history and clinical examination criteria for sciatica

1 Pain below knee	6 Quality descriptor of pain eg "burning"	11 Positive neural tension tests
2 Dermatomal distribution of symptoms	7 Stenotic aggravating /easing factors	12 Positive crossed straight leg raise
3 Positive cough/ sneeze	8 Sensory deficit in lower limb (LL) objectively	13 Aggravated with specific lumbar range of movement
4 Pins & needles/numbness: subjective reporting	9 Strength deficit in LL objectively	14 Other (e.g. questionnaires, quantitative sensory testing)
5 Leg pain worse than back pain	10 Altered LL reflexes	15 Positive findings from imaging e.g. MRI

4.4 Discussion

Following a comprehensive systematic search of the literature, 21 systems were identified that classified patients with back and leg pain. Three of the 21 systems focused specifically on LBLP patients (Fritz et al. 2007, Schafer et al. 2009a, Albert et al. 2012), the remaining 18 systems had some categories that included patients with leg pain within the LBP classification. There was a lack of consistency between classification systems when describing sciatica and its' clinical attributes. The definitions and diagnostic criteria for sciatica varied widely among the systems, which mirrors findings from recent reviews on eligibility criteria in studies involving LBLP patients (Lin et al. 2014, Genevay et al. 2010). Consensus on how to define leg pain due to sciatica and agreement on clinical criteria to distinguish patients with sciatica is needed. If eligibility criteria were more consistent across studies this would enhance communication with patients and among clinicians when discussing diagnosis and possible treatment outcomes (Konstantinou et al. 2008, Genevay et al. 2010, Lin et al. 2014).

The call to identify clinically relevant subgroups of LBP patients has been a top priority in primary care back pain research since the mid-1990s (Borkan et al. 1998). Researchers have responded accordingly with a proliferation of subgroup studies (Coste et al. 2013). Several published reviews have appraised LBP classification systems, each with a slightly different focus, but primarily the emphasis has been on non-specific LBP classification. This is the first review to look specifically at the classification of patients with LBLP.

The quality of the 21 systems varied, and those that scored higher on the appraisal tool were ones with evidence of more robust methods of development and more supporting published work on reliability, validity and generalisability. The majority of the systems used a judgement approach to development, ranging from authors' opinion to expert consensus panels.

Some systems used statistical methods to identify clusters of symptoms that best discriminate between patients, giving an objective means of identifying subgroups of patients which helps to

avoid author bias. However relying on statistical clustering in isolation can give rise to content validity issues, with subgroups not clinically recognisable or identifiable. Using a combined approach of judgement, preferably with group consensus, and statistical methods to identify subgroups, is recommended (Ford et al. 2007) but only one system did this (Smart et al. 2011). This is not unique to LBP classification system development. A review on the methodological properties of various classification criteria for the rheumatic diseases (Johnson et al. 2007), showed that over half of the systems were developed based on expert opinion and had not included patient data-sets. The authors considered this a significant deficiency which affects face validity and reliability of the classification criteria.

Among the systems classifying according to clinical features, the QTFC system (Spitzer et al. 1987) scored highest on the quality appraisal tool. The QTFC system has been extensively investigated, validated and adapted and has widespread application in research with many studies using the first four categories to explore differences among groups and investigate their prognosis. For these reasons, and considering its simplicity and brevity, it seems well placed for use in primary care. However to enhance consistency of this classification, it does require more detailed clarification of the clinical criteria for neurological involvement in category 4.

Pathoanatomical classification systems generally scored low on the appraisal tool, primarily because development was mainly based on authors' opinion. Many consider the pathoanatomical approach, which seeks to associate specific structures with symptoms, as outdated and unhelpful to patients. It is thought to lead to overuse of diagnostic procedures with subsequent implications on cost and patients' expectations if findings do not match clinical symptoms (Fourney et al. 2011). A review of recommendations for LBP clinical practice (Dagenais et al. 2010) noted that none of the guidelines recommend that clinicians should attempt to identify specific anatomical structures involved in LBP, once potentially serious spinal pathology, specific causes and substantial neurological involvement have been ruled out. Others argue that identification of a cause for LBP is important for patients and one of the main reasons for seeing a primary care

practitioner (Cassisi et al. 1993). Neglecting patients' expectations can impact negatively on patient satisfaction, and 'diagnostic uncertainty' or inadequate explanation of cause, can lead to higher levels of depression (Serbic and Pincus 2014) and fear avoidance beliefs (Waddell 2004) in LBP patients.

The treatment based approach classification systems included the McKenzie system which is a popular treatment based approach among clinicians, despite evidence that it is not superior to other treatments. An additional three papers were grouped under screening tools and prediction rules, where statistical methods were used to identify a cluster of items to assist diagnosis or predict response to treatment. Classification according to pain mechanisms is gaining popularity in musculoskeletal medicine and three papers applied this system to LBLP (Schafer et al. 2009a, Smart et al. 2011, Nijs et al. 2015). Schafer et al (2009a) designed their system specifically for LBLP patients. Some confusion arises comparing the nomenclature and criteria of the pain mechanism subgroups. Leg pain due to sciatica was categorised as denervation or peripheral sensitisation by Schafer et al. (2009a), as peripheral neuropathic pain by Smart et al. (2011) and predominantly neuropathic pain by Nijs et al. (2015). All had different clinical criteria. Work has been done to support the external validity of Schafer et al.'s system but they struggled to demonstrate discriminative validity of the categories (Schafer et al. 2014). This may reflect the judgement based development process of the system. A more robust method including statistical techniques and consensus could serve to improve the validity of the system and assign criteria that allow clearer differentiation between the subgroups.

Smart et al. (2012a) used statistical methods to identify items from history and physical examination items that were predictive of peripheral neuropathic pain. These were: history of nerve injury, pathology or compromise, pain in a dermatomal distribution and positive neural tension tests. The authors recognised that these items differ considerably from criteria found in neuropathic pain screening tools and reflect that this may be because their patients were

recruited from primary care settings with less severe presentations than the more severe pain populations in studies from which these questionnaires were derived.

Schafer et al. (2009a) defined their denervation group as patients with at least two neurological deficits (motor, sensory or reflex). Yet despite these neurological deficits indicative of nerve root compromise, this category also includes an s-LANSS screening tool score of less than 12, indicative of a low probability of neuropathic pain. Contrary to this, the categories of compressive and non-compressive radiculopathy in Vining et al.'s (2013) pathoanatomical system, have a LANSS score of 12 or over, suggesting all radicular pain has a high probability of being neuropathic. Nijs et al. (2015) have different screening criteria for neuropathic pain in LBP which includes confirmation of a nervous system abnormality with diagnostic testing e.g. electromyography (EMG) or imaging.

4.4.1 Strengths and limitations

This is the first review that has focused specifically on classification of LBLP. The search identified over 13,000 citations for initial screening. This large number reflects the breadth of the search strategy and large number of databases searched with minimal restrictions. The broad search strategy was deemed necessary to include all possible terms that could be used to describe LBLP and classification and avoid missing any systems. The search strategy was supplemented by first-author searches and hand searching reference lists. Identified systems were not excluded on the basis of quality and study appraisal was systematically and independently carried out by two reviewers. Other systems or supporting evidence may have been missed, e.g. unpublished student studies or cases of publication bias if findings were unsupportive of the system.

Twenty-one of the 49 papers in the systematic review were identified through supplementary search strategies rather than through the initial search. Despite the comprehensive search strategy which included up to 34 terms to describe LBLP, a possible reason for not identifying papers in the initial databases search, is because of the vast nomenclature used to describe and identify LBLP.

4.5 Conclusion

The classification of LBLP merits more attention, especially in primary care settings where most of these patients are assessed and managed. This should start with agreement on the criteria that reasonably distinguish sciatica from pain referred into the leg from structures in the back other than the nerve root. This PhD plans to develop a diagnostic model to identify criteria that distinguish sciatica from referred leg pain. Findings from this systematic review will help inform which items from clinical assessment to select for the diagnostic modelling process. Items from clinical assessment that were consistently identified in the classification systems for LBLP will be considered for inclusion in the diagnostic model.

An approach that uses data from large, unselected groups of primary care LBLP patients to classify them according to relevant characteristics from self-reported measures, clinical examination findings and perhaps even demographic information, deserves more attention to appreciate the characteristics of this subgroup of LBP patients. This methodology has been used more often in systems to subgroup psychosocial characteristics in chronic LBP patients (McCarthy et al. 2004) and it is equally applicable in LBLP classification. This PhD also plans to use the statistical approach of latent class modelling to identify subgroups of LBLP patients from a large primary care cohort of primary care consulters. The factors to include in the model will be guided by the findings in the review of clinical criteria consistently used among LBLP classification systems.

Prior to development of a diagnostic model and identification of classes of LBLP consulters, agreement on the diagnosis of LBLP among clinicians will be explored which is the next chapter of the thesis.

Chapter Five: Agreement and inter-rater reliability amongst clinicians diagnosing low back-related leg pain

5.1 Introduction

The clinical task of differentiating between sciatica and referred leg pain in LBLP patients can be difficult (Waddell 2004, Leffler and Hansson 2008, Bogduk 2009, Scholz et al. 2009), and clinicians may disagree as to its presence or absence in a patient with LBLP (Vroomen et al. 2000, Fairbank 2007). The majority of LBLP patients are first seen in the primary care setting and are assessed by clinicians such as GPs, physiotherapists, osteopaths, chiropractors, all of whom use a combination of findings from history and physical examination to evaluate the nature of the leg pain, reach a diagnostic decision and make management plans accordingly. However, items from history taking have failed to show both high sensitivity and specificity in patients with suspected sciatica symptoms due to disc herniation (Vroomen et al. 1999) and a Cochrane review of physical examination items for lumbar radiculopathy due to disc herniation showed that most had poor individual diagnostic performance (van der Windt et al. 2010). A review of the accuracy of diagnostic tests to detect lumbar spinal stenosis (which can be responsible for radicular symptoms) reported that no firm conclusions could be drawn about the diagnostic performance of tests due to the poor quality of the included studies (De Graaf et al. 2006).

There is no agreed “gold standard” for diagnosing sciatica (van der Windt et al. 2010). Literature suggests that in the absence of a well-accepted reference standard, expert clinical opinion may be considered an appropriate alternative to diagnosis providing that it is reasonably reliable (Felson and Anderson 1995, Katz et al. 2000, Coggon et al. 2005). However it is know that clinicians do not always agree on diagnosis, and there are numerous examples in the literature to illustrate this, such as diagnosis of carpal tunnel syndrome (Bachmann et al. 2005), shoulder disorders (de Winter et al. 1999) and neck and arm pain (Tampin et al. 2012).

Results from Freynhagen et al.'s (2008) study using QST on LBLP patients reflected the challenge faced by clinicians when trying to distinguish between the two entities of referred leg pain and sciatica. Although lower limb sensory changes in specific dermatomal distributions are typically associated with a sciatica presentation, in their LBLP sample they showed that it was possible to detect sensory deficits in the distal leg in patients diagnosed with referred leg pain.

Numerous studies have examined reliability of classification systems for LBP (see chapter four, systematic review), but only one study has reported on the reliability of a proposed pain mechanism based classification system designed for LBLP patients (Schafer et al. 2009b). Whilst reliability has been documented as mainly poor for individual clinical tests of radicular pain (van der Windt et al. 2010), the reliability of the overall diagnostic decision as to whether the clinical presentation in LBLP patients is sciatica or referred leg pain has received less attention. One study did investigate this, and showed that when neurologists consecutively examined patients with sciatica, they disagreed on the presence of nerve root involvement in one in four patients after history taking and one in five patients after physical examination (Vroomen et al. 2000). It is also known that studies of agreement on features of sciatica have generally not shown better than fair reliability (McCarthy et al. 2007, Smart et al. 2010).

Considering the recognised importance of differentiating between sciatica and referred leg pain, there is a lack of studies examining the reliability of this diagnostic decision.

5.1.1 Study aim and objectives

Aim: The overall aim of this study is to investigate the agreement and reliability amongst clinicians when diagnosing patients in primary care who present with symptoms of LBLP.

Objectives

(i) To investigate the agreement and reliability amongst experienced physiotherapists taking part in an observational cohort study of LBLP patients (the ATLAS study) when diagnosing patients presenting in primary care with symptoms of LBLP.

(ii) To gain a broader insight into current agreement on the clinical diagnosis of LBLP amongst healthcare professionals by investigating the agreement and reliability between the ATLAS study clinicians and other healthcare professionals (not involved in the ATLAS study) when diagnosing LBLP.

(iii) Investigate the relationship between different levels of confidence in diagnosis with agreement and reliability indices.

(iv) Identify elements of the assessment that led clinicians to their diagnosis, using a standardised proforma and use this information to gain insight into reasons for potential disagreement.

5.2 Methods

Methods of testing agreement and reliability in clinical studies involving patients vary in the literature, ranging from consecutive examination of patients (Vroomen et al. 2000), combined assessment with two or more raters (Smart et al. 2010), case notes review (Tampin et al. 2012) or video assessment (Fritz et al. 2000). Video-recording of the assessment was chosen in this study as it has the advantage of only having to assess the patient once, which reduces patient burden and aggravation of symptoms, and limits the potentially confounding effect of repeated testing on patient responses. It also allows more than one rater to view the patient assessments and was a practical and feasible method to allow the research clinics to run effectively.

There were two parts to this reliability study.

Part One

Raters for Part One were the study (ATLAS) trained experienced musculoskeletal physiotherapists. They each carried out assessments on patients and at a later date watched assessments on video of patients they had not examined. They are named **Group A** when assessing the patients and **Group B** when watching the patient assessments on videos.

Part Two

In Part Two, a group of healthcare professionals from varied clinical backgrounds, were recruited to watch the same patient assessments on video. The aim was to gain a broader insight into current agreement on the clinical diagnosis of LBLP amongst healthcare professionals locally and nationally. These raters are named **Group C.** See box 5.1 for definitions of the three groups of raters.

5.2.1 Sampling

Subjects for this reliability study were participants in the ATLAS study (see chapter three). Inclusion and exclusion criteria for subjects recruited to the reliability study were the same as for the ATLAS cohort study (see chapter three, section 3.2.3)

As part of the ATLAS study, patients underwent a standardised clinical assessment, full details of which are given in chapter three. Six of the seven physiotherapists who carried out these assessments were the raters involved in the reliability study (Group A and Group B). Raters in Group C had no involvement in the ATLAS study.

Ethical approval for the ATLAS study was granted by the South Birmingham Research Ethics Committee (REC ref. 10/H1207/82). Included in this ethics application was permission to ask the ATLAS participants if they would allow video recording of their clinical assessment for the reliability study.

Recruitment of patients to the reliability study started in August 2011 and ended in July 2012 when the requisite number of videos had been recorded (see sample size estimation, section 5.2.6). Patients were recruited in a sample of convenience. The thesis researcher attended the ATLAS research clinics until all physiotherapists involved in the study were videoed assessing at least six patients each. Video recording was only carried out on patients who agreed to the video recording of their clinical assessment, which was obtained during the main ATLAS study consent process with the clinic research nurse.

Box 5.1 Definitions of the three groups of raters

Group A = Physiotherapists (involved in the ATLAS study) performing the clinical assessment of the patients

Group B = Physiotherapists (involved in the ATLAS study) watching the assessments on video

Group C = Health professionals (no involvement in the ATLAS study) watching the assessments on video

Videos were subsequently edited by the thesis researcher to remove any dialogue between the patient and clinician where findings from the assessment were discussed. This was done to ensure that other clinicians watching the assessments on video remained unaware of the diagnosis made by the assessing physiotherapist.

5.2.2 Examiners and training

The physiotherapists who examined the patients (Group A) and subsequently watched patient assessments on video (Group B) took part in 2.5 days training for the ATLAS study as outlined in chapter three. The raters in the Part Two (Group C) did not participate in any prior training.

5.2.3 Assessment

Full details of the assessment schedule are given in chapter three. The physiotherapists (Group A) completed the assessment in a standardised format, asking all the questions in the history section and carrying out all relevant clinical tests in the assessment form (Appendix D). At the end of the assessment, the physiotherapist documented whether the leg pain was due to nerve root involvement (yes/no), how confident they were in their clinical diagnosis (0-100% scale) and listed the most relevant elements that led them to their clinical impression. (see box 3.2, page 24).

Diagnosis was decided by the individual assessing physiotherapist based on his/her clinical opinion and experience, and by reflecting on the guidelines and evidence presented during the training. The history taking and physical examination took approximately 30 minutes to complete.

5.2.4 Part One: Group A and Group B

When all the videos were collated, the physiotherapists involved in the study watched videoed assessments carried out by another study physiotherapist. Factors listed below were taken into consideration when allocating the videos to the study physiotherapists (Group B) to ensure:

- The physiotherapists viewing the video had not treated the patient they were watching, as in some cases the physiotherapists had follow-on contact with patients assessed by one of the other physiotherapists.
- There was no order effect for any of the physiotherapy raters (i.e. raters did not watch all patients recruited early to the study or all patients recruited later in the study).
- Physiotherapists viewed at least one video assessment from each of the other physiotherapists.
- An even spread of time to view the videos (i.e. approximately 180 minutes to view all six videos).

The order in which the physiotherapists were asked to view the videos was not predetermined. The six video files were copied into a folder on a laptop and could be watched in any order they chose.

Physiotherapists watching the assessments on video answered the same three questions (box 3.2, page 24). They did not have access to the clinical notes made by the assessing physiotherapist and were blind to that assessor's diagnostic decision. The watching of the videos was supervised by the thesis researcher and in any cases where a test outcome was not clear to the therapist watching the video, they were told by the researcher what had been documented in the original assessment notes.

5.2.5 Part Two: Group A and Group C

The health professionals (Group C) involved in Part Two, were recruited as a sample of convenience and were all approached individually by the thesis researcher. They were either sent

or given the videos on a password encrypted USB stick and watched the videos at their workplace or home. They were asked to watch the videos in an area where only they could see and hear the footage and confirmed in writing that they would not use the videos for any purpose other than for the study. It was suggested that videos could be paused at any time but ideally to try and watch an assessment in "one go" as if they were observing a colleague carrying out an assessment in a clinical setting. At the end of watching the full assessment, they were asked to answer the same three questions that raters in Group A and B had answered (box 5.2). The groupings of videos given to the raters in Group B were kept the same for the health professionals and allocated in no particular order once they agreed to take part in the study. The order in which the raters (Group C) were asked to view the videos was not predetermined, they could watch them on their computer in any order they chose.

5.2.6 Sample size

The sample size estimation used in this study is based on Sim and Wright's (2005) table for subject requirement in a two rater study. For this two rater inter-rater agreement and reliability evaluation at least 30 subjects were needed for analysis at 90% power in order to detect a kappa of 0.6 (from a null hypothesis value of 0 (α=0.05)) with a 95% confidence interval. This was based on an assumption that the proportion of positive ratings in the sample would be approximately 30% (i.e. the prevalence of sciatica). However a sample size of 30 allows for variation in prevalence ranging between 30% to 70%.

Multiple pairs of raters are recommended to enhance the quality and generalisability of reliability studies (May et al. 2006) hence the maximum possible number of raters were used for Part One. This was predetermined as six raters because this was the number of physiotherapists involved in assessing patients (excluding the thesis researcher who was videoing the assessments) who attended the ATLAS research clinics. A convenience sample of six raters was recruited to take part in Part Two so they could also each watch the same number of videos.

5.3 Analysis

The results were summarised using percentage agreements and kappa coefficients with two sided 95% confidence intervals. Kappa coefficients were computed using SPSS version 21. To allow for a more comprehensive interpretation of the kappa coefficient, the percentage of positive and negative agreements were calculated, the effect of bias and prevalence was reported and adjusted kappas were computed. Descriptions and formulae for these calculations are discussed and presented below.

5.3.1 Two-by-two tables

This study used a basic binary classification system. Raters diagnosed if a patient's leg pain was due to nerve root involvement (sciatica) or referred pain. Two-by-two (2x2) tables were used to assist visualisation and interpretation of data and took the form shown below (figure 5.1, adapted from Sim and Wright 2005).

		Raters in Group A		
		Sciatica	Referred	Total
Raters in Group B	Sciatica	a	b	g_1
	Referred	c	d	g_2
	Total	f_1	f_2	n

Diagonal cells (a and d) represent agreement; diagonal cells (b and c) represent disagreement; cells f_1, f_2, g_1, g_2 represent marginal totals; n=total number of subjects.

Figure 5.1 2x2 table for paired ratings on a two category nominal scale

5.3.2 Agreement

Percentage observed agreement ($\left(\frac{a+d}{n}\right) \times 100$) calculated the total percentage of agreement between raters. Percentage positive agreement ($\left(\frac{a}{n}\right) \times 100$) and percentage negative agreement ($\left(\frac{d}{n}\right) \times 100$) were calculated to give information on separate agreements on presence of sciatica or referred pain, respectively.

5.3.3 Reliability

Percentage observed agreement does not factor in agreement expected by chance. Agreement beyond that expected by chance was measured by calculating a single index called the kappa coefficient.

Kappa takes the form (Sim and Wright 2005).

$$\text{kappa} = \frac{observed\ agreement\ (Po) - chance\ agreement\ (Pc)}{1 - chance\ agreement\ (Pc)}$$

Using the notations from the 2x2 table (fig. 5.1) where:

$$P_o = \left(\frac{a+d}{n}\right) \text{ and } P_c = \frac{\left((f1 \times g1) \div n\right) + \left((f2 \times g2) \div n\right)}{n}$$

Interpretations of the kappa coefficients were based on distinctions outlined by Landis and Koch (1977) where ≤0 = poor, 0.01-0.2 = slight, 0.21-0.4 = fair, 0.41-0.6 = moderate, 0.61-0.80 = substantial and 0.81-1.0 = almost perfect. Although the kappa coefficient is widely recognised and used as the appropriate statistic of measuring reliability, issues may influence the interpretation of kappa coefficients, namely prevalence of the condition and bias from observers (Byrt et al. 1993, Sim and Wright 2005).

5.3.4 Prevalence

Evidence of a prevalence effect can cause the kappa coefficient to be unrepresentatively low (Hallgren 2012). A prevalence effect exist if raters tend to choose one diagnostic option more often that the other or if there are genuinely more frequent occurrences of one condition within the population under study (Hallgren 2012). It can be identified when the proportion of agreements on the positive diagnosis differs from the proportion of agreements on the negative diagnosis. This is expressed as the prevalence index (Sim and Wright 2005) and is calculated from the 2x2 table presented in figure 5.1 using the formula: Prevalence index *(PI)* $= \frac{a-d}{n}$

Its value can range from -1 to + 1 and would equal 0 when "positive" and "negative" ratings are equally probable (Byrt et al. 1993). Sometimes the situation arises where percentage agreement is high, but if there is a noticeable prevalence effect this will lead to a lower kappa value. This paradox of high agreement rates and low kappa values due to the trait prevalence in the population under consideration has led some authors to propose alternative or additional measurements (Cicchetti and Feinstein 1990, Byrt et al. 1993). To adjust for high or low prevalence, the average of cells a and d (figure 5.1) is used to replace the values in these cells (Sim and Wright 2005). The kappa statistic is recomputed and called the prevalence adjusted kappa (PAK). The PAK will be calculated in this study if there is any evidence of a prevalence effect.

Gwet (2002) demonstrated that agreement can occur with a fixed probability of 0.5 if a rating is random, hence a reasonable value for chance agreement probability (Pc) should not exceed 0.5. Worse kappa statistics can be expected when marginal totals f1 and g1 are either very small or very large. Gwet (2002) proposed an alternative chance corrected statistic (AC_1) where the probability of chance agreement (Pc) will always vary between 0 and 0.5. In the event of high agreement and low kappa values, the AC_1 scores will be calculated to assess any change in the kappa value.

$$AC_1 = P_0 - P_{ac} / 1 - P_{ac}$$

Where $P_{ac} = 2P_1 (1-P_1)$ and $P_1 = ((f_1+g_1)/2)/n$

5.3.5 Bias

It is also possible that bias may influence kappa values; evidence of bias can cause the kappa coefficient to be unrepresentatively high (Hallgren 2012). Bias exists when the two groups of raters assign different proportions of subjects as positive for the condition in question, in other words disagreements are asymmetrical (Byrt et al. 1993, Sim and Wright 2005). Bias may arise in a situation when raters have different professional backgrounds or training which leads them to interpret clinical presentations differently. It is the difference in proportions of "yes" in the two

groups and is reflected in the differences between cells b and c in figure 5.1. The equation for the bias index is: Bias index = $\frac{b-c}{n}$

The bias index ranges from 0 to 1. If the bias index is large, the kappa coefficient is higher than when bias index is low or zero (Byrt et al. 1993). As with prevalence, a bias adjusted kappa (BAK) can be calculated where the average of cells b and c is used to replace the values in these cells (Sim and Wright 2005). The BAK will be calculated if there is any evidence of a bias effect.

5.3.6 Confidence in diagnosis

It is recognised that the composition of a study sample can have a significant impact on kappa values (Vach 2005), as the greater the proportion of patients who have very clear symptoms or findings of a condition, the easier it is for different observers to agree (Vroomen et al. 2000). Separate analysis was planned to assess the relationship between confidence in diagnosis and agreement and reliability. It was hypothesised that as confidence in diagnosis increased, corresponding agreement and reliability indices would also increase. Confidence levels were also explored for groups of raters and in cases of disagreement.

5.3.7 Reasons for diagnosis

Percentage agreements and kappa values do not give any information about the various types or sources of disagreements (Viera and Garrett 2005). Clinicians listed reasons for their diagnostic decision. In cases where clinicians disagreed, reasons for their diagnostic decision were explored to give insights into why they made their decisions or why there was a difference in diagnostic opinion between the raters.

5.3.8 Combinations of analyses

Four combinations of analyses were carried out:

Analysis I — Group A and Group B: % agreement, kappa coefficient, prevalence and bias, adjusted kappa.

Analysis II — Group A and Group C: % agreement, kappa coefficient, prevalence and bias, adjusted kappa.

Analysis III — Group A and Group B: relationship between confidence in diagnosis and kappa coefficient for the subgroup of patients where both raters' confidence in diagnosis increased (in 5% increments) from >50% to >90%.

Analysis IV — Group A and Group C: relationship between confidence in diagnosis and kappa coefficient for the subgroup of patients where both raters' confidence in diagnosis increased (in 5% increments) from >50% to >90%

5.4 Results

Video recordings of the physiotherapy assessment for the reliability study were collected from 40 participants to account for possible drop outs. Of these 40 videos, 36 were used for analysis. Two videos were not used because the assessing physiotherapist determined that the patients were ineligible to be in the main ATLAS study following the clinical assessment. One of the patients did not have leg pain and the other had leg pain that was not low back-related. The other two videos were not used due to poor visual quality of the video, making it difficult to see some of the clinical tests being performed. In the analyses involving raters in Group C (Analysis II and IV) calculations were done on a sample size of 35 patients because there was missing data for one patient, as the clinician decided that the leg pain was neither sciatica nor referred pain.

The median age of the participating patients was 51 years (range 23-74 years) and 61% were female. Pain intensity was 5.3 (2.7 standard deviation (sd)) for leg pain and 5.6 (2.8 sd) for low

back pain. Over half the sample had pain below the knee (58%). Self-reported disability averaged 13.4 (6.2 sd) on the RMDQ. Using the cut-off of >11 on the HADS for probable/moderate/severe anxiety and depression, 33% reported anxiety and 28% reported depression. The mean Sciatica Bothersomeness Index score was 14.6 (5.4 sd). A summary of the characteristics of the 36 patients collected from the self-report information in questionnaires is given in table 5.1.

Six of the seven clinical assessors, described in chapter three (table 3.2, page 30), were the raters who performed the clinical assessments (Group A) and viewed the videos (group B).[12] Three were spinal specialist physiotherapists and three were senior musculoskeletal physiotherapists. The same numbers of health professionals were recruited to take part in Part Two. This allowed for a representation of a variety of health professions and meant that they could also watch six videos each. They included an NHS based extended scope practitioner physiotherapist, an MSK physiotherapist working in private practice, a specialist registrar in rheumatology, a GP, a chiropractor and an osteopath. Years in practice since qualification averaged 20 years (range 14-26 years) and the allied health professionals had an average of 20.5 years' experience (range 15-26 years) in predominately treating a musculoskeletal caseload. The rheumatologist, GP and NHS based physiotherapist were all practising in the Staffordshire area. The physiotherapist in private practice, osteopath and chiropractor were based in Brighton, Manchester and Hampshire respectively. One of the participants opted to come to the Research Institute for Primary Care and Health Sciences at Keele University to watch the videos in a quiet room, the other five participants watched the videos at their place of work or home.

[12] The thesis researcher carried out some of the clinical assessments for the ATLAS study but was not a rater in the reliability study.

Table 5.1 Descriptive characteristics of the reliability study patient sample

Study sample n=36	**n =36**	
Gender, Female	22 (61%)	
Age (years) median (range)	51 (23-74)	
Intensity back pain [a] (0-10) mean (SD)	5.6 (2.8)	
Intensity leg pain [a] (0-10) mean (SD)	5.3 (2.7)	
Bothersomeness [b] of leg symptoms in last week (0-24) mean (SD)	14.6 (5.4)	
Duration pain	Back	Leg
0-6 weeks	13 (36%)	16 (44%)
6-12 weeks	10 (28%)	8 (22%)
3-6 months	4 (11%)	5 (14%)
7-12 months	2 (6%)	2 (6%)
> 12 months	7 (19%)	4 (11%)
Missing	0 (0%)	1 (3%)
Below knee pain	21 (58%)	
Off work because of back/leg pain	4 (11%)	
Reduced hours/duties	3 (8%)	
RMDQ (0-23) mean (SD)	13.4 (6.2)	
HADS Anxiety subscale score (0-21) median (range)	9 (2,15)	
Anxiety cases [c] (score ≥ 11)	12 (33%)	
HADS Depression subscale score (0-21) median (range)	6 (0,17)	
Depression cases [c] (score ≥ 11)	10 (28%)	

SD, standard deviation; RMDQ, Roland Morris Disability Questionnaire; HADS, Hospital Anxiety and Depression Scale.
All figures are frequencies (percentages) unless stated otherwise.
[a] Pain intensity measured using the mean of three 0 to 10 numerical rating scales for least and usual pain (back or leg) over the previous 2 weeks and current pain intensity.
[b] Four questions (Sciatica Bothersomeness Index) relating to the bothersome in the last week of the (i)leg pain, (ii)numbness and tingling, (iii)weakness and (iv) pain whilst sitting was given a composite score out of 24.
[c] Using the cutoff of ≥11 on the HADS for probable/moderate/severe anxiety and depression.

5.4.1 Analysis I – Group A and Group B

Raters in Group A and Group B both diagnosed sciatica pain in 25 of the 36 patients. These were not all the same patients; there was disagreement in 10 cases. The overall percentage agreement between the pairs of raters was 72% with a kappa coefficient of 0.35 (CI 0.02, 0.68). This result is

considered fair reliability. The data for the paired ratings on this two category nominal scale are displayed in table 5.2.

Table 5.2 2x2 results table for Group A and Group B

		Raters in Group A		
		Sciatica	Referred	Total
Raters in Group B	Sciatica	20	5	25
	Referred	5	6	11
	Total	25	11	36

Prevalence index was 0.39. Although this is a low value, the PAK was calculated which gave an adjusted kappa value of 0.44. The AC_1 recomputed the kappa coefficient as 0.51. These small increases to the kappa value meant the reliability rating changed from "fair" to "moderate". The bias index was zero, it was not necessary to calculate the BAK as there was no evidence of bias among the raters.

5.4.2 Analysis II - Group A and Group C

Raters in Group A diagnosed sciatica in 25 of the 35 patients. Raters in Group C diagnosed sciatica in 23 of the 35 patients. They disagreed on 10 cases (table 5.3). The overall percentage agreement between the pairs of raters was 71% with a kappa coefficient of 0.34 (CI 0.02, 0.69) which is considered a fair agreement. The prevalence index was 0.37, the PAK was 0.49. The recomputed AC_1 kappa was 0.50. Bias was negligible at 0.06 hence it was not necessary to calculate the BAK.

Table 5.3 2x2 results table for Group A and Group C

		Raters in Group A		
		Sciatica	Referred	Total
Raters in Group C	Sciatica	19	4	23
	Referred	6	6	12
	Total	25	10	35

A summary of the results from Analysis I and II is presented in table 5.4

Table 5.4 Results from Analysis I and II

	Analysis I (Group A & B)	Analysis II (Group A & C)
Sample size (n)	n=36	n=35
Agreement (%)		
Overall	72%	71%
Positive	56%	54%
Negative	17%	17%
Reliability	Fair	Fair
kappa	0.35	0.34
95% Confidence Intervals (CI)	0.02, 0.68	0.02, 0.69
Prevalence index	0.39	0.37
Bias Index	0.00	0.06
Adjusted kappa		
Prevalence adjusted kappa	0.44	0.49
Alternative chance corrected kappa (AC_1)	0.51	0.50

5.4.3 Analysis III- Group A and B: relationship between confidence in diagnosis and kappa coefficient

Agreement and reliability indices were calculated for levels of confidence in diagnosis that ranged from >50% up to >90%. Levels of agreements (fig. 5.2) and kappa coefficients (fig. 5.3) were seen to increase as confidence in diagnosis of both raters increased. The trend of increasing agreement and reliability indices was noticeably evident once confidence in diagnosis of both raters was greater than 70%. In the 24 cases where both raters were at least 80% confident in their diagnosis the kappa coefficient increased to 0.68. This reliability is considered substantial between raters.

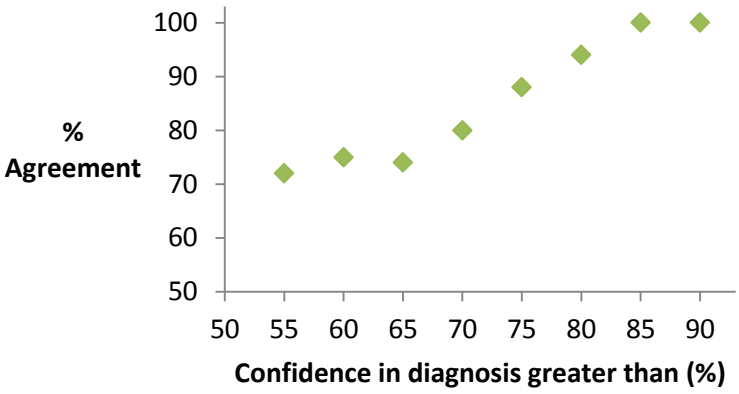


Figure 5.2 Relationship between confidence in diagnosis and overall percentage agreement between Group A and Group B raters

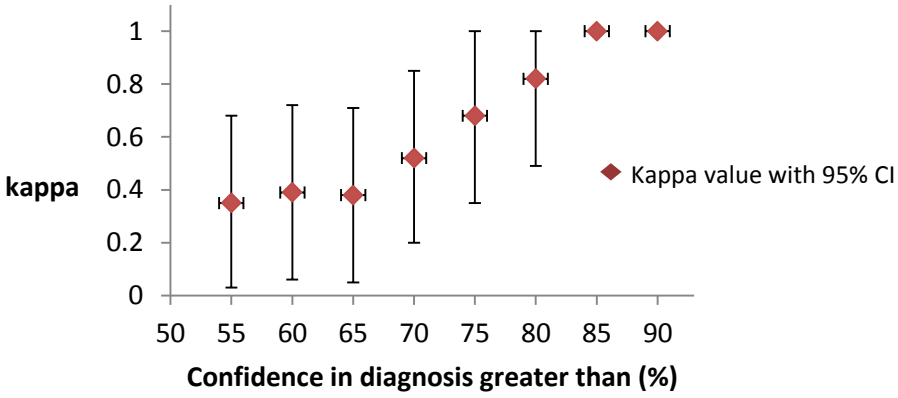


Figure 5.3 Relationship between confidence in diagnosis and kappa coefficient for Group A and Group B raters

5.4.4 Analysis IV- Group A and C: relationship between confidence in diagnosis and kappa coefficient

Results between raters in Group A and C were almost identical to those seen between raters in group A and B. Levels of agreements and kappa coefficients increased as confidence in diagnoses of both raters in group A and C increased. For example, in the subgroups of patients when both raters' confidence was 80% or above, the overall percentage agreement increased from 71% to 89% and the kappa coefficient increased from 0.34 to 0.68.

5.4.5 Confidence in diagnosis among raters and cases of disagreement.

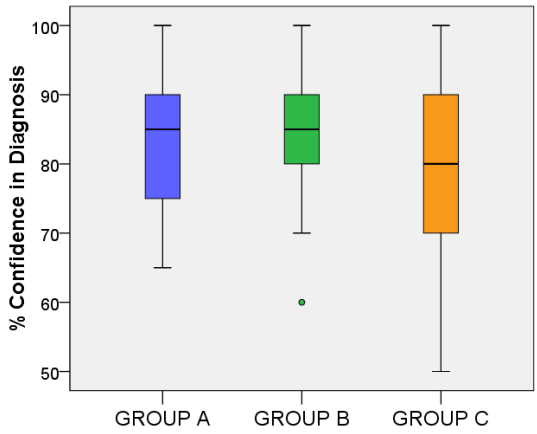
Confidence in overall diagnosis for the three groups of raters is displayed in the box plot graph (figure 5.4i). When the physiotherapists were performing the assessment, they had the same levels of confidence in their diagnosis as when they watched the assessments on video (85% median percentage confidence). Raters in group C were slightly lower at 80%. Confidence in diagnosis was examined separately for cases of agreement and disagreement.

Agreement versus Disagreement cases

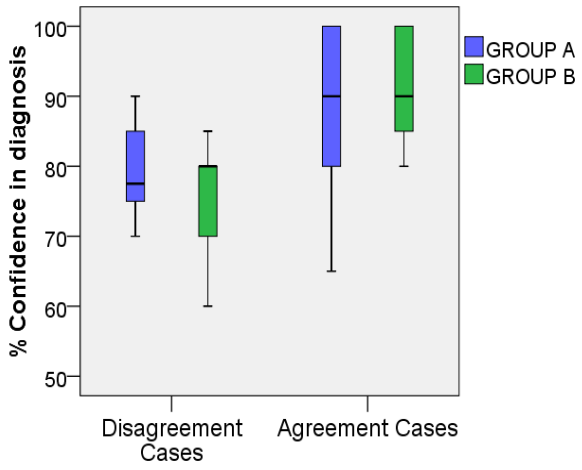
Group A and B: For the 10 cases of disagreement between raters in Groups A and B (Analysis I), median confidence in diagnosis was lower: 77.5% for Group A (range 70-90%) and 80% for Group B (range 60-85%) (Fig 5.4 ii). In seven out of 10 cases of disagreement, confidence of both raters was lower than 80%. For the 26 cases of agreement between raters in Group A and B (Analysis I), both groups' median confidence in diagnosis was 90% (range 65-100% for Group A, range 80-100% for Group B).

Group A and C: For the 10 cases of disagreement between raters in Group A and C (Analysis II), median confidence in diagnosis was lower: 75% (range 70-90%) for Group A and 70% (range 55-95%) for Group C (Fig 5.6iii). In eight out of 10 cases of disagreement, confidence of both raters was lower than 80%. For the 25 cases of agreement between raters in Group A and C (Analysis II) median confidence in diagnosis was 90% for group A (range 65 to 100%) and 80% for Group C (range 50 to 100%). The lower levels of confidence among all the raters in the disagreement cases suggests that these patients may have had signs and symptoms that were more difficult to interpret making diagnosis more difficult.

5.4 (i) Overall confidence: Raters in Group A, B and C



5.4 (ii) Group A v B
Disagreement and agreement cases



5.4 (iii) Group A v C
Disagreement and agreement cases

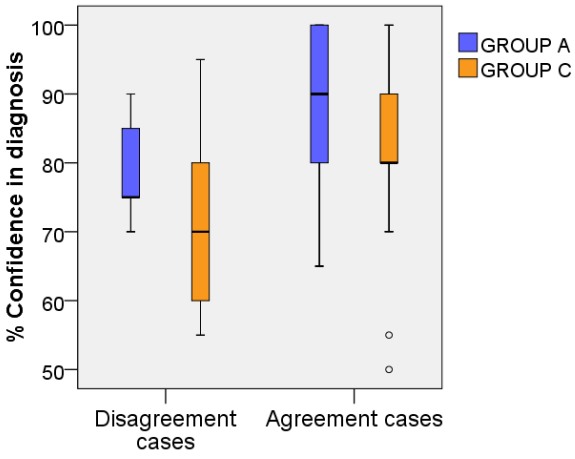


Figure 5.4 Box plots representing median, range and interquartile values comparing confidence in diagnosis amongst groups of raters in disagreement and agreement cases

5.4.6 Reasons for diagnosis

Therapists were asked to list up to four reasons for their diagnosis and this text was examined in the cases of disagreement. There were ten disagreement cases in both Analysis I and Analysis II. Nine of the ten disagreement cases were the same for Analysis I and Analysis II.

Answers were varied, no dominant theme emerged but the most commonly cited area for disagreement between the raters seemed to be interpretation of subjective and objective sensory findings. One example was a rater in Group A diagnosing referred pain and listing "reduced sensation objectively non dermatomal" whereas the rater in Group B wrote "decreased sensation

objectively" and the rater in Group C wrote "dulling pin prick sensation". In this example, raters in Group B and C both diagnosed sciatica. In another case a "positive cough sneeze" was one of the reasons given by the Group A rater who diagnosed sciatica, but it was not mentioned by the Group B or C rater, both of whom diagnosed referred leg pain. A case of differing interpretation of neural tension tests was demonstrated when a Group A rater diagnosed sciatica and recorded "the femoral nerve stretch was positive", whereas the Group B and C raters both recorded "negative neural tension tests" and diagnosed referred leg pain.

Raters with confidence of at least 80% were more likely to give more reasons for their decision. Up to two thirds of raters in Group A who had greater than 80% confidence in their diagnosis of sciatica listed three or four reasons for their diagnosis. Six of the seven raters in Group A who had confidence under 80% when diagnosing sciatica, had two or less reasons listed for their decision.

The table presented in Appendix H compares the reasons for diagnosis given by raters A, B and C in cases of disagreement.

5.5 Discussion

5.5.1 Summary of findings

This study showed that the reliability of diagnosing sciatica in LBLP patients with symptoms of any duration and severity is fair amongst experienced clinicians. Percentage agreement for both parts of the study, and reliability as measured by the kappa coefficient, were 72% and 71% and 0.35 and 0.34 respectively. The agreement percentage is reasonable but kappa values under 0.6 are considered below the minimum standards for reliability coefficients (Kottner and Streiner 2011). The range of diagnostic confidence in this study varied between 50% to 100% and further analysis showed that when both raters' confidence in clinical diagnosis was high (80% and above n=24), levels of agreement and reliability improved substantially (kappa = 0.68).

5.5.2 Comparison to other studies

Numerous studies have reported on reliability of multi category classification systems for LBP. These systems are based on specific algorithms which possibly make it easier to agree on categories (Schafer et al. 2009a, Smart et al. 2010). One other study looked specifically at the reliability of the overall clinical impression when assessing LBLP patients (Vroomen et al. 2000). Reliability was substantial (kappa of 0.66) amongst pairs of neurologists who consecutively examined 91 patients with a new episode of sciatica "of sufficient intensity to justify 14 days of bed rest". However, comparing kappa values between studies is considered to be limited due to the differences in methods and sample characteristics (Byrt et al. 1993, Vroomen et al. 2000). One explanation for the lower kappa value seen in this reported study is that subjects were an unselected group, recruited from primary care with symptoms of varying degrees of severity and duration. The greater the proportion of patients with very clear symptoms or findings indicative of the condition of interest, the easier it is for different observers to agree (Vroomen et al. 2000) and conversely agreement on diagnosis may decrease with a greater proportion of "difficult to decide on" patients (Vach 2005). This was reflected in this study by the levels of confidence in diagnosis. Confidence was lower in cases of disagreement, and higher levels of agreement and reliability were seen when diagnostic confidence increased.

The differing interpretations of clinical signs and symptoms amongst raters may also explain the kappa values. Despite consensus that comprehensive history taking and clinical examination are the cornerstones to a sound diagnostic process for LBLP (Freynhagen et al. 2008, Bogduk 2009), inconsistencies are evident in studies when it comes to defining the specific criteria for diagnosing sciatica. A review by Genevay et al. (2010) examined eligibility criteria in published RCTs for two specific LBLP presentations: radiculopathy due to lumbar herniated disc, and neurogenic claudication due to lumbar spinal stenosis (LSS). Of the 12 studies identified for radiculopathy due to lumbar herniated disc, no single diagnostic eligibility category was used in all studies. In the seven studies on neurogenic claudication due to LSS, the presence of positive imaging findings

consistent with LSS was the one criteria used in all studies. This finding was mirrored in a more recent and larger systematic review by Lin et al. (2014) who assessed eligibility criteria to define populations with "radiating"[13] leg pain or symptoms associated with back pain in RCTs of conservative treatments carried out in primary care. Among the 77 studies included in their review, they found there was no consistency in eligibility criteria for defining leg pain due to sciatica.

Some studies that include patients with LBLP require clinicians to make a diagnosis of nerve root pain if certain criteria are fulfilled (Laslett et al. 2005, Freynhagen et al. 2008, Kongsted et al. 2012) but these criteria tend to differ. In the Laslett el al. (2005) study, physiotherapists recorded nerve root pain when leg pain was provoked by nerve tension tests. Kongsted et al. (2012) defined nerve root involvement as the presence of at least one positive finding in the neurological examination. Freynhagen et al. (2008) tabled six criteria for patients to be classified as having radicular pain (due to nerve root involvement): dermatomal pain, motor, sensory or reflex deficit, abnormal straight leg raise and pain on positive femoral nerve stretch. Of the 12 patients in their study diagnosed with radicular pain, four patients had one of these criteria, four patients had two criteria, three patients had four criteria and one patient had none of the criteria which serves to illustrate the variability of clinical signs in these patients. Although diagnostic accuracy of individual items in clinical assessment of sciatica is poor (Vroomen et al. 1999, van der Windt et al. 2010), clinicians are likely to give more weight to certain positive signs when making a confident diagnosis. To improve reliability of this study, fulfilling predefined criteria to make a sciatica diagnosis as opposed to giving an overall clinical impression could have been specified. However as highlighted above, as of yet, clear diagnostic criteria for confidently identifying sciatica has not been agreed on.

[13] Radiating pain in Lin et al's (2014) review included leg pain defined as sciatica, radiculopathy, radicular pain, radicular syndrome and/or lumbar disc herniation, and did not include referred/non-nerve root leg pain

Training of assessors and standardisation of procedures aims to minimize bias in measurement and rating (Kottner et al. 2011). This reported thesis study sought to strike a balance between an appropriate level of standardisation and a setting that reflects current practice in primary care. Using multiple pairs of raters enhances generalisability and reduces the effect of rater bias. Although the physiotherapists were all experienced senior clinicians, very similar results were seen amongst clinicians from varied clinical backgrounds. Regardless of training, standardisation or professional background, reliability was merely fair when diagnosing LBLP, indicating that differentiating between some of these patients is a diagnostic challenge for clinicians in primary care. Irrespective of level of standardisation of procedures used, it is probably difficult to standardise the interpretation of a test result and this is probably where most of the variation in clinical diagnosis comes from. Not all patients are difficult to diagnose, but those cases that are, reduce the reliability indices as was shown in this study.

It is recognised that the composition of a study sample can have a significant impact on kappa values (Vach 2005), the greater the proportion of patients who have very clear symptoms or signs, the easier it is for different observers to agree. To compare reliability indices across studies it is recommended that the main focus should be on finding out if studies differ with regard to the number of subjects "difficult to decide on" (Vach 2005). Considering this sample were an unselected group of patients who consulted their GP with LBLP, it is likely that they will encompass a wider range of characteristics than a more selected sample, for example from secondary care with more severe and disabling symptoms. A possible indication of the number of subjects "difficult to decide on" is reflected in how confident clinicians were in making their diagnoses. This study found that confidence in diagnosis was lower in the cases of disagreement between the raters. In contrast, higher levels of agreement and reliability were seen when therapists' confidence in their diagnosis increased (Analysis III and IV). The therapists in this reliability study were also asked to document the reasons that led them to their diagnostic decision. A trend was seen whereby more reasons for diagnosis were given by the raters who had

higher confidence in their diagnosis. This is not a surprising finding, as the more findings that are related to a certain condition; the more confident the assessor is in identifying that condition (Vach 2005).

The apparent discrepancy between high agreement and low kappa values as seen in this study is commonly observed in reliability studies and much has been written about addressing the issue. Examining the influence of bias of examiners and prevalence is recommended. The bias index was zero or close to zero for Analysis I and II so is not considered an issue of concern. The proportion of agreements on the diagnosis of sciatica did differ from the proportion of agreements on the diagnosis of referred leg pain for all groups and the prevalence indices suggest a possible prevalence effect which can subsequently lower the kappa value. Adjusted kappa scores were calculated to account for prevalence (PAK and AC_1 table 5.4) and the new kappa value moved from the "fair" category to "moderate" reliability suggesting that the higher prevalence of sciatica in this sample did lower the kappa value slightly. It has been argued that adjusting for prevalence or bias should not be done as it masks useful information about the raters and the sample (Hoehler 2000). Instead Hoehler (2000) suggests that researchers should focus on aiming for populations with "trait prevalence near 50%". It is difficult to ascertain the degree to which high prevalence in a study sample reflects true prevalence in the population or the diagnostic behaviour of the clinicians (Sim and Wright 2005, Schafer et al. 2009b).

5.5.3 Strengths and limitations

Use of video lends strengths and limitations to the study design. It allows several raters to make independent diagnoses without burdening the patient with repeated assessment and potentially aggravating their symptoms. Although video recording is considered an established method of recording GP consultations for research purposes (Coleman 2000) it has not been used in studies involving LBP patients that investigate the reliability of clinical diagnosis.

Studies that have used video recording as a test-retest design for LBP patients have done so for functional movement testing only. Luomajoki et al. (2007) videoed patients performing 10 standardized active movements to assess movement control dysfunction; Fritz et al. (2000) videoed patients performing specific movements and to determine the inter-rater reliability of judgments regarding the pain effect of lumbar movement testing on LBP patients and Dankaerts et al. (2006) used video observations of postures and movements alongside subjective case reports to overcome geographical distance between assessors.

However, the use of video could lead to the Hawthorne effect i.e. that behaviour of patients or clinicians would alter due to being videoed, although a review of video recording in general practice found no conclusive evidence of the Hawthorne effect (Coleman 2000). Physiotherapists performing the assessment had the same levels of confidence in their diagnosis as when they watched the assessments on video, possibly indicating that their performance and decision making was not influenced by being video recorded. The two groups of raters who watched the assessments on video did make very similar diagnostic decisions. There were ten disagreement cases between both Group A and B and Group A and C, and nine of the disagreement cases were the same between both groups of video watching raters. To explore the difference in opinions of diagnosis, raters were asked to document reasons for their diagnosis and some of the reasons listed suggested that the raters watching the videos had interpreted some of the neurological testing differently to the rater performing the assessment. In the case of diagnostic disagreement it is not possible to know whether the method of watching a video of a clinical examination negatively influences the ability to interpret the results of a test which contributes to diagnostic decisions. The researcher was present for all the viewings of the videos by raters in Group B and was rarely asked to clarify outcomes of tests. Raters in Group C watched the videos in their own home or workplace and did not contact the researcher to discuss any of the video assessments. The non-standardisation of the method of video watching makes it difficult to draw conclusions about its robustness as a test-retest method for diagnostic decision making.

The study sample represented patients from primary care seen in daily clinical practice. The number of patients recruited in this reliability study is similar to the majority of published reliability studies on LBP classification systems (see chapter four, systematic review). However, the sample size calculation is based on specifying a zero value for kappa in the null hypothesis. The null hypothesis should ideally be set at a higher level, usually ≥ 0.4 which is considered more clinically acceptable (Sim and Wright 2005). However, to use this higher kappa cut-off as the sample size requirement, would require a sample size of 255 subjects (Sim and Wright 2005) which was not feasible for this study.

5.5.4 Clinical Implications

This current study reflects usual clinical practice where an overall clinical impression is made based on signs and symptoms. The results confirm that differentiating between sciatica and referred leg pain can be difficult. Diagnostic modelling in primary care LBLP populations, which assigns weights to various combinations of signs and symptoms has not been done. A clinical diagnostic tool could assist more timely identification of patients with sciatica, facilitate onward referral to appropriate management pathways, as well as allowing clinicians to give patients a clear and consistent opinion on the cause of their leg pain.

5.6 Conclusion

In this study, clinicians demonstrated different diagnostic impressions in just over a quarter of cases following assessments of LBLP patients, which led to a fair reliability rating of their diagnostic decision. Some of this variability may have come from the methodology of using video recording but the diversity of signs and symptoms that these patients present with and the lack of clear guidelines as to what are the strongest criteria for differentiating between sciatica and referred leg pain cannot be ignored. Ways of improving clinician agreement on diagnosis requires further exploration and one solution is to assist the diagnostic process by identifying the optimal

combination of items from the clinical assessment that best discriminate between these patients. The following chapters will set out to do this.

Chapter Six: Identification of sciatica in primary care consulters with low back-related leg pain: Development of a diagnostic model

6.1 Background

In medicine, diagnosis is "estimating the risk or probability that a specific condition is present" (Moons et al. 2015). When patients present with low back and leg pain, the clinician gathers information from the clinical examination to inform their diagnosis. Once the possibility of serious spinal pathology ('red flags') is ruled out, along with other non-spinal reasons for the leg pain (for example; hip pain or vascular problems) the differential diagnosis is between leg pain that is sciatica or non-specific referred pain from the low back. Making this diagnostic decision is recognised as difficult at times (Bogduk 2009), and the reliability study (chapter five) showed that clinicians can disagree on diagnosis. There is no universally agreed or accepted definition or clinical description of sciatica (Genevay et al. 2010, Lin et al. 2014, Germon et al. 2014), something that was also clearly reflected in the findings from the systematic review (chapter four) where subgroups of LBLP due to sciatica had widely varying diagnostic inclusion criteria (table 4.9, page 76). This variability in diagnostic criteria for sciatica can limit the generalisability of results from studies on this subgroup of LBP patients as the same condition may not be evaluated across studies.

In practice, approximately two thirds of patients with back pain will also have leg pain (Hill et al. 2011a, Kongsted et al. 2013). It is not always feasible or necessary to make specific diagnoses in a primary care setting but early identification and differentiation of symptoms of LBLP (sciatica versus referred leg pain) are important for communicating likely diagnosis and prognosis to patients, formulating treatment plans, and guiding the need for referrals to specialist services in a timely manner.

Individual items from history (Vroomen et al. 1999) and physical examination (van der Windt et al. 2010, Iversen et al. 2013, Al Nezari et al. 2013) in patients with sciatica have mostly shown poor diagnostic performance. Diagnostic performance measures reflect a test's ability to discriminate between and/or predict disease in an individual (Eusebi 2013). Combining clinical assessment items is recommended to improve diagnostic performance (van der Windt et al. 2010, Shultz et al. 2015). In primary care, and other settings, clinicians assessing LBP patients integrate several patient characteristics and symptoms to make a prediction about diagnosis. Diagnosis in this case is therefore inherently multivariable (Steyerberg 2009). Diagnostic models are tools that combine predictors to estimate the probability that a condition of interest is present in an individual with a certain predictor profile (Moons et al. 2012).

A review of the literature found six pre-existing diagnostic models to identify sciatica in patients with back and leg pain (table 6.1). These have mainly been developed in secondary care settings with conflicting methods of reference standard and predictor selection.

Three of the six studies selected history items only as predictors in their models (Beattie et al. 2000, Konstantinou et al. 2012c, Verwoerd et al. 2014) and all used different reference standards. Beattie et al. (2000) examined the association between MRI findings and three self-report items of pain location, dysesthesia or weakness in a secondary care population. Konstantinou et al. (2012c) used self-report items as predictors and the reference standard was clinical diagnosis. In their diagnostic model, Verwoerd et al. (2014) used six pre-selected history items, based on previous literature. Information was collected from a patient's interview with a research nurse. The reference standards were MRI findings of either nerve root compression or disc herniation, independently assessed by two neuroradiologists and a neurosurgeon, blinded to any clinical information. The population in this study was a highly selected group of 395 patients with severe sciatica lasting 6-12 weeks who had received a diagnosis of "incapacitating lumbosacral radicular syndrome" from a neurologist (25 patients had already undergone MRI). This highly selected group contrasts to the primary care population with symptoms of any pain severity and duration

in the study of Konstantinou et al. (2012c). Three studies reported diagnostic models using combinations of self-report, history and physical examination items as predictors (Vucetic et al. 1999, Vroomen et al. 2002, Coster et al. 2010). The earliest published diagnostic model used myelography (imaging using contrast dye in the spinal canal and plain x-rays/ computerised tomography (CT) to visualise the nerve roots) as the reference standard (Vucetic et al. 1999) and the population were patients based in secondary care awaiting surgery. Vroomen et al. (2002) investigated the diagnostic value of patient characteristics, history and physical examination items using MRI as the reference standard. Findings from needle electromyography (EMG), which can provide information about the localisation and degree of the nerve root involvement were included as a predictor in Coster et al.'s (2010) study and MRI was used as a reference standard.

The usefulness of MRI as a reference test has been questioned (van der Windt et al. 2010). As discussed in chapter one, positive MRI findings can be found in asymptomatic individuals (Jensen et al. 1994), patients with nerve root compression symptoms may have normal MRIs (Iversen et al. 2013) and MRI findings fail to distinguish sciatica patients in terms of symptom severity (Karppinen et al. 2001). Expert clinical opinion may be considered an appropriate alternative reference standard for diagnosis in the absence of a well-accepted reference standard, with the premise that it is reasonably reliable (Coggon et al. 2005).

To date there is no consensus on what cluster of items best identify sciatica. This study aims to identify the combination of items from clinical assessment that best identify sciatica in primary care consulters with LBLP. The objective of this study is to develop a diagnostic prediction model. The guidelines for transparent reporting of a multivariable model for individual prognosis or diagnosis (The TRIPOD statement, Moons et al. 2015) were followed.

Table 6.1 Multivariable models for identifying sciatica in the literature

Study and population	Index test/ Predictors	Reference test	Predictors in final multivariable model	Odds Ratios (95% CI) for multivariable logistic regression	Model performance (AUC and 95% CI)
Verwoerd et al. 2014 Secondary care n=365	History taking (6 pre -selected items)	MRI (blinded to clinical diagnosis[a])	• Male sex • Sensory loss	1.77 (1.05, 3.00)[b] 2.31 (1.10, 4.85)[b] 3.54 (1.64, 7.64)[c]	0.65 (0.58, 0.71) [b] 0.66 (0.58, 0.74)[c]
Konstantinou et al. 2012c Primary care n=511	Self report items (from questionnaire)	Clinical diagnosis from physiotherapists	• Pain below knee • Leg pain worse than back pain • Numbness/ pins & needles	2.61 (1.68, 4.06)[d] 2.82 (1.85, 4.29)[e] 1.99 (1.20, 3.31)[d] 2.88 (1.72, 4.82)[e] 1.68 (1.07, 2.63)[e]	0.74 (0.70, 0.79)[d] 0.76 (0.72, 0.80)[e]
Beattie et al. 2000 Secondary care n= 408	Self report items (three measures: presence of weakness; paraesthesia/numbness and pain location drawing	MRI	Distal leg pain and MRI findings: • nerve root compression • severe nerve root compression • lateral stenosis no compression Paraesthesia and MRI findings: • lateral stenosis no compression	 2.35 (1.36, 4.06) 2.19 (1.17, 4.12) 2.09 (1.03,4.26) 3.15 (1.36, 7.37)	AUC not reported
Coster et al. 2010 Primary care n=202	Patient characteristics, history items, clinical examination and EMG	MRI (not blinded)	• Dermatomal radiation; • Positive cough/sneeze; • Positive SLR; • Denervation on EMG	2.1 (1.3, 4.8) 2.4 (1.2, 4.7) 3.0 (1.6, 5.7) 4.5 (2.1, 9.5)	AUC not reported

Table 6.1 Multivariable models for identifying sciatica in the literature

Study and population	Index test/ Predictors	Reference test	Predictors in final multivariable model	Odds Ratios (95% CI) for multivariable logistic regression	Model performance (AUC and 95% CI)
Vroomen et al. 2002 Primary care n=274	Patient characteristics, history items and physical examination	MRI (blinded)	• Age (yrs) 41-50 v 16-40 51-81 v 16-40 • Symptom duration (days)15-30 v <15 • Paroxysmal spasm • Pain worse leg than back • Dermatomal distribution • Positive cough/sneeze/strain • Finger floor distance (cms) >24 v 0-4 • Paresis	1.8 (1.3, 2.6) 2.8 (1.9, 4.2) 2.2 (1.5, 3.3) 1.8 (1.3, 2.5) 4.5 (3.3, 6.2) 3.2 (2.2, 4.7) 2.0 (1.4, 2.7) 2.8 (1.9, 4.3) 5.2 (3.3, 11.6)	History findings alone: 0.80 (CI not reported) Adding physical examination findings: 0.83 (CI not reported)
Vucetic et al. 1999 Secondary care n=160	History items, physical examination and surgery findings	Myelography preoperatively	• Incapacitating pain • Crossed lasegue (SLR) sign • Education • Dislocation dura/root on myelogram • No comorbidity • No previous surgery • Lumbar sagittal ROM < 33°	30.5 (5.2, 179.9) 6.1 (1.1, 33.7) 5.5 (1.5, 19.5) 5.0 (1.6, 15.2) 3.6 (1.0, 12.2) 3.5 (1.1, 10.4) 3.1 (1.0, 9.3)	AUC not reported

CI, confidence interval; AUC, Area under the receiving operating characteristic curve; MRI, magnetic resonance imaging; EMG, electromyography; SLR, straight leg raise; ROM range of movement.

[a] "Blinded"; no knowledge of patient's clinical symptoms or clinical diagnosis when reading the MRI scans.

[b] Disc herniation on MRI was the reference standard.

[c] Nerve root compression on MRI was the reference standard.

[d] "Confirmatory" reference standard: excluding possible inconclusive cases.

[e] "Indicative" reference standard which included "possible" cases.

6.2 Methods

6.2.1 Source of data

This study is a cross-sectional diagnostic study. Data is sourced from participants in the ATLAS cohort (**A**ssessment and **T**reatment of **L**eg pain **A**ssociated with the **S**pine), a prospective observational study of primary care consulters who had visited their GP with LBLP of any duration and intensity. Full details of the ATLAS study methodology, including eligibility criteria are detailed in chapter three.

Data collection

At the ATLAS research clinic, all patients received a standardised clinical assessment by an experienced musculoskeletal physiotherapist. Full details of the assessment are detailed in chapter three. Self-reported measures were collected from questionnaires which participants completed prior to their clinical assessment. Clinical assessment findings were collected at the ATLAS research clinic. The seven physiotherapists who performed the assessments were experienced musculoskeletal clinicians who had been given training in the procedures of the study. At the end of the clinical assessment, the physiotherapists were asked to document (i) whether the presence of leg pain was due to nerve root involvement (NRI) and (ii) confidence (0-100%) in their clinical diagnosis/impression.

Clinicians made their diagnostic decision based on information from history and physical examination findings. They did not have any additional diagnostic information available (e.g. findings from advanced testing such as MRI) to assist with their diagnosis. Within 14 days of the clinical assessment, patients eligible and consenting to take part in the ATLAS study received an MRI scan of the lumbar spine as part of the research study, providing there were no clinical contraindications to the procedure. MRIs were scored by a senior consultant musculoskeletal radiologist, blind to any clinical information about the patient's presentation other than that the

patient had LBLP (not specifying which leg). The radiologist provided a clinical report indicating definite, possible or absence of nerve root compression.

6.2.2. Outcome

The outcome of interest in this study is a diagnosis of sciatica. Two reference standards were chosen for the diagnostic model (box 6.1). Justification for the two reference standard definitions is outlined below.

Box 6.1 Definitions for the two reference standards

Model one reference standard:
High confidence (≥ 80%) sciatica clinical diagnosis.

Model two reference standard:
High confidence (≥ 80%) sciatica clinical diagnosis with confirmatory MRI findings.

Model one reference standard: High confidence (≥ 80%) sciatica clinical diagnosis

A diagnosis of sciatica was concluded if the physiotherapist indicated the presence of leg pain was due to sciatica and if their confidence in their clinical diagnosis was 80% or above.

The objective of the study described in chapter five was to determine the reliability of the clinical diagnosis of LBLP amongst clinicians. The results showed that reliability was fair (kappa = 0.35), but improved substantially as confidence in diagnosis increased. A cut off point of ≥ 80% diagnostic confidence was used for this diagnostic model reference standard because at this criterion, reliability among clinicians was moderate (kappa =0.68).

Model two reference standard: High confidence (≥ 80%) sciatica clinical diagnosis with confirmatory MRI findings

The second reference standard combines the clinician's diagnosis of sciatica with confirmatory (positive) findings of NRI on MRI. Using clinical diagnosis alone as a reference standard necessitates use of information from the clinical assessment items (the predictors) by the clinicians, to make their diagnosis. Using this approach may leave diagnosis open to incorporation

bias as the reference standard is not blind to knowledge of the predictors under consideration (Knottnerus 2002).

Other models have used findings from imaging (MRI) which eliminates the potential for incorporation bias (Vroomen et al. 2002, Verwoerd et al. 2014). However relying on imaging alone is unsatisfactory and can lead to misclassification of patients, as described earlier in section 6.1. Therefore the Model two reference standard was high confidence clinical diagnosis plus nerve root involvement on MRI, in line with patient's symptoms.

6.2.3 Predictors

Nine items were initially chosen as potential candidate predictors for inclusion in the diagnostic model from the larger set of available self-report and clinical assessment findings. Predictor selection was guided by findings from the systematic review on LBLP classification systems (chapter four), clinical knowledge and published literature (Steyerberg 2009). Factors which influenced the predictor selection for this model were based on the following criteria:

a) Clinical criteria for sciatica consistently identified in the systematic review (chapter four) of LBLP classifications systems (table 4.9 page 76).
b) Expert consensus from a published Delphi study involving representatives from LBP disciplines on items from clinical assessment considered most important for distinguishing sciatica from referred leg pain in LBLP patients (Konstantinou et al. 2012b).
c) Items used in published multivariable diagnostic models shown to have acceptable diagnostic accuracy for identifying sciatica (table 6.1).
d) Items clinicians documented in the ATLAS clinical assessment as most important for diagnosing sciatica (see chapter three, box 3.2, question 3).

Eight of the nine predictors were items from the clinical assessment and recorded by the physiotherapist carrying out the assessment. One item (intensity of leg pain) was collected from the baseline self-report questionnaire that participants completed before their clinical

assessment. The selected variables are presented and cross referenced to the criteria above [a,b,c,d] that influenced their selection. More detailed description of the clinical assessment variables and method of performance of the physical examination tests is in chapter three, section 3.3 page 24.

Self-report/history items

Subjective sensory changes [a,b,c,d] (yes/no): A positive response is recorded if patients reported that they had noticed symptoms such as numbness, pins and needles or tingling in their leg.

Below knee pain [a,b,c,d] (yes/no): A positive response is recorded if the assessing physiotherapist marked any areas of pain below the knee on the body chart manikin, this could include areas of pins and needles or numbness below the knee.

Leg pain worse than back pain [a,b,c,d] (yes/no): A positive response is recorded if the patient reported that their leg pain is worse/ or bothers them more than their back pain.

Leg pain intensity [b,c] (0-10): This was measured using the mean of three 0 to 10 numerical rating scales for 'least' and usual' leg pain over the previous two weeks and 'current' leg pain.

Positive cough or sneeze [a,b,c,d] (yes/no): A positive response is recorded if patient's leg pain (or on occasion buttock pain) is reproduced or increased on coughing, sneezing or straining.

Physical examination items

Myotome deficit [a,b,c,d] (yes/no). A deficit was defined as less than normal muscle strength (grade 5/5 on oxford manual muscle testing scale) in any of the tested myotomal muscle groups of the symptomatic lower limb(s) (see table 6.2).

Reflex deficit [a,b,c,d] (yes/no): A deficit was defined as either a reduced/significantly reduced or absent tendon reflex (at the knee or the ankle) in the symptomatic lower limb(s).

Sensory deficit [a,b,c,d] (yes/no): A deficit was defined as either a reduced, significantly reduced or absent (anaesthesia) response to pin prick testing of dermatomal distribution areas in the symptomatic lower limb(s).

Positive neural tension tests [a,b,c,d] (yes/no): Straight leg raise (SLR) test, crossed SLR, femoral nerve stretch test, slump test. A positive neural tension test was defined as reproduction of the patient's leg pain during performance of any of the neural tension tests. These tests were performed sequentially (see chapter three, section 3.3.2, page 27), hence it was not appropriate to consider them as individual predictors. For example, if a SLR was positive, the clinician would not generally have performed a slump test as well.

6.2.4 Sample size

The sample size for this diagnostic model was not formally calculated as data from an existing cohort (ATLAS study) was used. However, despite consensus that adequate sample size is needed for developing diagnostic prediction models, there are no generally accepted approaches as to what constitutes an adequate sample size calculation (Steyerberg 2009). Sample size is usually discussed in relation to adequacy of size in relation to the number of candidate predictors. A rule of thumb often quoted in studies suggests a guide of 10 events per predictor variable degrees of freedom ($p<m/10$ (where p=parameter, and m=smallest number in either category of the outcome variable) to allow for a reliable model (Peduzzi et al. 1996). However, many authorities point out that this rule is not based on any convincing scientific reasoning (Steyerberg 2009, Moons et al. 2012). This model initially had nine predictors and the smallest outcome category had 100 events (100 patients diagnosed with referred pain for reference standard one; see section 6.3.2). Hence the guide of 10 events per predictor was adequately satisfied. The sample size (n=395) is similar to other published diagnostic model on LBLP patients (table 6.1).

6.2.5 Statistical analysis

Defining the study population

The characteristics of the population were presented in terms of sociodemographics, pain characteristics, disability, time off work, psychosocial factors, clinical findings and MRI findings. A comparison was made between the characteristics of the population used in the diagnostic modelling and those that were excluded from the analysis due to application of the reference standard criteria.

Predictor variable preparation

The self-report predictor of leg pain intensity was maintained as a continuous variable (0-10) for the multivariable analysis. Categorisation of continuous variables is a popular approach in prediction modelling but is discouraged due to loss of information (Moons et al. 2012). The four predictors from history taking required a 'yes/no' response from the patient (pain below the knee, back pain worse than leg pain, subjective sensory changes, positive cough/sneeze). Of the physical examination predictors, the neural tension tests were recorded as positive or negative reflecting clinical practice documentation.

The neurological examination items of myotomes, reflexes and sensation were ordinal measurements with several levels and were dichotomised as either positive or negative. See table 6.2 for description of the method used to measure and categorise the neurological examination variables.

Table 6.2 Predictor coding and categorisation for results of neurological examination

Predictor	Measurement	Coding	Coding categorisation 0= normal/ no deficit 1= abnormal/deficit
Myotomes	5/5	0	Myotome deficit was recorded with a code of 1 in any of the muscle groups assessed
	4/5	1	
	3/5	1	
	2/5	1	
	1/5	1	
	0/5	1	
Reflexes	Normal	0	Reflex deficit was recorded with a code of 1 for ankle jerk and/or knee jerk reflex
	Slightly reduced	1	
	Significantly reduced	1	
	Absent	1	
Sensation	Normal	0	Sensory deficit was recorded with a code of 1 for sensation testing
	Reduced pin prick	1	
	Loss of pin prick	1	
	Anaesthesia	1	

Revision of predictor categories

Keeping the categories for each of the neurological predictors was initially planned but on inspection of the cross tabulated frequency tables, on most occasions, any deficit other than normal was predominantly diagnosed by the clinicians as sciatica, resulting in small or zero frequencies in the "referred" leg pain cells. Hence it was not practical, clinically or statistically, to consider them as individual predictors. The decision to combine the neurological deficit variables in a clinically acceptable way was made (Hosmer and Lemeshow 2000). The three neurological tests of myotomes, reflexes and sensation were combined to make one variable— 'any deficit on neurological testing'. This left **seven predictors** for selection in the multivariable model.

Diagnostic accuracy of individual predictors

The diagnostic accuracy of the seven predictors was individually calculated for both reference standards using sensitivity, specificity, predictive values, likelihood ratios (LRs), diagnostic odds ratios and the area under the receiver operating characteristic curve (ROC) curve (c statistic, AUC) (see box 6.2 for definitions). Examining these values allows for comparison with previous

literature on diagnostic accuracy of individual items from clinical assessment and also allows comparison of diagnostic accuracy of individual items versus combination of items.

Univariable logistic regression

Univariable logistic regression analysis was carried out to quantify the relationship between each individual predictor variable and the presence of sciatica. Univariable analysis was not performed to determine which predictors go forward into the multivariable analysis based on their association with the outcome. This process of predictor selection is no longer recommended because important predictors with a non-statistical unadjusted association with the outcome could be falsely rejected due to potential confounding by other predictors (Harrell 2001, Steyerberg 2009, Moons et al. 2012).

Collinearity

Predictors may be strongly correlated to each other, this is known as collinearity. High correlation among predictors makes it difficult to get good estimates of their distinct contribution to the outcome variable (Midi et al. 2010). Predictors were checked for correlation and if evidence of correlation was present, the "variance inflation factors" (VIF) of the predictors were calculated. A VIF>10 indicates strong collinearity and could hamper the reliability of the regression coefficients in the multivariable model (Steyerberg 2009). If this happens, the first approach is to check if the correlated variables can be combined into a single variable that makes clinical sense. Failing this, variables likely to be highly correlated will be identified then individually removed and the VIF is recalculated until the maximum number of predictor variables can be retained whilst not exceeding a VIF of 10 (O' Brien 2007, Midi et al. 2010).

Box 6.2 Commonly used measures describing the accuracy of a diagnostic test

Consider a test (clinical examination item) with either a positive result (T+) or negative (T-) used to distinguish between people with a positive diagnosis (of sciatica) (D+) or a negative diagnosis (referred leg pain) (D-)

2x2 table for a diagnostic test

	Reference Standard Diagnosis	
Test	D+	D-
Positive (T+)	True positive (TP)	False positive (FP)
Negative (T-)	False negative (FN)	True Negative (TN)

Table of probabilities related to diagnostic tests (adapted from Hunink et al. (2014) pp125)

Name	Definition	Estimate from 2x2 table
Sensitivity	Probability of a positive test/response in patients with sciatica diagnosis	TP/(TP+FN)
Specificity	Probability of a negative test/response in patients without sciatica diagnosis	TN/(TN+FP)
Positive predictive value	Probability of presence of sciatica in patients with a positive test/response	TP/ (TP+ FP)
Negative predictive value	Probability of absence of sciatica in patients with a negative test/response	TN/ (FN+TN)
Positive Likelihood ratio	How many more (or less) times likely a person with a positive test/response will have the diagnosis	Sensitivity/1-specificity
Negative Likelihood ratio	How many more (or less) times likely a person with a negative test/response will not have the diagnosis	1-sensitivity/specificity

The receiver operating characteristic curve (ROC) represents the relationship between the sensitivity and specificity of a test with variable cut off points. An area under the curve (AUC) of 1 indicates a perfect test, whereas an AUC of 0.5 is a useless test with no discriminatory ability (Knottnerus 2002, Hunink et al. 2014).

Type of model

Binary logistic regression analysis was performed to determine which of the clinical predictors significantly contributed to the identification of patients with sciatica amongst the LBLP patients. All a priori selected predictor variables were included in the full multivariable model. Complete

case analysis was planned excluding individuals with missing values on any predictors (Moons et al. 2015).

Backwards stepwise selection procedure was used which is the preferred approach of automated selection procedures as it allows the modeller to assess the full model and consider the effect on the model as the variables are removed sequentially (Steyerberg 2009). The process starts with all potential predictors and then using the likelihood ratio test starts sequentially removing the predictor adding least to the model and refitting the model until all remaining predictors meet a predefined significance level. Usual practice of applying the standard significance level for testing of hypotheses, $p<0.05$, (Steyerberg 2009) was used in this model. The contribution made by each predictor variable within the final model was presented as beta coefficients and odds ratios with their 95% CIs.

Performance measures

Assessing the predictive performance of a model involves examining measures of calibration and discrimination. Calibration is the agreement between observed outcomes and prediction (Steyerberg et al. 2010). For this model, calibration is the agreement between the probability of receiving a diagnosis of sciatica using the constructed diagnostic model and the observed sciatica diagnosis according to the reference standard definition. Graphical assessment of calibration is recommended (Harrell 2001, Moons et al. 2015). The observed outcome was plotted against the predicted probability of the outcome obtained from the fitted logistic regression model using the Lowess smoothing curve technique (Austin and Steyerberg 2014). Perfect calibration shows a slope on the 45 degree line hence deviation of the line from the diagonal indicates lack of calibration. The plot was supplemented with the Hosmer and Lemeshow goodness of fit test (Hosmer and Lemeshow 2000). P value equal or greater than 0.05 supports the goodness of fit.

Discrimination is the ability of the model to distinguish between those who do and do not have the sciatica diagnosis. Discrimination was summarised using the area under the curve (AUC) of the receiver operating characteristic (ROC) curve. The ROC curve plots the true positive rate

(sensitivity) against the false positivity rate (1–specificity) at any given cut off value. The curve shows the model's ability to discriminate between patients with and without sciatica at subsequent cut off points across the range of predicted probabilities (Kuijpers et al. 2007). An AUC of 0.5 indicates no discrimination whereas AUC of 1.0 indicates perfect discrimination (Hosmer and Lemeshow 2000). A visual display of discrimination was done using simple histograms comparing the distribution and potential overlap in predicted probabilities generated from the logistic regression model (Royston and Altman 2010).

A measure of the usefulness of the model or how well it fits the data is the statistic R^2 which is an estimate of the proportion of variance explained by the model and reflects the correlation between the models predicted outcome and the observed outcome. Nagelkerke's R^2 is preferable over Cox and Snell's R^2 because it can achieve a value of one when the model predicts the data perfectly (Field 2005).

Internal validity

A prediction model is expected to perform optimally in the sample in which it was developed, but can become less accurate if tested in a new sample. This issue of overfitting can be evaluated quantitatively by using internal validation techniques. Bootstrapping is considered the most efficient means of internal validation because all the available data is used compared to split-sample analyses or cross validation techniques (Steyerberg 2001). Samples are drawn with replacement from the original sample. Large numbers of replications are advisable and the bias corrected percentile method is preferable for smaller sample sizes which accounts for bias in estimation of the distribution, based on the difference between the median of the bootstrap estimates and the sample estimate (Steyerberg 2009, Kim 2005).

Bootstrapping was performed on both models using 1000 samples. Output estimates were checked in samples of 2,000, 5,000 and 10,000 replications to confirm that estimates did not improve with greater replications. An adjusted AUC was calculated for the bootstrapped model to reflect the discriminative performance of the internally validated model.

Scoring tool

To illustrate how the diagnostic model could be used in the clinical setting, a simplified scoring tool for the best performing model was derived, which would give a LBLP patient their probability of having sciatica. The regression coefficients for each predictor in the final model were converted to a whole number by dividing each item coefficient by the lowest value coefficient (Steyerberg 2009). Scores were presented alongside their associated outcome probabilities.

Sensitivity analyses

Considering that the complete ATLAS dataset (n=609) was not used in the main analysis due to the reference standard criteria, sensitivity analyses were performed. Additional multivariable logistic regression models were derived using less restrictive confidence in diagnosis criteria. A model using MRI results only as a reference standard was also constructed and a model removing all patients who had a diagnosis of stenosis (clinician diagnosis or MRI findings) was performed to compare with the original Model one. Stenosis is associated with certain clinical features and older age hence an analysis removing this group of patients was done to assess its impact on the clinical diagnosis model. The log odds ratios, corresponding confidence intervals and AUCs of these additional models were reported.

The following alternative reference standards were used to define the outcome of sciatica:

Model one (a): Confidence > 70% Sciatica clinical diagnosis (and had an MRI scan)

Model one (b): Sciatica clinical diagnosis, all ranges of confidence (and had an MRI scan)

Model one (c): Sciatica clinical diagnosis, all ranges of confidence

Model one (d): Sciatica clinical diagnosis, all ranges of confidence, excluding 41 patients with stenosis (and had an MRI scan)

Model two (a): Confidence > 70% sciatica clinical diagnosis plus confirmatory MRI

Model two (b): Sciatica clinical diagnosis, all ranges of confidence plus confirmatory MRI

Model three: Confirmatory MRI only

For Model one (a-d) diagnosis was made irrespective of MRI results. Model one (c) included the full dataset so not all patients had an MRI scan.

Finally, a different predictor entry selection was done to mirror a method used in a published lumbosacral nerve root compression diagnostic model which showed that adding physical examination items only minimally improved the AUC performance of the model (Vroomen et al. 2002). History and self-report items were entered in a block using the reference standards in Model one and Model two, and the AUC of the models were compared to the AUC of Model one and two.

The majority of the analysis for this chapter was completed in SPSS version 21. Stata version 5 was used to produce the calibration plots, perform the bootstrapping and calculate the confidence intervals for the area under the curves (AUCs) of the bootstrapped model. Some computations (likelihood ratios) were carried out in an on-line statistical programme (http:// vassarstats.net/clin1.html Accessed 03/03/2016).

6.3 Results

6.3.1 Participants

Of the 1310 participants who initially attended the ATLAS research clinics, 609 participated in the ATLAS study and were eligible for inclusion in this diagnostic model analysis (figure 6.1). 214 patients were excluded from the diagnostic model analysis when (i) clinician confidence in diagnosis was <80% (n=173), (ii) patients who did not have an MRI scan (an additional 41 patients). 395 patients were included in the diagnostic model development analysis.

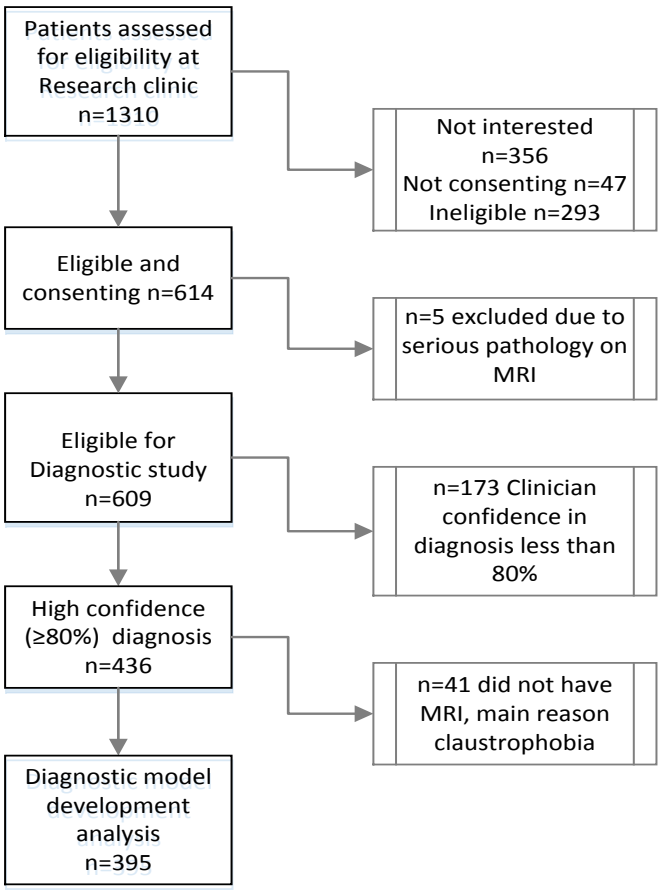


Figure 6.1 Participant flow diagram for diagnostic model development

Table 6.3 (a,b,c) displays characteristics of: (i) patients in the diagnostic model development sample (n=395), (ii) all patients in the ATLAS study (n=609), and (iii) those who were not included in the model building analysis (n=214) following application of the reference standard criteria of clinician confidence in diagnosis of ≥ 80% and having an MRI scan. Groups were similar in sociodemographic profiles. Compared to the patients included in the diagnosis model development analysis (n=395), the group excluded from the analysis (n=214) had a greater proportion of patients over 65 years of age (17.8% v 13.7%), a higher proportion of females (68.2% v 60%) and fewer were currently in a paid job (56.5% v 62.3%).

Table 6.3 (a,b,c) Comparing diagnostic model population (n=395) to eligible population of full data set (n=609) and eligible population excluded from the analysis due to the reference standard criteria (n=214) (diagnosis confidence < 80% or confidence ≥ 80% but no MRI)

Table 6.3(a) Sociodemographic characteristics

Socio-demographics	n=395 Diagnostic model cases	n=609 All Eligible cases	n=214 Excluded cases
Age (years) mean (SD)	49.8 (13.9)	50.2 (13.9)	50.9 (13.9)
Age categories			
18-34 years	64 (16.2)	91 (14.9)	27 (12.6)
35-44 years	82 (20.8)	136 (22.3)	54 (25.2)
45-54 years	102 (25.8)	152 (25.0)	50 (23.4)
55-64 years	93 (23.5)	138 (22.7)	45 (21.0)
65+ years	54 (13.7)	92 (15.1)	38 (17.8)
Gender, Female	237 (60.0)	381 (62.6)	146 (68.2)
Current smoker	128 (32.4)	194 (31.9)	67 (31.3)
BMI categories			
Normal/underweight	87 (22.0)	136 (22.4)	49 (22.9)
Overweight	146 (37.0)	223 (36.7)	77 (36.0)
Obese/ Morbidly obese	160 (40.5)	248 (40.9)	88 (41.1)
Socio-economic status [a]			
Higher	90 (23.3)	129 (21.8)	39 (18.9)
Intermediate	99 (25.6)	158 (26.2)	59 (28.6)
Routine	183 (47.3)	283 (47.7)	100 (46.7)
Never worked / long term unemployed	15 (3.9)	23 (3.9)	8 (3.7)
Currently in paid job [a]	266 (62.3)	367 (60.7)	121 (56.5)
Self-certified time off work or given sick note due to current episode [b]	99 (40.9)	144 (39.7)	45 (37.2)

SD, standard deviation; BMI, body mass index.
All figures are frequencies (percentages) unless stated otherwise as mean (standard deviation (SD)).
[a] Denominator varies for some variables due to missing data/not applicable cases. Range of missing data for column *Diagnostic model cases*: 1,8 (<5%); Column *All eligible cases*: 1,16 (<5%); Column *Excluded cases:* 1,8 (<5%).
[b] Applicable to only those working n=365.

Self-report measures (Table 6.3b) of pain and disability were very similar across groups. A greater proportion of patients in the excluded group (n=214) had their leg symptoms for more than 3 months (41.6% v 33.5%). There were a higher proportion of depressed patients in the sample used for the diagnostic model analysis (18.0% v 12.6%) but fewer comorbidities were reported by

this group compared to the 214 patients excluded from the analysis (10.9 % ≥ 2 comorbidities v 17.3%).

Table 6.3(b) Self report, physical, psychological and health characteristics

Self-report physical, psychological and health measures	**n=395 Diagnostic model cases**	**n=609 All eligible cases**	**n=214 Excluded cases**
RMDQ disability score (0-23) mean (SD)	12.8 (5.7)	12.7 (5.7)	12.4 (5.8)
Back pain intensity, mean (SD)	5.5 (2.2)	5.6 (2.2)	5.7 (2.1)
ǂ Leg pain intensity, mean (SD) [a]	5.3 (2.4)	5.2 (2.4)	5.1 (2.4)
Duration of back symptoms [a]			
Less than 6 weeks	132 (33.6)	218 (35.9)	86 (40.1)
6-12 weeks	94 (23.9)	126 (20.8)	32 (15.0)
>3 months	167 (42.5)	263 (43.3)	96 (44.9)
Duration of leg symptoms [a]			
Less than 6 weeks	165 (43.9)	251 (43.1)	86 (41.5)
6-12 weeks	85 (22.6)	120 (20.6)	35 (16.9)
>3 months	126 (33.5)	212 (36.4)	86 (41.6)
Sciatica Bothersomeness Index (0-24),mean (SD) [a]	14.2 (5.4)	14.2 (5.4)	14.3 (5.3)
s-LANSS, neuropathic pain score (≥12) [a]	194 (49.4)	293 (48.1)	99 (46.3)
STarT Back subgroup [a]			
Low risk	53 (13.9)	82 (13.9)	29 (13.9)
Medium risk	180 (47.4)	276 (46.9)	96 (45.9)
High risk	147 (38.7)	231 (39.2)	84 (40.2)
HADs anxiety subscale [a]			
Normal	208 (52.8)	316 (52.1)	108 (50.7)
Mild/possible case	80 (20.3)	120 (19.8)	40 (18.8)
Probable/moderate/severe case	106 (26.9)	171 (28.2)	65 (30.5)
HADs depression subscale [a]			
Normal	247 (62.5)	392 (64.4)	145 (67.8)
Mild/possible case	77 (19.5)	119 (19.5)	42 (19.6)
Probable/moderate/severe case	71 (18.0)	98 (16.1)	27 (12.6)
Pain self-efficacy score, mean (SD) [a]	34.0 (14.7)	34.1 (14.6)	34.5 (14.4)
Illness perceptions questionnaire (IPQ-R)			
Timeline [b]	221 (55.9)	345 (56.7)	124 (57.9)

Table 6.3(b) Self report, physical, psychological and health characteristics

Self-report physical, psychological and health measures	**n=395 Diagnostic model cases**	**n=609 All eligible cases**	**n=214 Excluded cases**
Personal control [a, c]	245 (62.0)	367 (60.7)	122 (57.5)
Identity score (0-7), mean (SD) [a]	5.9 (1.3)	5.9 (1.3)	6.0 (1.3)
General Health [a]			
Excellent / very good/ good	251 (63.7)	387 (63.5)	135 (63.1)
Fair/poor	143 (36.3)	222 (36.5)	79 (36.9)
Co-morbidities [d]			
None	250 (63.3)	371 (60.9)	121 (56.5)
One other health problem	102 (25.8)	158 (25.9)	56 (26.2)
Two or more other health problems	43 (10.9)	80 (13.1)	37 (17.3)
EQ—5D summary index [a]	0.4 (0.3)	0.4 (0.3)	0.5 (0.3)

RMDQ, Roland Morris Disability Questionnaire; SD, standard deviation; s-LANSS, self-report Leeds Assessment of Neuropathic Symptoms and Signs; HADS, Hospital Anxiety and Depression Scale.
All figures are frequencies (percentages) unless stated otherwise as mean (SD).
[ł] Predictors used in the multivariable diagnostic model.
[a] Denominator varies for some variables due to missing data/not applicable cases. Range of missing data for Column *Diagnostic model cases*: 1,20 (<5%); Column *All eligible cases*: 1,25 (<5%); Column *Excluded cases:* 1,12 (<6%).
[b] My back/leg pain will last for a long time" (agree or strongly agree).
[c] "What I do can determine whether my back/leg pain gets better" (agree or strongly agree).
[d] Comorbidities health problems included: chest problems, heart problems, raised blood pressure, diabetes, circulation problems in the legs.

Clinical characteristics (table 6.2c) for the diagnostic model group (n=395) differed considerably to the group excluded from the analysis (n=214). A greater proportion of patients in the diagnostic model group had a positive cough/sneeze (25.8% v 12 .6%), leg pain worse than back pain (50.1% v 38.3%), neurological deficits (57.7% v 46.3%) and positive neural tension tests (60.8% v 44.4%). Subjective sensory symptoms (pins and needles/numbness) were slightly lower in the diagnostic model analysis group (61.5% v 65.0%).

Table 6.3(c) History and physical examination items from clinical assessment

History items	n=395 Diagnostic model cases	n=609 All eligible cases	n=214 Excluded cases
ǂ Positive cough/ sneeze	102 (25.8)	129 (21.2)	27 (12.6)
ǂ Below knee pain	278 (70.4)	430 (70.6)	152 (71.0)
ǂ Leg pain worse than back pain	198 (50.1)	280 (46.0)	82 (38.3)
ǂ Subjective sensory changes in leg	243 (61.5)	382 (62.7)	139 (65.0)
Physical Examination items			
ǂ Neurological tests deficit (any positive test)	228 (57.7)	327 (53.7)	99 (46.3)
Myotomes [a]	81 (20.5)	105 (17.2)	24 (11.2)
Sensation	173 (43.8)	253 (41.5)	80 (37.4)
Reflexes	91 (23.0)	119 (19.5)	28 (13.3)
ǂ Neural tension tests (any positive test)	240 (60.8)	335 (55.0)	95 (44.4)
Straight leg raise (SLR)	221 (55.9)	297 (48.8)	76 (35.5)
Crossed SLR	21 (5.3)	22 (3.6)	1 (0.5)
Slump	64 (16.2)	84 (13.8)	20 (9.3)
Femoral nerve stretch	27 (6.8)	41 (6.7)	14 (6.5)
Imaging [a]			
Positive MRI findings for nerve root compression	231 (58.5)	297 (53.7)	66 (30.8)
Clinical diagnosis; Stenosis	34 (8.6)	48 (7.9)	14 (6.5)

All figures are frequencies (percentages) unless stated otherwise as mean (SD).
ǂ Predictors used in the multivariable diagnostic model.
[a] Denominator varies for some variables due to missing data/not applicable cases. Total missing data for myotomes n=1 (<1%); for Imaging *Diagnostic model cases*: n= 1(<1%); *All eligible cases*: n=56 (9%); *Excluded cases:* n=55 (26%).

6.3.2 Outcome

Of the 395 patients who fulfilled the reference standard criteria, 75% (n=295) were diagnosed with sciatica using reference standard one. Using reference standard two, where clinical diagnosis was corroborated by positive MRI findings, 51% (n=200) were diagnosed with sciatica as shown in the table 6.4.

Table 6.4 Cross-tabulation between clinical diagnosis and MRI findings

		Clinical diagnosis of sciatica		
		Yes	**No**	**Total**
Nerve root involvement on MRI	**Yes**	200	31	231
	No	95	69	164
	Total	295	100	395

6.3.3 Predictors

Table 6.5 presents the cross-tabulated frequency data of the six categorical variables for both reference standards (leg pain intensity was a continuous variable). Following re-categorisation of the variables by combining neural tension tests as one variable and results from neurological testing as one variable, all cells had at least 5 events (see Appendix I for 2x2 table prior to re-categorisation of variables). Comparing the "yes" responses across the two references standards, it is clear that for reference standard one, the majority of cases had a sciatica clinical diagnosis when they had a 'yes' response to the any of the history or clinical examination items. This was not the case for reference standard two; there were a higher proportion of false positive cases.

Table 6.5 Descriptive cross-tabulated frequency data of categorical predictors for diagnosis of sciatica in patients with low back-related leg pain with Model one reference standard: high confidence (≥ 80%) clinical diagnosis and Model two reference standard: high confidence (≥ 80%) clinical diagnosis plus confirmatory MRI

		Model one		Model two	
History items		Sciatica (*n*=295)	Referred (*n*=100)	Sciatica (*n*=200)	Referred (*n*=195)
Sensory changes	Yes	206 (69.8%)	37 (37.0%)	138 (69.0%)	105 (53.8%)
	No	89 (30.2%)	63 (63.0%)	62 (31.0%)	90 (46.2%)
Below knee pain	Yes	250 (84.7%)	28 (28.0%)	173 (86.5%)	105 (53.8%)
	No	45 (15.3%)	72 (72.0%)	27 (13.5%)	90 (46.2%)
Leg pain worse than back pain	Yes	183 (62.0%)	15 (15.0%)	142 (71.0%)	56 (28.7%)
	No	112 (38.0%)	85 (85.0%)	58 (29.0%)	139 (71.3%)
Cough/Sneeze	Yes	95 (32.2%)	7 (7.0%)	79 (39.5%)	23 (11.8%)
	No	200 (67.8%)	93 (93.0%)	121 (60.5%)	172 (88.2%)
Clinical Examination items					
Positive neural tension	Yes	230 (78.0%)	10 (10.0%)	157 (78.5%)	83 (42.6%)
	No	65 (22.0%)	90 (90.0%)	43 (21.5%)	112 (57.4%)
Neurological deficit	Yes	212 (71.9%)	16 (16%)	149 (74.5%)	79 (40.5%)
	No	83 (28.1%)	84 (84%)	51 (25.5%)	116 (59.5%)

6.3.4 Diagnostic accuracy of individual items

The diagnostic properties of the individual predictors were calculated and presented (table 6.6) using the information from the cross-tabulated frequency tables of the predictor items against both reference standards (table 6.5).

The predictor with the greatest sensitivity was 'pain below knee' (0.85 and 0.87 for reference standard one and two respectively). For Model one reference standard, with the exception of 'pain below knee', specificity values were greater than sensitivity values. This is reflected in the cross-tabulated table (table 6.5) with very few false positive counts in the cells. High specificity is considered a proxy for "ruling in" the condition of interest. Sensitivity was lower because not all patients had these positive tests. For example 102 participants out of 395 reported leg pain on cough/sneeze or strain and only seven of these were diagnosed by clinicians as not having sciatica. Two hundred patients with sciatica did not have this positive indicator, indicating the test does not have good sensitivity for identifying those who do have the condition of interest.

Average sensitivity and specificity is reflected in the AUC value. 'Leg pain worse than back pain' had an AUC greater than 0.7 for both reference standards. The best performing individual test overall was neural tension test for Model one reference standard (AUC 0.84) but it performed less well (AUC 0.68) with MRI in the reference standard for Model two.

For Model one reference standard, positive predictive values (PPV) were high ranging from 0.90 to 0.96, meaning the likelihood of having a positive diagnosis of sciatica is very high if the index test is positive. This range was considerably lower for Model two reference standard (0.64 to 0.77). However, predictive values need to be interpreted in relation to the pre-test probability of the diagnosis (i.e. prevalence) in this case 0.75 for Model one reference standard and 0.51 for Model two reference standard, hence the difference between the pre and post-test probabilities is similar for the two reference standards.

Likelihood ratios (LR+/-) are often considered more clinically applicable as they help gauge an individual's post-test probability of disease and can be more intuitive to a patient (i.e. "you are x times more/less likely…"). In this analysis, positive likelihood ratios (LR+) were all above one and all negative likelihood ratios (LR-) were below one, indicating statistical significance in relation to a null hypothesis of LR=1 (Grimes and Schulz 2005). LR+ is more relevant to this modelling as the aim is to diagnose according to a reference standard. LR- are considered more applicable if the aim is to 'rule out' a diagnosis or condition (Deeks and Altman 2004). For Model one reference standard, five variables had a LR+ within the range considered to represent "small" increases in post-test probability of sciatica. Neural tension tests had the highest LR+ of 7.8, considered a moderate increase in probability of having sciatica, or put another way, those with sciatica are nearly 8 times more likely to have positive neural tension. However for Model two reference standard, the LR+ for neural tension was less than 2 (1.84; CI 1.54, 2.20). For Model two reference standard, all items had weaker post predictive probabilities, with only two predictors representing small increases in post-test probability.

Table 6.6 Diagnostic accuracy of individual history and physical assessment items to identify sciatica in patients with low back-related leg pain.

Item	Reference Standard for Model	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)	LR+ve (95% CI)	LR-ve (95% CI)	AUC (95% CI)
Subjective sensory changes	One	0.70 (0.64 ,0.75)	0.63 (0.53, 0.72)	0.85 (0.79, 0.89)	0.41 (0.34, 0.50)	1.89 (1.44, 2.46)	0.48 (0.40, 0.58)	0.66 (0.60, 0.73)
	Two	0.69 (0.62, 0.75)	0.46 (0.39, 0.53)	0.57 (0.50, 0.63)	0.59 (0.51, 0.67)	1.28 (1.09, 1.50)	0.67 (0.54, 0.84)	0.58 (0.52, 0.63)
Below knee pain	One	0.85 (0.80, 0.86)	0.72 (0.62, 0.80)	0.90 (0.86, 0.93)	0.62 (0.52, 0.70)	3.03 (2.20, 4.16)	0.21 (0.16, 0.28)	0.78 (0.72, 0.84)
	Two	0.87 (0.81, 0.91)	0.46 (0.40, 0.53)	0.62 (0.56, 0.68)	0.77 (0.68, 0.84)	1.61 (1.40, 1.85)	0.29 (0.20, 0.42)	0.66 (0.61, 0.72)
Leg pain worse than back pain	One	0.62 (0.56, 0.68)	0.85 (0.76, 0.91)	0.92 (0.88, 0.96)	0.43 (0.36, 0.50)	4.14 (2.57, 6.65)	0.45 (0.38, 0.52)	0.74 (0.68, 0.79)
	Two	0.71 (0.64, 0.77)	0.71 (0.64, 0.77)	0.72 (0.65, 0.78)	0.71 (0.64, 0.77)	2.47 (1.94, 3.14)	0.41 (0.33, 0.51)	0.71 (0.66, 0.76)
Positive cough/sneeze	One	0.32 (0.27, 0.38)	0.93 (0.86, 0.97)	0.93 (0.86, 0.97)	0.32 (0.27, 0.37)	4.60 (2.20, 9.58)	0.73 (0.67, 0.79)	0.63 (0.57, 0.68)
	Two	0.40 (0.33, 0.47)	0.88 (0.83, 0.92)	0.77 (0.68, 0.85)	0.58 (0.53, 0.64)	3.35 (2.20, 5.10)	0.69 (0.61, 0.77)	0.64 (0.58, 0.69)
Neural tension tests	One	0.78 (0.73, 0.82)	0.90 (0.82, 0.95)	0.96 (0.92, 0.98)	0.58 (0.40, 0.66)	7.80 (4.32, 14.08)	0.24 (0.20,0.30)	0.84 (0.80, 0.88)
	Two	0.79 (0.72, 0.84)	0.57 (0.50, 0.64)	0.65 (0.59, 0.71)	0.72 (0.64, 0.79)	1.84 (1.54, 2.20)	0.37 (0.29, 0.49)	0.68 (0.63, 0.73)
Neurological deficit [a]	One	0.72 (0.66, 0.77)	0.84 (0.75, 0.90)	0.93 (0.89, 0.96)	0.50 (0.43, 0.58)	4.49 (2.85, 7.08)	0.33 (0.28, 0.40)	0.78 (0.73, 0.83)
	Two	0.75 (0.68, 0.80)	0.59 (0.52, 0.66)	0.65 (0.59, 0.71)	0.69 (0.62, 0.76)	1.84 (1.52, 2.22)	0.43 (0.34, 0.55)	0.67 (0.62, 0.72)

CI, confidence interval; PPV, positive predictive value; NPV, negative predictive value; LR+ve, positive likelihood ratio; LR-ve, negative likelihood ratio; AUC, area under the receiver operating characteristic curve.

Model one reference standard: High confidence (≥ 80%) clinical diagnosis of sciatica.

Model two reference standard: High confidence (≥ 80%) clinical diagnosis of sciatica plus confirmatory MRI.

[a] n=394 (missing data on n=1).

6.3.5 Univariable analysis

In the univariable analysis, all predictor items were significantly associated with both reference standard outcomes ($p<0.001$) (table 6.7). The odds ratios for the individual items in Model one were all higher than Model two. The greatest strength of association with Model one reference standard was 'positive neural tension tests' with very high odds ratios of 31.85 and a wide confidence interval (15.67, 64.71). For Model two reference standard, 'leg pain worse than back pain' had the highest association with the reference standard diagnosis (6.08; CI 3.93, 9.39).

Table 6.7 Univariable associations between candidate predictors and sciatica for both reference standards

Items	Model One OR (95% CI)	Model Two OR (95% CI)
Subjective sensory changes	3.94 (2.45, 6.34)	1.91 (1.26, 2.88)
Below knee pain	14.29 (8.33, 24.51)	5.50 (3.35, 9.00)
Leg pain worse than back pain	9.26 (5.10, 16.82)	6.08 (3.93, 9.39)
Leg pain intensity	1.52 (1.35, 1.70)	1.39 (1.26, 1.53)
Positive cough / sneeze	6.31 (2.82, 14.13)	4.88 (2.91, 8.21)
Neural tension tests [a]	31.85 (15.67, 64.71)	4.93 (3.17, 7.66)
Neurological deficit [b]	13.41 (7.42, 24.23)	4.29 (2.80, 6.58)

OR, Odds ratios; CI, confidence interval
Model one reference standard high confidence ≥80% clinical diagnosis.
Model two reference standard high confidence ≥80% clinical diagnosis plus confirmatory diagnosis.
[a] Positive straight leg raise or slump or femoral nerve stretch.
[b] Myotome or reflex or sensory deficit.

6.3.6 Collinearity

Correlation was evident among most predictors with the exception of subjective sensory changes with leg pain worse than back pain, and positive cough sneeze with neurological deficit. Collinearity diagnostics on all seven predictors showed that the VIF factors were below 10, ranging between 1.15 and 1.37. This confirmed that predictor coefficients were not changed considerably due to high correlations amongst predictors and all predictors were eligible for use in the multivariable model.

6.3.7 Missing data

There was missing data on one patient for the neurological deficit predictor (the clinical assessor documented they were unable to assess myotomes due to severe pain). Imputation for the multivariable model was not necessary because there was only one case of missing data and it was considered a case of "missing completely at random"[14]. Omitting this patient's data from the analysis was not considered to pose a threat of selection bias (Sterne et al. 2009).

6.3.8 Model specification

Multivariable modelling with backward predictor selection was performed on complete data from 394 participants, using the seven included predictor variables. Model one, with clinical diagnosis as the reference standard, produced a final model with five significantly independent items (at the level $p<0.05$). Positive cough/sneeze and intensity of leg pain were eliminated. Six items were retained in Model two which included MRI in the reference standard. Subjective sensory changes was the only predictor eliminated from Model two. Four items were retained in both models: below knee pain, leg pain worse than back pain, positive neural tension tests, neurological deficit (motor, reflex or sensory deficit). The odds ratios for Model one were all higher than those for Model two. The multivariable models are presented in table 6.8.

[14] "Missing completely at random"- *'there are no systematic differences between the missing values and the observed values. For example blood pressure measurements may be missing because of a breakdown of an automatic sphygmomanometer'* (Sterne et al. 2009).

Table 6.8 Multivariable associations between the clinical assessment items and sciatica for the two reference standards

	Model One		Model Two	
Item	***Beta***	***OR (95% CI)***	***Beta***	***OR (95% CI)***
Subjective sensory changes	0.98	2.66 (1.20, 5.90)	NS	NS
Below knee pain	1.83	6.25 (2.80, 13.94)	0.76	2.13 (1.19, 3.83)
Leg pain worse than back pain	1.52	4.55 (1.89, 10.99)	1.08	2.94 (1.77, 4.89)
Leg pain intensity	NS	NS	0.14	1.15 (1.03, 1.29)
Positive cough / sneeze	NS	NS	0.92	2.50 (1.34, 4.65)
Neural Tension tests [a]	3.07	21.63 (9.00,51.97)	0.56	1.76 (1.03, 3.00)
Neurological deficit [b]	2.14	8.50 (3.80,19.01)	1.04	2.81 (1.69, 4.69)
Intercept	***-3.25***		***-2.98***	
ROC AUC		***0.95 (0.93, 0.98)***		***0.82 (0.78, 0.86)***

OR, odds ratios; CI, confidence intervals; ROC AUC, area under the receiver operating characteristic curve; NS, non-significant variable.
[a] Positive straight leg raise or slump or femoral nerve stretch.
[b] Myotome or reflex or sensory deficit.

The full prediction formulae for both models are presented below which allows for prediction of the probability of sciatica for individuals (Kjupers et al. 2007).

Model one: Probability (sciatica) =

1/1+exp – [-3.25 + (subjective sensory changes x 0.98) + (below knee pain x 1.83) + (leg pain worse than back pain x 1.52) + (neural tension x 3.07) + (neurological deficit x 2.14)]

For example, in a LBLP patient with below knee pain but none of the other items, their probability of having sciatica according to Model one reference standard is:

$$= \frac{1}{1 + \text{Exp} - [-3.25 + (0 \text{ x } 0.98) + (1 \text{ x } 1.83) + (0 \text{ x } 1.52) + (0 \text{ x } 3.07) + (0 \text{ x } 2.14)]}$$

Probability = 0.195

% Probability = 19.5%.

Or in a LBLP patient with below knee pain, leg pain worse than back pain and neurological deficit:

$$= \frac{1}{1 + \text{Exp} - [-3.25 + (0 \text{ x } 0.98) + (1 \text{ x } 1.83) + (1 \text{ x } 1.52) + (0 \text{ x } 3.07) + (1 \text{ x } 2.14)]}$$

The % probability of having sciatica would be 90%.

Model two: Probability (sciatica) =

1/ 1+exp – [-2.98 + (below knee pain x 0.76) + (leg pain worse than back pain x 1.08) + (intensity leg pain x 0.14) + (positive cough/sneeze x 0.92) + (neural tension x 0.56) + (neurological deficit x 1.04)]

Using Model two formula, in a LBLP patient with, for example, below knee pain and 6/10 leg pain intensity, their probability of having sciatica would be:

$$= \frac{1}{1 + \text{Exp} - [-2.98 + (1 \text{ x } 0.76) + (0 \text{ x } 1.08) + (6 \text{ x } 0.14) + (0 \text{ x } 0.92) + (0 \text{ x } 0.56) + (0 \text{ x } 1.04)]}$$

Probability = 0.20

% Probability = 20%.

Or in a LBLP patient with pain below knee, leg pain worse than back pain, neurological deficit and leg pain intensity 6/10:

$$\frac{1}{1 + \text{Exp} - [-2.98 + (1 \text{ x } 0.76) + (1 \text{ x } 1.08) + (6 \text{ x } 0.14) + (0 \text{ x } 0.92) + (0 \text{ x } 0.56) + (1 \text{ x } 1.04)]}$$

The % probability of having sciatica would be 68%.

6.3.9 Model performance

Calibration

Figure 6.2 shows the calibration plots for both models. Perfect prediction should lie on the 45 degree best fit line for agreement with the outcome. Visually, model one is well calibrated with a slope very close to the perfect fit line. Model two is not as well calibrated as can be seen from the shape of the slope. The Hosmer and Lemeshow statistical test for the observed data for model

one supported the goodness of fit of the model (χ^2 =11.4, p=0.18) whereas Model two showed poor calibration (χ^2 =22.4 p=0.004).

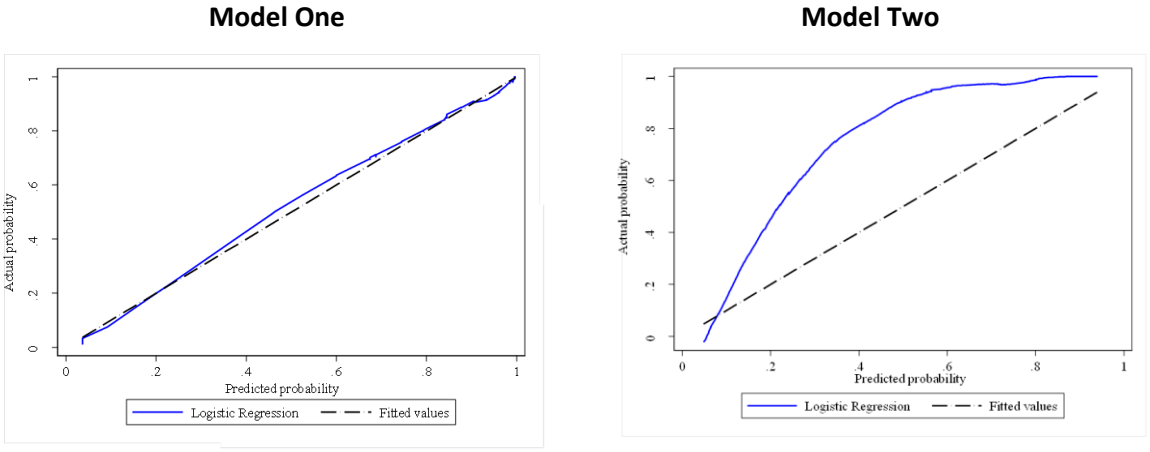


Figure 6.2 Calibration plots for Model one and Model two

Discrimination

The area under the ROC curve for Model one was 0.95 (CI 0.93, 0.98) hence the model showed an almost perfect ability to discriminate between those who do and do not have sciatica. Area under the ROC curve for Model two was 0.82 (CI 0.78, 0.86), indicating excellent discrimination (Hosmer and Lemeshow 2000). It is clear that there is greater overlap of probabilities in model two from visually inspecting the histograms (figure 6.3) of the predicted probabilities distribution. The Nagelkerke R^2 value for model one and two was 0.73 and 0.42 respectively. The greater the magnitude of the correlation between the predicted values and the observed values, the greater the R-squared. Values closer to 1 reflect a more valid tool.

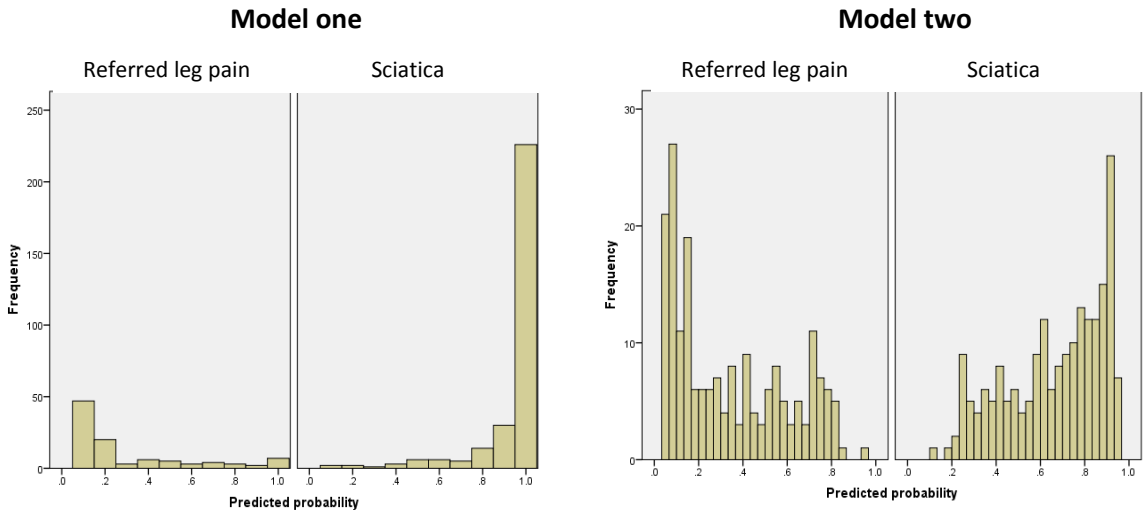


Figure 6.3 Histograms of predicted probabilities distribution for Model one and two

6.3.10 Internal validation

Minimal overfitting was seen in both models following internal validation using bootstrapping. Odds ratios and their 95% confidence intervals for the bootstrapped model are shown in table 6.9. In model two, the lower confidence interval for neural tension test dropped just below 1 (0.97). The adjusted AUCs for both models were unaltered following bootstrapping.

Table 6.9 Multivariable associations between the clinical assessment items and sciatica for the two reference standards following bootstrapping (1000 replications)

Item	Model one Multivariable OR (95% CI)	Model two Multivariable OR (95% CI)
Subjective sensory changes	2.66 (1.26, 6.20)	NS
Below knee pain	6.25 (2.94, 14.11)	2.14 (1.19, 4.36)
Leg pain worse than back pain	4.55 (1.79, 10.31)	2.95 (1.61, 4.69)
Leg pain intensity	NS	1.15 (1.03, 1.29)
Positive cough / sneeze	NS	2.51 (1.41, 4.85)
Neural tension tests [a]	21.63 (8.55, 55.51)	1.77 (0.97, 2.98)
Neurological deficit [b]	8.50 (3.51, 19.17)	2.83 (1.60, 4.76)
AUC ROC	**0.95 (0.93, 0.98)**	**0.82 (0.78, 0.86)**

OR odds ratios; CI confidence intervals; ROC AUC, area under the receiver operating characteristic curve; NS, non-significant variable.
[a] Positive straight leg raise or slump or femoral nerve stretch.
[b] Myotome or reflex or sensory deficit.

6.3.11 Scoring tool

A simple scoring method for Model one, which was the better performing model, was developed by converting the beta coefficient values into easy to use numbers. A total score of 10 could be achieved (table 6.10).

Table 6.10 Scoring tool development for Model one

Predictors	Beta	Beta/0.98	Rounded score
Subjective sensory changes	0.98	1	1
Below knee pain	1.83	1.87	2
Leg pain worse than back pain	1.52	1.55	2
Neural tension tests (Straight leg raise or slump or femoral nerve stretch)	3.07	3.13	3
Neurological deficit (myotomes or reflexes or sensation)	2.14	2.18	2

The corresponding predicted probability of sciatica for each sum score is shown in table 6.11.

Table 6.11 Sum score and corresponding predicted probability of sciatica diagnosis

Sum Score (Prevalence 75%)	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Number	36	19	19	20	21	47	29	41	61	25	76
Observed sciatica (%)	3	11	11	50	67	85	86	100	97	100	100
Mean Predicted probability of sciatica (%)	4	9	19	42	63	83	93	96	99	100	100

It could also be represented as ranges (arbitrarily chosen), corresponding to the probability of sciatica:

- Scores 0-2 low probability (< 20%)
- Score 3 or 4: moderate probability of sciatica (42-63%)
- Score 5: high probability (83%)
- Score 6 or more: very high probability (93-100%)

Using this clinical diagnostic model, (with high confidence clinical diagnosis as the reference standard) a threshold score of 5 or above suggests high likelihood of sciatica (at least 83% likely to be diagnosed with sciatica). Using coordinates from the ROC curve, at this threshold, the model has a sensitivity of 0.85 and specificity of 0.88.

6.3.12 Sensitivity analyses

Odds ratios, corresponding 95% confidence intervals and AUC for additional multivariable models are shown in table 6.12. The model was repeated for Model one reference standard (clinical diagnosis) using less restrictive confidence in diagnosis criteria (Model one (a), (b) and (c)) leading to inclusion of more of the available sample. With no restriction in diagnostic confidence (Model one (b) and (c)) the predictor 'positive cough/sneeze' remained in the final models and odds ratios for 'leg pain worse than back pain' and 'neurological deficit' were slightly lower in the larger sample sizes models. Removing patients with a diagnosis of stenosis from the analysis (Model one(d)) and comparing directly to Model one showed minimal variation to the model outcome.

For Model two reference standard, the effect of changing confidence in diagnosis had little influence on odds ratios. The larger sample size (n=552) had a slightly lower AUC. When MRI alone was the reference standard (Model three), positive neural tension tests, below knee pain and subjective sensory changes were not in the final model. This model had the lowest AUC value (0.70; CI 0.65, 0.74). The two predictors common to all nine models were leg pain worse than back pain and neurological deficit.

Using a different predictor entry selection, history and self-report items only were entered in a block and the AUC of the model was calculated. For Model one, the AUC of the model reduced from 0.95 (CI 0.93, 0.98) to 0.89 (CI 0.86, 0.93) and for Model two, the AUC of the model reduced from 0.82 (CI 0.78, 0.86) to 0.79 (CI 0.75, 0.84). This showed that model performance was very good using history and self-report items only, but improved even more with the inclusion of the physical examination items.

The performance of the models reported in this chapter, irrespective of reference standard, exceed all other multivariable models for LBLP reported in the literature (table 6.1). Reasons for this, and issues involved in comparing reported published diagnostic models, will be explored in the next section.

Table 6.12 Multivariable associations between the clinical assessment items and sciatica comparing seven additional reference standards variations

Item	**Model one** **n=394** OR (95% CI)	**Model one (a)** **n=447** OR (95% CI)	**Model one (b)** **n=552** OR (95% CI)	**Model one (c)** **n=607** OR (95% CI)	**Model one (d)** **n=354** OR (95% CI)	**Model two** **n=394** OR (95% CI)	**Model two (a)** **n=447** OR (95% CI)	**Model two (b)** **n=552** OR (95% CI)	**Model three** **n=552** OR (95% CI)
Subjective sensory changes	2.66 (1.20, 5.90)	2.68 (1.32, 5.47)	2.38 (1.32, 4.29)	2.46 (1.40, 4.32)	2.80 (1.20, 6.55)	NS	NS	NS	NS
Below knee pain	6.25 (2.80, 13.94)	5.43 (2.67, 11.06)	6.05 (3.30, 11.11)	6.43 (3.60, 11.51)	6.57 (2.83, 15.23)	2.13 (1.19, 3.83)	2.08 (1.22, 3.54)	2.17 (1.35, 3.50)	NS
Leg pain worse than back pain	4.55 (1.89, 10.99)	3.43 (1.56, 7.53)	2.68 (1.39, 5.15)	2.95 (1.59, 5.47)	4.24 (1.66, 10.82)	2.94 (1.77, 4.89)	2.84 (1.77, 4.54)	2.65 (1.77, 3.99)	2.35 (1.61, 3.42)
Intensity leg pain	NS	NS	NS	NS	NS	1.15 (1.03, 1.29)	1.18 (1.06, 1.31)	1.13 (1.03, 1.23)	1.10 (1.02, 1.20)
Positive cough / sneeze	NS	NS	2.87 (1.05, 7.82)	3.02 (1.09, 8.37)	NS	2.50 (1.34, 4.65)	2.08 (1.18, 3.66)	2.00 (1.21, 3.32)	2.00 (1.25, 3.18)
Positive neural tension tests	21.63 (9.00, 51.97)	25.42 (11.1, 58.24)	20.33 (9.62, 42.98)	24.71 (11.73, 52.07)	21.12 (8.57, 52.09)	1.76 (1.03, 3.00)	2.00 (1.23, 3.28)	2.00 (1.32, 3.05)	NS
Neurological deficit	8.50 (3.80, 19.01)	7.23 (3.52, 14.84)	4.53 (2.51, 8.18)	4.63 (2.62, 8.18)	8.41 (3.58, 19.76)	2.81 (1.69, 4.69)	2.65 (1.65, 4.24)	1.94 (1.30, 2.90)	1.66 (1.15, 2.38)
AUC	0.95 (0.93, 0.98)	0.95 (0.92, 0.97)	0.93 (0.90, 0.95)	0.94 (0.91, 0.96)	0.96 (0.94, 0.98)	0.82 (0.78, 0.86)	0.82 (0.78, 0.86)	0.79 (0.75, 0.82)	0.70 (0.65, 0.74)

OR, Odds Ratio; CI, confidence intervals; AUC, area under the receiver operating characteristic curve; NS, non-significant variable.
Model one: Confidence ≥80% sciatica clinical diagnosis (and had an MRI scan) n= 394.
Model one (a): Confidence > 70% sciatica clinical diagnosis (and had an MRI scan) n= 447.
Model one (b): Sciatica clinical diagnosis, all ranges of confidence (and had an MRI scan) n=552.
Model one (c): Sciatica clinical diagnosis, all ranges of confidence n=607.
Model one (d): Sciatica clinical diagnosis, all ranges of confidence, excluding stenosis patients (and had an MRI scan) n=354.
Model two: Confidence ≥80% sciatica clinical diagnosis plus confirmatory MRI n= 394.
Model two (a): Confidence > 70% sciatica clinical diagnosis plus confirmatory MRI n= 447.
Model two (b): Sciatica clinical diagnosis, all rages of confidence plus confirmatory MRI n= 552.
Model three: Confirmatory MRI only n= 552.

6.4 Discussion

The aim of the study presented in this chapter was to explore the contribution of items from clinical assessment that best identify sciatica in primary care consulters with LBLP. The results show that a distinct cluster of items identify sciatica.

6.4.1 Summary of main findings

In the absence of a gold standard for diagnosing sciatica, two reference standards were used. Model one, using the high confidence in clinician diagnosis reference standard, retained five items and Model two, with the addition of confirmatory MRI in the reference standard, retained six items. Four items remained in both models: below knee pain, leg pain worse than back pain, positive neural tension tests, neurological deficit (myotome, reflex or sensory). Model one (clinical diagnosis) was well calibrated (p=0.18) with "almost perfect" discrimination of AUC 0.95 (CI 0.93, 0.98). Performance measures for model two (clinical diagnosis plus confirmatory MRI) showed good discrimination (AUC 0.82; CI 0.78, 0.86) but poor calibration (p=0.004). Bootstrapping revealed minimal overfitting in both models.

6.4.2 Comparison to other studies

The predictors that were retained in both models were unsurprising from a clinical perspective. Pain below the knee is suggested a proxy for sciatica (Dionne et al. 2008). Leg pain worse than back pain performed strongly in univariable and multivariable analyses in both models. In Model two, it was the item with the strongest association with sciatica (OR 2.94; CI 1.77, 4.89). Although in everyday clinical practice it is undoubtedly used to guide diagnosis, it has perhaps received less attention in the literature for diagnosing or selecting patients with sciatica for relevant studies. An example being a recently completed RCT evaluating medication (pregabalin) for sciatica which does not includes leg pain worse than back pain as an inclusion criterion to identify sciatica patients (Mathieson et al. 2013).

Neural tension test remained in both models with considerable difference in the magnitude of the odds ratios between the two models. In Model one; almost all patients who had a positive SLR were diagnosed by clinicians as having sciatica. Hence in the multivariable model it had a very high odds ratio of 21.63 (CI 9.51, 51.97). However when MRI is included in the reference standard (Model two) considerably less weight was seen in the association between positive neural tension and sciatica as the odds ratio in Model two was much lower, (OR 1.76; CI 1.03, 3.00). Neurological signs of myotome, reflex or sensory deficit were associated with sciatica in both models. The predictors not included in the multivariable Model one were; positive cough/sneeze and intensity of leg pain. Positive cough/sneeze did have high univariable association with the outcome for Model one, but when combined with the other predictors it lost its significance. Results from the multivariable Model two did not include subjective descriptors of pins and needles/numbness.

Beattie et al. (2000) found pain below the knee was associated with severe nerve root compression in their diagnostic model. Self-report symptoms of weakness or numbness showed minimal association with MRI findings, which was similar to the findings for Model two. Paraesthesia showed weak association with MRI finding of stenosis or multilevel nerve root compression. Using self-report items from a primary care population of LBLP patients, Konstantinou et al. (2012c) found three items (pain below knee, leg pain worse than back pain and feeling of numbness or pins and needles in the leg) associated with a clinical diagnosis of sciatica. However because of their two models' performance (AUC 0.72 for only definite cases of sciatica; AUC 0.74 for definite and possible cases of sciatica indicated by clinical diagnosis) the authors concluded that clinical assessment is needed to identify sciatica when accurate case definition is important.

Using six history items as predictors, Verwoerd et al. (2014) found gender and sensory loss remained significant predictors in their model, but performance was poor with an AUC of 0.65 (CI 0.58, 0.71) with nerve root compression on MRI as the reference standard (AUC 0.66; CI 0.58, 0.74) for disc herniation on MRI). Vroomen et al. (2002) used MRI findings (of nerve root

compression due to disc or stenosis) as the reference standard and found two patient characteristics (age and gender), four history items (spasmodic pain, pain worse in leg than back, pain in a dermatomal distribution and positive cough/sneeze) and two physical examination items (restricted forward bending and myotomal muscle weakness), were associated with nerve root compression on MRI. Items were entered in blocks to reflect the sequence of a clinical examination with performance evaluated at each stage. Before physical examination, the AUC of the model was 0.8 and improved only slightly to 0.83 when information from physical examination was added. The authors concluded that much of the diagnostic information revealed by physical examination had already been revealed by history items. Because of these results, for this thesis a sensitivity analysis was carried out to compare the diagnostic models' performance using just history items before adding physical items. Results showed that adding the physical examination items to the history items improved the diagnostic performance of both models.

Verwoerd et al. (2014) chose to validate the history items used in the model published by Vroomen et al. (2002) in their dataset, which resulted in a much lower AUC of 0.58. The authors suggested that omitting one of the original model variables and the difference in patient setting and selection could explain the difference in model performances. The results do however highlight the relevance of validating a model in an external population as models will invariably perform better in their development dataset (Steyerberg 2009).

The model in this thesis did not include patient demographics. These were not considered important to differentiate leg pain amongst experts in the field of LBP (Konstantinou et al. 2012b) and are more likely to be informative when considering prognosis as opposed to diagnosis (Haugen et al. 2012). The predictor of paroxysmal spasm was not a question included in the ATLAS study clinical assessment as it was not an item considered relevant based on findings from a Delphi study (Konstantinou et al. 2012b). In lieu of "dermatomal distribution", the variable "pain below the knee" was used which includes dermatomal distribution of lower nerve roots (L4, L5, S1). SLR was not predictive of nerve root compression in Vroomen et al.'s (2002) model which was

considered a surprising finding. However restricted finger to floor distance was a positive predictor and the authors point out that the action of bending over to touch toes stretches the spinal nerve roots and sciatic nerve. In the model presented in this thesis, neural tension tests which included SLR, was a strong predictor in the clinical diagnosis model, but much less so when MRI was included in the reference standard.

The reason why MRI findings are not associated with positive SLR findings in the combined model (Model two) or using MRI alone as a reference standard (Model three) merits further exploration, especially considering the emphasis that is given in the literature to positive neural tension tests as a diagnostic criterion for sciatica. Amongst the twenty two LBLP classification systems identified in chapter four, seventeen of them included positive neural tension tests in their clinical criteria for sciatica (table 4.9, page 76). It is suggested that neural tension tests may cause pain due to chemical mediators e.g. substance P irritating a nerve root but not generating detectable signal on MRI (Beattie et al. 2000).

In the earliest published diagnostic model which used myelography as the reference standard (Vucetic et al. 1999), six predictors remained in their final model. Incapacitating pain had the largest odds ratio and was defined as patients mostly bedridden because of pain and in need of constant analgesia. This predictor reflects the more severe population of the study who were selected for surgery.

In Coster et al.'s (2010) study, after multivariable analysis, positive SLR, positive cough/sneeze, dermatomal distribution and ongoing denervation from EMG readings (indicating ongoing NRI) were significant predictors of nerve root compression using MRI as a reference standard. Despite the use of MRI to diagnose the condition of interest, SLR remained in the model which is in contrast to Vroomen et al.'s (2002) findings. However, the neuroradiologists who reported the MRIs, had clinical information about the patients, which could be a source of diagnostic bias. The

authors did not report any performance measures for their model hence making it difficult to make comparisons to other models.

Considerable variations in reporting standards are seen in the previously published literature. Few mention the calibration of their models and the recommended ways of displaying calibration and discrimination graphically were not often seen. None of the studies attempted to create a scoring system or tool, most likely due to the poor performance of most of the models. Internal validation was reported in one model (Verwoerd et al. 2014). The report in this thesis has followed the recently published guidelines from the TRIPOD group (Moons et al. 2015). A greater uptake of these reporting standards may be seen in the future.

6.4.3 Strength and limitations of the study

As there is no gold standard for diagnosing sciatica, selection of a reference standard is always a challenge and can be subject to criticism. In this study, for Model one, expert clinical opinion was chosen as a reference standard, which is considered by some epidemiologists, as appropriate in the development of diagnostic criteria in the absence of a gold standard (Coggon et al. 2005). It also reflects current practice in primary care when in the majority of cases, diagnosis and initial management plans are put into place without access to imaging, at least initially.

However, because the clinicians unavoidably were using information from the assessment predictor variables to make their diagnosis, this contributes to incorporation bias. Ideally the reference standard and the predictors should be independent of one another to avoid inflation of accuracy estimates (Worster and Carpenter 2008, Reitsma et al. 2009). The cross-tabulated frequency table for Model one did suggest incorporation bias because clinicians diagnosed most patients as having sciatica if any of the history or physical assessment items were positive.

A second reference standard was chosen which combined confirmatory MRI findings with the high confidence clinical diagnosis, in order to address to some extent the issue of incorporation bias. Prior to performing the multivariable analysis, it was obvious from examining the 2x2 tables

(table 6.5) and diagnostic accuracy tables (table 6.6) for individual items, that MRI and positive clinical assessment findings did not have the same level of agreement. As several studies have used MRI alone as the reference standard, a sensitivity analysis using MRI only as a reference standard was performed. It had the poorest model performance (AUC 0.70) and did not retain the predictors "pain below the knee" or "positive neural tension tests". Excluding these variables, especially "pain below the knee" which is considered a crude proxy for sciatica (Dionne et al. 2008, Germon et al. 2014), does not agree with clinical opinion or evidence in the literature and reflects the mismatch seen in numerous studies between clinical presentation and MRI findings (Karppinen et al. 2001).

Alternative approaches to deal with an "imperfect reference standard" include using a combination of reference standards in a sequential manner to diagnose patients (Reitsma et al. 2009). For example firstly interpreting clinical information, then, if needed, combining this information with further diagnostic tests (e.g. MRI). Another recommended means of limiting bias with reference standard selection is the use of consensus so more than two raters agree on a diagnosis (Reitsma et al. 2009). However, both these methods can result in selection bias as the "easier to identify" cases are selected therefore losing the heterogeneity of patients seen in normal clinical life. Reliability of diagnosis was examined using a sample of 36 patients (chapter five) which led to introducing a high confidence in diagnosis criteria for the reference standards. Using the criteria and excluding patients who did not have an MRI reduced the sample size from a potential 609 participants to 395 eligible participants. The clinical characteristics for the patients eligible for the diagnostic model analysis did differ considerably from the clinical profile of the 214 patients excluded from the analysis. There was a much higher proportion of positive history and physical examination item in the group with higher confidence in diagnosis. This reflects the results in the reliability study (chapter five), where patients with more clear-cut symptoms are potentially easier to agree on, and subsequently have higher clinician confidence in diagnosis. The sensitivity analysis was performed to assess the impact of the selection criteria. Reassuringly the

model which included all participants performed very similarly to the models with fewer participants (table 6.12).

Identifying predictors to enter into a multivariable model is subject to debate, and accounts for much of the variability when comparing results from other sciatica modelling studies. Recommended methods on predictor selection have changed in recent times and univariable modelling to select significant predictors for inclusion in multivariable modelling is no longer recommended (Steyerberg 2009, Moons et al. 2015). The choice for this model was primarily based on consensus from a multidisciplinary team of 42 experts in the area of LBP who rated items from clinical assessment that contribute most to the diagnosis LBLP due to sciatica (Konstantinou et al. 2012b). A criticism of this model is that too few candidate predictors were initially selected for analysis. One of the issues identified during preparation of the variables was small or zero frequencies within 2x2 table cells so the decision to collapse the variables was made which reduced the candidate predictors from nine to seven. A variable that was not included in the analysis was patients' descriptors of leg pain, for example burning, shooting, electric type of pain. This was a question considered important in the Delphi study (Konstantinou et al. 2012b). The reason for its exclusion was because of the large amount of missing data on this item (25%).

6.4.4 Clinical implications

A simple scoring tool was developed from Model one, taking into consideration the varying magnitude of the regression coefficients. This tool could be useful to clinicians and researchers wishing to support their clinical judgement regarding the probability of whether a patient's leg pain is due to sciatica. In addition, considering the excellent performance of model one, it could potentially be recommended as an eligibility criteria tool for research studies on patients with sciatica in primary care, to enable more optimum identification of homogenous groups for research purposes.

A clinical study compared the clinical presentation of surgical candidates having nerve root decompression surgery, to the NICE guidelines (2009) for identifying nerve root compression (sciatica) (Germon et al. 2014). In the guidelines, lumbosacral nerve root compression was described as "unilateral leg pain worse than back pain, pain radiating to the foot or toes, and numbness or paraesthesia in the same distribution, which is associated with motor neurological deficit". Germon et al. (2014) concluded that the guidelines did not describe 99% of the surgical patients seen and described in their study, and posed the question *"Can specialists be provided with guidance that enables them to identify the vast majority of people with symptomatic lumbosacral root compression so that they can be managed appropriately?"*

This model is another step in the right direction to improving the assessment and timely identification of patients with sciatica. The effect on clinical practice of using this tool warrants further investigation. Future research may investigate the impact of applying this diagnostic model on referral decision and treatment options to gauge if it leads to more efficient management of LBLP patients.

6.5 Conclusion

This model used information from clinical assessment to estimate the likelihood of sciatica in patients with LBLP. This is the first study that explores the challenges of reference standard selection in identifying sciatica and compares several models with different reference standards. A clear cluster of items was found that consistently identified sciatica: pain below the knee, leg pain worse than back pain, positive neural tension and neurological deficit. External validation of this model is the next logical step.

This study did highlight the considerable challenge, implications and sources of bias inherent with reference standard selection in diagnostic research. An alternative approach is the use of a statistical technique called latent class modelling to combine multiple test results and relate the observed patterns of test results to unknown or latent categories of patients (Rutjes et al. 2007,

Reitsma et al. 2009). In latent class models the outcome of interest is not defined by clinical opinion, but is a statistical based entity. This circumvents the need for a reference standard and avoids the issue of incorporation bias. The following chapter will examine diagnostic classification of LBLP using this alternative approach.

Chapter Seven: Classification of low back-related leg pain using latent class modelling

7.1 Introduction

The previous chapter reported a diagnostic model for LBLP and highlighted the considerable challenges, implications and sources of bias inherent with reference standard selection in diagnostic modelling. An alternative method is the use of a statistical approach called latent class (LC) modelling (also known as finite mixture modelling). This method combines multiple test results and relates the observed patterns of test results to unknown or latent classes of patients (Reitsma et al. 2009). The LC modelling analysis places people into the single cluster that they are most likely to belong to, but the indicators used in the modelling procedure can be either continuous or categorical measurements (Muthen and Muthen 2010).

Mixture models aim to identify unobserved heterogeneity in a population and to find meaningful groups of patients that are similar in their responses to measured variables (Muthen 2004). The identified groups are not overtly measured, but are latent, hence the name of the technique. The observed data is used to create classes with minimal within-class variation and maximum between-class variation (Kongsted et al. 2015).

LC modelling is a form of clustering that has the advantages of: (i) providing statistical evaluation to help the researcher identify the optimum number of clusters (ii) using variables of mixed measurement (iii) allowing the use of datasets with missing data because allocation to clusters is based on probabilities and (iv) providing superior accuracy of classification by giving classification probabilities for individual patients (Magidson and Vermunt 2002).

The objective of this analysis is to identify subgroups of LBLP patients using statistical techniques that do not require the use of a reference standard which may have problems of incorporation bias.

This chapter describes the analyses in two parts. The first part mirrors the diagnostic model described in chapter six, while the second part explores the possibility of unobserved subgroups within the overall LBLP cohort.

In part one, the analysis uses LC modelling as an independent statistical approach to compare agreement between the clinical diagnosis reference standard and a statistically derived "diagnosis" (Taylor et al. 2006). There is no clinical definition of sciatica in LC modelling thus patients will be classified according to their response to the same clinical assessment items used in the diagnostic model (chapter six). Part one uses the sample of the ATLAS cohort with high confidence (≥ 80%) in diagnosis (n=395) and the same seven variables from self-report and clinical assessment, that were used in the diagnostic model (figure 7.1).

Aims specific to the analysis described in part one:

(i) Classify patients into two groups using LC modelling.

(ii) Compare the agreement between the two statistically derived groups identified by LC modelling and the clinically defined groups with and without a diagnosis of sciatica.

(iii) Explore if additional statistically derived classes within this sample can be identified.

(iv) Compare the characteristics of the statistically derived classes identified by LC modelling to the clinically defined groups with and without a diagnosis of sciatica.

In part two, the analysis aims to identify classes of LBLP using the whole ATLAS sample (n=609) and includes further available variables from self-report and clinical assessment (figure 7.1). Variables for the analysis in part two will be explained in more detail later in this chapter. This will allow further classification of LBLP patients into subgroups which may be potentially clinically relevant.

Aims specific to the analyses described in part two:

(i) Identify diagnostic classes of LBLP patients with statistically distinct characteristics, using LC modelling.

(ii) Describe these classes and determine if they differ on selected demographic, pain, physical function, psychosocial and work features, risk of persistent disability and MRI findings.

(iii) Compare the statistically derived classes to the clinically defined groups with and without a diagnosis of sciatica

7.2 Methods

7.2.1 Sample

The ATLAS cohort was used for this analysis. There are no definitive sample size specifications for LC modelling. Simulation studies suggest at least 200 participants for LC modelling with continuous variables (Nyland et al. 2007) and 300 participants for LC modelling with dichotomous variables (Swanson et al. 2012). Based on results from several simulation studies, Finch and Bronk (2011) recommends LC modelling requires samples "well into the hundreds" and to ideally aim for a sample size of 500. The two samples for the analyses in part one and part two each have at least 300 participants. Part two had a sample size of just over 600 participants.

7.2.2 Variables selected for inclusion in the latent class modelling

Part one used the seven variables from the clinical assessment that were selected for the diagnostic model analysis outlined in chapter six (leg pain intensity; subjective sensory changes; below knee pain; leg pain worse than back pain; positive cough/sneeze/strain; positive neural tension tests; neurological deficit). These variables, as previously described in chapter six (section 6.2.3, page 116), were selected based on clinical assessment items identified in the systematic review (chapter four), from expert consensus (Konstantinou et al. 2012b), and items clinicians documented in the ATLAS study clinical assessment as being important for distinguishing sciatica from referred leg pain in LBLP patients.

For the analysis in part two, twelve variables were included in the LC model (figure 7.1). There is no restriction in latent modelling on the number of variables to model the classes. Some additional items were included for the following reasons. The self-report item of back pain

intensity was included to gauge severity of back pain and how it compared to leg pain. Response to movement has been included in previous diagnostic models for LBLP (Vroomen et al. 2002) and mentioned in classification system for LBLP (table 4.9 page 76), hence the response of leg pain to both lumbar spine flexion (bending forward) and extension (bending backwards) was included in this latent model. Instead of the combined "neurological deficit" variable, the information was subdivided into reflex, myotome and sensory deficit to test whether individual neurological findings contributed to distinguishing between patients. In classification systems for LBLP, neurological deficit items were individually listed in over half of the identified systems (table 4.9, page 76). In the diagnostic model it was unfeasible to enter these variables individually due to small and zero cell count frequencies (see chapter six, section 6.2.5, page 119).

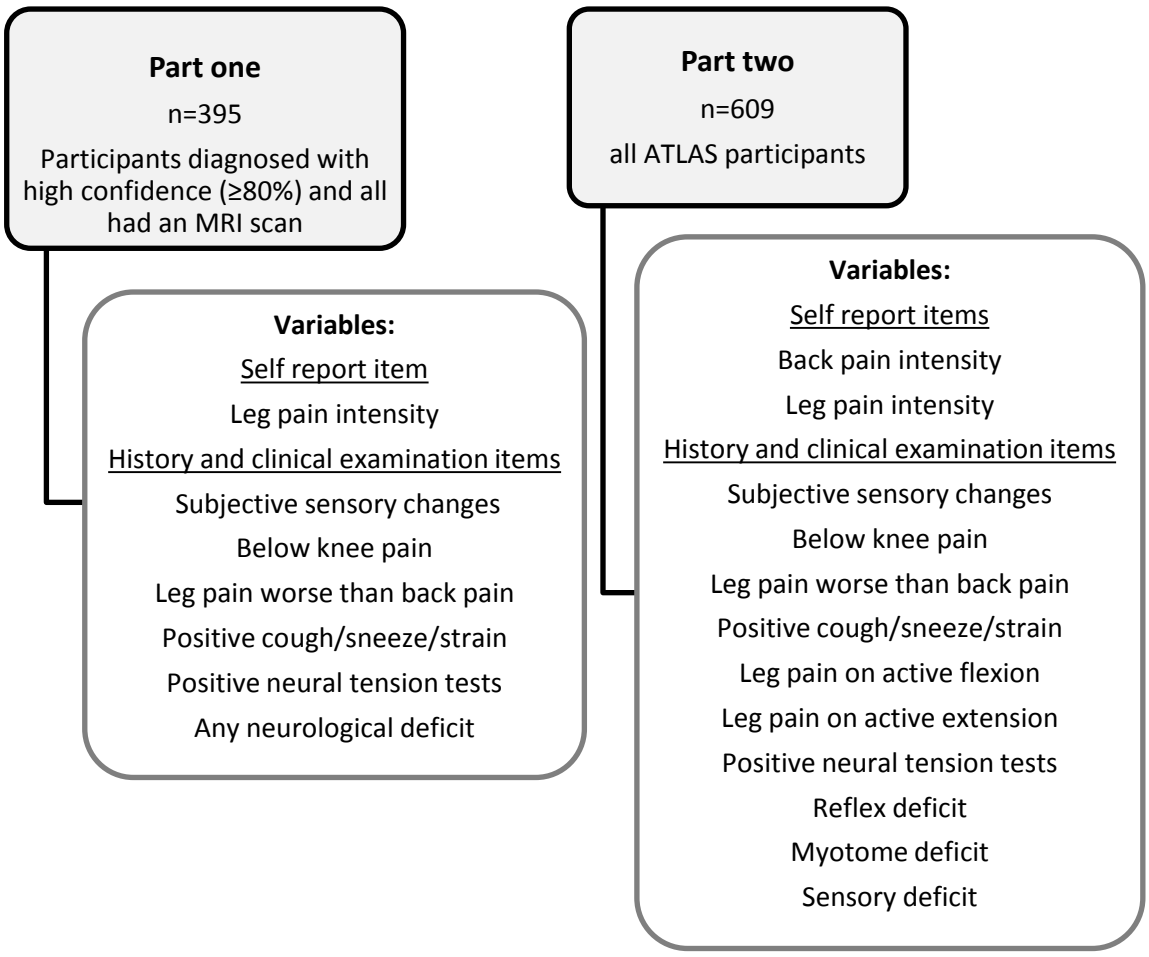


Figure 7.1 Variables for latent class modelling for Part one and Part two

7.3 Statistical analysis

7.3.1 Agreement between two latent classes (part one) and clinical diagnosis reference standard

Concordance between the high confidence clinical diagnosis groups (with and without sciatica) and the two class latent model was assessed using percentage agreements and kappa coefficients with two sided 95% CIs. Interpretation of the kappa coefficient was based on ranges from 0 to 1 as outlined in chapter five (section 5.3.3, page 90). The analysis was repeated comparing the latent classes with: (i) high confidence clinical diagnosis and confirmatory MRI findings and (ii) MRI only.

7.3.2 Latent class model development

For part one, a two solution LC model was first fitted in order to compare the two statistically derived classes to the clinically defined groups. To determine the optimum number of classes for both part one and part two, LC models were fitted consecutively starting with a two class solution and then adding another class for each successive model. The optimal number of classes was determined by a combination of the following:

(i) Goodness of fit statistics: Bayesian Information Criterion (BIC) and Akaike Information Criteria (AIC) with models of different number of classes where a lower number is optimal, and the Lo-Mendall - Rubin (LMR) adjusted likelihood ratio test (LRT) and bootstrapped parametric LRT which assesses whether adding one further class significantly improves the model fit (Nyland et al. 2007, www.ats.ucla.edu/stat/mplus accessed 08/08/2015).

(ii) Entropy: a measure of distinction and amount of overlap between the classes, ranging from 0-1 where a number closer to one is optimal (Collins and Lanza 2010).

(iii) Size of each class: at least 5% of the sample should be in each class (Yang 2006).

(iv) Class distinction: patients should be allocated to the class for which their probability is highest. Average posterior probabilities for individuals allocated to a class should exceed 0.7 which indicates clear separation (Clark et al. 2006).

(v) Clinical relevance and interpretability of each class: well-fitting models should make clinical sense and the classes should differ as may be expected on variables not used in the generation of the model (Green et al. 2015, Rathod et al. 2015). Visual inspection of the graphical output displaying the item response probabilities of the categorical values and mean pain values was planned to assist interpretation of the groups.

7.3.3 Class characteristics

The identified optimum LC solutions in part two were labelled to reflect the average pain intensity and the probabilities of a positive 'response' (range 0 to 1) to the categorical clinical assessment items entered in the LC modelling. In the context of the LC analysis, a 'response' means having an abnormal test or answering the history item question in a way which is considered to possibly be related to pathology. A probability of 1 means that patients in that class all responded "yes" to that item e.g. all had 'pain below the knee'. Probabilities closer to 0.5 reflect more ambiguity in distinguishing classes (Green et al. 2015). Agreement with clinical diagnosis was considered when labelling the identified classes.

Distribution of LBLP classes according to demographic, self-report pain, function, psychosocial characteristics, and clinical features are presented as frequencies and percentages, or mean and standard deviations, alongside the number of participants with complete data for each variable of interest. Each characteristic was compared across the number of identified classes using ANOVA for continuous variables (Kruskall Wallis test when normality and homogeneity of variance assumptions were not met) and Pearson's Chi squared test (Fisher's exact test used for cell frequencies <5) for categorical variables.

Key characteristics of the statistically derived LBLP classes were compared to the clinically defined groups with and without a diagnosis of sciatica.

The variables used to describe classes included age, gender, smoking status, BMI, work interference and time off work, back and leg pain intensity scores, neuropathic pain score, low

back and leg pain related disability, Sciatica Bothersomeness Index score, anxiety and depression, sleep disturbance, STarT Back tool score and MRI findings.

LC modelling was performed in Mplus version 5. All other analyses were performed in SPSS version 21. Analyses were two tailed and considered statistically significant if $p<0.05$.

7.4 Results part one

7.4.1 Two solution latent class model

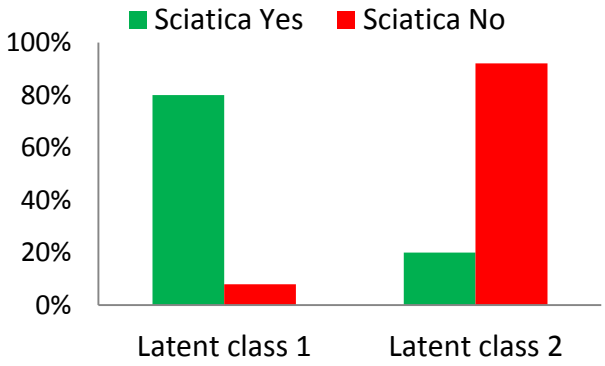
The concordance between the two class latent model solutions and the clinical diagnosis groups diagnosed with and without sciatica according to three different reference standards, is displayed in table 7.1.

The overall percentage agreement between the clinical diagnosis groups and the two latent classes was 83% with a kappa coefficient of 0.62 (CI 0.54, 0.70) considered indicative of substantial agreement (Landis and Koch 1977). Comparing the two class latent model to groups diagnosed by high confidence clinical diagnosis and confirmatory MRI findings, showed agreement of 72% and kappa 0.43 (CI 0.35, 0.52) considered indicative of moderate agreement. Comparing against MRI only, agreement was 64% with a kappa of 0.26 (CI 0.16, 0.35) which is fair agreement.

Table 7.1 Agreement between the diagnostic model reference standards and the two class latent model solution: 2x2 tables (left column) and bar charts (right column)

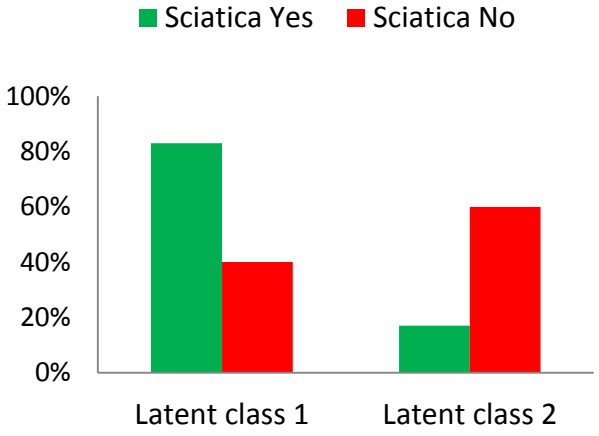
High confidence clinical diagnosis reference standard

		Sciatica diagnosis		
		Yes	No	Total
Latent class	**1**	236	8	244
	2	59	92	151
Total		295	100	395



High confidence clinical diagnosis plus confirmatory MRI reference standard

		Sciatica diagnosis		
		Yes	No	Total
Latent class	**1**	166	78	244
	2	34	117	151
	Total	200	195	395



Confirmatory MRI reference standard

		Sciatica diagnosis		
		Yes	No	Total
Latent class	**1**	167	77	244
	2	64	87	151
	Total	231	164	395

7.4.2 Model development with additional class solutions

Additional classes were added to the two class solution. The optimum LC model had three classes of LBLP patients (table 7.2). Although the entropy value was the lowest for this three class model, all other factors were optimal. The three class solution had the lowest BIC and AIC. The change in p value for the LMR adjusted LRT changed from being significant to non-significant when the solution changed from two to three classes, however the bootstrapped parametric LRT indicated that three classes were significant and an additional fourth class did not improve the LC model. Simulation studies have shown that the bootstrapped LRT is more consistent than the LMR test for identifying the correct number of classes (Nyland et al. 2007). The three class solution had sufficient numbers in each class whereas the four class solution had a group with <3% of the sample. There was a high probability of individuals in the three class solution being classified in their allocated group (all average probabilities > 0.80) (see table 7.3). Figure 7.2 gives a visual display of the item response probabilities of the categorical variables for the two, three and four class solution and their corresponding leg pain intensity.

Table 7.2 Fit indices of the latent class models of LBLP patients (n=395)

Number of classes	BIC	AIC	LMR adjusted LRT p value	Bootstrapped parametric LRT p value	Entropy	Smallest sample size [a] (%)
2	4652.13	4588.46	p=0.000	0.000	0.731	153 (38.7)
3	**4649.59**	**4554.10**	**p=0.176**	**0.000**	**0.651**	**83 (21.0)**
4	4681.54	4554.22	p=0.4322	0.6667	0.726	11 (2.8)

AIC, Akaike Information Criteria; BIC, Bayesian Information criteria; LMR, Lo-Mendall-Rubin; LRT likelihood ratio test.
[a] The number (proportion) of patients in the smallest class, at least 5% of sample should be in each class
The bold text indicates the model that was selected as having the optimal number of classes.

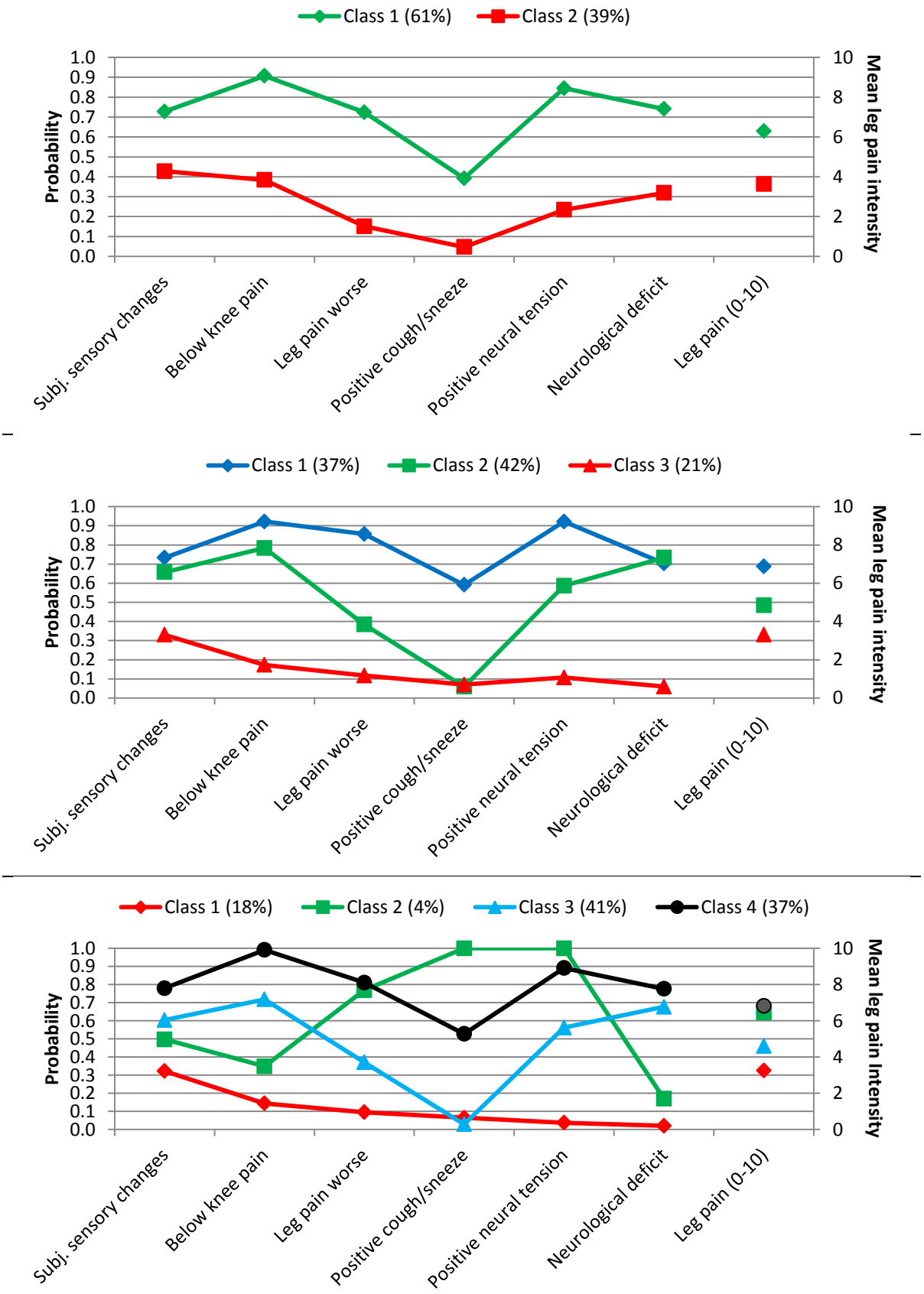


Figure 7.2 Graphs of item response probabilities of categorical variables (left vertical axis) and mean leg pain intensity (right vertical axis) for two, three and four latent class solutions

7.4.3 Class characteristics

The description of the three classes was based on the individual items response (see section 7.3.3) probabilities and their corresponding leg pain intensity displayed in table 7.3. Class 1 (n=147) had high leg pain intensity (6.9/10), and very high probability (≥0.85) of three clinical indicators being positive (leg pain worse (p=0.86), below knee pain (p=0.92), positive neural tension (p=0.92)). Class 2 (n=165) had moderate leg pain intensity (4.8/10), and high probability (>0.7) of below knee pain (p=0.78), and neurological deficit (p=0.73). Class 3 (n=83) had the lowest leg pain score (3.3/10) and low probability (≤0.33) of any clinical indicators being positive.

Table 7.3 Latent classes of LBLP: Class specific characteristics and positive item probabilities

	Class1	**Class 2**	**Class 3**
Class size based on most likely latent class membership	**147 (37.2%)**	**165 (41.8%)**	**83 (21.0%)**
Average posterior probabilities for most likely latent class membership	0.842	0.801	0.887
Mean leg pain intensity (0-10) [a]	6.88	4.83	3.31
Subjective sensory changes	0.733	0.658	0.330
Below knee pain	0.922	0.784	0.173
Leg pain worse than back pain	0.857	0.384	0.117
Positive cough/sneeze/strain	0.591	0.059	0.070
Positive neural tension tests	0.922	0.587	0.107
Neurological deficit	0.703	0.733	0.060

[a] Leg pain intensity measured using the mean of three 0 to 10 numerical rating scales for least and usual leg pain intensity over the previous two weeks and current leg pain intensity.

To explore any differences between classes, the three identified latent classes were compared on demographic, self-report and clinical characteristics (Table 7.4 a,b,c). In class 3 a greater proportion of patients were of higher socio-economic status. Otherwise there were no significant differences in demographic and work status characteristics between the three classes (table 7.4 a).

Table 7.4(a) Socio-demographics characteristics of the three latent classes

Socio-demographics (Denominator)[a]	Class 1 (n=147)	Class 2 (n=165)	Class 3 (n=83)	p value[Δ]
Age (years) mean (SD)	50.2 (13.1)	50.1 (14.2)	48.7 (14.7)	0.693
Gender, Female	82 (55.8)	100 (60.6)	55 (66.3)	0.291
Current smoker	59 (40.1)	50 (30.3)	19 (22.9)	0.210
BMI categories (393)				0.302
Normal/underweight	32 (21.8)	36 (22.0)	19 (23.2)	
Overweight	54 (36.7)	55 (33.5)	37 (45.1)	
Obese/morbidly obese	61 (41.5)	73 (44.5)	26 (31.7)	
Socioeconomic status (387)				0.010
Higher	27 (18.0)	33 (20.4)	30 (36.6)	
Intermediate	45 (31.5)	34 (21.0)	20 (24.4)	
Routine	64 (44.8)	88 (54.3)	31 (37.8)	
Never worked/long term unemployed	7 (4.9)	7 (4.3)	1 (1.2)	
Currently in paid job (392)	96 (65.8)	96 (58.5)	54 (65.9)	0.342
Self-certified time off work or sick note due to current episode (100)	40 (27.2)	38 (23.0)	22 (26.5)	0.348

SD, standard deviation; BMI, body mass index.
All figures are frequencies (percentages) unless otherwise stated as mean (SD).
[Δ] Significance p-value (α=0.05) for the difference between patients in the three latent classes on ANOVA for continuous variables (Kruskill Wallis for variable BMI) and Chi squared test for categorical variables.
[a] Denominator varies for some participants due to missing data or non-applicable cases

It can be seen from table 7.4(b) that disability and pain were associated with class membership; patients in class 1 had the highest levels of disability and pain and scored highest on the Sciatica Bothersomeness Index. Class 3 were least affected. Class 1 and 2 had higher proportions of

patients with possible neuropathic pain (as defined by s-LANSS), but differences between classes 1 and 2 were minimal. In class 1, 2.8 % of the patients were categorised low risk on the STarT Back tool, compared to 19% and 22.8% in classes 2 and 3 respectively.

There was no difference between classes on anxiety measures but class 1 had more probable cases of depression. Pain self-efficacy and quality of life scores were lower for class 1 and patients in class 1 had the highest proportion of patients with reported sleep disturbance (table 7.4 b).

Table 7.4(b) Self-report physical, psychological and health measures of the three latent classes

Self-report physical, psychological and health measures (Denominator [a])	**Class 1 (n=147)**	**Class 2 (n=165)**	**Class 3 (n=83)**	**p value[Δ]**
RMDQ disability score (0-23) mean (SD)	14.8 (4.7)	12.0 (5.9)	10.9 (5.7)	<0.001
Back pain intensity, mean (SD)	6.3 (2.2)	5.2 (2.0)	4.9 (2.0)	<0.001
Leg pain intensity, mean (SD)	7.1 (1.8)	4.7 (2.0)	3.3 (1.9)	<0.001
Duration (376)				
Back pain <6 weeks	48 (32.9)	58 (35.4)	26 (31.3)	0.796
Back pain>6 +weeks	98 (67.1)	106 (64.6)	57 (68.7)	
Leg pain < 6 weeks	58 (40.6)	70 (45.5)	37 (46.8)	0.584
Leg pain > 6+ weeks	85 (59.4)	84 (54.5)	42 (53.2)	
Sciatica Bothersomeness Index (SBI)	17.2 (4.4)	13.1 (4.8)	10.8 (5.3)	<0.001
(0-24), mean (SD)				
s-LANSS, neuropathic pain score (≥12) predominantly neuropathic	83 (56.8)	85 (51.5)	26 (31.7)	0.001
STarT Back subgroup				<0.001
Low risk	4 (2.9)	31 (19.0)	18 (22.8)	
Medium risk	71 (51.4)	74 (45.4)	35 (44.3)	
High risk	63 (45.7)	58 (35.6)	26 (32.9)	
HADS Anxiety (394) subscale				0.532
Normal	71 (48.6)	95 (57.6)	42 (50.6)	
Mild/possible	32 (21.9)	32 (19.4)	16 (19.3)	
Probable/moderate/severe	43 (29.5)	38 (23.0)	25 (30.1)	

Table 7.4(b) Self-report physical, psychological and health measures of the three latent classes

Self-report physical, psychological and health measures (Denominator [a])	Class 1 (n=147)	Class 2 (n=165)	Class 3 (n=83)	p value[Δ]
HADS Depression subscale				0.036
Normal	81 (55.1)	111 (67.3)	55 (66.3)	
Mild/possible	28 (19.0)	33 (20.0)	16 (19.3)	
Probable/moderate/severe	38 (25.9)	21 (12.7)	12 (14.5)	
Pain self-efficacy score (0-60), mean (SD)	29.2 (14.3)	35.9 (14.5)	38.8 (13.3)	<0.001
IPQ-R				
Timeline [b]	81 (55.1)	94 (57.0)	46 (55.4)	0.941
Personal control [c] (393)	88 (60.7)	108 (65.5)	49 (59.0)	0.539
Identity score (0-7) (377), mean (SD)	5.9 (1.2)	5.9 (1.4)	5.7 (1.4)	0.359
Co-morbidities				0.233
None	95 (64.6)	100 (60.6)	55 (66.3)	
One other health problem	40 (27.2)	47 (28.5)	15 (18.1)	
Two or more other health problems	12 (8.2)	18 (10.9)	13 (15.7)	
EQ—5D summary index	0.32 (0.3)	0.51 (0.3)	0.51 (0.3)	<0.001
General Health (394)				0.927
Excellent/very good/good	95 (65.2)	107 (64.8)	49 (59.1)	
Fair/poor	51(34.8)	58 (35.2)	34 (40.9)	
Sleep Disturbance [d] (yes)	115 (78.2)	112 (67.9)	50 (60.2)	0.012

RMDQ, Roland Morris Disability Questionnaire; SD, standard deviation; s-LANSS, self-report Leeds Assessment of Neuropathic Symptoms and Signs; HADS, Hospital Anxiety and Depression Scale; IPQ-R Revised Illness Perceptions Questionnaire.
All figures are frequencies (percentages) unless otherwise stated as mean (SD).
[Δ] Significance p-value *(α=0.05)* for the difference between patients in the three latent classes on ANOVA for continuous variables (Kruskill wallis for variable EQ-5D) and Chi squared test for categorical variables.
[a] Denominator varies for some participants due to missing data or non applicable cases.
[b] "My back/leg pain will last for a long time" (agree or strongly agree).
[c] "What I do can determine whether my back/leg pain gets better" (agree or strongly agree).
[d] Question on back and/or leg pain associated sleep disturbance was asked during the clinical assessment.

As expected, based on item probabilities for class membership, class 1 had the highest proportion of patients who responded positively to the clinical assessment history items. For physical examination items, class 3's response to neurological testing was almost 100% normal. Class 2 and 1 had reasonably similar proportions of patients with neurological deficits. The majority of patients in class 1 had a positive SLR compared to only half in class 2 and 7% in class 3 (table 7.4 c)

Table 7.4(c) Clinical characteristics of the three latent classes

Items from clinical assessment (Denominator[a])	CLass 1 (n=147)	Class 2 (n=165)	Group 3 (n=83)	p value[Δ]
History items				
Positive cough/ sneeze/strain	90 (61.2)	6 (3.6)	6 (7.2)	<0.001
Below knee pain	130 (90.3)	105 (68.2)	24 (30.4)	<0.001
Leg pain worse than back pain	130 (88.4)	59 (35.8)	9 (10.8)	<0.001
Subjective sensory changes	109 (74.1)	109 (66.1)	25 (30.1)	<0.001
Physical examination items				
Myotomes (393)				<0.001
Normal	107 (72.8)	124 (75.2)	82 (100.0)	
Mild weakness	36 (24.5)	34 (20.6)	0 (0.0)	
Severe weakness	4 (2.7)	7 (4.2)	0 (0.0)	
Reflexes				<0.001
Normal	101 (68.7)	121 (73.3)	82 (98.8)	
Slightly reduced	12 (8.2)	12 (7.3)	0 (0.0)	
Absent/significantly reduced	34 (23.2)	32 (19.3)	1 (1.2)	
Sensation				<0.001
Normal	74 (50.3)	66 (40.0)	82 (98.8)	
Reduced pin prick	55 (37.4)	75 (45.5)	1 (1.2)	
Loss of pin-prick	18 (12.2)	24 (14.5)	0 (0.0)	
Allodynia	17 (11.6)	14 (8.5)	6 (7.2)	0.488
SLR Positive	131 (89.1)	84 (50.9)	6 (7.2)	<0.001
Crossover SLR	16 (10.9)	4 (2.4)	1 (1.2)	0.001
Femoral nerve stretch positive	9 (6.1)	18 (10.9)	0 (0.0)	0.005
Slump test positive	35 (23.8)	27 (16.4)	2 (2.4)	<0.001

SLR, straight leg raise.
All figures are frequencies (percentages) unless otherwise stated as mean (SD).
[Δ] Significance p-value (α=0.05) for the difference between patients in the three latent classes on ANOVA for continuous variables and Chi squared test for categorical variables.
[a]Denominator varies for some participants due to missing data.

MRI findings of nerve root involvement were seen in 77.5% of patients in class 1, 55.5 % for class 2 and 31.3% in class 3. The same proportion of patients (10%) in class I and 2 had a diagnosis of stenosis made by the clinician. Clinicians' confidence in diagnosis was highest for class 1 (table 7.4d).

Table 7.4(d) MRI and clinical diagnosis characteristics of the three latent classes

Diagnosis and confidence in diagnosis	Group 1 (n=147)	Group 2 (n=165)	Group 3 (n=83)	p value[Δ]
Clinical diagnosis sciatica	143 (97.3)	142 (86.1)	10 (12.0)	<0.001
MRI (394)				
Clear nerve root compression	90 (61.2)	65 (39.6)	16 (19.3)	<0.001
Possible nerve root compression	24 (16.3)	26 (15.9)	10 (12.0)	0.658
Clinical diagnosis n=295[a]				
Disc	99 (69.2)	82 (57.5)	3 (30.0)	
Stenosis	15 (10.5)	15 (10.6)	4 (40.0)	
Not sure	29 (20.3)	45 (31.7)	3 (30.0)	
Clinician confidence in diagnosis (80-100%) mean (SD)	90.8 (6.9)	86.6 (6.1)	87.6 (7.3)	<0.001

MRI, magnetic resonance imaging; SD, standard deviation.
[Δ] Significance p-value (α=0.05) for the difference between patients in the three latent classes on ANOVA for continuous variables and Chi squared test for categorical variables
[a] Clinicians asked to document specific diagnosis in sciatica presentations such as *disc/stenosis.* An option *"not sure*" was also available.

Concordance between the statistically derived LBLP classes and patients with and without high confidence clinical diagnosis of sciatica are highlighted in table 7.5. In class 1, 97% (143/147) of the patients had a clinical diagnosis of sciatica. In class 2, 86% had a clinical diagnosis of sciatica. In class 3, 88% had a clinical diagnosis of referred leg pain.

Table 7.5 Concordance between latent classes and clinical diagnosis groups: 2x2 table

		Latent class			Total
		1	2	3	
Sciatica diagnosis	Yes	143	142	10	295
	No	4	23	73	100
	Total	147	165	83	395

7.4.4. Comparing clinical groups and latent classes

Inspection and comparison of the key characteristics between latent class 3 and the referred pain clinical diagnosis group were very similar (Appendix J). This was unsurprising as the agreement between the groups was high (88%), hence the groups represented almost the same patients.

The two statistically derived latent classes 1 and 2 had a very high proportion of people with sciatica according to clinical diagnosis (97% and 86% for class 1 and 2 respectively). The focus will thus be on comparing key characteristics of the statistically derived classes 1 and 2 to the high confidence clinically defined subgroup of sciatica patients. Inspection of table 7.6 highlights some features that were notably different.

In every domain, mean scores or proportion of patients with a clinical diagnosis of sciatica lie between the higher values of class 1 and the lower values of class 2. The latent modelling has identified a two tier classification of sciatica based on severity of leg pain and impact on physical, health and psychosocial characteristics. The next step was to apply LC modelling to the whole cohort of LBLP patients to explore if the same three class solution was apparent or if more latent classes were identified.

Table 7.6 Comparison of key characteristics for latent classes 1 and 2 and high confidence clinical diagnosis sciatica patients

	Class 1 N=147	Class 2 N=165	Sciatica clinical diagnosis n=295
Age (years), mean (SD)	50.2 (13.1)	50.1 (14.2)	50.3 (13.9)
Gender, Female	82 (55.8)	100 (60.6)	170 (57.6)
Current smoker	59 (40.1)	50 (30.3)	103 (34.9)
BMI categories			
Normal/underweight	32 (21.8)	36 (22.0)	62 (21.1)
Overweight	54 (36.7)	55 (33.5)	104 (35.4)
Obese/morbidly obese	61 (41.5)	73 (44.5)	128 (43.5)
Self-certified time off work or sick note with current episode	40 (27.2)	38 (23.0)	76 (25.8)
Back/leg interference with work performance, mean (SD)	6.5 (2.9)	5.6 (3.1)	6.1 (2.9)
RMDQ disability score, (0-23) mean (SD)	14.8 (4.7)	12.0 (5.9)	13.3 (5.6)
Back pain intensity, (0-10) mean (SD)	6.3 (2.2)	5.2 (2.0)	5.6 (2.2)
Leg pain intensity, (0-10) mean (SD)	7.1 (1.8)	4.7 (2.0)	5.8 (2.3)
Sciatica Bothersomeness Index (SBI)(0-24), mean (SD)	17.2 (4.4)	13.1 (4.8)	15.1 (5.1)
s-LANSS, neuropathic pain score (≥12) predominantly neuropathic	83 (56.8)	85 (51.5)	159 (54.1)
EQ—5D summary index mean (SD)	0.3 (0.3)	0.5 (0.3)	0.4 (0.3)
Sleep disturbance due to back/leg pain	115 (78.2)	112 (67.9)	216 (73.2)
STarT Back subgroup			
Low risk	4 (2.9)	31 (19.0)	31 (10.9)
Medium risk	71 (51.4)	74 (45.4)	143 (50.2)
High risk	63 (45.7)	58 (35.6)	111 (38.9)
HADS Anxiety subscale			
Mild/possible	32 (21.9)	32 (19.4)	56 (19.0)
Probable/moderate/severe	43 (29.5)	38 (23.0)	76 (25.9)
HADS Depression subscale			
Mild/possible	28 (19.0)	33 (20.0)	58 (19.7)
Probable/moderate/severe	38 (25.9)	21 (12.7)	57 (19.3)
MRI findings			
Clear nerve root compression	90 (61.2)	65 (39.6)	153 (52.0)
Possible nerve root compression	24 (16.3)	26 (15.9)	47 (16.0)
Clinician confidence in diagnosis (80-100%) mean (SD))	90.8 (6.9)	86.6 (6.1)	88.8 (6.9)

SD, standard deviation; BMI, body mass index; RMDQ, Roland Morris Disability Questionnaire; s-LANSS, self-report Leeds Assessment of Neuropathic Symptoms and Signs; HADS, Hospital Anxiety and Depression Scale.
All figures are frequencies (percentages) unless otherwise stated as mean (SD).

7.5 Results: part two

7.5.1 Model development

Using data from the whole ATLAS cohort (609) LC modelling was conducted for models of two up to seven classes. Figures 7.3 to 7.7 display graphs of item response probabilities and mean leg and back pain intensity for two up to six class solutions. The seven class solution was not considered further as one group had a sample size of less than 5%. The optimum LC model had five classes of LBLP patients (table 7.7). The BIC was lowest for the five class solution and compared to two, three and four class solution, the entropy was highest. Although the LMR p value suggested the three class solution was sufficient (p=0.035), the bootstrapped LRT p value remained non-significant for all class solutions. Entropy did improve in the six class solution but the BIC was higher. With seven classes the sample size of the smallest class was below 4%.

Table 7.7 Fit indices of the latent class analysis models of LBLP patients (n=609)

Number of classes	BIC	AIC	LMR adjusted LRT p value	Bootstrapped parametric LRT p value	Entropy	Smallest sample size [a] (%)
2	12101.838	11982.719	0.0001	0.000	0.714	281 (46.3)
3	12005.723	11829.250	0.035	0.000	0.738	147 (24.1)
4	11951.353	11717.527	0.180	0.000	0.728	121 (19.9)
5	**11941.422**	**11650.242**	**0.121**	**0.000**	**0.742**	**69 (11.3)**
6	11974.379	11625.845	0.108	0.000	0.791	51 (8.4)
7	12002.221	11596.334	0.040	0.000	0.802	24 (3.9)

BIC Bayesian Information criteria; AIC Akaike Information Criteria; LMR Lo Mendall Rubin; LRT likelihood ratio test.
The bold text indicates the model selected as having the optimal number of classes.
[a] The number (proportion) of patients in the smallest class; at least 5% of sample should be in each class.

Examining the graphs (figures 7.3-7.7) depicting the probabilities to a positive response on clinical assessment items, suggests that the four class solution has more distinct classes than the five class solution. Classes 2 and 5 in the five class solution have an almost identical profile with regards to response to categorical items. There were very clear differences in pain intensity (mean leg, back pain) in classes 2 and 5, in this five class solution.

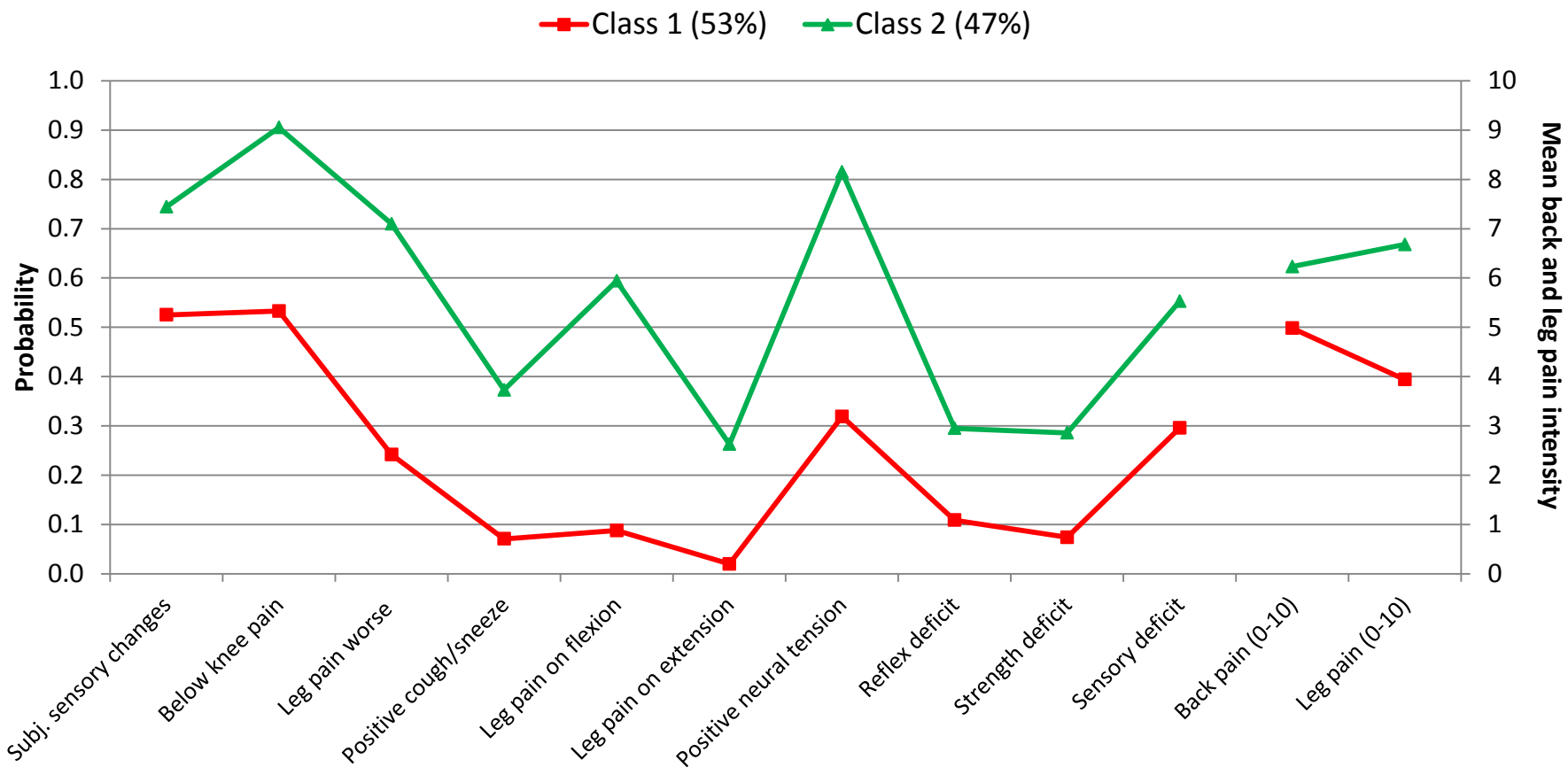


Figure 7.3 Two class solution

Item response probabilities of categorical variables (left vertical axis) and mean leg and back pain intensity (right vertical axis)

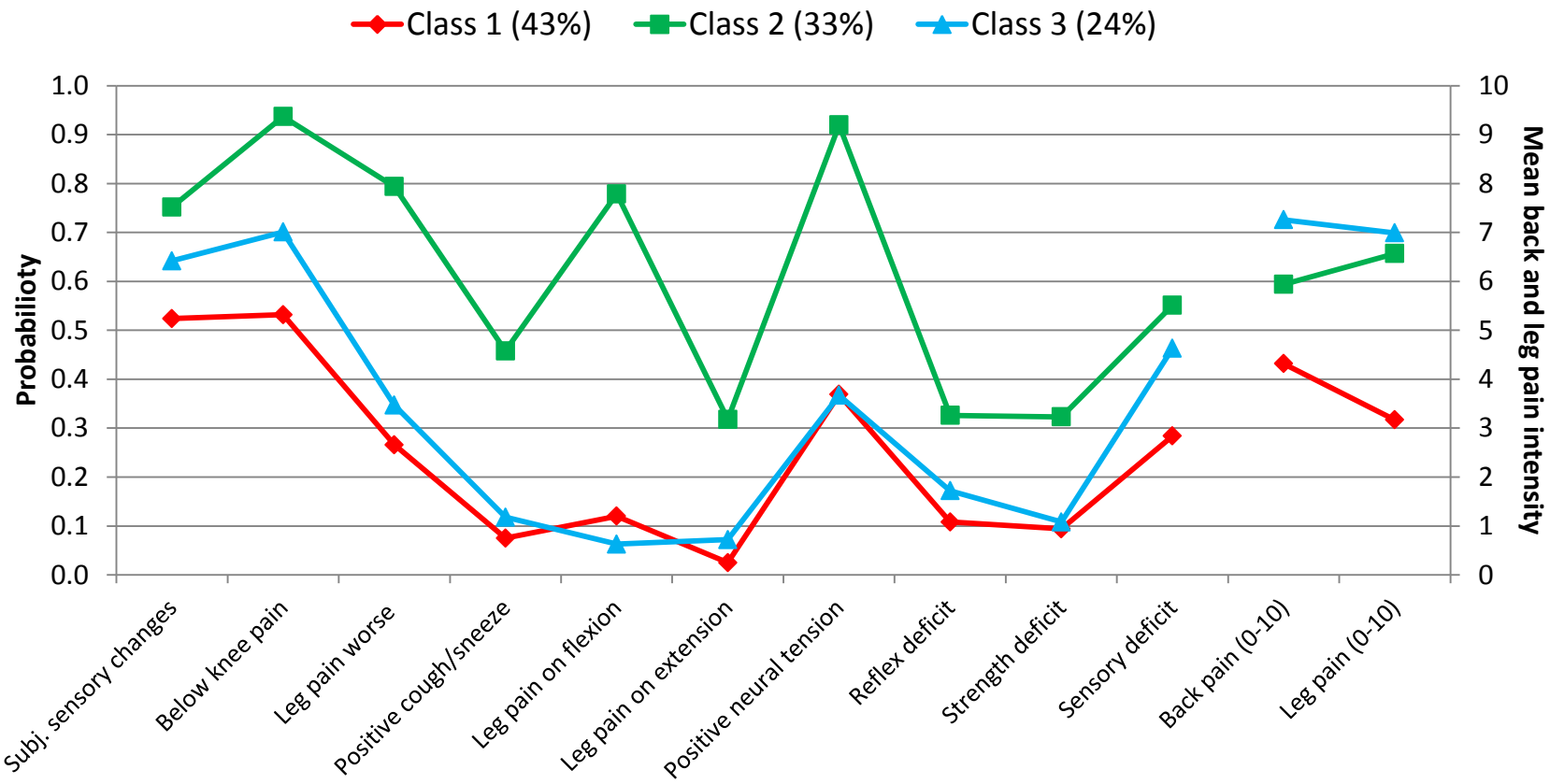


Figure 7.4 Three class solution

Item response probabilities of categorical variables (left vertical axis) and mean leg and back pain intensity (right vertical axis)

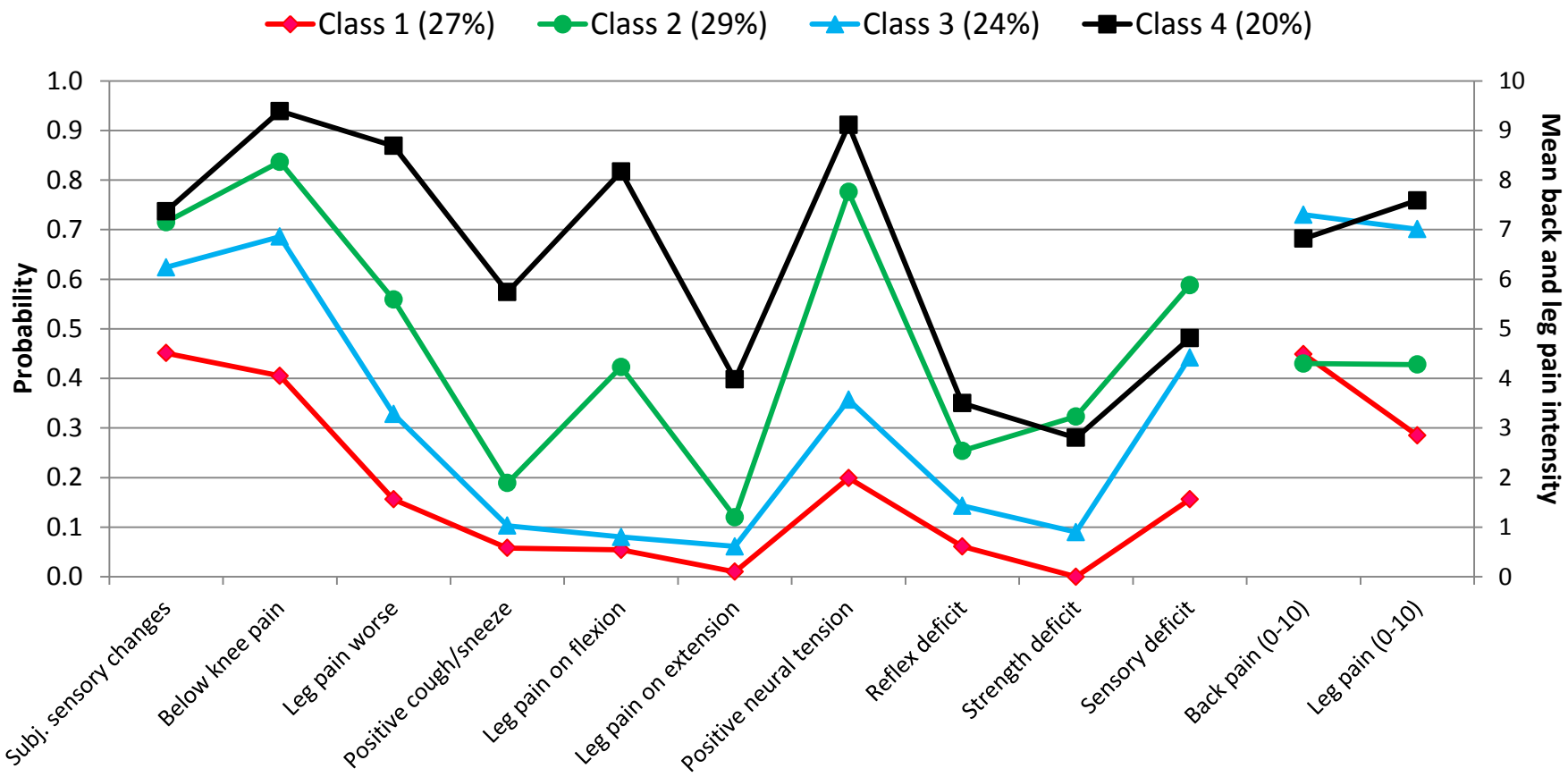


Figure 7.5 Four class solution

Item response probabilities of categorical variables (left vertical axis) and mean leg and back pain intensity (right vertical axis)

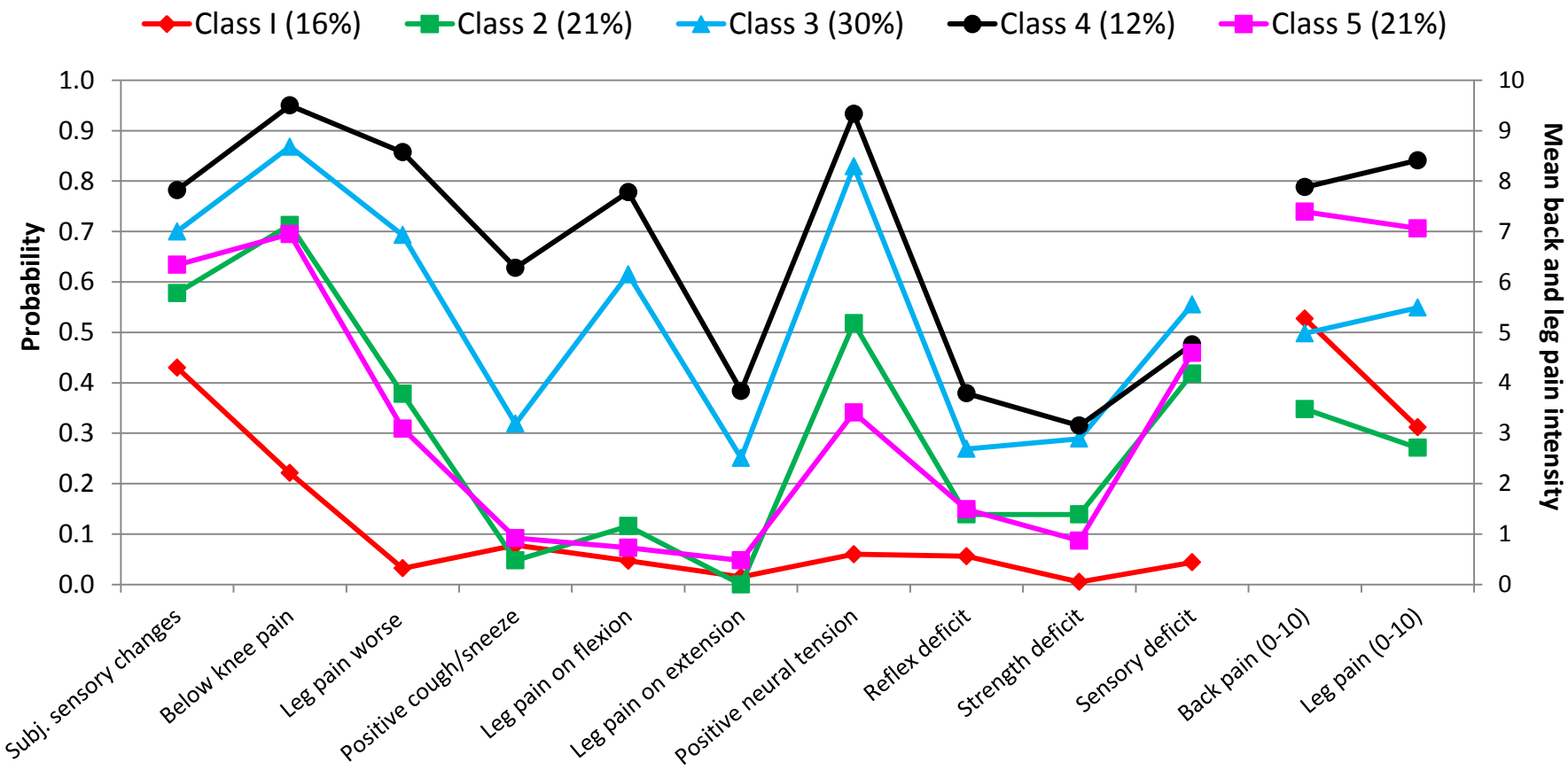


Figure 7.6 Five class solution

Item response probabilities of categorical variables (left vertical axis) and mean leg and back pain intensity (right vertical axis)

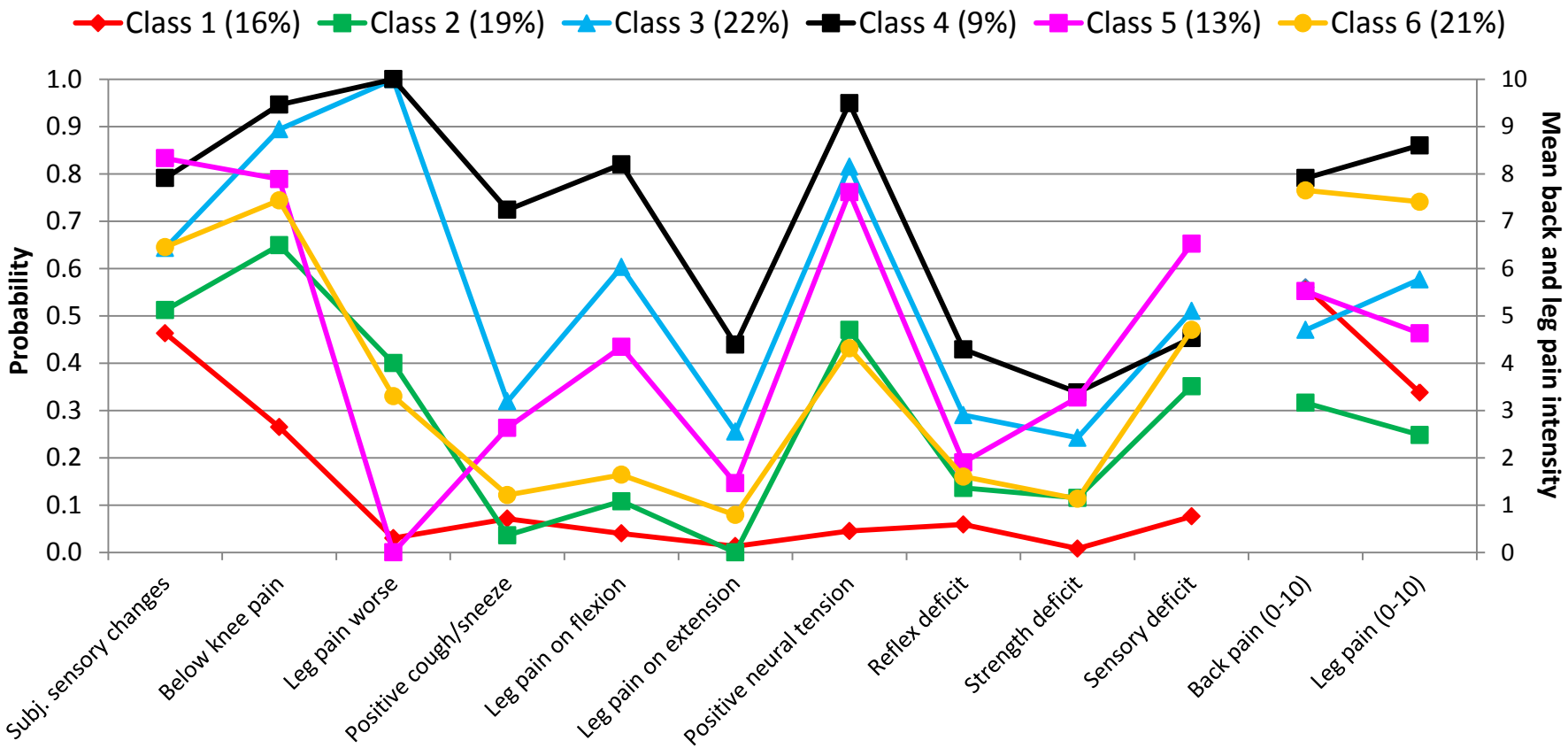


Figure 7.7 Six class solution

Item response probabilities of categorical variables (left vertical axis) and mean leg and back pain intensity (right vertical axis)

There was a high probability of individuals in the five class solution being classified in their allocated group, with all average probabilities > 0.80 (see table 7.8).

Table 7.8 Five latent classes of LBLP: Class specific characteristics and positive item probabilities

	Class1	Class 2	Class 3	Class 4	Class 5
Class size based on most likely latent class membership	**104 (17%)**	**122 (20%)**	**188 (31%)**	**69 (11%)**	**126 (21%)**
Average posterior probabilities for most likely latent class membership	0.804	0.852	0.825	0.857	0.844
Mean back pain intensity (0-10) [a]	5.27	3.47	4.98	7.88	7.39
Mean leg pain intensity (0-10) [a]	3.12	2.71	5.49	8.41	7.06
Subjective sensory changes	0.430	0.578	0.700	0.782	0.634
Below knee pain	0.221	0.713	0.868	0.950	0.695
Leg pain worse than back pain	0.032	0.378	0.693	0.857	0.309
Positive cough/sneeze/strain	0.078	0.048	0.319	0.628	0.092
Leg pain on flexion	0.047	0.166	0.615	0.778	0.073
Leg pain on extension	0.015	0.000	0.251	0.384	0.048
Positive neural tension tests	0.060	0.518	0.829	0.933	0.341
Reflex deficit	0.056	0.139	0.269	0.379	0.149
Myotome deficit	0.005	0.139	0.289	0.315	0.087
Sensory deficit	0.044	0.418	0.555	0.476	0.459

[a] Pain intensity measured using the mean of three 0 to 10 numerical rating scales for least and usual pain (back or leg) intensity over the previous two weeks and current pain intensity.

7.5.2 Description of classes

The description of the five classes was based on the individual items response probabilities and their corresponding back and leg pain intensity displayed in table 7.8. These values are also displayed graphically in figure 7.6.

Class 1: low intensity leg pain with some subjective sensory changes. All other clinical items had very low probability of being positive (≤0.22). Back pain is worse than leg pain. Agreement with clinical diagnosis: 80.8% of this group were given a clinical diagnosis of referred leg pain. Mean confidence in diagnosis: 83%. Suggested name for class 1: *Referred leg pain*.

Class 2: low intensity back and leg pain with back pain slightly worse than leg pain. High probability of below knee pain (0.7) and moderate probability of subjective (0.57) and objective (0.42) sensory deficits. Moderate probability of a positive SLR (0.52). Agreement with clinical diagnosis: 81.1% of this group were diagnosed with sciatica. Mean confidence in diagnosis: 81%. Suggested name for class 2: *Mild sciatica*.

Class 3: moderate leg pain, slightly higher leg pain than back pain intensity. Very high likelihood of below knee pain (0.86) and positive neural tension (0.83). Low probability of reflex or myotome deficit (<0.3), higher probability of sensory deficit (0.56). Agreement with clinical diagnosis: 93.1% diagnosed with sciatica. Mean confidence in diagnosis: 86%. Suggested name for class 3: *Moderate sciatica*

Class 4: High intensity back and leg pain. High probability of most clinical assessment items being positive, especially neural tension, leg pain worse than back pain and below knee pain. Highest probability among all the classes of neurological deficit and positive cough/sneeze. Agreement with clinical diagnosis: 100% diagnosed with sciatica. Mean confidence in diagnosis 90%. Suggested name for class 4: *Severe sciatica*.

Class 5: High intensity back and leg pain, back pain slightly greater than leg pain. High probability (0.7) of pain below the knee. Not likely to have positive neural tension. Strong profile of subjective and objective sensory changes compared to other groups. This class's profile is similar to class 2 but with much higher pain severity. Agreement with clinical diagnosis: 71.0% diagnosed with sciatica. Mean confidence in diagnosis 79%. Suggested name for class 5: *Atypical sciatica*

The five classes were compared on demographic, self-report and clinical characteristics (table 7.9 a). The classes did not differ significantly in age, gender or BMI. There was a greater proportion of smokers in classes 4 (severe sciatica) and 5 (atypical sciatica) and these two classes had fewer patients of higher and intermediate socioeconomic status.

Table 7.9(a) Socio-demographics characteristics of the five latent classes

Socio-demographics **Denominator** [a]	**Class 1** **104**	**Class 2** **122**	**Class 3** **188**	**Class 4** **69**	**Class 5** **126**	**p value** [Δ]
Age categories						0.238
18-34	21 (20.2)	16 (13.1)	29 (15.4)	9 (13.0)	16 (12.7)	
35-44	24 (23.1)	31 (25.4)	34 (18.1)	22 (31.9)	25 (19.8)	
45-54	32 (30.8)	25 (20.5)	50 (26.6)	14 (20.3)	31 (24.6)	
55-64	14 (13.5)	33 (27.0)	42 (22.3)	17 (24.6)	32 (25.4)	
65+	13 (12.5)	17 (13.9)	33 (17.6)	7 (10.1)	22 (17.5)	
Age (years) mean (SD)	47.2 (13.8)	50.4 (13.3)	50.9 (14.4)	49.2 (12.7)	51.9 (14.1)	0.111
Gender, Female	76 (73.1)	72 (59.0)	113 (60.1)	42 (60.9)	80 (63.5)	0.187
Current smoker	27 (26.0)	29 (23.8)	52 (27.7)	30 (43.5)	56 (44.4)	0.000
BMI (607) categories:						
Normal/Underweight	28 (26.9)	28 (23.1)	35 (18.6)	17 (24.6)	28 (22.4)	0.233
Overweight	45 (43.3)	44 (36.4)	75 (39.9)	16 (23.2)	43 (34.4)	0.107
Obese/Morbidly obese	31 (29.8)	49 (40.5)	78 (41.5)	36 (52.2)	54 (43.2)	0.056
Socioeconomic status (593)						<0.001
Higher	28 (26.9)	43 (36.1)	34 (18.6)	9 (13.8)	15 (12.3)	
Intermediate	33 (31.7)	31 (26.1)	58 (31.7)	12 (18.5)	24 (19.7)	
Routine	41 (39.4)	43 (36.1)	85 (46.4)	36 (55.4)	78 (63.9)	
Never worked/long term unemployed	2 (1.9)	2 (1.7)	6 (3.3)	8 (12.3)	5 (4.1)	
Currently in paid job (605)	71 (68.9)	79 (65.3)	121 (64.7)	37 (53.6)	59 (47.2)	0.003
Self-certified time off work (363)	25 (35.7)	20 (25.6)	42 (35.0)	11 (29.7)	8 (13.8)	0.032
Current sick note (365)	22 (31.4)	16 (20.3)	34 (28.3)	14 (37.8)	14 (16.2)	0.279

BMI, body mass index; SD, standard deviation.
All figures are frequencies (percentages) unless otherwise stated as mean (SD).
[Δ] Significance *p-value (α=0.05) for the difference between patients in the five latent classes on ANOVA for continuous variables (Kruskill Wallis for variables BMI) and Chi squared test for categorical variables (Fisher's exact test for variable socioeconomic class).*
[a]Denominator varies for some participants due to missing data or non-applicable cases

A similar trend across the five classes was seen for disability scores (RMDQ), leg pain intensity, sciatica bothersomeness and neuropathic pain scores (table 7.9 b). In ascending order of score severity was: class 2 (mild sciatica), class 1 (referred leg pain), class 3 (moderate sciatica), class 5 (atypical sciatica) and class 4 (severe sciatica). Class 5 had the greatest proportion of patients with back and leg pain for over 6 weeks. Further analysis of duration of leg pain revealed 56% of patients in class 5 had leg pain for over 3 months whilst the other classes ranged from 22% (class 2) to 38% (class 3). Nearly a quarter of patients in class five had leg pain for over one year compared to 13% and below for the other 4 classes.

The STarT Back tool grouped 69% and 64% in classes 4 and 5 respectively as being at high risk of poor prognosis in terms of disability. Only 13% of patients in class 2 were at high risk. Classes 1 and 3 had approximately one third of patients categorised as high risk. On inspection of one of the individual questions in the STarT Back tool "I have had pain in the neck or shoulder at some time in the last two weeks", a higher proportion (60%) in classes 1 and 5 had neck or shoulder pain. Anxiety and depression cases and mean scores were highest for class 4, followed by class 5. Class 2 had the lowest proportion of patients categorised as anxious or depressed. Class one (referred leg pain) had higher anxiety levels than classes 2 and 3. Pain self-efficacy was lowest for class 4 and highest for class 2. The highest proportion of patients who felt their pain was going to last a long time was in class five and the same class had the least proportion of patients who felt what they do determines if their pain gets better. A higher proportion in class five had poorer general health and two or more other health problems. EQ5D summary index was 0.13 and 0.29 for groups 4 and 5 respectively, classes 1, 2 and 3 scores ranged from 0.48 to 0.66 (class 2).

Table 7.9(b) Self report physical, psychological and health measures of the five latent classes

Self-report physical, psychological and health measures (Denominator)[a]	**Class 1** **104**	**Class 2** **122**	**Class 3** **188**	**Class 4** **69**	**Class 5** **126**	**p value [Δ]**
RMDQ disability score	11.5 (5.6)	8.6 (5.0)	12.8 (4.7)	16.7 (5.1)	15.1 (5.5)	<0.001
(0-23) mean (SD) (607)						
Back pain intensity, mean (SD)	5.3 (1.7)	3.4 (1.4)	5.0 (1.5)	8.0 (1.3)	7.5 (1.4)	<0.001
Leg pain intensity, mean (SD)	3.1 (1.4)	2.6 (1.2)	5.5 (1.3)	8.5 (1.1)	7.2 (1.4)	<0.001
Duration of symptoms						
Back pain (607)						
<6 weeks	40 (38.5)	50 (41.0)	70 (37.4)	22 (31.9)	36 (28.8)	0.279
>6 +weeks	64 (61.5)	72 (59.0)	117 (62.6)	47 (68.1)	89 (71.2)	
Leg pain (583)						
< 6 weeks	49 (49.5)	63 (54.8)	77 (42.3)	28 (42.4)	34 (28.1)	0.001
> 6+ weeks	50 (50.5)	52 (45.2)	105 (57.7)	38 (57.6)	87 (71.9)	
> 3 months	31 (31.3)	24 (20.9)	69 (37.9)	20 (30.3)	68 (56.2)	<0.001
>12 months	15 (15.2)	10 (8.7)	24 (13.2)	3 (4.5)	29 (24.0)	
Sciatica Bothersomeness Index (0-24) mean (SD) (582)	11.1 (4.9)	10.0 (4.4)	14.7 (4.0)	19.8 (3.5)	17.2 (4.4)	<0.001
S-LANSS, neuropathic pain score (≥12) (606)	37 (35.6)	44 (36.4)	100 (53.2)	45 (66.2)	67 (53.6)	<0.001
STarT Back subgroup (589)						<0.001
Low risk	17 (17.0)	44 (37.0)	16 (8.8)	0 (0.0)	5 (4.0)	
Medium risk	52 (52.0)	59 (49.6)	105 (58.0)	20 (30.8)	40 (32.3)	
High risk	31 (31.0)	16 (13.4)	60 (33.1)	45 (69.2)	79 (63.7)	
Pain in neck/shoulder in last two weeks [b]	62 (60.20	58 (47.9)	87 (46.3)	30 (43.5)	76 (60.3)	<0.001
HADS Anxiety (607) mean (SD)	8.0 (3.8)	6.1 (4.0)	7.1 (3.5)	9.8 (4.1)	9.4 (4.6)	<0.001
HADS Anxiety subscale						
Normal	47 (45.6)	84 (68.9)	115 (61.5)	23 (33.3)	47 (37.3)	<0.001
Mild/possible	24 (23.3)	18 (14.8)	38 (20.3)	13 (18.8)	27 (21.4)	
Probable/moderate/severe	32 (31.1)	20 (16.4)	34 (18.2)	33 (47.8)	52 (41.3)	
HADS Depression mean (SD)	6.1 (3.5)	4.7 (3.6)	5.8 (3.4)	8.8 (4.3)	7.7 (4.5)	<0.001
HADS Depression subscale						
Normal	68 (65.4)	98 (80.3)	134 (71.3)	29 (42.0)	63 (50.0)	<0.001
Mild/possible	24 (23.1)	15 (12.3)	33 (17.6)	14 (20.3)	33 (26.2)	
Probable/moderate/severe	12 (11.5)	9 (7.4)	21 (11.2)	26 (37.7)	30 (23.8)	

Table 7.9(b) Self report physical, psychological and health measures of the five latent classes

Self-report physical, psychological and health measures (Denominator)[a]	**Class 1** **104**	**Class 2** **122**	**Class 3** **188**	**Class 4** **69**	**Class 5** **126**	**p value [Δ]**
Pain self-efficacy score (0-60), mean (SD) (593)	37.6 (12.4)	42.9 (12.5)	34.7 (12.3)	22.5 (15.6)	28.4 (14.3)	<0.001
IPQ-R						
Timeline [c] (609)	62 (59.6)	55 (45.1)	98 (52.1)	44 (63.8)	86 (68.3)	0.002
Personal control [d] (605)	69 (66.3)	32 (73.8)	119 (63.6)	34 (50.7)	55 (44.0)	<0.001
Identity score (IPQ-R) (0-7) (584) mean (SD)	6.0 (1.3)	5.5 (1.5)	5.9 (1.2)	6.1 (1.2)	6.2 (1.1)	0.001
EQ—5D summary index (590)	0.54 (0.3)	0.66 (0.2)	0.48 (0.3)	0.13 (0.3)	0.29 (0.3)	<0.001
Co-morbidities						0.139
Two or more other health problems	16 (15.4)	15 (12.3)	21 (11.2)	5 (7.2)	23 (18.3)	
General Health (608)						
Excellent/very good/good	66 (63.5)	91 (74.5)	129 (68.6)	36 (52.9)	64 (50.8)	<0.001
Fair/poor	38 (36.5)	31 (25.5)	59 (31.4)	32 (47.1)	62 (49.2)	
Sleep Disturbance (yes) [e]	69 (66.3)	73 (59.8)	133 (70.7)	61 (88.4)	92 (73.0)	0.001

RMDQ, Roland Morris Disability Questionnaire; SD, standard deviation; s-LANSS, self-report Leeds Assessment of Neuropathic Symptoms and Signs; HADS, Hospital Anxiety and Depression Scale; IPQ-R, Illness perceptions questionnaire-short form.
All figures are frequencies (percentages) unless otherwise stated as mean (SD).
[Δ] Significance p-value *(α=0.05)* for the difference between patients in the five latent classes on ANOVA for continuous variables (Kruskill Wallis for variables HADS (depression), EQ-5D, IPQ-R) and Chi squared test for categorical variables (Fishers exact test for variable general health).
[a] Denominator varies for some participants due to missing data or non-applicable cases.
[b] Question from StarT Back tool.
[c] "My back/leg pain will last for a long time" (agree or strongly agree).
[d] "What I do can determines whether my back/leg pain gets better" (agree or strongly agree).
[e] Question on back and/or leg pain associated sleep disturbance was asked during the clinical assessment.

Table 7.9(c) reflects much of the information from table 7.8 which showed the probability of a positive response to the clinical assessment items used to model the classes. Class 1 (referred leg pain) had very low proportions of patients with positive response to most of the items, in particular neurological deficit and positive neural tension tests. The majority of patients in class 4 had positive responses to clinical assessment items. Numbers of patients documented as having movement fear avoidance (at the clinical interview/assessment) were higher in classes 4 and 5.

Half of the patients in class 5 reported feeling pessimistic about the future with regards to improvement of their LBLP

Table 7.9(c) Clinical characteristics of the five latent classes

Items from clinical assessment **Denominator** [a]	**Class 1** **104**	**Class 2** **122**	**Class 3** **188**	**Class 4** **69**	**Class 5** **126**	**p value**[Δ]
Positive cough/ sneeze/strain	6 (5.8)	5 (4.1)	61 (32.4)	46 (66.7)	11 (8.7)	<0.001
Below knee pain	22 (21.2)	91 (74.6)	163 (86.7)	66 (95.7)	88 (69.8)	<0.001
Leg pain worse than back pain	2 (1.9)	50 (41.0)	133 (70.7)	59 (85.5)	36 (28.6)	< 0.001
Subjective sensory changes i.e. pins & needles/numbness	47 (45.2)	70 (57.4)	130 (69.1)	54 (78.3)	81 (64.3)	<0.001
Evidence of:						
Fear avoidance [b]	26 (25.0)	20 (16.4)	39 (20.7)	24 (34.8)	48 (38.1)	<0.001
Distress [b]	17 (16.3)	16 (13.1)	25 (13.3)	28 (40.6)	40 (31.7)	<0.001
Work issues [b] (145)	10 (31.3)	6 (13.0)	14 (29.2)	4 (57.0)	12 (100.0)	<0.001
Low mood [b]	35 (33.7)	28 (23.1)	60 (31.9)	32 (46.4)	52 (41.3)	0.016
Depression [b]	8 (7.7)	8 (6.6)	13 (6.9)	10 (14.5)	22 (17.5)	0.025
Passive coping [b]	11 (10.7)	13 (10.7)	21 (11.2)	21 (30.4)	26 (20.6)	0.001
Pessimistic outlook [b]	40 (38.8)	29 (23.8)	54 (28.7)	21 (30.4)	63 (50.4)	<0.001
Spinal flexion increases leg pain	5 (4.8)	12 (9.8)	118 (62.8)	55 (79.7)	7 (5.6)	<0.001
Spinal extension increases leg pain	1 (1.0)	0 (0.0)	49 (26.1)	27 (39.1)	4 (3.2)	<0.001
Myotome deficit (607)	1 (1.0)	17 (13.9)	55 (29.3)	23 (33.3)	9 (7.3)	<0.001
Normal	103 (99.0)	105 (86.1)	133 (70.7)	46 (66.7)	116 (92.8)	
Mild weakness	1 (1.0)	14 (11.5)	50 (26.6)	21 (30.4)	6 (4.8)	
Severe weakness	0 (0.0)	3 (2.5)	5 (2.7)	2 (2.9)	3 (2.4)	
Reflex deficit	6 (5.8)	16 (13.1)	53 (28.2)	27 (39.1)	17 (13.5)	<0.001
Normal	98 (94.2)	106 (86.9)	135 (71.8)	42 (60.9)	109 (86.5)	
Slightly reduced	1 (1.0)	4 (3.3)	17 (9.0)	5 (7.2)	3 (2.4)	
Absent/significantly reduced	5 (4.8)	12 (9.9)	36 (19.2)	22 (31.9)	14 (11.1)	

Table 7.9(c) Clinical characteristics of the five latent classes

Items from clinical assessment Denominator [a]	Class 1 104	Class 2 122	Class 3 188	Class 4 69	Class 5 126	p value[Δ]
Sensory deficit	3 (2.9)	52 (42.6)	107 (59.6)	32 (46.4)	59 (46.8)	<0.001
Normal	101 (97.1)	70 (57.4)	81 (43.1)	37 (53.6)	67 (53.2)	<0.001
Reduced pin prick	3 (2.9)	46 (37.7)	81 (43.1)	26 (37.7)	45 (35.7)	
Loss of pin-prick	0 (0.0)	6 (4.9)	26 (13.8)	6 (8.7)	14 (11.1)	
Allodynia	9 (8.7)	13 (10.7)	17 (9.0)	10 (14.5)	9 (7.1)	0.536
Positive neural tension	3 (2.9)	70 (57.4)	157 (83.5)	65 (94.2)	40 (31.7)	<0.001
SLR Positive	2 (1.9)	64 (52.5)	137 (72.9)	61 (88.4)	33 (26.2)	<0.001
Crossover SLR	0 (0.0)	1 (0.8)	13 (6.9)	8 (11.6)	0 (0.0)	<0.001
Femoral test positive	1 (1.0)	7 (5.7)	20 (10.6)	3 (4.3)	10 (7.9)	0.024
Slump positive	0 (0.0)	23 (18.9)	37 (19.7)	14 (20.3)	10 (7.9)	<0.001

All figures are frequencies (percentages) unless otherwise stated as mean (SD).

[Δ] Significance p-value (α=0.05) for the difference between patients in the five latent classes on ANOVA for continuous variables and Chi squared test for categorical variables (Fishers exact test for variable work issues, extension increases leg pain, myotomes categorised; reflexes categorised; dermatomes categorised, SLR positive, femoral test positive, slump positive)

[a] Denominator varies for some participants due to missing data or non-applicable cases

[b] These were ascertained by clinical questioning, not from questionnaires.

A similar pattern emerges when looking at the proportion of patients in each class where clinicians were at least 80% confidence in their diagnosis (table 7.9 d). In over 90% of patients in class 4, clinicians had high confidence in diagnosis whereas class 5 had just over half of the group with a high confidence clinical diagnosis. Class 2 ('mild sciatica') had the second lowest proportion of patients diagnosed with high confidence.

Table 7.9(d) MRI and clinical diagnosis characteristics of the five latent classes

Diagnosis Denominator [a]	Class 1 104	Class 2 122	Class 3 188	Class 4 69	Class 5 126	p value*
Clinical diagnosis sciatica	20 (19.2)	99 (81.1)	175 (93.1)	69 (100.0)	89 (70.6)	<0.001
MRI (554)						
Clear or possible nerve root compression	25 (26.3)	56 (50.5)	106 (63.1)	57 (89.1)	53 (45.7)	<0.001
Clinical diagnosis (451)[b]						0.108
Disc	9 (45.0)	51 (51.5)	104 (59.4)	47 (68.1)	31 (35.2)	
Stenosis	2 (10.0)	8 (8.1)	19 (10.9)	6 (8.7)	13 (14.8)	
Not sure	9 (45.0)	40 (40.4)	52 (29.7)	16 (23.2)	44 (50.0)	
Clinician confidence in diagnosis mean (SD)	82.6 (9.7)	80.5 (10.3)	85.8 (10.1)	90.1 (8.9)	78.7 (10.3)	<0.001
Clinician confidence ≥80%	72 (69.2)	75 (61.4)	156 (83.0)	63 (91.3)	70 (55.6)	<0.001

MRI, magnetic resonance imaging; SD, standard deviation; IQR, inter-quartile range.
All figures are frequencies (percentages) unless otherwise stated as mean (SD).
Δ Significance *p-value (α=0.05) for the difference between patients in the five latent classes on ANOVA for continuous variables and Chi squared test for categorical variables*.
[a] Denominator varies for some participants due to missing data or non-applicable cases.
[b] Clinicians asked to document specific diagnosis in sciatica presentations such as *disc/ stenosis*. An option *"not sure*" was also available.

7.5.3 Comparing clinical diagnosis groups and the five latent classes

The charts in figure 7.8 and 7.9 contrast the five latent classes (coloured lines or bars) with the two clinical diagnosis groups (dotted lines). Each graph represents a key characteristic summarised in tables 7.9 (a-d) and the dotted lines represent the equivalent values for the groups diagnosed with and without sciatica (Appendix K; ATLAS baseline results table Konstantinou et al. 2015).

The 'moderate sciatica' class (class 3) and the sciatica clinical diagnosis group have very similar sociodemographic, psychosocial, pain, disability and work profiles. The referred leg pain class (class 1) has very similar scores to the clinical diagnosis group of referred leg pain patients. Class 4 (severe sciatica) and class five (atypical sciatica) show consistently higher scores on most domains. Class 2 (mild sciatica) generally has lower scores than all other classes and the clinical diagnosis 'referred leg pain' group.

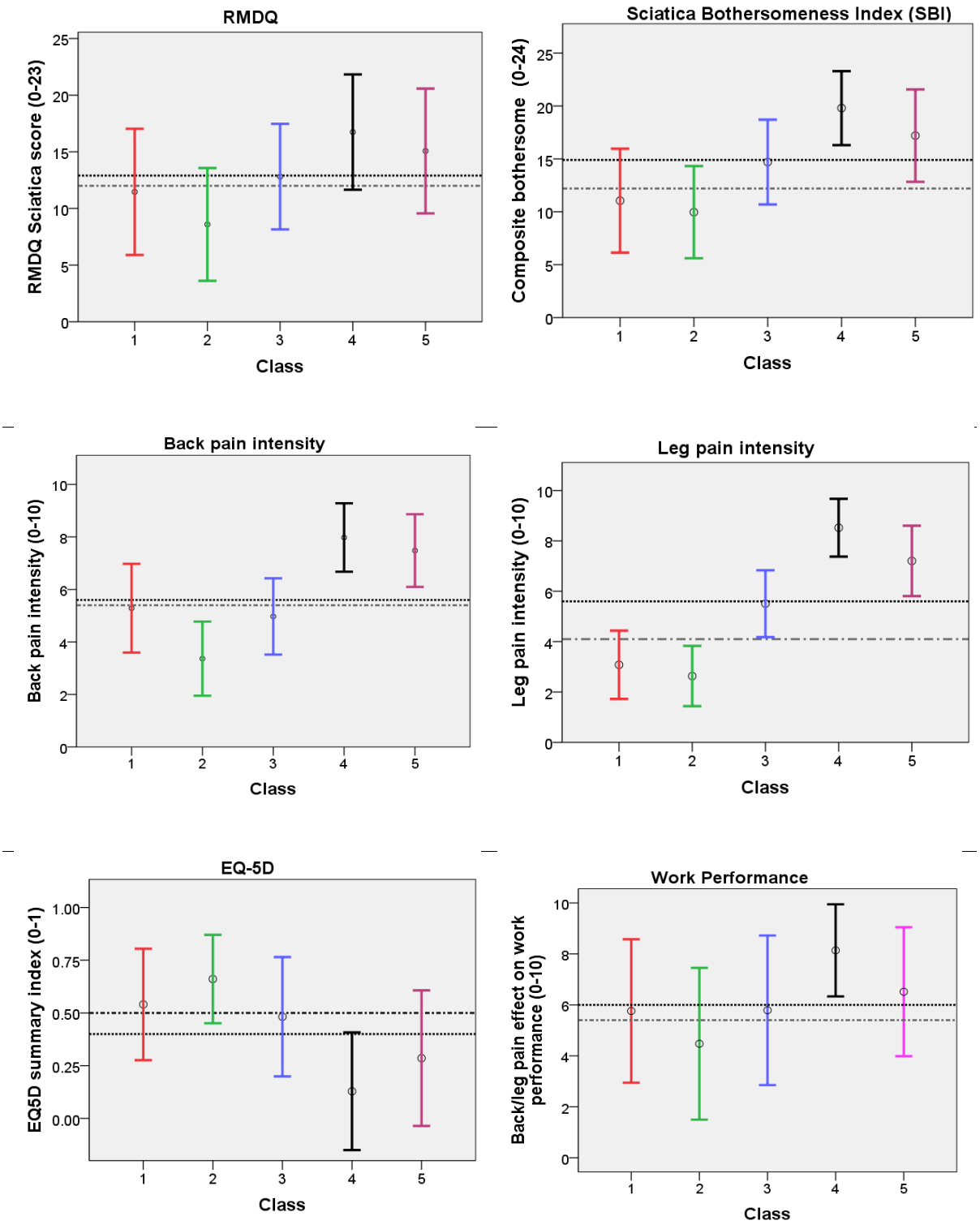


Figure 7.8 Error bar charts comparing mean scores of key characteristics (continuous measures) between five latent class groups and clinical diagnosis subgroups with sciatica (·····) and without sciatica (-·-·-·-). Error bars represent +/- 1 SD

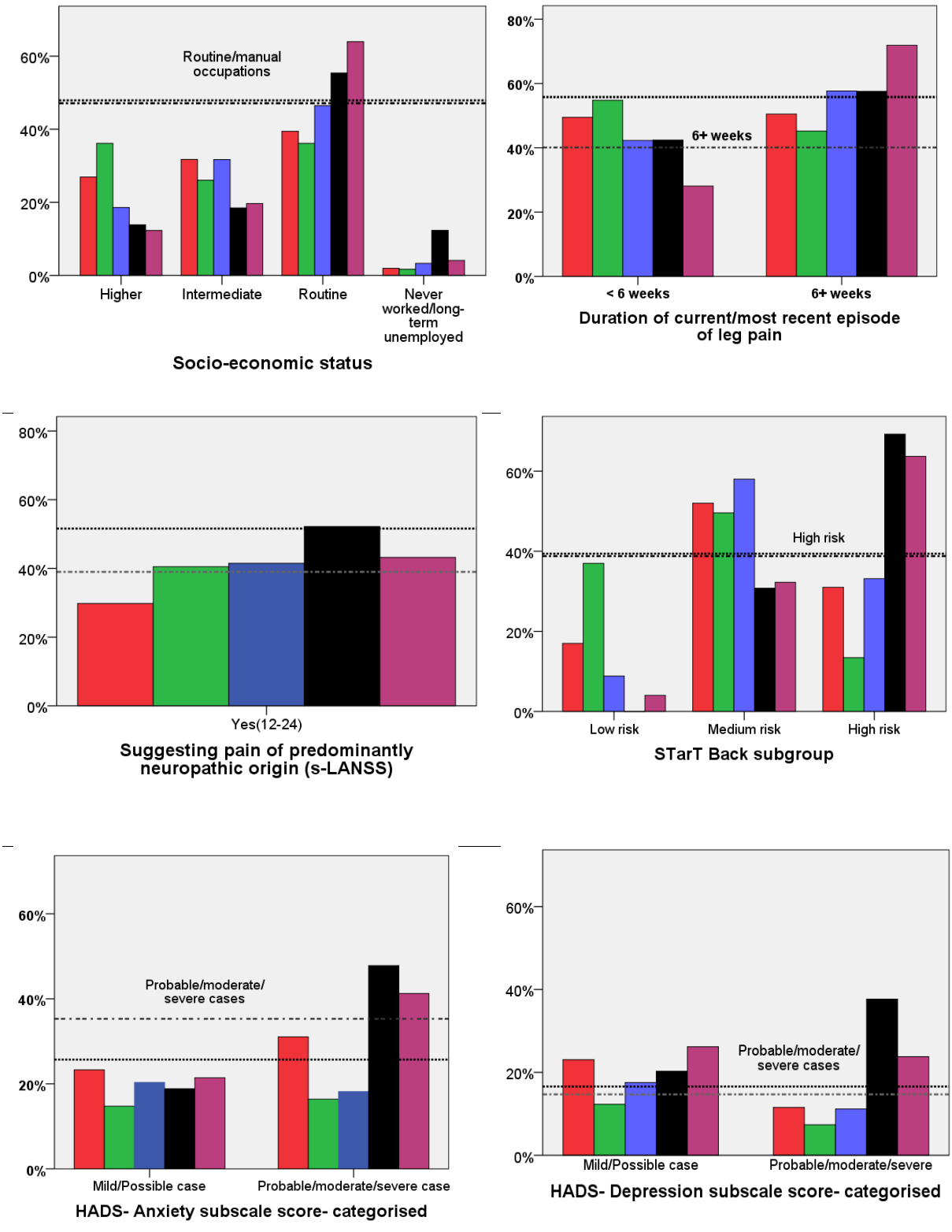


Figure 7.9 Bar charts comparing proportions of positive responses on key characteristics (categorical measures) between five latent classes and clinical diagnosis subgroups with sciatica (·····) and without sciatica (- - - -).

class 1 ***class 2*** ***class 3*** ***class 4*** ***class 5***

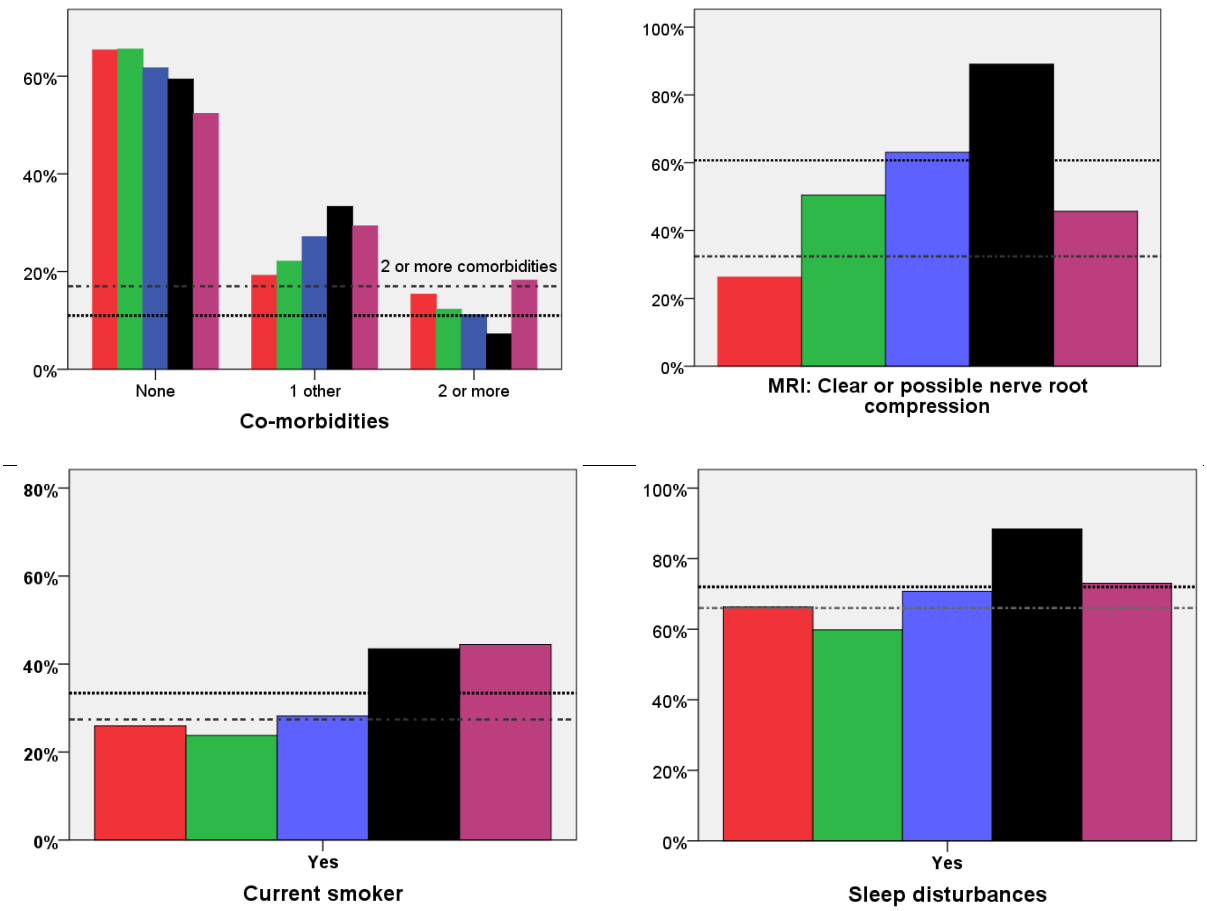


Figure 7.9 Bar charts (continued) comparing proportions of positive responses on key characteristics (categorical measures) between five latent classes and clinical diagnosis subgroups with sciatica (·····) and without sciatica (- - - -).

class 1 ***class 2*** ***class 3*** ***class 4*** ***class 5***

An insight into patient's pain over the last year was gained from one of the self-report questions "How has you back/leg pain been over the last year?" (figure 7.10). Patients in class 4 and 5 had the greatest proportion with "severe pain all the time" and none in these classes had "pain that gradually improved". Class 5 had the lowest proportion of patients with a "first ever episode" of pain or "a few episodes, mostly painfree periods in between". The seven responses to the pictorial questions were dichotomised to mild or moderate/severe (see chapter three, section 3.7.2, page 33) pain in the previous year. The proportions in classes with moderate/severe pain was lowest for class 1 (30%); Classes 2, 3 and 4 had 42%, 48% and 55% respectively with moderate/severe pain and class 5 had the highest proportion with 71% categorised as having a moderate/severe pain trajectory over the previous year.

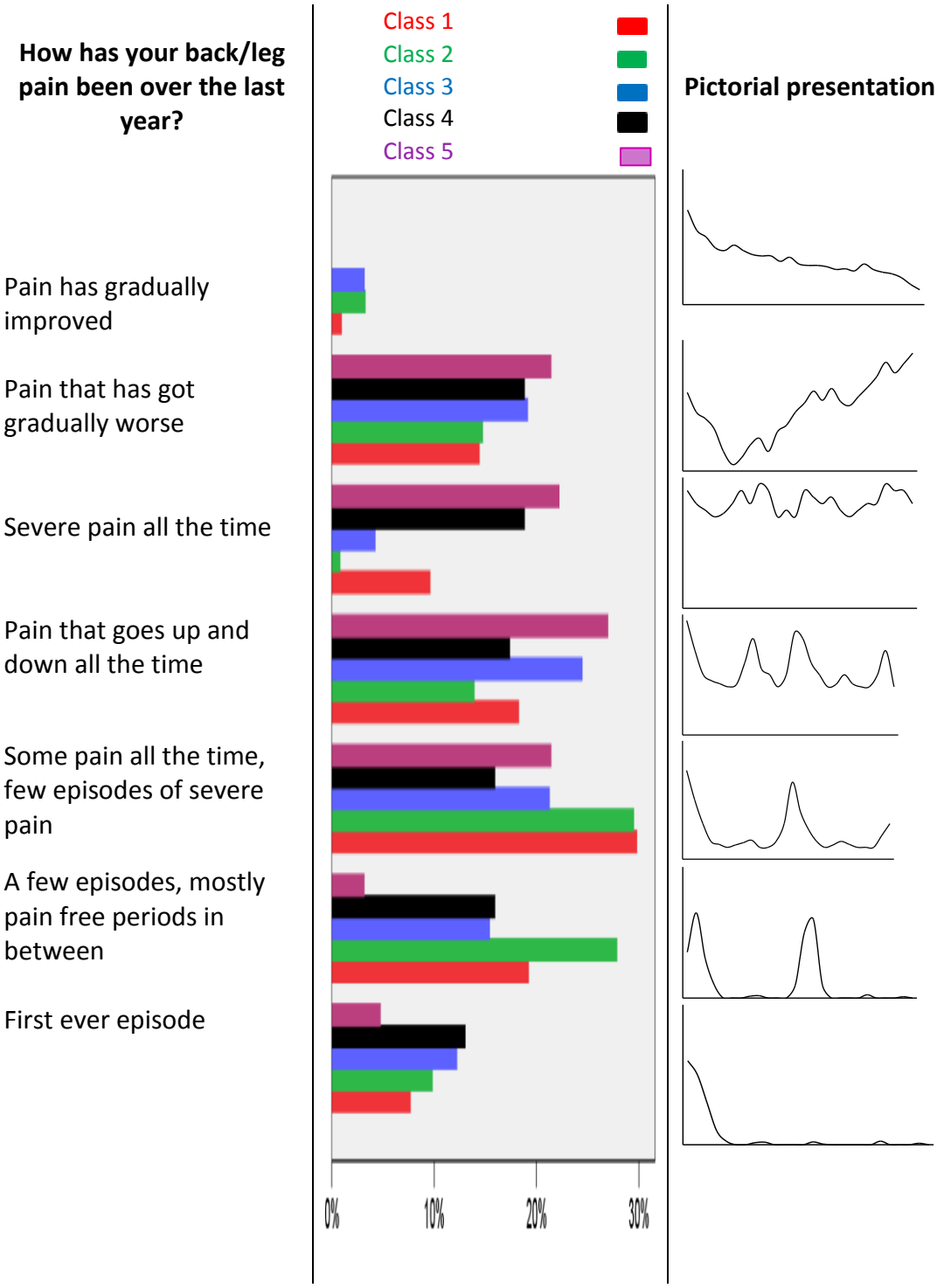


Figure 7.10 Bar chart comparing the five classes' response to "how has your back/leg pain been over the last year?"

7.6 Discussion

7.6.1 Principal findings

Using LC modelling, this study classified primary care consulters with LBLP into distinct classes based on their response to clinical assessment items used to guide diagnosis. Within the classes, patients could be distinguished as having either 'referred leg pain' or 'sciatica', based on their probability responses to the clinical assessment items. The statistical technique of LC modelling revealed further subgroups with distinct clinical profiles within this broad two-fold classification. This study is the first to have used LC modelling to identify potentially clinically relevant classes of primary care consulters with symptoms of back and leg pain.

Part one

The analysis described in part one used a selected sample from the ATLAS cohort in whom the assessing clinician had high confidence in their clinical diagnosis, and all patients had an MRI scan. Latent classes were identified using the same seven variables that were used in the diagnostic model described in chapter six. The reason for using the same sample and variables was to reflect on (i) the appropriateness of the reference standards used in the clinical diagnostic model and (ii) to compare findings from the diagnostic model to those from the LC model to assess which variables contributed most to a clear distinction of diagnostic groups.

The two classes identified in LC analysis, showed high agreement with the two groups defined according to the clinical diagnosis reference standard. Agreement and reliability indices of the two LC solution were highest with the clinical diagnosis reference standard and performed less well against the reference standards of clinical diagnosis plus confirmatory MRI, and MRI only. This supports the use of clinical diagnosis as a satisfactory reference standard for the clinical diagnostic tool developed and presented in chapter six.

When further classes were added to the latent model, it showed that a three class solution was a better fit for the data, suggesting this sample population could be classified into further groups. The LC model identified a class which had very low probability of a positive response to any of the

clinical assessment items and this class reflected a group not as severely effected in terms of pain, disability, work and psychosocial factors. This class had high agreement with the group of patients with a clinical diagnosis of 'referred leg pain'. Class 1 and 2 had very clear clinical characteristics one would expect of patients with sciatica. Their overall profile of pain, disability, work and psychosocial factors suggested that class 1 was a severe sciatica group and class 2 less severe.

From a clinical assessment perspective, both the sciatica latent classes had high probability of below knee pain and neurological deficit hence these were not helpful clinical features for distinguishing between the groups. Two items were clearly different between the groups: leg pain worse than back pain and positive neural tension tests. Both these items had very high probability of being positive (>85%) in the severe sciatica class. In the diagnostic multivariable model, these two clinical items had strong association with a diagnosis of sciatica, in particular positive neural tension (odds ratios of 21.6) (see table 6.8, chapter six, page 138). It suggests clearer signs to help clinicians make a more confident diagnosis.

Considering this three class solution was based on the sample of patients with high confidence in clinical diagnosis, the next step was to apply the same technique to the whole cohort of LBLP patients (n=609), to explore if the same three class solution was apparent or if more subgroups were identifiable when the full spectrum of LBLP was used to model the data. The rest of the discussion will focus on these results.

Part two

In the analysis presented in part two, five classes were identified. One class was clearly a referred leg pain group with mild leg pain severity. Three classes were considered to represent varying severity of sciatica (mild, moderate and severe). The fifth class was more difficult to define, presenting with a very similar response to clinical assessment items as the mild sciatica class (below knee pain and moderate probability of positive neural tension) but with much higher back and leg pain severity and a worse profile on pain, function, health and psychosocial measures.

Some additional items from the clinical assessment were included as variables, to model the full dataset. Back pain intensity contributed to distinguishing class 2 from class 5 and could also shed light on why the referred leg pain class had a slightly more severe overall profile compared to the mild sciatica class whose mean back pain intensity was lower. For example, the difference between a person's leg and back pain has been shown to be a significant predictor of the outcome of decompressive surgery for symptoms due to spinal stenosis (Kleinstück et al. 2009) or disc herniation (Kleinstück et al. 2011) with greater back pain intensity relative to leg pain, showing worse outcomes after surgery.

The probability of having leg pain on flexion (forward bending) was higher for patients in classes 3 and 4. This could be explained by its similarity to the mechanics of performing a straight leg raise, a test which was also highly probable of being positive in these two classes. Leg pain on extension did not add much information to distinguishing the classes. The elements of the neurological examination were added individually to see if they provided any further value to the classification but this was not observed as all probabilities across the classes remained below 55%. This possibly reflects the lower prevalence of these findings across the groups. Unsurprisingly, the highest probability for reflex and myotome deficit was seen in the high severity sciatica class. Although the probabilities for sensory deficit among the groups did not exceed 0.56, there is a clear distinction between class 1 which has minimum probability of sensory deficit (0.04) and the remaining four classes which have similar profiles on sensory testing.

Significant differences were seen between the classes on a range of key characteristics. 11% of the cohort who were classified as 'severe sciatica', presented with the highest BMI and greatest proportion of smokers and patients on sick leave from work. Their mean disability level as measured by RMDQ was 16.7, comparable to secondary care sciatica populations taking part in clinical trials involving surgery (16.4) (Peul et al. 2007). The high scores for the 'severe sciatica' class are seen across all domains, many exceeding values seen in secondary care sciatica cohorts. By comparison, the 'mild sciatica' class had the lowest level of disability (8.6) comparable to other

primary care LBP research cohorts with and without leg pain (8.8) (Hill et al. 2011a). Use of the STarT Back tool to screen patients who present to their GP with LBP +/- leg pain has shown that just over half (56.8%) are at low risk of a poor outcome (Hill et al. 2011b). This reduces to 26% in those receiving physiotherapy treatments for the same problem (Hill et al. 2011b). The 'mild sciatica' class in this analysis had the greatest numbers at low risk of poor outcome (37%) with only 14% considered high risk. This contrasts to class 4 and 5 where over two thirds of patients in both classes were categorised potentially at high risk of poor outcome.

When comparing the characteristics of patients in the ATLAS cohort according to the clinical diagnosis of 'sciatica' and 'referred leg pain' (Konstantinou et al. 2015), no significant differences were found between these groups in characteristics such as back pain intensity, disability, proportion of smokers, time off work, sleep disturbance, quality of life, future risk of poor outcome and depression (see Appendix K for table reproduced from paper). In this LC modelling, the most striking aspect of the analysis is the wide variation in clinical and other characteristics seen among patients predominantly classified as having sciatica. The results show that the class named 'moderate sciatica', consistently reflects the profile of the overall group of patients with a clinical diagnosis of sciatica. The 'referred leg pain' class mirrors the clinically diagnosed group of referred leg pain patients. However, two classes ('severe sciatica' and 'atypical sciatica) present with considerably greater severity in terms of pain, disability, risk of poor outcome, impact on work and psychosocial and health related issues. For example, the diagnostic groups of sciatica and referred leg pain had similar depression scores (Konstantinou et al. 2015). However class 4 had three times as many patients with depression and class 5 had twice as many. The proportion of anxious patients in class 4 was double that seen in the overall diagnostic sciatica group (Konstantinou et al. 2015).

Class 5, 'atypical sciatica', did not have a clear cut diagnostic pattern. Patients in this class had a 70% probability of being diagnosed by the clinician as having sciatica but confidence in this diagnosis was lowest compared to all other groups, with just over half in the class with 80% or

above confidence in diagnosis. Despite their challenging presentation in terms of diagnosis, patients in this class had a very similar profile to the 'severe sciatica' class with regards to levels of pain, disability, health and psychosocial measures. Notably, this was the group with the most comorbidities, highest proportion of smokers, and greater numbers with duration of leg pain of more than three months. Anxiety and depression were similar to class 4 and markedly greater than class 2 ('mild sciatica) and class 1 ('referred leg pain'). Over two thirds of patients in class 5 believed their back/leg pain would last a long time. They had the smallest proportion (44%) who felt they had personal control over their symptoms. There was also suggestion of having pain elsewhere as over 60% of them reported pain in the neck or shoulders in the last two weeks, at consultation. Class 1 ('referred leg pain) had the next highest proportion of pessimistic patients and generally this class had more psychosocial issues (self-report and history items from clinical assessment) compared to the mild and moderate sciatica classes.

The latent classes shed more light on the profile of patients more likely to have possible neuropathic pain characteristics. Just over half of the 'moderate sciatica' class (class 3) and the 'atypical sciatica' class (class 5) had neuropathic pain but class 4 ('severe sciatica') stood out as having the highest proportion, with two thirds of patients having possible neuropathic pain.

7.6.2 Comparison to other studies

Identification of subgroups of LBP patients is considered a number one research priority and much work has been done in this area. The majority of classification systems have focused on non-specific LBP. The systematic review presented in chapter four, identified 21 classification systems that included patients with low back and leg pain with only three focusing specifically on applying classification techniques to LBLP patients only. The quality of the methodology varied widely and most were based on clinical opinion (judgement). The method of classification used in this study was a statistical approach with clinical judgement used to help interpret the classes. This recommended approach to classification (Ford et al. 2007) has not been used to classify LBLP

patients before. The identified classes in this study can be compared to some of the classification systems identified in the systematic review (chapter four).

Schafer et al. (2009a) used a judgement approach to classify LBLP into four groups based on pain mechanisms (see chapter four, section 4.3.3, page 70). Any patient with a score over 12 on the s-LANSS neuropathic pain scale questionnaire was classified in the 'neuropathic sensitisation' group. All classes derived from this latent class modelling had proportions of patients with neuropathic pain according to the s-LANSS questionnaire score, however, classes 3, 4 and 5 with the more severe profiles had considerably greater proportions of patients with possible neuropathic pain. Schafer et al's 'denervation' group had marked sensory and motor deficits. This group was not identifiable among the latent classes of LBLP. Schafer et al.'s 'peripheral nerve sensitisation' group is characterised by marked nerve mechanosensitivity (i.e. positive neural tension) in the absence of neurological deficit. A positive response to nerve trunk palpation was also used to assign patients to this group. A further study to investigate if Schafer et al.'s four groups differed in terms of disability and psychosocial factors (Walsh and Hall 2009), showed that the 'peripheral nerve sensitisation' group had greater disability than all other groups and more fear avoidance beliefs compared to 'neuropathic sensitisation' and 'denervation' groups. This was considered a surprising finding at the time, a result more expected from the 'neuropathic sensitisation' group. These findings may mirror the relevance of the positive neurodynamic tests. In this LC study, the most severe sciatica class (class 4) had the highest (>90%) probability of positive neural tension. Schafer et al.'s fourth group called "nociceptive musculoskeletal pain", does not have neurological deficit or positive neural tension and is similar to class 1 in this LC study.

In Smart et al.'s (2011) classification system, multivariable analysis identified discriminatory clusters of signs and symptoms associated with a clinically determined dominance of each of their three pain groups (i) 'central sensitisation pain' (ii) 'peripheral neuropathic pain' and (iii) 'nociceptive pain'. The clinical assessment items associated with their 'peripheral neuropathic

pain' group were history of nerve injury, dermatomal leg pain and pain provocation with neural tension testing. Similar to the results from the LC modelling, response to neural tension tests in Smart et al.'s study seemed the characteristic distinguishing the group with 'peripheral neuropathic pain' from the nociceptive pain group and central sensitisation group.

The first four categories of the widely adapted Quebec Task Force Classification (QTFC) (Spitzer et al. 1987) system are based on the location of the back and leg pain and whether the pain is accompanied by neurological deficit. Several validation studies have shown that patients with leg pain were more severely affected than those with localised back pain, and those with signs of nerve root involvement were the ones most severely effected in terms of disability and work ability. One of the shortcomings of the QTFC is the broad description for category four (LBP with signs of nerve root involvement) and studies have used different criteria to describe this group thus making comparisons across studies more difficult. Ben Debba et al. (2000) adapted the QTFC system and choose the SLR test to differentiate between groups. In their study they showed that patients classified with distal leg pain and positive SLR (category four) were 13 times more likely to be treated surgically than patients with back pain only. This finding highlights a positive SLR as a marker of condition severity and could help explain why class 4 could be distinguished by the very high probability of having positive neural tension.

7.6.3 Strengths and limitations

LC modelling as a statistical technique for subgroup identification, has reasonably widespread application in musculoskeletal pain research and has been used to identify phenotypes from cross-sectional data in cohorts with pain in the foot (Rathod et al. 2015), hand (Green et al. 2015), shoulder (Groenier et al. 2006), knee (Kittelson et al. 2016) and multiple pain sites (Lacey et al. 2015). LC modelling has been used in developing classification systems for arthritis related conditions, including psoriatic arthritis (Taylor et al. 2006) and juvenile idiopathic arthritis (Thomas et al. 2000).

The application of LC modelling is becoming more common in LBP research. Examples include using LC modelling to identify subgroups of adolescents at risk of developing LBP (Mikkonen et al. 2016), matching LBP patients as likely responders to cognitive behavioural therapy (Barons et al. 2014) and mapping trajectories of LBP (Dunn et al. 2013, Deyo et al. 2015, Kongsted et al. 2015). Within some of these datasets, patients with LBLP have been included but the focus has never been on reporting their profiles or outcomes separately.

This is the first study that has applied the statistical technique of LC modelling to classify patients with LBLP, including sciatica. Previous classification systems for non-specific LBP have used clustering techniques to identify subgroups and are considered high quality as they derived their groups based on results from statistical analysis of patients' data, as opposed to an a priori judgement approach (McCarthy et al. 2004). There isn't a perfect method of deciding on the optimal number of classes and these five classes may not reflect the precise clustering of LBLP patients among primary care consulters. However, the five class solution was chosen based on the optimal statistical fit of the data and the classes seemed to represent distinguishable subgroups of LBLP. The sample used in the analysis represented a true primary care population presenting initially to their GP, with variable severity and duration of symptoms. This contrasts with many studies which select the most severe cases of sciatica recruited in secondary care settings. The analysis revealed two out of the five classes within the spectrum of LBLP with severe pain and disability profiles.

7.6.4 Clinical implications

One of the goals of this PhD work is to provide a more valid assessment, feasible and appropriate for primary care, which distinguishes between low back-related non-specific leg pain and sciatica.

The clinical diagnostic model provided this and its validity is further supported by the two class latent model which showed good agreement with the clinical diagnosis using the same clinical assessment items. However, this latent modelling has gone one step further to show that this

group of LBLP patients can be further classified into distinct subgroups. Statistical modelling revealed five classes of LBLP patients with different patterns of response to clinical assessment items and clear differences in pain, disability, health, work and psychosocial profile. The next obvious question is whether this information can be utilised for prognostic purposes. To answer this, a prospective look at outcomes over time is required. It may also matter in terms of how these patients are managed initially and later on. The five identified classes may have implications for clinical research. Heterogeneous study populations in clinical research can potentially confound outcomes (Taylor et al. 2006). The classes identified in this study may be more homogenous groups that may represent uniquely different responders to specific interventions. The next step would be to consider optimum management pathways for these classes and formally test whether different management options improve outcomes. This classification system of LBLP may enable practitioners to offer more appropriate treatment options for each group.

7.7 Conclusion

Sciatica is considered a risk factor for poor prognosis in LBP presentations and may also require a different therapeutic approach to non-specific LBP (Fairbank 2007, Freynhagen et al. 2008). Accurate definitions are important and this work shows it may be necessary to rethink current definitions of sciatica. It is not only important to distinguish between referred leg pain and sciatica, but it may also be important for clinical decision making, to recognise and identify the different subgroups within the sciatica pain group.

This work used statistical techniques to identify subgroups of LBLP and as such clinical judgement is needed to interpret the groups in the context of patients' presentations seeking care. The interpretation of the groups was primarily done by the PhD candidate and her supervisory team. To strengthen the interpretation of the classes and improve or confirm their face validity, the next

stage is to present these to clinicians working with spinal pain patients for evaluation and feedback.

Chapter Eight: Workshop with clinicians to evaluate thesis findings

8.1 Introduction

This chapter gives a narrative synthesis of the opinions and feedback of clinicians who took part in a workshop where the thesis researcher presented the diagnostic model and latent classes. The findings will further contribute to informing the interpretation and clinical relevance of the thesis research findings.

8.1.1 Aims and objectives

Aim: To explore the clinical relevance of the main thesis findings on diagnosis and classification of LBLP, with clinicians.

Specific objectives

(i) To gauge clinicians' opinions on the clinical usefulness and relevance of the diagnostic tool for sciatica.

(ii) To investigate the face and content validity of the LBLP classification system derived from LC modelling and its perceived usefulness in clinical practice.

8.2 Invitation of participants

A workshop was arranged with clinicians to evaluate the clinical relevance of the main PhD research findings. Physiotherapy researcher facilitators based at the Institute for Primary Care and Health Sciences, Keele University, identified clinical leads in the surrounding NHS trusts of Cheshire East, Staffordshire and Stoke on Trent, and Shropshire. Invites were sent by the thesis researcher via email (Appendix L) to the clinical leads outlining the aims of the workshop and asking them to identify clinicians who predominantly assessed and treated spinal patients and would be able and interested in attending the workshop.

Fifteen clinicians attended the workshop on June 30th 2016. They included 14 physiotherapists and one osteopath. The osteopath worked in private practice and the physiotherapists had clinical roles in an NHS setting, ranging from recently qualified therapists to experienced extended scope practitioners (see figure 1 for breakdown of job titles).

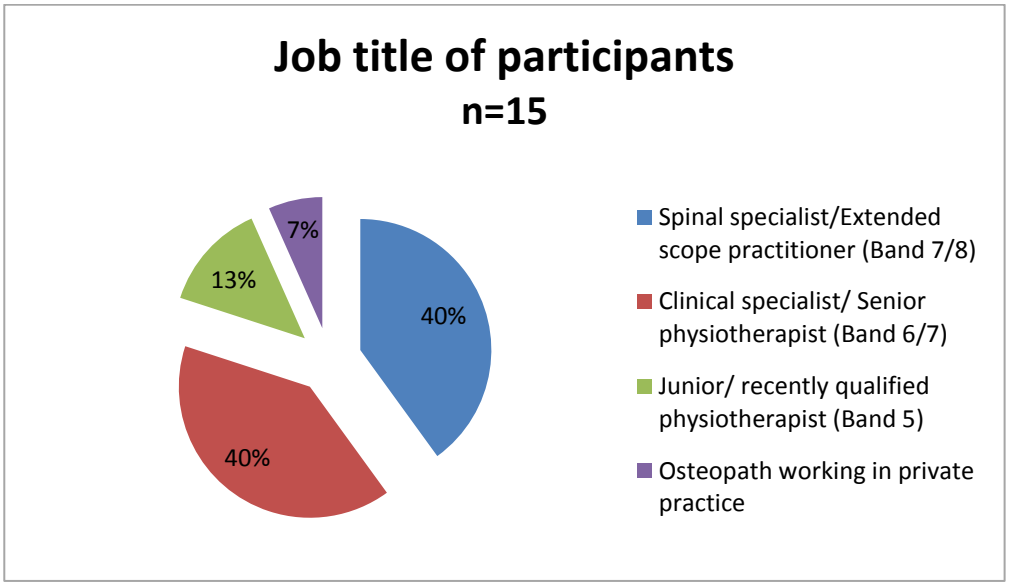


Figure 8.1 Breakdown of clinical role/job title of clinicians

The clinicians were qualified on average 9.7 years (range 0.8 to 25 years) and they had an average of 8.3 years' experience (range 0.6 to 22 years) in treating a predominantly musculoskeletal caseload. Six of the 15 clinicians worked part time in their clinical role. On average, clinicians reported 53% of their clinical caseload was spinal patients (range 30-90%).

8.3 Workshop format

The workshop ran for approximately three hours. The thesis researcher gave two presentations, the first one focused on the clinical diagnostic model and the second on findings from the LC modelling analysis. Following the first presentation on the diagnostic tool, clinicians were asked, as a group, to consider the clinical relevance of the tool. The second presentation was followed by group discussions on potential descriptors for the latent classes and their clinical relevance. For

the discussions on the results of the latent class analysis, the clinicians initially worked in two small groups, followed by a whole group discussion.

The discussions were captured using a digital recorder and the thesis researcher used the recording to prompt recall on the main points and themes of the discussions to help prepare this narrative summary. On arrival, the clinicians were asked to complete a form (Appendix M) asking them about their job role and clinical experience. Using the same form they also gave signed consent for the content of the workshop discussions to be used in the thesis write-up or in presentations relating to the thesis.

As described in chapter four, an appraisal of LBLP classification systems in the literature was carried out as part of this thesis research. The questions appraising content and face validity of a classification system (box 8.1) were considered when reflecting on the thoughts and suggestions made about the five classes, by the clinicians who attended the workshop.

Box 8.1 Appraisal of content and face validity in classification systems (from Buchbinder et al. 1996)

Content validity

- Is the domain and all specific exclusions from the domain clearly specified?
- Are all relevant categories included?
- Is the breakdown of categories appropriate, considering the purpose?
- Are the categories mutually exclusive?
- Was the method of development appropriate?

Face validity

- Is the nomenclature used to label the categories satisfactory?
- Are the terms used based upon empirical (directly observable) evidence?
- Are the criteria for determining inclusion into each category clearly specified?
- If yes do these criteria appear reasonable?
- Have the criteria been demonstrated to have reliability or validity?
- Are the definitions of criteria clearly specified?

8.4 Diagnostic model: Discussion points

As previously mentioned, the first session of the workshop focused on the diagnostic model (chapter six). The thesis researcher presented the results of the study (presentation slides in Appendix N) and clinicians were invited to comment on the clinical relevance of the tool and ask any question relating to the details of the model. The main themes identified in the discussion are outlined below.

8.4.1 Clinicians response to diagnostic tool

The clinicians agreed that the tool seemed to make clinical sense and there was nothing unsurprising when matching the scores from the scoring tool with the corresponding probability of a sciatica diagnosis. One of the clinicians commented that she felt the model represented what she would expect to see clinically e.g. a patient with a score of 6 would strongly be suspected of having sciatica. One clinician queried how the tool would help if someone had a lower score, for example 3 or 4.

One clinician queried why dermatomal pain was not included in the tool because it was a sign that would increase his confidence in a sciatica diagnosis. The same clinician pointed out that he looks for a "*very positive SLR*" before he is convinced of a sciatica diagnosis.

> "*SLR would have to be very positive, not necessarily 60 degrees, more like 20 degrees"*

In terms of neurological deficit, it was asked why the neurological items were not used as split individual components of myotomes, dermatomes and reflexes and if analysing the components individually would change the diagnostic model.

Clarification was requested by a participant as to whether the tool would be preferred to a clinician's decision, and did we know if using the tool is superior to a clinician's opinion?

Some of the clinicians felt it would be useful in primary care, in particular to help GPs to more readily prescribe neuropathic pain (NP) medication. The prescription of NP medication was a topic

discussed by the clinicians on several occasions during the workshop. Some felt it was not being done promptly enough by GPs and others felt use of the tool would increase their confidence to ask the GP to prescribe NP medication if a sciatica diagnosis was suspected.

Some clinicians wondered if adoption of the tool in clinical practice would act as a deterrent for ordering MRI scans and thus lead to cost savings. They felt clinicians could be reassured by the tool's agreement with MRI findings (as shown in this thesis) and this may lead to decreased use of MRI scans in cases where an MRI was perhaps being used to help with low confidence in diagnosis.

Use of the word sciatica was also discussed, some clinicians felt it was "*an umbrella term*" and not helpful for diagnosis. They would consider a more specific term like "*disc, piriformis syndrome, SIJ"* to be a diagnosis. One clinician mentioned that in her workplace, they were strongly discouraged to give a patient a diagnosis. It was suggested by another clinician that often patients just want a diagnosis, because they ask "*what have I got*?" The reply to this was:

"*you can never be 100% sure of a diagnosis particularly without an MRI".*

Several participants commented that the tool would be useful for more junior clinicians.

8.5 Classification of LBLP using latent class modelling: Discussion points

The second part of the workshop focused on the presentation of the LBLP classes identified by LC modelling (Appendix N for presentation slides). The five classes were initially presented by focusing on their response to the clinical assessment items. Using the graph depicting the probabilities to a positive response on clinical assessment items and pain intensity (see chapter seven, figure 7.6, page 178), they were given some time to work in pairs to look at the five classes and consider what they would call these groups or how they would "diagnose" them.

Subsequently, the thesis researcher presented more information about the five classes, including how the patients in the classes had been diagnosed by clinicians, overall MRI findings for the five

different classes and a range of characteristics for each class including pain, disability, neuropathic pain, psychosocial factors, work interference, previous year of back and leg pain and risk of poor outcome (STarT Back subgroup). The clinicians were asked to consider the following questions:

- Do you recognise these classes?
- Do you perceive that these classes could add useful information to your clinical practice?
- Can you give a description/name for each class?

8.5.1 Learning tool

Clinicians felt the classes could be a useful "learning tool" for more junior clinicians or clinicians less experienced in managing spinal patients. The classification system could be used to signpost certain patients to a service or clinician with the appropriate skills to offer or deliver more invasive treatments. For example it was suggested that categorising a patient in class 4 or 5 may prompt a clinician to discuss this patient with more experienced clinicians, or refer them onwards to a more specialised spinal service. The classes could also be used to identify patients that may be suitable for clinicians wishing to gain more experience in managing more complex patients e.g. class 5 patients, or varying a clinician's caseload so they do not see all patients from one particular class.

8.5.2 Treatment

Although the workshop participants were not specifically asked to consider treatment strategies for the classes, treatment options were a significant theme in their discussions. It was suggested by one of the extended scope clinicians that a patient in class 4 or 5 is more likely to be investigated based on the high proportions in these classes that were at high risk of poor outcome based on the STarT Back tool. "*Do something with them quickly*" was expressed by another clinician. It was suggested that Class 4 patients could be treated more intensively with interventions such as injections or surgery, because of their obvious sciatica clinical presentation, high pain severity and severe impact in terms of disability, impact on work and function and

psychosocial factors. The issue of when to "*hold on*" to a patient, and when to refer on for more invasive treatment was discussed in relation to a potential use of the classes.

The theme of encouraging GPs to prescribe NP medication was discussed again and some suggested the classification system could be used to communicate with GPs or consultants. Concerns were raised about a classification system being overly prescriptive and that this might have implications for services based on responses from commissioners funding clinical services. The example given was being encouraged to *"get rid of"* class 2 patients because of their low pain and disability profile. In other words encourage self-management following the initial assessment but not offering any further follow-up treatment or reviews. "*Getting the right patients to the right person*" was a possible advantage of the system. Auditing a service that uses the classification system was suggested as a helpful way of identifying resources and training needs. An example given was identifying a service that needs more spinal specialist clinicians.

8.5.3 How to classify a patient?

A clear focus among some clinicians was how to assign patients to the classes. Suggestions given were the use of an algorithm, a tick box, or linking the diagnostic tool to the classes, e.g. a cumulative score from the diagnostic tool would indicate membership to a particular class.

"To apply it clinically I need to have a way of actually classifying the patients, and then apply the best treatment."

Some clinicians queried whether five classes were needed and wondered if the system could be reduced to three classes which would be an easier system to use. Other clinicians felt this was "*dumbing down*" the classes and would lose important information about the patients by ignoring the results of the analysis.

8.5.4 Labelling the classes

The clinicians were asked how they would label the classes and the thesis researcher did not initially share the labels of the classes that have been described in chapter seven. The majority of

the clinicians suggested mild, moderate and severe sciatica for classes 2, 3 and 4 respectively. A suggestion for class 1 was “clinically silent”, “not sciatica”, “referred leg pain” and “simple leg pain”. Class 5 prompted the most discussion and the following terms and descriptors were suggested: “Central sensitisation”, “yellow flag group”, “non copers” (as opposed to class 2 who were called “the copers”), “more going on”, “neuropathic more than nociceptive pain”, “complex with more comorbidities”, “central sensitised sciatica”, “pseudo sciatica”, “suspected sciatica”, “we don’t know what’s going on sciatica”.

One clinician was not keen on labelling the classes and queried why they needed labels or names at all. He argued the overall system should be given a name and the classes should stay as numbers with descriptors for the classes corresponding to the numbers. He felt this would avoid “muddying things” and adding further nomenclature to the literature.

Another clinician suggested classes 2, 3 and 4 should be labelled “nerve root”, and class 5 should be “sciatica”, based on her previous point that sciatica was a non-specific umbrella term for leg pain.

The thesis researcher shared the labels that she had given to the classes. Although ‘atypical sciatica’ had not been mentioned by the clinicians, they agreed it could be added to their list.

There was no clear agreement on the optimum label for class 5.

8.6 Summary

A strong theme that emerged from the workshop was use of both the diagnostic model and LBLP classes as a learning tool and a guide for directing resources within existing services. Potential for improving care for patients was considered, including quicker access to NP medication, signposting patients to appropriate management pathways and less reliance on MRI.

Clinicians appreciated that this was a “snapshot” of patient characteristics, based on cross-sectional data, and they were keen to know more about the clinical course of these patients.

Feedback from the clinicians supported the content validity of the LBLP classification. The method of development was considered appropriate following clear explanation as to how the classes were chosen. The classes were seen to represent all relevant categories of LBLP patients and clinicians felt the classes were distinct from each other. Some were keen to make them less distinct and merge some of the classes. As regards face validity, the main question concerned the labelling of the classes. There was no clear agreement on how to label class 5, and even whether the classes should have a descriptive name at all. The idea of not labelling the classes was an interesting suggestion and can be appreciated as a means of avoiding preconceived opinions as to how the patients should present or respond based on a label associated with a diagnosis. Criteria to determine inclusion into each category was not clear to clinicians. The classes identified by LCA are based on probabilities and further work is needed to consider how to allocate patients to a class and what, if any, characteristics other than clinical findings should be included in this process.

8.6.1 Reflections on the workshop

This evaluation was planned as an informal discussion with clinicians and not designed as a focus group which would require more structured thematic analysis. Based on themes that emerged from these informal discussions, a next step would be to set up a more formal process, and consider involving clinicians in planning management strategies for LBLP patients within the context of the research findings, with the addition of further findings on the clinical course of these patients. From a personal perspective, organising and facilitating this workshop further highlighted the importance of engaging clinicians in the process of research. Clinicians were keen to attend and be involved in the dissemination of this research.

Some clinicians expressed interest in attending the workshop but were unable due to busy workloads. Running a workshop at the clinicians' workplace could be explored in the future. However clinicians may feel more at ease to offer opinions out of their work environments and

not within earshot of other colleagues and/or managers. The spread of clinical experience and expertise amongst the workshop participants may have meant the more junior clinicians did not feel comfortable offering their opinion. This is not investigated as clinicians were not asked for specific feedback about the experience of the workshop.

Some clinicians requested fewer classes and this highlighted perhaps the usefulness of engaging further clinical input and agreement when choosing the number of classes in the first instance, particularly in cases when there isn't a clear statistical fit to the data.

The diagnostic tool was suggested as potentially useful for GPs, hence gauging GP opinion on this tool would be informative.

8.7 Conclusion

The overall response from the clinicians to the work presented was positive and there was an obvious desire to know how both the tool and the classes could be used in the clinical setting and whether it would improve outcome for patients. This workshop was targeted at clinicians, predominantly NHS based physiotherapists. To gain a broader insight, feedback from other healthcare professionals involved in the assessment and management of spinal patients would be ideal.

Chapter Nine: Discussion and conclusion

9.1 Introduction

The overall aim of this PhD was to establish clinical criteria that identify sciatica in patients who present with LBLP. This chapter provides a summary and synthesis of the principal findings and discusses the strengths and limitations of the work. Comparison with existing literature has been addressed in the preceding chapters. This chapter will reflect on whether the principal aim of the thesis was achieved and highlight challenges that were encountered along the way. Implications of the findings are explored in relation to implementation in clinical practice and potential ideas for future research are considered.

9.2 Principal findings

Classification of LBLP in the literature: Following a systematic search of the literature, very few identified papers specifically addressed the classification of LBLP. Within the systems, varying definitions of sciatica and associated clinical features were found. Methods of classification system development mainly relied on clinical opinion as opposed to also including statistical approaches. Using a combined clinical judgement and statistical approach is optimal.

Reliability amongst clinicians diagnosing LBLP: Regardless of training, assessment standardisation or professional background, reliability was merely fair amongst clinicians when diagnosing sciatica in LBLP patients with symptoms of any duration and severity. As confidence in diagnosis increased, agreement and reliability indices improved considerably. Not all patients with LBLP are difficult to diagnose, but inclusion of those cases that are reduces the reliability indices as was shown in this study. Ways of improving clinician agreement on diagnosis may be to assist the diagnosis process by identifying the optimal combination of items from the clinical assessment that best discriminate between these patients.

Development of a diagnostic model to identify sciatica: Findings from the systematic review helped to inform which variables to use for the diagnostic modelling. Based on results from the reliability study, high confidence clinical diagnosis was used as the reference standard for the diagnostic model. Results were compared to a model that used high confidence clinical diagnosis and confirmatory MRI findings as the reference standard. Four items from the clinical assessment were common to both final models: pain below the knee, leg pain worse than back pain, neurological deficit findings and positive neural tension. A scoring tool was derived, based on the clinical diagnosis model, which may be useful in clinical practice to identify patients with sciatica.

Classification of LBLP using the statistical approach of latent class modelling: Classes of LBLP were identified through latent class modelling. Five distinct classes were evident: a referred leg pain group, mild, moderate and severe sciatica groups and an atypical sciatica group with severe pain and disability but fewer of the classic signs and symptoms of sciatica. These classes potentially represent clinically relevant subgroups of LBLP that could respond differently to certain management pathways.

Clinicians' evaluation of the diagnostic tool and the LBLP classes: The diagnostic model and latent classes were presented to clinicians for evaluation of their clinical relevance and perceived usefulness in clinical practice. Use of the scoring tool to identify sciatica, in both research and clinical settings, was generally favoured by the clinicians. More guidance on how to practically classify patients into the five classes identified by LC modelling was suggested in order to help consider treatment options.

9.3 Strengths and limitations of the thesis

Strengths and limitations specific to each objective have been discussed individually within the chapters. This section reflects more broadly on the thesis. In particular, issues which affect the interpretation of the diagnostic model and the identified latent classes are considered.

9.3.1 Sample selection

This is one of the few studies in the literature about patients presenting with back-related leg pain that focuses on patients who have consulted in primary care. Patients with relevant symptoms of any severity and duration were recruited to the ATLAS cohort. In addition, no strict diagnostic criteria were imposed before entry into the cohort. There may still have been some element of selection bias, for example patients being missed by the GP or patients with more or less severe symptoms choosing to participate (or not) in the study. Therefore the invited patients who did not respond to the invite or were not interested in taking part in the study once they presented at clinic, may have differed from those that did take part. Data on age and gender and level of deprivation was available on those that did attend clinic but were ineligible or not interested in taking part, and these were similar to the recruited sample (Konstantinou et al. 2015); however no other measurements were available for comparison. Unknown response can have significant implications for estimating prevalence (van Loon et al. 2003) but this thesis did not aim to determine prevalence of LBLP or sciatica. The recruited sample represents a broad spectrum of LBLP presentation in terms of symptom severity and duration, which makes the findings more generalisable to clinical practice, in particular primary care, where most of these patients are initially seen and managed. Selection was not confined to more severe cases, as is often seen in studies of sciatica, and this was particularly apparent when latent classes of LBLP were identified with a range of mild to severe profiles in terms of pain intensity, disability and psychosocial factors.

9.3.2 Clinical assessment

Clinical assessment of patients with LBLP is routinely performed by primary care clinicians and is the cornerstone of diagnosis. Findings from the clinical assessment informed the bulk of the analysis carried out in this thesis. The reliability study was based on the diagnosis made by the clinicians, the diagnostic model used the variables collected in the clinical assessment including the clinical diagnosis and latent classes were modelled also using information from the clinical

assessment. Hence it is important to be reassured about the quality, validity and robustness of the assessment process. One strength of the clinical assessment was its evidence based development phase (carried out prior to the ATLAS study recruitment phase and not part of this thesis) which included a Delphi study with LBP experts (Konstantinou et al. 2012b).

At the ATLAS research clinics, patients with LBLP were diagnosed by clinicians as having either sciatica or referred leg pain. Clinicians with similar training carried out the assessments to optimise standardisation of the assessment procedure and the diagnostic decision (used as a reference standard for the clinical diagnosis model). This perhaps could also be viewed as a weakness of the study design as the findings may not be generalizable to clinicians with less experience. However, clinicians who participated in the workshop (chapter eight) agreed that the diagnostic model and the LBLP classes would be useful tools for more junior clinicians, which suggested that the findings are applicable to more inexperienced clinicians.

When the clinicians made their diagnosis, they also recorded confidence in their diagnostic decision. There is recognition of "diagnostic uncertainty" when assessing LBP patients (Serbic and Pincus 2014) and the potential implications of this, such as referring for diagnostic scans (Chou et al. 2011, McCullough et al. 2012). Hence this confidence rating proved very insightful when reflecting on reliability amongst clinicians when diagnosing sciatica and subsequently was used to choose a more optimal reference standard for the clinical diagnostic model. Confidence in clinical diagnosis also aided interpretation of the five LBLP classes identified with the method of latent class modelling. All patients in the ATLAS study underwent a comprehensive standardised clinical assessment therefore a wide range of clinical items/variables were available for selection in the diagnostic model and latent class analysis. No important variables in the assessment schedule were missing or overlooked. One of the clinicians taking part in the workshop evaluation (chapter 8) queried why dermatomal leg pain (leg pain corresponding to a lumbar nerve root, see figure 1.2 chapter one) was not used in the diagnostic model. Previous diagnostic models for sciatica have included pain in a dermatomal distribution and shown its association with disc herniation on

MRI (Vroomen et al. 2002, Coster et al. 2000). During the development phase of the clinical assessment (Konstantinou et al. 2012b), experts agreed the question "area or distribution of pain in the legs" should be included; however "dermatomal" distribution was not specified. A proxy for pain in a dermatomal distribution is "pain below the knee" (Dionne et al. 2008) which was used for this analysis. It is not known whether being able to distinguish between pain in a dermatomal distribution as opposed to more diffuse non dermatomal pain below the knee may have had some added relevance when describing the four latent classes that had high probability of the 'below knee pain' feature (classes 2 to 5), in particular class 5. However, evidence suggests that sciatica symptoms do not necessarily follow a clear dermatomal pattern (Murphy et al. 2009). Descriptors of pain such as burning, sharp or electric shock were not included in any of the analyses due to a large amount of missing data. It was a question in the clinical assessment that was not recorded by the assessing physiotherapist in over 25% of cases. This may have been a helpful item to describe the latent classes. Some of the descriptors (electric shock, burning) are included in the s-LANSS questionnaire, which was used to compare the five classes.

9.3.3 Clinical diagnosis of sciatica

The physiotherapists who carried out the standardised clinical assessments made a diagnosis of either referred leg pain or sciatica. This diagnosis does not reflect the specific pathoanatomical source of the nerve root involvement, i.e. sciatica due to a disc prolapse or stenosis. Some clinicians who attended the workshop (chapter eight) felt this distinction was a more appropriate diagnosis. The literature has considered diagnostic models specifically for stenosis (Konno et al. 2007) and certain features from the patient's presentation are considered indicative of stenosis, e.g. increase of pain with spinal extension activity, and ease of symptoms with sitting or a forward bending position of the lumbar spine. The diagnostic model in this thesis did not aim to distinguish between disc or stenotic symptoms and no questions or items from the clinical assessment specific to stenosis were included in any of the diagnostic model analyses.

Among the patients included in the clinical diagnostic model (n=394), 41 patients had a diagnosis of stenosis made by either the assessing clinician or according to MRI findings. Hence the model was repeated excluding patients with a diagnosis of stenosis, to see if this changed the model output in any way. No change was seen in the final output of the diagnostic model, confirming that the clinical items identified sciatica without a need for specifying a pathoanatomical cause. When describing and interpreting the latent classes, the patients with a diagnosis of stenosis were considered. They were not seen to directly influence the formation of the groups, i.e. no group consisted predominantly of patients with a clinical diagnosis of stenosis. Hence the diagnostic model developed in this thesis can be used to identify sciatica in patient with LBLP, irrespective of the pathoanatomical cause. Its use is not advocated as a tool to help distinguish between disc or stenotic symptoms.

9.3.4 Reference standard for clinical diagnostic model

There is no agreed reference standard for diagnosing sciatica (van der Windt et al. 2010). Much consideration was given to the choice of a suitable reference standard for the clinical diagnostic model prior to performing the diagnostic model analysis. All opportunities were availed of to discuss the reference standard selection or case definition of “sciatica” with experts in the field, which included my supervisory team, clinicians, members of the Spinal Research Group at the Institute for Primary Care and Health Sciences at Keele (where the thesis researcher is based), and experts in the field of LBP research external to Keele. The limitations and challenges of selecting a reference standard were acknowledged. Some colleagues argued MRI findings should be used, others agreed clinical diagnosis was acceptable. Following discussions, the thesis researcher made the decision to use the “high confidence clinical diagnosis” as the primary reference standard and compare this with the reference standard “high confidence clinical diagnosis and confirmatory MRI findings”. Several sensitivity analyses were also performed including MRI only as a reference standard and clinical diagnosis with no restriction in confidence in diagnosis. Following the analysis, findings from the diagnostic model analysis were presented at several forums and

conferences at a local level (Post Graduate Research Symposium, Keele University) and international level (Society of Back Pain Research, International Federation of Orthopaedic Manipulative Physical Therapists, British Society of Rheumatology). Limitations of an "imperfect reference standard" were regularly acknowledged by conference participants and discussions primarily focused on the commonly observed clinical scenario of discordant findings between clinical diagnosis and MRI findings. LBP guidelines advocate that, with the exception of "red flags" suggesting serious pathology, imaging for sciatica should only be done when it is likely to influence further management of the condition (e.g. if epidural injections or spinal surgery are being considered), not in response to diagnostic uncertainty (Chou et al. 2007, van Tulder et al. 2010, Webster et al. 2013). The results from the diagnostic models which compared reference standard of clinical diagnosis to clinical diagnosis plus confirmatory MRI were interesting. Although the weighting for clinical assessment items identified in the clinical diagnostic model was much stronger than that of the items identified by the model that included MRI, four items were identical. The model using MRI only as a reference standard did not perform very well with regards to calibration and discrimination. Positive neural tension test did not feature in this final model, yet in the clinical diagnostic model this item had the highest odds ratio of 21.6, and a positive SLR was highly probable in the "moderate" and "severe" sciatica latent classes. It is suggested that neural tension tests may cause pain due to chemical substances around the nerve root but not generating detectable signal on MRI (Beattie et al. 2000). Disc disruption that does not result in prolapse can result in as much leg pain as discs showing more severe disruption (Ohnmeiss et al. 1997).

Currently the majority of MRI scanners in the country require the patients to lie supine with a pillow under the knee, which is often a position of ease for patients with sciatica, as opposed to a more symptom provocative position of standing or sitting. A previous review on lumbar spine position for MRI scanning (Alyas et al. 2008) recommended re-imaging in an upright position with the addition of lumbar spine flexion and extension, when conventional MRI shows no evidence of

nerve root compression, in cases of a convincing clinical presentation. Another factor to consider is that the state of current MRI technology may not be sensitive enough to detect nerve root involvement in all cases, bearing in mind that it gives a two dimension (2D) picture as opposed to 3D. A prospective observational study in sciatica patients showed the superior effectiveness of 3D imaging to identify nerve root compression in cases with a clinical diagnosis of sciatica (Zhang et al. 2009). This technology is likely to advance in future years.

9.3.5 Classification of LBLP

The LC modelling was an option in order to circumvent the need for a reference standard and provided an opportunity for considering unidentified/hidden/latent groups among the full spectrum of LBLP patients. Despite removing the need for a predefined clinical diagnosis, the modelling still involves subjective decision making for choosing the variables for the model and subsequently interpreting the model solutions. The variables used to model the classes were informed by findings from the systematic review (chapter four), other diagnostic models in the literature and results from a previous Delphi study (Konstantinou et al. 2012b), so as to make the selection as robust as possible. Similarly, selection of the optimal number of classes took into consideration many factors as outlined in chapter seven including the statistical indices of fit, clinical interpretation of the classes and their agreement with clinical diagnosis and MRI findings.

When the latent classes were presented and described to clinicians attending the workshop, the feedback was that they recognise patients they see and treat in clinical practice, within the classes. A topic that arose during the workshop discussions was to perhaps have a classification system with fewer classes to make it more user-friendly. Others argued that this would lose important information. On reflection it may have been beneficial to include some additional clinicians in the decision making process when selecting the optimal number of classes.

Recognising these potential limitations, the resultant classes did give some new insights into the clinical spectrum of LBLP, which have not been highlighted previously in the literature.

9.4 Implications for clinical practice

The reliability study showed that distinguishing between referred leg pain and sciatica can be difficult, but agreement on diagnosis improves amongst clinicians as confidence in clinical diagnosis increases. A clinical diagnostic model to identify patients with sciatica was developed and a simple scoring tool was devised which could be used by clinicians to assist in the diagnostic process. The model identified key items from the clinical assessment that suggest a very high likelihood of sciatica. Conversely, absence of these symptoms suggests the presence of referred leg pain. A comment in a paper by Waddell and colleagues in 1982 observed that many items of clinical information duplicate other data and recommended "careful concentration on limited quantity of information" in clinical assessment to avoid confusion (Waddell et al. 1982). This concept was reiterated in the clinician workshop when participants suggested that the tool could help clinicians focus on specific items of the clinical assessment to assist with their diagnostic triage. Assuming that serious spinal pathology and leg pain attributed to non-spinal aetiologies are excluded, focusing on these five items could assist timely and confident identification of sciatica. Longer duration between onset of sciatica and effective treatment can have an unfavourable impact on symptoms and treatment outcomes (Lewis et al. 2011). Timely identification could potentially improve outcomes for these patients by facilitating quicker access to appropriate management pathways. This hypothesis would need to be tested in a study that compares current usual care to the use of the tool in primary care settings and follows up patients over time to see if outcomes (for example pain, disability measures or use of secondary care services) are improved with use of the tool.

The tool involves five simple items— three history questions and two physical examination tests, and could potentially be feasible for GPs to use as they are usually the first point of contact for these patients. If a patient scores at least five out of ten on the scoring tool, there is a high probability that their leg pain is sciatica. Use of the tool may expedite the prescribing of appropriate medication such as neuromodulating medication as opposed to what often happens

in clinical practice: a stepped approach to analgesia, a referral to physiotherapy and then a visit back to the GP to request more appropriate pain medication. Clinicians at the workshop also commented on this issue, and felt if the tool was used by GPs, it could facilitate earlier appropriate pain management for patients with sciatica. The next logical step is to have these discussions with GPs to see if they would consider the tool a useful addition to their practice and whether they would use it.

When patients consult with sciatica, they seek a diagnosis and a legitimisation of their pain and symptoms (Hopayian and Notley 2014). Research has shown that perceived diagnostic uncertainty in patients can negatively influence their subsequent beliefs, behaviours and outcomes (Serbic and Pincus 2014). GPs also perceive they gain trust from their patients who consult with musculoskeletal pain by giving a diagnosis or referring for tests (Parsons et al. 2007). Using a single recognised phrase or word to describe a combination of symptoms and clinical findings, can assist effective communication between clinicians and their patients and with other healthcare professionals (Croft et al. 2015). Giving patients a clear informed diagnosis has the advantage of minimising the effect of "iatrogenic disability" where clinicians unintentionally cause persistent or worsening disability for their patients when patients are given inconsistent diagnoses by different health professionals. To make this work, the healthcare community also needs to agree on the "recognised phrase or word" to describe leg pain due to nerve root involvement when giving a diagnosis.

Evidence suggests that classifying patients based on signs and symptoms, and matching treatment accordingly, seems to produce better outcomes when compared with treatment not matched to a classification method (Childs et al. 2004, Fritz et al. 2007, Wilgenbusch et al. 2014, Ford et al. 2016). Many of the clinicians at the workshop felt that if LBLP patients were allocated to one of the five classes identified in the latent class modelling, it could be helpful to both patients and clinicians to signpost patients to appropriate services. This was particularly in relation to patients in class 5 whom they felt may require more of a cognitive behavioural therapy approach to

address their pain management and could be directed to specific clinicians who could deliver such an intervention. Matching treatment to the classification groups requires further work and is an area for future research study.

9.5 Recommendations for future research

The reliability study showed that when clinicians' confidence in diagnosis is lower, they are less likely to agree on a diagnosis of LBLP. The subsequent stage of the thesis developed a clinical diagnostic tool that identified items that would help discriminate between referred leg pain and sciatica. The next logical step is to repeat the reliability study and evaluate if reliability improved when clinicians use the clinical diagnostic tool to assist their diagnostic decision.

The systematic review highlighted the wide variation in clinical criteria used to describe patients classified with sciatica. This mirrored findings from previous reviews on eligibility criteria in studies that involved patients with sciatica (Genevay et al. 2010, Lin et al. 2014). To facilitate research on homogenous groups, as recommended by LBP researchers in primary care (Costa et al. 2013), the clinical assessment items identified in the diagnostic tool could be used to select eligible patients with sciatica for participation in trials or other studies. The LC modelling revealed five distinct groups of LBLP; these classes could also be taken into consideration when selecting participants for intervention studies, as the classes may represent groups likely to need a different management approach.

Would use of the clinical diagnostic tool in clinical settings influence clinical practice, for example reduce imaging referrals, or enhance patient outcomes in any way? Before this could be considered, external validation of the tool is necessary in another LBLP population using the same clinical assessment variables. This work is currently being planned and collaborations are being developed with the University of Southern Denmark and the FORMI institute in Oslo Norway, both of whom have LBLP cohorts and have collected similar clinical assessment items.

Perhaps the most exciting prospect for future research is to capture change over time in key outcomes, such as pain and disability, for the LBLP classes identified with latent class modelling. The thesis describes the classes using cross-sectional data; prognostic information requires a longitudinal design. Such longitudinal studies would help to shed more light on the clinical course of these classes. Results from such longitudinal analysis could help to inform the design of further potential intervention studies. Replicating the LC modelling in this thesis, in similar LBLP cohorts (Formi Oslo, Norway), is also planned to see if similar classes are observed.

Overall, literature suggests that outcomes of sciatica are reasonably favourable but not as good as that of LBP alone, especially in cases of severe symptoms (Koes et al. 2007, Lequin et al. 2013). The spectrum of LBLP presentations seen in the latent classes suggests that this poor outcome in more severe cases may be for different reasons. Both class 4 ("severe sciatica") and class 5 ("atypical sciatica") had high back and leg pain intensity and disability. Based on the clinical presentation and associated characteristics, it is possible that their management should be different. For example, some of the clinicians attending the workshop suggested that class 4 are a group they would want to promptly refer onwards to a more specialist spinal/interface service for further evaluation or opinion on spinal injection or surgery.

At the moment, this is speculation based on observation of cross-sectional data, but further work is planned with clinicians and LBP researchers to discuss potential treatment strategies for these classes. This will be aided by knowledge of their prognosis, bearing in mind that the participants in the ATLAS cohort were well managed in primary care with prompt onward referral to physiotherapy and secondary care services where needed.

Currently the approach to management of sciatica is a stepped care approach in those who are not deteriorating or presenting with signs suggestive of sinister pathology, starting with non-invasive treatments and progressing to more invasive treatment options (Lewis et al. 2011). Similarly, moving up the analgesic ladder with medication management is recommended and

offering drugs to treat neuropathic pain if the “pain remains uncontrolled” (NICE Clinical knowledge summary on sciatica, 2015). Timing and when to move to the next step in a stepped approach to sciatica is not clear, particularly in those with higher pain levels (Lee et al. 2013). Possibly this stepped approach could be counter-productive for some groups of sciatica patients and a more aggressive approach may be needed earlier in the presentation. In recent work with patient user groups at Keele University in preparation for designing a trial on stratified care for sciatica, all felt that early pain relief is a key outcome.

9.6 Conclusion

There is no current agreement on which features better identify sciatica in patients with LBLP. This perhaps explains why recognising LBLP patients with sciatica can be difficult at times, especially in the primary care setting. A diagnostic model, developed as part of this thesis, that performs well in a primary care cohort may help with this diagnostic decision. Further classification of LBLP revealed distinct classes that reflect the spectrum of primary care LBLP and represent groups that may have different clinical courses.

This PhD thesis set out to explore the clinical presentation of a cohort of LBLP patients in sufficient detail, with a view to providing shape and guidance for future studies to test treatment effectiveness and prognosis for LBLP consulters. The results have given plenty of scope for exciting future research to help improve our understanding of LBLP and improve our ability to recognise these patients, guide management and inform on prognosis, to ultimately improve outcomes for patients with LBLP.

References

Aggarwal, R., Ringold, S., Khanna, D., Neogi, T., Johnson, S. R., Miller, A., Brunner, H. I., Ogawa, R., Felson, D., Ogdie, A., Aletaha, D., Feldman, B. M. (2015). Distinctions between diagnostic and classification criteria? *Arthritis Care and Research,* 67(7), 891-7.

Al Nezari, N. H., Schneiders, A. G., Hendrick, P. A. (2013). Neurological examination of the peripheral nervous system to diagnose lumbar spinal disc herniation with suspected radiculopathy: a systematic review and meta-analysis. *The Spine Journal,* 13(6), 657-74.

Albert, H., Sorensen, J., Christensen, B., Manniche, C. (2013). Antibiotic treatment in patients with chronic low back pain and vertebral bone edema (Modic type 1 changes): a double-blind randomized clinical controlled trial of efficacy. *European Spine Journal,* 22(4), 697-707.

Albert, H. B., Hauge, E., Manniche, C. (2012). Centralization in patients with sciatica: are pain responses to repeated movement and positioning associated with outcome or types of disc lesions? *European Spine Journal,* 21(4), 630-6.

Alyas, F., Connell, D., Saifuddin, A. (2008). Upright positional MRI of the lumbar spine. *Clinical Radiology,* 63(9), 1035-48.

Andersson, G. B. (1997). Low back pain. *Journal of Rehabilitation Research and Development,* 34(4), ix-x.

Atlas, S. J., Deyo, R. A., Patrick, D. L., Convery, K., Keller, R. B., Singer, D. E. (1996). The Quebec Task Force classification for Spinal Disorders and the severity, treatment, and outcomes of sciatica and lumbar spinal stenosis. *Spine, 21*(24), 2885-92.

Atlas, S. J., Keller, R. B., Wu, Y. A., Deyo, R. A., Singer, D. E. (2005). Long-term outcomes of surgical and nonsurgical management of sciatica secondary to a lumbar disc herniation: 10 year results from the maine lumbar spine study. *Spine,* 30(8), 927-35.

Attal, N., Perrot, S., Fermanian, J., Bouhassira, D. (2011). The neuropathic components of chronic low back pain: a prospective multicenter study using the DN4 Questionnaire. *Journal of Pain,* 12(10), 1080-87.

Austin, P. C., Steyerberg, E. W. (2014). Graphical assessment of internal and external calibration of logistic regression models by using loess smoothers. *Statistics in Medicine,* 33(3), 517-35.

Austin, P. C., Tu, J. V., Ho, J. E., Levy, D., Lee, D. S. (2013). Using methods from the data-mining and machine-learning literature for disease classification and prediction: a case study examining classification of heart failure subtypes. *Journal of Clinical Epidemiology,* 66(4), 398-407.

Bachmann, L. M., Jüni, P., Reichenbach, S., Ziswiler, H.-R., Kessels, A. G., Vögelin, E. (2005). Consequences of different diagnostic 'gold standards' in test accuracy research: Carpal Tunnel Syndrome as an example. *International Journal of Epidemiology,* 34(4), 953-55.

Barker, M. E. (1990). A practical classification of spinal pains based on a study of patients seen in British general practice over a five year period. In: Fairbank J. C., Pynsent, P. B. (eds) *Back pain: classification of syndromes.* Manchester, Manchester University Press**.**

Baron, R., Binder, A., Attal, N., Casale, R., Dickenson, A. H., Treede, R. D. (2016). Neuropathic low back pain in clinical practice. *European Journal of Pain,* 20(6), 861-73.

Barons, M. J., Griffiths, F. E., Parsons, N., Alba, A., Thorogood, M., Medley, G. F., Lamb, S. E. (2014). Matching patients to an intervention for back pain: classifying patients using a latent class approach. *Journal of Evaluation in Clinical Practice,* 20(4), 544-50.

Beattie, P. F., Meyers, S. P., Stratford, P., Millard, R. W., Hollenberg, G. M. (2000). Associations between patient report of symptoms and anatomic impairment visible on lumbar magnetic resonance imaging. *Spine,* 25(7), 819-28.

Bendebba, M., Torgerson, W. S., Long, D. M. (2000). A validated, practical classification procedure for many persistent low back pain patients. *Pain,* 87(1), 89-97.

Bennett, M. I., Smith, B. H., Torrance, N., Potter, J. (2005). The S-LANSS score for identifying pain of predominantly neuropathic origin: validation for use in clinical and postal research. *The Journal of Pain,* 6(3), 149-58.

Bernard, Jr., T.N, Kirkaldy-Willis, W. H. (1987). Recognizing specific characteristics of nonspecific low back pain. *Clinical Orthopaedics and Related Research,* 217, 266-80.

Billis, E. V., Mccarthy, C. J., Oldham, J. A. (2007). Subclassification of low back pain: a cross-country comparison. *European Spine Journal,* 16(7), 865-79.

Boden, S. D., Davis, D. O., Dina, T. S., Patronas, N. J., Wiesel, S. W. (1990). Abnormal magnetic-resonance scans of the lumbar spine in asymptomatic subjects. A prospective investigation. *Journal of Bone and Joint Surgery American,* 72(3), 403-8.

Bogduk, N. (2009). On the definitions and physiology of back pain, referred pain, and radicular pain. *Pain,* 147(1-3), 17-19.

Bogduk, N., Mcguirk, B. (2002). *Medical management of acute and chronic low back pain: an evidence-based approach.* Amsterdam, Elsevier.

Boos, N., Semmer, N., Elfering, A., Schade, V., Gal, I., Zanetti, M., Kissling, R., Buchegger, N., Hodler, J., Main, C. J. (2000). Natural history of individuals with asymptomatic disc abnormalities in magnetic resonance imaging: predictors of low back pain-related medical consultation and work incapacity. *Spine,* 25(12), 1484-92.

Borkan, J. M., Koes, B., Reis, S., Cherkin, D. C. (1998). A report from the Second International Forum for Primary Care Research on Low Back Pain. Reexamining priorities. *Spine,* 23(18), 1992-6.

Brisby, H., Balagué, F., Schafer, D., Sheikhzadeh, A., Lekman, A., Nordin, M., Rydevik, B., Fredman, P. (2002). Glycosphingolipid antibodies in serum in patients with sciatica. *Spine,* 27(4), 380-6.

Buchbinder, R., Blyth, F. M., March, L. M., Brooks, P., Woolf, A. D., Hoy, D. G. (2013). Placing the global burden of low back pain in context. *Best Practice & Research. Clinical Rheumatology,* 27(5), 575-89.

Buchbinder, R., Goel, V., Bombardier, C., Hogg-Johnson, S. (1996). Classification systems of soft tissue disorders of the neck and upper limb: Do they satisfy methodological guidelines? *Journal of Clinical Epidemiology,* 49(2), 141-9.

Burton, A. K., Tillotson, K. M., Main, C., Hollis, S. (1995). Psychosocial predictors of outcome in acute and subchronic low back trouble. *Spine,* 20(6), 722-8.

Bush, K., Cowan, N., Katz, D. E., Gishen, P. (1992). The natural history of sciatica associated with disc pathology. A prospective study with clinical and independent radiologic follow-up. *Spine,* 17(10), 1205-12.

Butler D. S. (2000). *The Sensitive nervous system.* Adelaide, Noigroup Publications.

Byrt, T; Bishop, J; Carlin, J (1993). Bias, prevalence and kappa. *Journal of Clinical Epidemiology,* 46(5): 423-429.

Cassisi, J. E., Sypert, G. W., Lagana, L., Friedman, E. M., Robinson, M. E. (1993). Pain, disability, and psychological functioning in chronic low back pain subgroups: myofascial versus herniated disc syndrome. *Neurosurgery,* 33(3), 379-85.

Chandrasoma P., Taylor C. R. (1998). *Concise Pathology*. New York, McGraw Hill.

Cherkin, D. C., Deyo, R. A., Street, J. H., Barlow, W. (1996). Predicting poor outcomes for back pain seen in primary care using patients' own criteria. *Spine,* 21(24), 2900-7.

Childs, J. D., Fritz, J. M., Flynn, T. W., Irrgang, J. J., Johnson, K. K., Majkowski, G. R., Delitto, A. (2004). A clinical prediction rule to identify patients with low back pain most likely to benefit from spinal manipulation: A validation study. *Annals of Internal Medicine,* 141(12), 920-8.

Chou, R., Qaseem, A., Owens, D. K., Shekelle, P., Clinical Guidelines Committee of the American College of Physicians. (2011). Diagnostic imaging for low back pain: advice for high-value health care from the American College of Physicians. *Annals of Internal Medicine,* 154(3), 181-9.

Chou, R., Qaseem, A., Snow, V., Casey, D., Cross Jr, J. T., Shekelle, P., Owens, D. K. (2007). Diagnosis and treatment of low back pain: A Joint Clinical Practice Guideline from the American College of Physicians and the American Pain Society. *Annals of Internal Medicine,* 147(7), 478-W120.

Cicchetti, D and Feinstein, A (1990). High agreement but low kappa: II. resolving the paradoxes. *Journal of Clinical Epidemiology,* 43(6): 551-558.

Clark, D. B., Jones, B. L., Wood, D. S., Cornelius, J. R. (2006). Substance use disorder trajectory classes: diachronic integration of onset age, severity, and course. *Addictive Behaviors,* 31(6), 995-1009.

Coggon, D., Martyn, C., Palmer, K. T., Evanoff, B. (2005). Assessing case definitions in the absence of a diagnostic gold standard. *International Journal of Epidemiology,* 34(4), 949-52.

Coleman, T. (2000). Using video-recorded consultations for research in primary care: advantages and limitations. *Family Practice,* 17(5), 422-7.

Collins, L. M., Lanza, S. T. (2010). *Latent class and latent transition analysis: With applications in the social, behavioural, and health sciences.* New York, John Wiley & Sons.

Costa Lda, C., Koes, B. W., Pransky, G., Borkan, J., Maher, C. G., Smeets, R. J. (2013). Primary care research priorities in low back pain: an update. *Spine,* 38(2), 148-56.

Coster, S., De Bruijn, S. F., Tavy, D. L. (2010). Diagnostic value of history, physical examination and needle electromyography in diagnosing lumbosacral radiculopathy. *Journal of Neurology,* 257(3), 332-37.

Croft, P., Altman, D. G., Deeks, J. J., Dunn, K. M., Hay, A. D., Hemingway, H., Leresche, L., Peat, G., Perel, P., Petersen, S. E., Riley, R. D., Roberts, I., Sharpe, M., Stevens, R. J., Van Der Windt, D. A., Von Korff, M., Timmis, A. (2015). The science of clinical practice: disease diagnosis or patient prognosis? Evidence about "what is likely to happen" should shape clinical practice. *BMC Medicine,* 13:20.

Dagenais, S., Tricco, A. C., Haldeman, S. (2010). Synthesis of recommendations for the assessment and management of low back pain from recent clinical practice guidelines. *The Spine Journal,* 10(6), 514-29.

Dankaerts, W., O'Sullivan, P. B., Straker, L.M., Burnett, A.F., Skouen, J. S. (2006). The inter-examiner reliability of a classification method for non-specific chronic low back pain patients with motor control impairment. *Manual Therapy,* 11, 28-39.

De Graaf, I., Prak, A., Bierma-Zienstra, S., Thomas, S., Peul W., Koes, B. (2006) Diagnosis of lumbar spinal stenosis. A systematic review of the accuracy of diagnostic tests. *Spine* 31(10), 1168-76.

De Winter, A. F., Jans, M. P., Scholten, R. J., Deville, W., Van Schaardenburg, D., Bouter, L. M. (1999). Diagnostic classification of shoulder disorders: interobserver agreement and determinants of disagreement. *Annals of the Rheumatic Diseases,* 58(5), 272-77.

Deeks, J. J., Altman, D. G. (2004). Diagnostic tests 4: likelihood ratios. *British Medical Journal,* 329(7458), 168-9.

Delitto, A., Erhard, R. E., Bowling, R. W. (1995). A treatment-based classification approach to low back syndrome: identifying and staging patients for conservative treatment. *Physical Therapy,* 75(6), 470-85.

Delitto A., George, S. Z., Van Dillen, L. R., Whitman, J. M., Sowa, G., Shekelle, P., Denninger, T. R., Godges, J. J. (2012). Low back pain. *Journal of Orthopaedic and Sports Physical Therapy,* 42(4), A1-A57.

Delwaide, P. J., Toulouse, P. (1981). Facilitation of monosynaptic reflexes by voluntary contraction of muscle in remote parts of the body. Mechanisms involved in the Jendrassik Manoeuvre. *Brain,* 104(Pt 4), 701-9.

DeRosa, C. P., Porterfield, J. A. (1992). A physical therapy model for the treatment of low back pain. *Physical Therapy,* 72(4), 261-9.

Deyo, R. A., Bryan, M., Comstock, B. A., Turner, J. A., Heagerty, P., Friedly, J., Avins, A. L., Nedeljkovic, S. S., Nerenz, D. R., Jarvik, J. G. (2015). Trajectories of symptoms and function in older adults with low back disorders. *Spine,* 40(170), 1352-62.

Deyo, R. A., Weinstein, J. N. (2001). Low back pain. *New England Journal of Medicine,* 344(5), 363-70.

Dionne, C. E., Dunn, K. M., Croft, P. R., Nachemson, A. L., Buchbinder, R., Walker, B. F., Wyatt, M., Cassidy, J. D., Rossignol, M., Leboeuf-Yde, C., Hartvigsen, J., Leino-Arjas, P., Latza, U., Reis, S., Gil Del Real, M. T., Kovacs, F. M., Oberg, B., Cedraschi, C., Bouter, L. M., Koes, B. W., Picavet, H. S., Van Tulder, M. W., Burton, K., Foster, N. E., Macfarlane, G. J., Thomas, E., Underwood, M., Waddell, G., Shekelle, P., Volinn, E., Von Korff, M. (2008). A consensus approach toward the standardization of back pain definitions for use in prevalence studies. *Spine,* 33(1), 95-103.

Dunn, K. M., Campbell, P., Jordan, K. P. (2013). Long-term trajectories of back pain: cohort study with 7-year follow-up. *BMJ Open,* 3(12), e003838.

Dunn, K. M., Croft, P. R. (2006). The importance of symptom duration in determining prognosis. *Pain,* 121(1-2), 126-32.

Dunn, K. M., Jordan, K. P., Croft, P. R. (2010). Recall of medication use, self-care activities and pain intensity: a comparison of daily diaries and self-report questionnaires among low back pain patients. *Primary Health Care Research & Development,* 11(01), 93-102.

El Barzouhi, A., Vleggeert-Lankamp, C., Lycklama Ã Nijeholt, G. J., Van, D. K., Van, D. H., Koes, B. W., Peul, W. C. (2014). Influence of low back pain and prognostic value of MRI in sciatica patients in relation to back pain. *PLoS One,* 9(3), e90800.

El Sissi, W., Arnaout, A., Chaarani, M. W., Fouad, M., El Assuity, W., Zalzala, M., Dershaby, Y. E., Youseif, E. (2010). Prevalence of neuropathic pain among patients with chronic low-back pain in the Arabian Gulf Region assessed using the leeds assessment of neuropathic symptoms and signs pain scale. *The Journal of International Medical Research,* 38(6), 2135-45.

EuroQuol Group. (1990). EuroQol - a new facility for the measurement of health-related quality of life. *Health Policy*, 16(3), 199-208.

EuroQuol Research Foundation (2016). http://www.euroqol.org/ (accessed 08.02.2016).

Eusebi, P. (2013). Diagnostic accuracy measures. *Cerebrovascular Diseases,* 36(4), 267-72.

Fairbank, J. (2007). Sciatica: an archaic term. *British Medical Journal,* 335(7611), 112.

Fairbank, J., Gwilym, S. E., France, J. C., Daffner, S. D., Dettori, J., Hermsmeyer, J., Andersson, G. (2011). The role of classification of chronic low back pain. *Spine,* 36(21 Suppl), S19-42.

Felson, D. T., Anderson, J. J. (1995). Methodological and statistical approaches to criteria development in rheumatic diseases. *Classification and Assessment of Rheumatic Diseases: Part I,* 9(2), 253-266.

Field, A. (2005). *Discovering Statistics using SPSS* 3rd edition. London, Sage Publications.

Finch, W. H., Bronk, K. C. (2011). Conducting Confirmatory Latent Class Analysis Using Mplus. *Structural Equation Modeling: A Multidisciplinary Journal*, 18(1), 132-151.

Ford, J., Story, I., O' Sullivan, P., Mcmeeken, J. (2007). Classification systems for low back pain: a review of the methodology for development and validation. *Physical Therapy Reviews,* 12(1), 33-42.

Ford, J. J., Hahne, A. J., Surkitt, L. D., Chan, A. Y., Richards, M. C., Slater, S. L., Hinman, R. S., Pizzari, T., Davidson, M., Taylor, N. F. (2016). Individualised physiotherapy as an adjunct to guideline-based advice for low back disorders in primary care: a randomised controlled trial. *British Journal of Sports Medicine,* 50(4), 237-45.

Foster, N. E., Thomas, E., Bishop, A., Dunn, K. M., Main, C. J. (2010). Distinctiveness of psychological obstacles to recovery in low back pain patients in primary care. *Pain,* 148(3), 398-406.

Fourney, D. R., Dettori, J. R., Hall, H., Hartl, R., Mcgirt, M. J., Daubs, M. D. (2011). A systematic review of clinical pathways for lower back pain and introduction of the Saskatchewan spine pathway. *Spine,* 36(21 Suppl), S164-S171.

Frank, A. O., De Souza, L. H., McAuley, J. H., Sharma, V., Main, C. J. (2000). A cross-sectional survey of the clinical and psychological features of low back pain and consequent work handicap: use of the Quebec Task Force classification. *International Journal of Clinical Practice,* 54(10), 639-45.

Fransen, M., Woodward, M., Norton, R., Coggan, C., Dawe, M., Sheridan, N. (2002). Risk factors associated with the transition from acute to chronic occupational back pain. *Spine,* 27(1), 92-8.

Freynhagen, R., Rolke, R., Baron, R., Tölle, T. R., Rutjes, A.-K., Schu, S., Treede, R.-D. (2008). Pseudoradicular and radicular low-back pain – A disease continuum rather than different entities? Answers from quantitative sensory testing. *Pain,* 135(1-2), 65-74.

Fritz, J. M., Delitto, A., Vignovic, M., Busse, R. G. (2000). Interrater reliability of judgments of the centralization phenomenon and status change during movement testing in patients with low back pain. *Archives of Physical Medicine and Rehabilitation,* 81(1), 57-61.

Fritz, J. M., Lindsay, W., Matheson, J. W., Brennan, G. P., Hunter, S. J., Moffit, S. D., Swalberg, A., Rodriquez, B. (2007). Is there a subgroup of patients with low back pain likely to benefit from mechanical traction? Results of a randomized clinical trial and subgrouping analysis. *Spine,* 32, E793-800.

Ganko, R., Rao, P. J., Phan, K., Mobbs, R. J. (2015). Can bacterial infection by low virulent organisms be a plausible cause for symptomatic disc degeneration? A systematic review. *Spine,* 40(10), E587-92.

Garfin, S. R., Rydevik, B. L., Brown, R. A. (1991). Compressive neuropathy of spinal nerve roots. A mechanical or biological problem? *Spine,* 16(2), 162-6.

Genevay, S., Atlas, S. J., Katz, J. N. (2010). Variation in eligibility criteria from studies of radiculopathy due to a herniated disc and of neurogenic claudication due to lumbar spinal stenosis: A structured literature review. *Spine,* 35(7), 803-811.

Genevay, S., Finckh, A., Mezin, F., Tessitore, E., Guerne, P. A. (2009). Influence of cytokine inhibitors on concentration and activity of MMP-1 and MMP-3 in disc herniation. *Arthritis Research and Therapy,* 11(6), R169.

Germon, T., Singleton, W., Hobart, J. (2014). Is NICE guidance for identifying lumbar nerve root compression misguided? *European Spine Journal,* 23(Suppl 1), S20-4.

Glassman, S. D., Carreon, L. Y., Anderson, P. A., Resnick, D. K. (2011). A diagnostic classification for lumbar spine registry development. *The Spine Journal* 11(12), 1108-16.

Green, D. J., Jordan, K. P., Protheroe, J., Van Der Windt, D. (2015). Development of hand phenotypes and changes in hand pain and problems over time in older people. *Pain,* 157(3), 569-76.

Grimes, D. A., Schulz, K. F. (2005). Refining clinical diagnosis with likelihood ratios. *Lancet,* 365(9469), 1500-5.

Groenier, K. H., Winters, J. C., Van Schuur, W. H., De Winter, A. F., Meyboom-De Jong, B. (2006). A simple classification system was recommended for patients with restricted shoulder or neck range of motion. *Journal of Clinical Epidemiology,* 59(6), 599-607.

Grotle, M., Brox, J. I., Veierod, M. B., Glomsrod, B., Lonn, J. H., Vollestad, N. K. (2005). Clinical course and prognostic factors in acute low back pain: Patients consulting primary care for the first time. *Spine,* 30(8), 976-82.

Grovle, L., Haugen, A. J., Keller, A., Natvig, B., Brox, J. I., Grotle, M. (2010). The bothersomeness of sciatica: patients' self-report of paresthesia, weakness and leg pain. *European Spine Journal,* 19(2), 263-9.

Gwet, K. (2002) Kappa statistic is not satisfactory for assessing the extent of agreement between raters. *Series: Statistical methods for Inter-rater Reliability Assessment* 1(6): 1-6.

Hahne, A. J., Ford, J. J., Surkitt, L. D., Richards, M. C., Chan, A. Y., Thompson, S. L., Hinman, R. S., Taylor, N. F. (2011). Specific treatment of problems of the spine (STOPS): design of a randomised controlled trial comparing specific physiotherapy versus advice for people with subacute low back disorders. *BMC Musculoskeletal Disorders,* 12: 104.

Hall, H., McIntosh, G., Melles, T. (1994). A different approach to back pain diagnosis: identifying a pattern of pain. *The Canadian journal of Continuing Medical Education*, 6, 31-41.

Hall, H., McIntosh, G., Boyle, C. (2009). Effectiveness of a low back pain classification system. *The Spine Journal*, 9(8), 648-57.

Hallgren KA (2012) Computing inter-rater reliability for observational data: An overview and tutorial. *Tutorials in Quantitative Methods for Psychology,* 8(1): 23–34.

Harrell, F. E. (2001). *Regression modeling strategies: with applications to linear models, logistic regression, and survival analysis.* New York, Springer.

Hassey, A., Gerrett, D., Wilson, A. (2001). A survey of validity and utility of electronic patient records in a general practice. *British Medical Journal,* 322(7299), 1401-5.

Haswell, K., Gilmour, J., Moore, B. (2008). Clinical decision rules for identification of low back pain patients with neurologic involvement in primary care. *Spine,* 33(1), 68-73.

Haugen, A. J., Brox, J. I., Grovle, L., Keller, A., Natvig, B., Soldal, D., Grotle, M. (2012). Prognostic factors for non-success in patients with sciatica and disc herniation. *BMC Musculoskeletal Disorders,* 13: 183.

Hearne, M. A. (1997). Physical therapy treatment of low back pain: A report of outcomes according to three types of patients classification. *Journal of Rehabilitation Outcome Measures,* 1, 23-31.

Hefford, C. (2008). McKenzie classification of mechanical spinal pain: Profile of syndromes and directions of preference. *Manual Therapy,* 13(1), 75-81.

Heikkilä, J. K., Koskenvuo, M., Heliovaara, M., Kurppa, K., Riihimäki, H., Heikkilä, K., Rita, H., Videman, T. (1989). Genetic and environmental factors in sciatica evidence from a nationwide panel of 9365 adult twin pairs. *Annals of Medicine,* 21(5), 393-8.

Hicks, G., Gaines, J., Shardell, M., Simonsick, E. (2008). Associations of back and leg pain with health status and functional capacity of older adults: findings from the retirement community back pain study. *Arthritis and Rheumatism,* 59(9), 1306-13.

Hill, J. C., Dunn, K. M., Lewis, M., Mullis, R., Main, C. J., Foster, N. E., Hay, E. M. (2008). A primary care back pain screening tool: identifying patient subgroups for initial treatment. *Arthritis and Rheumatism,* 59(5), 632.

Hill, J. C., Konstantinou, K., Egbewale, B. E., Dunn, K. M., Lewis, M., Van Der Windt, D. (2011a). Clinical outcomes among low back pain consulters with referred leg pain in primary care. *Spine,* 36(25), 2168-75.

Hill, J. C., Whitehurst, D. G., Lewis, M., Bryan, S., Dunn, K. M., Foster, N. E., Konstantinou, K., Main, C. J., Mason, E., Somerville, S., Sowden, G., Vohora, K., Hay, E. M. (2011b). Comparison of stratified primary care management for low back pain with current best practice (STarT Back): a randomised controlled trial. *Lancet,* 378(9802), 1560-71.

Hoehler, F.K. (2000) Bias and prevalence effects on kappa viewed in terms of sensitivity and specificity. *Journal of Clinical Epidemiology,* 53, 499-503.

Hopayian, K., Notley, C. (2014). A systematic review of low back pain and sciatica patients' expectations and experiences of health care. *Spine Journal,* 14(8), 1769-80.

Hosmer, D. W., Lemeshow, S. (2000). *Applied logistic regression*. 2nd edition, New York, Wiley.

Hunink, M., Weinstein, M. C., Wittenberg, E., Drummond, M., Pliskin, J., Wong, J., Glasziou, P. (2014). *Decision making in health and medicine Integrating evidence and values*. 2nd edition, Cambridge, University Press.

Iversen, T., Solberg, T. K., Romner, B., Wilsgaard, T., Nygaard, O., Waterloo, K., Brox, J. I., Ingebrigtsen, T. (2013). Accuracy of physical examination for chronic lumbar radiculopathy. *BMC Musculoskeletal Disorders,* 14: 206.

Jensen, M. C., Kelly, A. P., Brant-Zawadzki, M. N. (1994). MRI of degenerative disease of the lumbar spine. *Magnetic Resonance Quarterly,* 10(3), 173-90.

Jensen, T. S., Finnerup, N. B. (2014). Allodynia and hyperalgesia in neuropathic pain: clinical manifestations and mechanisms. *The Lancet Neurology,* 13(9), 924-935.

Johnson, S.R., Goek, O. N., Singh-Grewal, D., Vlad, S.C., Feldman, B.M., Felson, D.T., Hawker, G.A., Singh, J.A., Solomon, D.H. (2007). Classification criteria in rheumatic diseases: a review of methodologic properties. *Arthritis and Rheumatism*, 15(7), 1119-33.

Jordan, K. P., Kadam, U. T., Hayward, R., Porcheret, M., Young, C., Croft, P. (2010). Annual consultation prevalence of regional musculoskeletal problems in primary care: an observational study. *BMC Musculoskelet Disorders,* 11: 144.

Karppinen, J., Malmivaara, A., Tervonen, O., Paakko, E., Kurunlahti, M., Syrjala, P., Vasari, P., Vanharanta, H. (2001). Severity of symptoms and signs in relation to magnetic resonance imaging findings among sciatic patients. *Spine,* 26(7), E149-54.

Karppinen, J., Daavittila, I., Noponen, N., Haapea, M., Taimela, S., Vanharanta, H., Ala-Kokko, L., Mannikko, M. (2008). Is the interleukin-6 haplotype a prognostic factor for sciatica? *European Journal of Pain,* 12(8), 1018-25.

Katz, J. N., Stock, S. R., Evanoff, B. A., Rempel, D., Moore, J. S., Franzblau, A., Gray, R. H. (2000). Classification criteria and severity assessment in work-associated upper extremity disorders: methods matter. *American Journal of Industrial Medicine,* 38(4), 369-72.

Kellgren, J. H. (1939). On the distribution of pain arising from deep somatic structures with pain areas. *Clinical Science*, 4, 35 -46.

Kendall, F. P., McCreary, E. K., Provance, P.G. (2005). *Muscles, testing and function: with posture and pain.* Baltimore, Md, Lippincott Williams & Wilkins.

Kim, J. (2005). Bias corrected bootstrap inferences for regression models with auto-correlated errors. *Economics Bulletin*, 3(44), 1-8.

Kindrachuk, D. R., Fourney, D. R. (2014). Spine surgery referrals redirected through a multidisciplinary care pathway: effects of nonsurgeon triage including MRI utilization. *Journal of neurosurgery. Spine,* 20(1), 87-92.

Kittelson, A. J., Stevens-Lapsley, J. E., Schmiege, S. J. (2016). Determination of pain phenotypes in knee osteoarthritis: A latent class analysis using data from the Osteoarthritis Initiative. *Arthritis Care & Research*, 68(5), 612-20.

Kleinstück, F. S., Fekete, T., Jeszenszky, D., Mannion, A. F., Grob, D., Lattig, F., Mutter, U., Porchet, F. (2011). The outcome of decompression surgery for lumbar herniated disc is influenced by the level of concomitant preoperative low back pain. *European Spine Journal,* 20(7), 1166-73.

Kleinstück, F. S., Grob, D., Lattig, F., Bartanusz, V., Porchet, F., Jeszenszky, D., O' Riordan, D., Mannion, A. F. (2009). The influence of preoperative back pain on the outcome of lumbar decompression surgery. *Spine,* 34(11), 1198-203.

Knottnerus, J. A. (2002). *The evidence base of clinical diagnosis,* London, BMJ Books.

Koes, B. W., Tulder, M., Lin, C. W., Macedo, L. G., Mcauley, J., Maher, C. (2010). An updated overview of clinical guidelines for the management of non-specific low back pain in primary care. *European Spine Journal,* 19(12), 2075-94.

Koes, B. W., Van Tulder, M. W., Peul, W. C. (2007). Diagnosis and treatment of sciatica. *British Medical Journal,* 334(7607), 1313-7.

Kongsted, A., Kent, P., Albert, H., Jensen, T. S., Manniche, C. (2012). Patients with low back pain differ from those who also have leg pain or signs of nerve root involvement - a cross-sectional study. *BMC Musculoskeletal Disorders,* 13: 236.

Kongsted, A., Kent, P., Hestbaek, L., Vach, W. (2015). Patients with low back pain had distinct clinical course patterns that were typically neither complete recovery nor constant pain. A latent class analysis of longitudinal data. *The Spine Journal,* 15(5), 885-94.

Kongsted, A., Kent, P., Jensen, T. S., Albert, H., Manniche, C. (2013). Prognostic implications of the Quebec Task Force classification of back-related leg pain: an analysis of longitudinal routine clinical data. *BMC Musculoskeletal Disorders,* 14: 171.

Konno, S., Hayashino, Y., Fukuhara, S., Kikuchi, S., Kaneda, K., Seichi, A., Chiba, K., Satomi, K., Nagata, K., Kawai, S. (2007). Development of a clinical diagnosis support tool to identify patients with lumbar spinal stenosis. *European Spine Journal,* 16(11), 1951-7.

Konstantinou, K., Beardmore, R., Dunn, K. M., Lewis, M., Hider, S. L., Sanders, T., Jowett, S., Somerville, S., Stynes, S., Van Der Windt, D. A., Vogel, S., Hay, E. M. (2012a). Clinical course, characteristics and prognostic indicators in patients presenting with back and leg pain in primary care. The ATLAS study protocol. *BMC Musculoskeletal Disorders,* 13: 4.

Konstantinou, K., Dunn, K. M. (2008). Sciatica: review of epidemiological studies and prevalence estimates. *Spine,* 33(22), 2464-72.

Konstantinou, K., Dunn, K. M., Ogollah, R., Vogel, S., Hay, E. M., Atlas Study Research Team. (2015). Characteristics of patients with low back and leg pain seeking treatment in primary care: baseline results from the ATLAS cohort study. *BMC Musculoskeletal Disorders,* 16: 332.

Konstantinou, K., Hider, S. L., Vogel, S., Beardmore, R., Somerville, S. (2012b). Development of an assessment schedule for patients with low back-associated leg pain in primary care: a Delphi consensus study. *European Spine Journal,* 21(7), 1241-9.

Konstantinou, K., Lewis, M., Dunn, K. (2012c). Agreement of self-reported items and clinically assessed nerve root involvement (or sciatica) in a primary care setting. *European Spine Journal,* 21(11), 2306-15.

Kottner, J., Streiner, D. L. (2011). The difference between reliability and agreement. *Journal of Clinical Epidemiology,* 64(6), 701-2.

Kuijpers, T., Van Der Heijden, G. J., Vergouwe, Y., Twisk, J. W., Boeke, A. J., Bouter, L. M., Van Der Windt, D. A. (2007). Good generalizability of a prediction rule for prediction of persistent shoulder pain in the short term. *Journal of Clinical Epidemiology,* 60(9), 947-53.

Kumar, M., Garg, G., Singh, L. R., Sing, T., Tyagi, L. K. (2011). Epidemiology, pathophysiology and symptomatic treatment of sciatica: a review. *International Journal of Pharmaceutical and Biological Archives,* 2(4), 1050-1061.

Lacey, R. J., Strauss, V. Y., Rathod, T., Belcher, J., Croft, P. R., Natvig, B., Wilkie, R., Mcbeth, J. (2015). Clustering of pain and its associations with health in people aged 50 years and

older: cross-sectional results from the North Staffordshire Osteoarthritis Project. *BMJ Open,* 5(11), e008389.

Landis, J. R., Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*. 33(1), 159-74.

Laslett, M., Mcdonald, B., Tropp, H., Aprill, C. N., Oberg, B. (2005). Agreement between diagnoses reached by clinical examination and available reference standards: a prospective study of 216 patients with lumbopelvic pain. *BMC Musculoskeletal Disorders,* 6: 28.

Laslett, M., van Wijmen, P. (1999). Low back and referred pain: diagnosis and a proposed new system of classification. *New Zealand Journal of Physiotherapy*, 27(2), 5-15.

Lee, J., Gupta, S., Price, C., Baranowski, A. P., British Pain, S. (2013). Low back and radicular pain: a pathway for care developed by the British Pain Society. *British Journal of Anaesthesia,* 111(1), 112-20.

Leffler, A. S., Hansson, P. (2008). Letter to the Editor of Pain on Freynhagen et al.: Pseudoradicular and radicular low-back pain – A disease continuum rather than different entities? Answers from quantitative sensory testing. Pain 2007;135:65–74. *Pain,* 135(3), 312-3.

Lequin, M. B., Verbaan, D., Jacobs. W. C. H., Brand, R., Bouma, G. J., Vandertop, W. P., Peul, W. C. (2013). Surgery versus prolonged conservative treatment for sciatica: 5-year results of a randomised controlled trial. *BMJ Open*, 3(5).

Lewis, R., Williams, N., Matar, H. E., Din, N., Fitzsimmons, D., Phillips, C., Jones, M., Sutton, A., Burton, K., Nafees, S., Hendry, M., Rickard, I., Chakraverty, R., Wilkinson, C. (2011). The clinical effectiveness and cost-effectiveness of management strategies for sciatica: systematic review and economic model. *Health Technology Assessment,* 15(39), 1-578.

Li, Y., Fredrickson, V., Resnick, D. K. (2015). How should we grade lumbar disc herniation and nerve root compression? A systematic review. *Clinical Orthopaedics and Related Research,* 473(6), 1896-1902.

Lidbeck, J. (2002) Central hyperexcitability in chronic musculoskeletal pain: a conceptual breakthrough with multiple clinical implications. *Pain Research Management*, 7, 81-92.

Lin, C. W., Verwoerd, A. J., Maher, C. G., Verhagen, A. P., Pinto, R. Z., Luijsterburg, P. A., Hancock, M. J. (2014). How is radiating leg pain defined in randomized controlled trials of conservative treatments in primary care? A systematic review. *European Journal of Pain,* 18(4), 455-64.

Lindahl, O., Rexed, B. (1951). Histologic changes in spinal nerve roots of operated cases of sciatica. *Acta Orthopaedica Scandinavica,* 20(3), 215-25.

Loisel, P., Vachon, B., Lemaire, J., Durand, M. J., Poitras, S., Stock, S., Tremblay, C. (2002). Discriminative and predictive validity assessment of the Quebec Task Force classification. *Spine,* 27(8), 851-7.

Luomajoki, H., Kool, J., de Bruin, E., Airaksinen, O. (2007) Reliability of movement control tests in the lumbar spine. *BMC Musculoskeletal Disorders,* 8, 90.

Machado, L. a. C., De Souza, M. S., Ferreira, P. H., Ferreira, M. L. (2006). The McKenzie method for low back pain: a systematic review of the literature with a meta-analysis approach. *Spine,* 31(9), E254-62.

Magidson, J., Vermunt, J. K. (2002). Latent class models for clustering: A comparison with K-means. C*anadian Journal of Marketing Research*, 20, 37-44.

Maniadakis, N., Gray, A. (2000). The economic burden of back pain in the UK. *Pain,* 84(1), 95-103.

Marras, W. S., Parnianpour, M., Ferguson, S. A., Kim, J. Y., Crowell, R. R., Bose, S., Simon, S. R (1995). The classification of anatomic and symptom based low back disorders using motion measure models. *Spine,* 20(23), 2531-46.

Mathieson, S., Maher, C. G., Mclachlan, A. J., Latimer, J., Koes, B. W., Hancock, M. J., Harris, I., Day, R. O., Pik, J., Jan, S., Billot, L., Lin, C. W. (2013). PRECISE - pregabalin in addition to usual care for sciatica: study protocol for a randomised controlled trial. *Trials,* 14: 213.

May, S., Littlewood, C., Bishop, A. (2006). Reliability of procedures used in the physical examination of non-specific low back pain: a systematic review. *The Australian Journal of Physiotherapy,* 52(2), 91-102.

McCarthy, C. J., Arnall, F. A., Strimpakos, N., Freemont, A., Oldham, J. A. (2004). The biopsychosocial classification of non-specific low back pain: a systematic review. *Physical Therapy Reviews,* 9(1), 17-30.

McCarthy, C. J., Gittins, M., Roberts, C., Oldham, J. A. (2007). The reliability of the clinical tests and questions recommended in international guidelines for low back pain. *Spine,* 32(8), 921-6.

McCullough, B. J., Johnson, G. R., Martin, B. I., Jarvik, J. G. (2012). Lumbar MR imaging and reporting epidemiology: do epidemiologic data in reports affect clinical management? *Radiology,* 262(3), 941-6.

Mckenzie, R.A. (1981). *The Lumbar Spine: Mechanical Diagnosis and Therapy.* Waikanae, New Zealand, Spinal Publications.

Merskey, H., Bogduk, N. (1994). *Classification of chronic pain,* Seattle, IASP Press.

Midi, H., Sarkar, S. K., Rana, S. (2010) Collinearity diagnostics of binary logistic regression model. *Journal of Interdisciplinary Mathematics*, 13(3), 253-67.

Mikkonen, P., Heikkala, E., Paananen, M., Remes, J., Taimela, S., Auvinen, J., Karppinen, J. (2016). Accumulation of psychosocial and lifestyle factors and risk of low back pain in adolescence: a cohort study. *European Spine Journal,* 25(2), 635-42.

Miranda, H., Viikari-Juntura, E., Martikainen, R., Takala, E. P., Riihimaki, H. (2002). Individual factors, occupational loading, and physical exercise as predictors of sciatic pain. *Spine,* 27(10), 1102-9.

Mixter , W. J., Barr , J. S. (1934). Rupture of the intervertebral disc with involvement of the spinal canal. *New England Journal of Medicine,* 211(5), 210-15.

Moher, D., Shamseer, L., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., Shekelle, P., Stewart, L. A. (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Systematic Reviews,* 4(1), 1-9.

Mooney, V., Robertson, J. (1976). The facet syndrome. *Clinical Orthopaedics Related Research,* 115, 149-56.

Moons, K. G., Altman, D. G., Reitsma, J. B., Ioannidis, J. P., Macaskill, P., Steyerberg, E. W., Vickers, A. J., Ransohoff, D. F., Collins, G. S. (2015). Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): explanation and elaboration. *Annals of Internal Medicine,* 162(1), W1-73.

Moons, K. G., Kengne, A. P., Woodward, M., Royston, P., Vergouwe, Y., Altman, D. G., Grobbee, D. E. (2012). Risk prediction models: I. Development, internal validation, and assessing the incremental value of a new (bio)marker. *Heart,* 98(9), 683-90.

Moss-Morris, R., Weinman, J., Petrie, K., Horne, R., Cameron, L., Buick, D. (2002). The revised illness perception questionnaire (IPQ-R). *Psychology and Health,* 17(1), 1-16.

Moynihan, R., Doust, J., Henry, D. (2012). Preventing overdiagnosis: how to stop harming the healthy. *British Medical Journal,* 344, e3502.

Murphy, D. R., Hurwitz, E. L., Gerrard, J. K., .Clary, R. (2009). Pain patterns and descriptions in patients with radicular pain: Does the pain necessarily follow a specific dermatome? *Chiropractic & Osteopathy*, 17, 9.

Murray, C. J. L., Richards, M. A., Newton, J. N., Fenton, K. A., Anderson, H. R., Atkinson, C., Bennett, D., Bernabé, E., Blencowe, H., Bourne, R., Braithwaite, T., Brayne, C., Bruce, N. G., Brugha, T. S., Burney, P., Dherani, M., Dolk, H., Edmond, K., Ezzati, M., Flaxman, A. D., Fleming, T. D., Freedman, G., Gunnell, D., Hay, R. J., Hutchings, S. J., Ohno, S. L., Lozano, R., Lyons, R. A., Marcenes, W., Naghavi, M., Newton, C. R., Pearce, N., Pope, D., Rushton, L., Salomon, J. A., Shibuya, K., Vos, T., Wang, H., Williams, H. C., Woolf, A. D., Lopez, A. D., Davis, A. (2013). UK health performance: findings of the Global Burden of Disease Study 2010. *The Lancet,* 381(9871), 997-1020.

Muthen, L., Muthen, B. (2010) *MPlus User's Guide.* 6th edition. Los Angeles CA, Muthen & Muthen.

Muthen, B. (2004). Latent variable analysis: Growth mixture modelling and related techniques for longitudinal data. In Kaplan, D. (Ed.), *Handbook of quantitative methodology for the social sciences* (pp. 345–368). Newbury Park, CA, Sage.

Nachemson, A.L., Andersson, G. B. (1982) Classification of low back pain *Scandinavian Journal of Work Environment and Health* 8, 134-6.

Nazareth, I., Landau, S., Yardley, L., Luxon, L. (2006). Patterns of presentations of dizziness in primary care—a cross-sectional cluster analysis study. *Journal of Psychosomatic Research,* 60(4), 395-401.

National Institute for Health and Clinican Excellence (2016). Low back pain and sciatica, NICE in development (GID-CGWAVE0681). www.nice.org.uk/guidance/indevelopment/gid-cgwave0681 (Accessed 26/07/2016).

National Institute for Health and Clinican Excellence (2009). *Low back pain: early management of persistent non specific low back pain. (Clinical guidelines 88)*. London, Royal College of Physicians London.www.nice.org.uk/CG88. (Accessed 25/04/2016).

National Institute for Health and Clinican Excellence (2015). *Clinical knowledge summary on sciatica.* http://cks.nice.org.uk/sciatica-lumbar-radiculopathy#!scenario. Accessed 23/04/2016.

Nicholas, M. K. (2007). The pain self-efficacy questionnaire: Taking pain into account. *European Journal of Pain,* 11(2), 153-63.

Nijs, J., Apeldoorn, A., Hallegraeff, H., Clark, J., Smeets, R., Malfliet, A., Girbes, E. L., De Kooning, M., Ickmans, K. (2015). Low back pain: guidelines for the clinical classification of predominant neuropathic, nociceptive, or central sensitization pain. *Pain Physician,* 18(3), E333-46.

Nitta, H., Tajima, T., Sugiyama, H., Moriyama, A. (1993). Study on dermatomes by means of selective lumbar spinal nerve block. *Spine,* 18(13), 1782-6.

Nyland, K. L., Asparouhov, T., Muthen, B. O. (2007). Deciding on the number of classes in latent class analysis and growth mixture modelling: A Monte Carlo simulation study. Structural Equation Modeling, 14, 535-69.

Office of National Statistics. (2010). *Standard Occupational Classification 2010. Volume 1: Structure and description of unit.* Hampshire, Palgrave Macmillan.

O' Brien, M. R. (2007). A caution regarding rules of thumb for variance inflation factors. *Quality and Quantity,* 41(5), 673-90.

Office for National Statistics. (2016a). www.ons.gov.uk/peoplepopulationandcommunity (Accessed 11/07/2016).

Office for National Statistics. (2016b). SOC2010 volume 3: the National Statistics Socio-economic classification (NS-SEC rebased on SOC2010). http://webarchive.nationalarchives.gov.uk/20160105160709/http://www.ons.gov.uk/ons/guide-method/classifications/current-standard-classifications/soc2010/soc2010-volume-3-ns-sec--rebased-on-soc2010--user-manual/index.html#6 (Accessed 05/04/2016).

O'Hearn, M., Lowry, C., Emerson-Kavchak, A., Courtney, C. (2009). Letter to the editor Re: "Interrater reliability of a new classification system for patients with neural low back-related leg pain". Journal of Manual & Manipulative Therapy 17(2), 109-118. *Journal of Manual & Manipulative Therapy*, 4, E115-E116.

O'Neill, C. W., Kurgansky, M. E., Derby, R., Ryan, D. P. (2002). Disc stimulation and patterns of referred pain. *Spine,* 27(24), 2776-81.

Ohnmeiss, D. D., Vanharanta, H., Ekholm, J. (1997). Degree of disc disruption and lower extremity pain. *Spine,* 22(14), 1600-5.

Ohtori, S., Suzuki, M., Koshi, T., Takaso, M., Yamashita, M., Inoue, G., Yamauchi, K., Orita, S., Eguchi, Y., Kuniyoshi, K., Ochiai, N., Kishida, S., Nakamura, J., Aoki, Y., Ishikawa, T., Arai, G., Miyagi, M., Kamoda, H., Suzuki, M., Toyone, T., Takahashi, K. (2011). Proinflammatory

cytokines in the cerebrospinal fluid of patients with lumbar radiculopathy. *European Spine Journal,* 20(6), 942-6.

Patrick, D. L., Deyo, R. A., Atlas, S. J., Singer, D. E., Chapin, A., Keller, R. B. (1995). Assessing health-related quality of life in patients with sciatica. *Spine,* 20(17), 1899-908.

Paatelma, M., Karvonen, E., Heinonen, A. (2009). Inter-tester reliability in classifying acute and subacute low back pain patients into clinical subgroups: a comparison of specialists and non-specialists. A pilot study. *Journal of Manual & Manipulative Therapy, 17*(4), 221-30.

Parsons, S., Harding, G., Breen, A., Foster, N., Pincus, T., Vogel, S., Underwood, M. (2007). The influence of patients' and primary care practitioners' beliefs and expectations about chronic musculoskeletal pain on the process of care: a systematic review of qualitative studies. *Clinical Journal of Pain,* 23(1), 91-8.

Pearce, J. M. (2007). A brief history of sciatica. *Spinal Cord,* 45(9), 592-6.

Peduzzi, P., Concato, J., Kemper, E., Holford, T. R., Feinstein, A. R. (1996). A simulation study of the number of events per variable in logistic regression analysis. *Journal of Clinical Epidemiology,* 49(12), 1373-9.

Petersen, T., Laslett, M., Thorsen, H., Manniche, C., Ekdahl, C., Jacobsen, S. (2003). Diagnostic classification of non-specific low back pain. A new system integrating patho-anatomic and clinical categories. *Physiotherapy Theory and Practice,* 19(4), 213-37.

Petersen, T., Thorsen, H., Manniche, C., Ekdahl, C. (1999). Classification of non-specific low back pain: a review of the literature on classifications systems relevant to physiotherapy. *Physical Therapy Reviews,* 4(4), 265-82.

Petersen, T., Olsen, S., Laslett, M., Thorsen, H., Manniche, C., Ekdahl, C., & Jacobsen, S. (2004). Inter-tester reliability of a new diagnostic classification system for patients with non-specific low back pain. *Australian Journal of Physiotherapy, 50*(2), 85-94.

Petersen, T., Larsen, K., Nordsteen, J., Olsen, S., Fournier, G., Jacobsen, S. (2011). The McKenzie method compared with manipulation when used adjunctive to information and advice in low back pain patients presenting with centralization or peripheralization: a randomized controlled trial. *Spine,* 36(24), 1999-2010.

Peul, W. C., Van Houwelingen, H. C., Van Den Hout, W. B., Brand, R., Eekhof, J. A., Tans, J. T., Thomeer, R. T., Koes, B. W., for the Leiden-the Hague Spine Intervention Prognostic Study Group. (2007). Surgery versus prolonged conservative treatment for sciatica. *New England Journal of Medicine,* 356(22), 2245-56.

Porchet, F. O., Wietlisbach, V., Burnand, B., Daeppen, K., Villemure, J.-G., Vader, J. P. (2002). Relationship between severity of lumbar disc disease and disability scores in sciatica patients. *Neurosurgery,* 50(6), 1253-9.

Rathod, T., Marshall, M., Thomas, M. J., Menz, H. B., Myers, H. L., Thomas, E., Downes, T., Peat, G., Roddy, E. (2015). Investigation of potential phenotypes of foot osteoarthritis: Cross-sectional analysis from the Clinical Assessment Study of the Foot. *Arthritis Care & Research,* 68(2), 217-27.

Reitsma, J. B., Rutjes, A. W. S., Khan, K. S., Coomarasamy, A., Bossuyt, P. M. (2009). A review of solutions for diagnostic accuracy studies with an imperfect or missing reference standard. *Journal of Clinical Epidemiology,* 62(8), 797-806.

Riddle, D. L. (1998). Classification and low back pain: a review of the literature and critical analysis of selected systems. *Physical Therapy,* 78(7), 708-37.

Roach, K. E., Brown, M. D., Albin, R. D., Delaney, K. G., Lipprandi, H. M., Rangelli, D. (1997). The sensitivity and specificity of pain response to activity and position in categorizing patients with low back pain. *Physical Therapy,* 77(7), 730-8.

Roland, M., Morris, R. (1983). A study of the natural history of back pain: Part I: Development of a reliable and sensitive measure of disability in low-back pain. *Spine,* 8(2), 141-4.

Royston, P., Altman, D. G. (2010). Visualizing and assessing discrimination in the logistic regression model. *Statistics in Medicine,* 29(24), 2508-20.

Rutjes, A. W., Reitsma, J. B., Coomarasamy, A., Khan, K. S., Bossuyt, P. M. (2007). Evaluation of diagnostic tests when there is no gold standard. A review of methods. *Health Technology Assessment,* 11(50), iii, ix-51.

Saal, J. S., Franson, R. C., Dobrow, R., Saal, J. A., White, A. H., Goldthwaite, N. (1990). High levels of inflammatory phospholipase A2 activity in lumbar disc herniations. *Spine,* 15(7), 674-8.

Schäfer, A., Hall, T., Briffa, K. (2009a). Classification of low back-related leg pain—A proposed patho-mechanism-based approach. *Manual Therapy,* 14(2), 222-31.

Schäfer, A., Hall, T., Mueller, G., Briffa, K. (2011). Outcomes differ between subgroups of patients with low back and leg pain following neural manual therapy: a prospective cohort study. *European Spine Journal,* 20(3), 482-90.

Schäfer, A., Hall, T. M., Lüdtke, K., Mallwitz, J., Briffa, N. K. (2009b). Interrater reliability of a new classification system for patients with neural low back-related leg pain. *Journal of Manual & Manipulative Therapy,* 17(2), 109-17.

Schäfer, A. G., Hall, T. M., Rolke, R., Treede, R. D., Ludtke, K., Mallwitz, J., Briffa, K. N. (2014). Low back related leg pain: an investigation of construct validity of a new classification system. *Journal of Back and Musculoskeletal Rehabilitation,* 27(4), 409-18.

Schepper, E. I. T., Koes, B. W., Oei, E. H. G., Bierma-Zeinstra, S., Luijsterburg, P. a. J. (2016). The added prognostic value of MRI findings for recovery in patients with low back pain in primary care: a 1-year follow-up cohort study. *European Spine Journal,* 25(4), 1234-41.

Scholz, J., Mannion, R. J., Hord, D. E., Griffin, R. S., Rawal, B., Zheng, H., Scoffings, D., Phillips, A., Guo, J., Laing, R. J. C., Abdi, S., Decosterd, I., Woolf, C. J. (2009). A novel tool for the assessment of pain: Validation in low back pain. *PLoS Medicine,* 6(4), e1000047.

Selim, A. J., Ren, X. S., Fincke, G., Deyo, R. A., Rogers, W., Miller, D., Linzer, M., Kazis, L. (1998). The importance of radiating leg pain in assessing health outcomes among patients with low back pain: Results from the Veterans Health Study. *Spine,* 23(4), 470-4.

Serbic, D., Pincus, T. (2014). Diagnostic uncertainty and recall bias in chronic low back pain. *Pain,* 155(8), 1540-6.

Shaw, W. S., Pransky, G., Fitzgerald, T. E. (2001). Early prognosis for low back disability: intervention strategies for health care providers. *Disability and Rehabilitation,* 23(18), 815-28.

Shultz, S., Averell, K., Eickelman, A., Sanker, H., Donaldson, M. B. (2015). Diagnostic accuracy of self-report and subjective history in the diagnosis of low back pain with non-specific lower extremity symptoms: A systematic review. *Manual Therapy,* 20(1), 18-27.

Sim, J., Wright, C. C. (2005). The kappa statistic in reliability studies: use, interpretation, and sample size requirements. *Physical Therapy,* 85(3), 257-68.

Smart, K. M., Blake, C., Staines, A., Doody, C. (2011). The discriminative validity of "nociceptive", "peripheral neuropathic" and "central sensitization" as mechanisms-based classifications of musculoskeletal pain. *Clinical Journal of Pain,* 27(8), 655-63.

Smart, K. M., Curley, A., Blake, C., Staines, A., Doody, C. (2010). The reliability of clinical judgments and criteria associated with mechanisms-based classifications of pain in patients with low back pain disorders: a preliminary reliability study. *Journal of Manual & Manipulative Therapy,* 18(2), 102-10.

Smart, K. M., Blake, C., Staines, A., Thacker, M., Doody C. (2012a) Mechanisms-based classifications of musculoskeletal pain: Part 2 of 3: Symptoms and signs of peripheral neuropathic pain in patients with low back (±leg) pain. *Manual Therapy,* 17(4), 345-51.

Smart, K. M., Blake, C., Staines, A., Doody C. (2012b) 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. *Manual Therapy*, 17(2):119-25.

Spitzer, W. O., LeBlanc, F. E., Dupuis, M. (1987). Scientific approach to the assessment and management of activity related spinal disorders. A monograph for clinicians. Report of the Quebec Task Force on Spinal Disorders. *Spine*, 12(7 Suppl), S1-59.

Stafford, M. A., Peng, P., Hill, D. A. (2007). Sciatica: a review of history, epidemiology, pathogenesis, and the role of epidural steroid injection in management. *British journal of Anaesthesia,* 99(4), 461-73.

Sterne, J. A., White, I. R., Carlin, J. B., Spratt, M., Royston, P., Kenward, M. G., Wood, A. M., Carpenter, J. R. (2009). Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *British Medical Journal,* 338, b2393.

Steyerberg, E. W. (2009). *Clinical prediction models: a practical approach to development, validation, and updating.* New York, Springer.

Steyerberg, E. W., Vickers, A. J., Cook, N. R., Gerds, T., Gonen, M., Obuchowski, N., Pencina, M. J., Kattan, M. W. (2010). Assessing the performance of prediction models: a framework for some traditional and novel measures. *Epidemiology,* 21(1), 128-38.

Stirling, A., Worthington, T., Rafiq, M., Lambert, P. A., Elliott, T. S. J. (2001). Association between sciatica and Propionibacterium acnes. *The Lancet,* 357(9273), 2024-5.

Swanson, S. A., Lindenberg, K., Bauer, S., Crosby, R. D. (2012). A Monte Carlo investigation of factors influencing latent class analysis: an application to eating disorder research. *The International Journal of Eating Disorders,* 45(5), 677-84.

Sweetman, B. J., Heinrich, I., Anderson, J. a. D. (1992). Clinical tests and patterns of low back pain. *Journal of Orthopaedic Rheumatology,* 5(4), 209-22.

Tampin, B., Briffa, N. K., Hall, T., Lee, G., Slater, H. (2012). Inter-therapist agreement in classifying patients with cervical radiculopathy and patients with non-specific neck-arm pain. *Manual Therapy,* 17(5), 445-50.

Taylor, W., Gladman, D., Helliwell, P., Marchesoni, A., Mease, P., Mielants, H. (2006). Classification criteria for psoriatic arthritis: Development of new criteria from a large international study. *Arthritis & Rheumatism,* 54(8), 2665-73.

Temple, L. K. F., Mcleon, R. S., Gallinger, S., Wright, J. G. (2001). Essays on science and society. Defining disease in the genomics era. *Science,* 293(5531), 807-8.

Thomas, E., Barrett, J. H., Donn, R. P., Thomson, W., Southwood, T. R. (2000). Subtyping of juvenile idiopathic arthritis using latent class analysis. *Arthritis & Rheumatism,* 43(7), 1496-503.

Tubach, F., Beauté, J., Leclerc, A. (2004). Natural history and prognostic indicators of sciatica. *Journal of Clinical Epidemiology,* 57(2), 174-9.

UCLA: Statistical Consulting Group. (2016). *Analyzing Data: Latent Class and Other Mixture Models.* http://www.ats.ucla.edu/stst/mplus/seminars/introMplus_part2/lca.htm.Accessed 13/01/2016.

Urquhart, D. M., Zheng, Y., Cheng, A. C., Rosenfeld, J. V., Chan, P., Liew, S., Hussain, S. M., Cicuttini, F. M. (2015). Could low grade bacterial infection contribute to low back pain? A systematic review. *BMC Medicine,* 13(1), 1-13.

Vach, W. (2005). The dependence of Cohen's kappa on the prevalence does not matter. *Journal of Clinical Epidemiology,* 58(7), 655-61.

Valat, J. P., Genevay, S., Marty, M., Rozenberg, S., Koes, B. (2010). Sciatica. *Best Practice and Research. Clinical Rheumatology,* 24(2), 241-52.

Van Der Windt, D. A., Simons, E., Riphagen, Ii, Ammendolia, C., Verhagen, A. P., Laslett, M., Deville, W., Deyo, R. A., Bouter, L. M., De Vet, H. C., Aertgeerts, B. (2010). Physical examination for lumbar radiculopathy due to disc herniation in patients with low-back pain. *Cochrane Database of Systematic Reviews,* (2), Cd007431.

Van Loon, A. J., Tijhuis, M., Picavet, H. S., Surtees, P. G., Ormel, J. (2003). Survey non-response in the Netherlands: effects on prevalence estimates and associations. *Annals of Epidemiology,* 13(2), 105-10.

Van Tulder, M., Peul, W., Koes, B. (2010). Sciatica: what the rheumatologist needs to know. *Nature Reviews Rheumatology,* 6(3), 139-45.

VassarStats: Website for Statistical Computation. http:// vassarstats.net/clin1.html. Accessed 03/03/2016.

Verwoerd, A. J., Peul, W. C., Willemsen, S. P., Koes, B. W., Vleggeert-Lankamp, C. L., El Barzouhi, A., Luijsterburg, P. A., Verhagen, A. P. (2014). Diagnostic accuracy of history taking to assess lumbosacral nerve root compression. *The Spine Journal,* 14(9), 2028-37.

Viera, A. J., Garrett, J. M. (2005) Understanding interobserver agreement: the kappa statistic. *Family Medicine,* 37(5), 360-3.

Vining, R., Potocki, P., Seidman, M., Morgenthal, A.P. (2013). An evidence based diagnostic classification system for low back pain. *Journal of Canadian Chiropractic Association,* 57(3), 189-204.

Vroomen, P. C., De Krom, M. C., Knottnerus, J. A. (2000). Consistency of history taking and physical examination in patients with suspected lumbar nerve root involvement. *Spine,* 25(1), 91-6.

Vroomen, P. C., De Krom, M. C., Wilmink, J. T., Kester, A. D., Knottnerus, J. A. (2002). Diagnostic value of history and physical examination in patients suspected of lumbosacral nerve root compression. *Journal of Neurology, Neurosurgery & Psychiatry,* 72(5), 630-34.

Vroomen, P. C. A., De Krom, M. C. T. F. M., Knottnerus, J. A. (1999). Diagnostic value of history and physical examination in patients suspected of sciatica due to disc herniation: a systematic review. *Journal of neurology,* 246(10), 899-906.

Vucetic, N., Astrand, P., Guntner, P., Svensson, O. (1999). Diagnosis and prognosis in lumbar disc herniation. *Clinical Orthopaedics and Related Research,* (361), 116-22.

Waddell, G. (2004). *The back pain revolution.* Edinburgh, Churchill Livingstone.

Waddell, G., Main, C. J., Morris, E. W., Venner, R. M., Rae, P. S., Sharmy, S. H., Galloway, H. (1982). Normality and reliability in the clinical assessment of backache. *British Medical Journal (Clinical Research Ed),* 284(6328), 1519-23.

Walsh, J., Hall, T. (2009). Classification of low back-related leg pain: do subgroups differ in disability and psychosocial factors? *Journal of Manual & Manipulative Therapy,* 17(2), 118-23.

Ware, J. E., Jr. (2000). SF-36 health survey update. *Spine,* 25(24), 3130-9.

Wassenaar, M., Van Rijn, R. M., Van Tulder, M. W., Verhagen, A. P., Van Der Windt, D. A., Koes, B. W., De Boer, M. R., Ginai, A. Z., Ostelo, R. W. (2012). Magnetic resonance imaging for diagnosing lumbar spinal pathology in adult patients with low back pain or sciatica: a diagnostic systematic review. *European spine journal,* 21(2), 220-7.

Weber, H., Holme, I., Amlie, E. (1993). The natural course of acute sciatica with nerve root symptoms in a double-blind placebo-controlled trial evaluating the effect of piroxicam. *Spine,* 18(11), 1433-8.

Webster, B. S., Bauer, A. Z., Choi, Y. S., Cifuentes, M., Pransky, G. S. (2013). Iatrogenic consequences of early magnetic resonance imaging in acute, work-related, disabling low back pain. *Spine,* 38(22), 1939-46.

Werneke, M., Deutscher, D., Hart, D., Stratford, P., Ladin, J., Weinberg, J., Herbowy, S., Resnik, L. (2014). McKenzie lumbar classification: Inter-rater agreement by physical therapists with different levels of formal McKenzie postgraduate training. *Spine,* 39(3), E182-E190.

Werneke, M. W., Hart, D. L. (2004). Categorizing patients with occupational low back pain by use of the Quebec Task Force Classification system versus pain pattern classification procedures: discriminant and predictive validity. *Physical Therapy,* 84(3), 243-54.

Wilgenbusch, C. S., Wu, A. S., Fourney, D. R. (2014). Triage of spine surgery referrals through a multidisciplinary care pathway: a value-based comparison with conventional referral processes. *Spine,* 39(22 Suppl 1), S129-35.

Wilson, L., Hall, H., McIntosh, G., Melles, T. (1999). Intertester reliability of a low back pain classification system. *Spine, 24*(3), 248-54.

Woolf, C. J. (2004) Pain: moving from symptom control toward mechanism-specific pharmacological management. *Annals of Internal Medicine*, 140, 441-41.

Worster, A., Carpenter, C. (2008). Incorporation bias in studies of diagnostic tests: how to avoid being biased about bias. *Canadian Journal of Emergency Medicine,* 10(2), 174-5.

Yang, C. C. (2006). Evaluating latent class analysis models in qualitative phenotype identification. *Computational Statistics and Data Analysis*, 50, 1090-1104

Yazici, H. (2009). Diagnostic versus classification criteria - a continuum. *Bulletin of the NYU Hospital for Joint Diseases,* 67(2), 206-8.

Yunus, M. B. (2005). The concept of central sensitivity syndromes. In Wallace, D. J., Clauw, D. J., editors. *Fibromyalgia and other central pain syndromes*. Philadelphia, Lippincott Williams and Watkins.

Zhang, Z., Song, L., Meng, Q., Li, Z., Pan, B., Yang, Z., Pei, Z. (2009). Morphological analysis in patients with sciatica: a magnetic resonance imaging study using three-dimensional high-resolution diffusion-weighted magnetic resonance neurography techniques. *Spine,* 34(7), E245-50.

Zigmond, A. S., Snaith, R. P. (1983). The hospital anxiety and depression scale. *Acta Psychiatrica Scandinavica,* 67(6), 361-70.

Appendices

KEELE UNIVERSITY

Arthritis Research UK
Providing answers today and tomorrow

NHS National Institute for Health Research

PRIMARY CARE SCIENCES

ARTHRITIS RESEARCH UK PRIMARY CARE CENTRE

The ATLAS Study

Assessment and **T**reatment of **L**eg pain **A**ssociated with the **S**pine

ATLAS

Patient Information Sheet

We would like to invite you to take part in our research study. Before you decide, it is important for you to understand why the research is being done and what it would involve for you. **Please take time to read the following information carefully.** Talk to others about the study if you wish.

Part 1 tells you the purpose of this study and what will happen to you if you take part.
Part 2 tells you more about the conduct of the study.

One of our team will go through this information sheet with you at the clinic. Please ask if there is anything that is not clear or if you would like more information. Please take time to decide whether or not you wish to take part. Your decision will not affect the care or treatment you receive.

Part 1

What is the purpose of the study?

People with a back problem who experience pain spreading to their legs often have worse symptoms and take longer to recover than people with no leg pain. This study is looking at the reasons for this. We want to find out how people who visit their GP with back and leg pain do over the following year, and if we can identify the people who don't recover as well as others do. This information will

be used to develop ways to improve the care of people with such a problem. We are not testing any new treatments in this study.

Why have I been invited?

You have received this information because you recently visited your GP or contacted the Physiotherapy Direct service about a back or leg problem, and you have an appointment at the Community Low Back and Leg Pain Clinic. If you do not wish to take part in the research study, a physiotherapist at the clinic will still see you. Please still attend your clinic appointment; you are under no obligation to take part in the study.

Do I have to take part?

No. It is up to you to decide. If you agree to take part we will then ask you to sign a consent form. You are free to withdraw from the study at any time, without giving a reason. This would not affect the standard of care you receive.

What will happen to me if I take part?

If you are interested in taking part in the study you will need to fill in the questionnaire enclosed with this information sheet and bring it with you to the clinic. When you arrive for your appointment at the clinic, a nurse will explain the study to you in detail. You will be given the chance to ask questions. If you agree to take part and meet the study eligibility criteria, you will be asked to fill in a consent form. The nurse will collect and check your questionnaire and ask you to complete another questionnaire during your appointment.

You will then see a physiotherapist who will assess you and give you advice about your back and leg problem. You may or may not be referred for further treatment. Your first visit to the clinic will last about one and a half hours. If you need more treatment after this, your appointments will last about 30 - 45 minutes. The results of your assessment will be used for the research study.

In some sessions there may be a second physiotherapist present who will complete part of your assessment, and you may be asked if you agree to your assessment being video-recorded. In some other sessions a researcher may ask to observe and audio-record your consultation with the physiotherapist. It is up to you to decide whether or not to agree to your session being recorded. You will still be able to take part in the study if you decide not to be recorded. **This would not affect the standard of care you receive.**

You may also be asked to come to the hospital for an MRI scan of your back as part of the research project. If there is any reason that prevents you from having an MRI scan (such as a pacemaker, or the presence of loose metalwork such as surgical artery clips or foreign bodies) your physiotherapist will say so. The scan will take about 30 minutes. If you suffer from claustrophobia while in the scanner you will be able to press a button to notify the radiographer and the scan will be stopped. There are no risks to your health from the radio waves used in an MRI scan. You won't need to have a scan if you have already had one in the past 6 months. Your physiotherapist will provide you with the results of your scan at your

next appointment. If you do not need to see the physiotherapist again they will arrange with you how to let you know the results of your scan.

You will be sent a short questionnaire every month for one year after your first visit to the clinic. The questionnaires at 4 and 12 months will be longer. You will be asked to fill in and return these questionnaires in a pre-paid envelope. You may receive up to two reminders for the longer 4 and 12 month questionnaires.

You may be invited to take part in an interview with a researcher about your experiences with back and leg pain and about your treatments. This interview would take place either at the clinic or at your home, whichever is more convenient for you. The interview would last about 45 minutes and would be audio-recorded. The recording will be typed out in full, and if you request a copy we will send it to you by post.

What if I decide not to take part?
If you decide not to take part in the study, you can still make an appointment to see a physiotherapist at the clinic. They will assess your back or leg problem and give advice on how to manage your back or leg pain. The physiotherapist may choose to refer you for further treatment.

What are the benefits or risks of taking part?
There are no direct benefits or risks for you from taking part. The care you receive from your doctor will not be affected. However, the information we get from this study may well help to improve the future treatment of people with back and leg pain.

This completes Part 1.

If the information in Part 1 has interested you, and you are considering taking part, please read Part 2 before making any decision.

Part 2

Will my taking part in the study be kept confidential?
All information collected about you during the research study will be kept strictly confidential. Any information about you which leaves the clinic will have your name and address removed so that you cannot be recognised. Any information you give us on the questionnaires will not be identified to you. On this basis, the data may be used in other research studies. Any video or audio-recordings will be transferred and stored securely.

What will happen if I don't want to carry on with the study?
You can withdraw from the study by telephoning us on 01782 733921. Withdrawing means that we would no longer contact you directly, but we would still keep and use the information you have provided up to the point of your withdrawal. If you contact us to withdraw from the study we will check whether you also want us to stop reviewing your medical records.

What will happen to the results of the study?
We intend to publish the broad scientific results of the study. We will make these results available to you unless you request otherwise. You will not be identified in any report or publication. If you take part in an interview or your consultation with the physiotherapist is observed, quotations may be used in reports of the study. Your identity or that of any third party will be hidden in any such report.

Who is organising and funding the research?
The Arthritis Research UK Primary Care Centre at Keele University is organising this research. Website: http://www.keele.ac.uk/research/pchs/pcmrc. The National Institute for Health Research, which is linked to the NHS, is funding the study with the Primary Care Research West Midlands North. Website: http://www.nihr.ac.uk

Who has reviewed the study?
All research in the NHS is looked at by an independent group of people, called a Research Ethics Committee, to protect your interests. This study has been reviewed and given favourable opinion by South Birmingham Research Ethics Committee (ref no: 10/H1207/82).

What if there is a problem?
If you have a concern about any aspect of this study, you should ask to speak to the researchers who will do their best to answer your questions. Please contact the study coordinator, Dr Ruth Beardmore, at Keele University on 01782 733921. If you remain unhappy and wish to complain formally, you can do this through the normal NHS complaints procedure. Details of which can be obtained from the Patient Advice and Liaison Service (PALS). You can contact PALS if you have any general questions or concerns about taking part in this research. Contact PALS at Stoke-on-Trent PCT on 0800 783 2865. Contact PALS at North Staffordshire PCT on 0800 389 9676. Website: http://www.pals.nhs.uk

Further information:
If you have any questions about the study, or if you have problems booking a clinic appointment, then please telephone **01782 733921** and ask to speak to the study nurse, **Shirley Caldwell.**

Thank you for taking time to read this information sheet.

KEELE UNIVERSITY

ATLAS

Assessment and Treatment of Leg pain Associated with the Spine

Baseline Questionnaire

- When completing this questionnaire, please try to be as accurate and honest as you can throughout. There are no ‘correct’ or ‘incorrect’ answers. Answer according to your own feelings, rather than how you think most people will answer.
- Try not to let your answer to one question influence your answers to other questions.
- **Please bring your completed questionnaire with you to your clinic appointment.**
- If you have any further questions about this questionnaire or the study in general, you can telephone Keele University on 01782 733921 during office hours, and ask to speak to Shirley Caldwell, the study nurse.

Thank you very much for your help with this research study

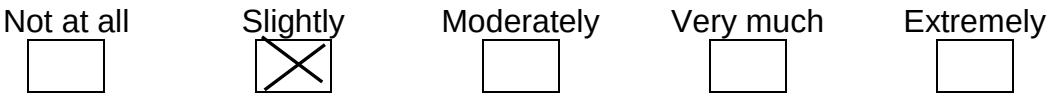
Appendix B

INSTRUCTIONS FOR THIS QUESTIONNAIRE

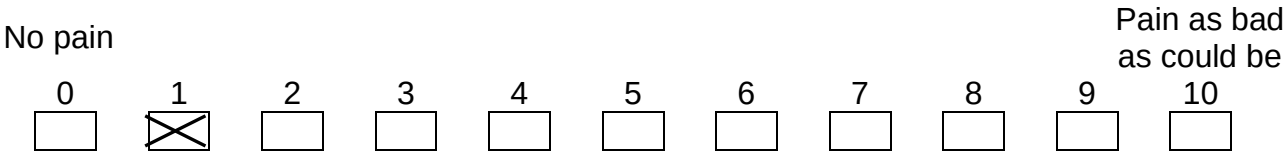
Please answer **all** of the questions, even if you feel that they do not apply to you. Some of the questions are arranged in sections according to the period of time that they ask about. Some questions may look like others, but it is important that you fill all of them in.

Some of the questions are about your back and / or leg problem and how you feel about your back or leg problem. Other questions are about you and your general health. Please take the time to read and answer each question carefully.

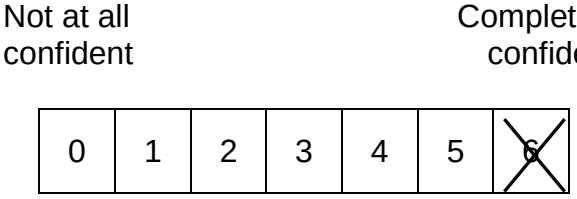
Most of the questions can be answered by putting a **cross** in a box next to or under your answer. For example, if you wish to answer 'Not at all', **cross** the box like this:



Here is an example of how to answer a question if you **don't** have any pain:



Here is an example of how to answer a question if you are **completely confident**:



Now please continue and fill in this questionnaire

Appendix B

Section A

The first few questions are about your **back pain**.

Some people with back pain tell us that they have **distinct bouts / episodes** of pain, with periods in between when they have no pain. For the first question we would like you to think about your **most recent bout / episode** of **back pain**.

*You do not need to be exact, please cross the **one** box nearest to your answer.*

1. Have you had this current bout / episode of **back pain** for…

Less than 2 weeks	2 to 6 weeks	6 to 12 weeks	3 to 6 months	7 to 12 months	More than 12 months
☐	☐	☐	☐	☐	☐

For the next few questions in this section, please think about the **last 2 weeks**.

2. In the **last 2 weeks**, on **average**, how intense was your **usual** back pain rated on a 0–10 scale, where 0 is 'no pain' and 10 is 'pain as bad as could be'?
   *(Please cross **one** box)*

No pain — Pain as bad as could be

0	1	2	3	4	5	6	7	8	9	10
☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

3. In the **last 2 weeks**, how intense was your **least** painful back pain rated on a 0–10 scale, where 0 is 'no pain' and 10 is 'pain as bad as could be'?
   *(Please cross **one** box)*

No pain — Pain as bad as could be

0	1	2	3	4	5	6	7	8	9	10
☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

4. Has the pain from your back **spread down your leg** or legs in the **last 2 weeks**?

Yes …… ☐ *Please continue with **question 5**.*

No…… ☐ *Please turn to **section C** on **page 4**.*

Appendix B

The next few questions ask about your **leg pain**.

5. **How far down your leg** or legs has the pain spread in the **last 2 weeks**?

(*Please tick **all** boxes that apply*)

	Above the knee	Below the knee
Right leg……..	☐	☐
Left leg…….	☐	☐

6. In the **last 2 weeks**, on **average**, how intense was your **usual <u>leg pain</u>** rated on a 0–10 scale, where 0 is 'no pain' and 10 is 'pain as bad as could be'?

*(Please cross **one** box)*

No pain — Pain as bad as could be

0	1	2	3	4	5	6	7	8	9	10
☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

7. In the **last 2 weeks**, how intense was your **least** painful **<u>leg pain</u>** rated on a 0–10 scale, where 0 is 'no pain' and 10 is 'pain as bad as could be'?

*(Please cross **one** box)*

No pain — Pain as bad as could be

0	1	2	3	4	5	6	7	8	9	10
☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

For the next question we would like you to think about your **most recent bout / episode** of **leg pain**.

*You do not need to be exact, cross the **one** box nearest to your answer.*

8. Have you had this current bout / episode of **leg pain** for….

Less than 2 weeks	2 to 6 weeks	6 to 12 weeks	3 to 6 months	7 to 12 months	More than 12 months
☐	☐	☐	☐	☐	☐

Appendix B

Section B

For these questions, please think about the **past week**.

Please rate the following symptoms on a 0-6 point scale, according to how **bothersome** they were in the **past week**, where 0 is 'not bothersome' and 6 is 'extremely bothersome'.

1. **Leg pain** (sciatica)….. *(Please cross **one** box)*

Not bothersome			Somewhat bothersome			Extremely bothersome
0	1	2	3	4	5	6
☐	☐	☐	☐	☐	☐	☐

2. **Numbness or tingling in leg, foot or groin** ….. *(Please cross **one** box)*

Not bothersome			Somewhat bothersome			Extremely bothersome
0	1	2	3	4	5	6
☐	☐	☐	☐	☐	☐	☐

3. **Weakness in leg or foot** (e.g. difficulty lifting foot)….. *(Please cross **one** box)*

Not bothersome			Somewhat bothersome			Extremely bothersome
0	1	2	3	4	5	6
☐	☐	☐	☐	☐	☐	☐

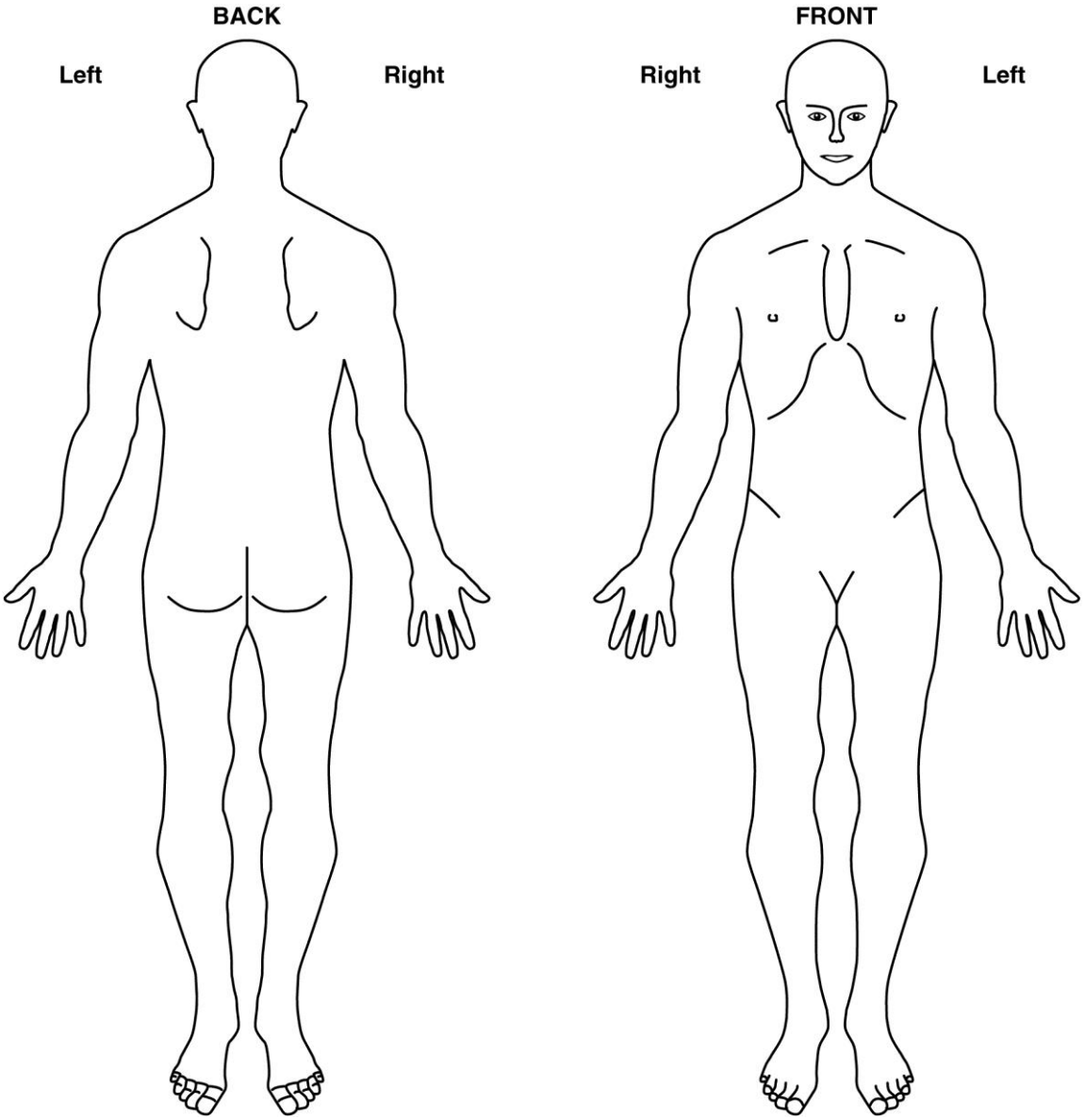
4. **Back or leg pain while sitting** ….. *(Please cross **one** box)*

Not bothersome			Somewhat bothersome			Extremely bothersome
0	1	2	3	4	5	6
☐	☐	☐	☐	☐	☐	☐

Appendix B

Section C

This question is about **recent pain** you may have had in **any part of your body**; it does not only refer to your back or legs. Please *shade in the diagram* below any pain that has lasted for **one day or longer** in the **last 4 weeks**. By pain we also mean ache, discomfort or stiffness. Please **do not** include pain due to feverish illness such as flu.



If you have ***not*** **had any body pain** that has lasted for one day or longer in the last 4 weeks, please put a cross in this box……

Appendix B

Section D

For these questions, please think about your back and / or leg pain over the **last 2 weeks.**

1. Overall, how **bothersome** has your back pain been in the **last 2 weeks**?

Not at all	Slightly	Moderately	Very much	Extremely
☐	☐	☐	☐	☐

For each of the following, please cross one box to show whether you agree or disagree with the statement, thinking about the **last 2 weeks.**

2. My back pain has **spread down my leg**(**s**) at some time in the last 2 weeks.
   Agree ☐ Disagree ☐

3. I have had pain in the **shoulder** or **neck** at some time in the last 2 weeks.
   Agree ☐ Disagree ☐

4. It's **really not safe** for a person with a condition like mine to be **physically active.**
   Agree ☐ Disagree ☐

5. In the last 2 weeks, I have **dressed more slowly** than usual because of my back pain.
   Agree ☐ Disagree ☐

6. In the last 2 weeks, I have only **walked short distances** because of my back pain.
   Agree ☐ Disagree ☐

7. **Worrying thoughts** have been going through my mind a lot of the time in the last 2 weeks.
   Agree ☐ Disagree ☐

8. I feel that **my back pain is terrible** and that **it is never going to get any better**.
   Agree ☐ Disagree ☐

9. In general, in the last 2 weeks, I have **not enjoyed** all the things I used to enjoy.
   Agree ☐ Disagree ☐

Appendix B

Section E

This set of questions is about you **today**.

When your back or leg hurts, you may find it difficult to do some of the things you normally do. This list contains sentences that people have used to describe themselves when they have back and / or leg pain. When you read them, you may find that some stand out because they describe you **today**. As you read the list, think of yourself **today**.

When you read a sentence that describes you **today**, put **a cross** in the box next to it. If the sentence does not describe you today, then leave the box empty and go on to the next sentence. Remember, only cross the box next to the sentence if you are sure that it describes you **today**.

1. I stay at home most of the time because of my back or leg problem ☐
2. I change position frequently to try and get my back or leg comfortable…… ☐
3. I walk more slowly than usual because of my back or leg problem…….. ☐
4. Because of my back or leg problem, I am not doing any of the jobs that I usually do around the house…… ☐
5. Because of my back or leg problem, I use a handrail to get upstairs…… ☐
6. Because of my back or leg problem, I lie down to rest more often…… ☐
7. Because of my back or leg problem, I have to hold on to something to get out of an easy chair…….. ☐
8. Because of my back or leg problem, I try to get other people to do things for me……. ☐
9. I get dressed more slowly than usual because of my back or leg problem.. ☐
10. I only stand for short periods of time because of my back or leg problem… ☐
11. Because of my back or leg problem, I try not to bend or kneel down…… ☐
12. I find it difficult to get out of a chair because of my back or leg problem….. ☐

Appendix B

Remember these questions are about you **today**.

13. My back or leg is painful almost all the time……. ☐
14. I find it difficult to turn over in bed because of my back or leg problem…… ☐
15. My appetite is not very good because of my back or leg pain…… ☐
16. I have trouble putting on my socks (or tights) because of the pain in my back or leg…….. ☐
17. I only walk short distances because of my back or leg pain…….. ☐
18. I sleep less well because of my back or leg problem…….. ☐
19. Because of my back or leg pain, I get dressed with help from someone else…….. ☐
20. I sit down for most of the day because of my back or leg problem…… ☐
21. I avoid heavy jobs around the house because of my back or leg problem... ☐
22. Because of my back or leg pain, I am more irritable and bad tempered with people than usual…….. ☐
23. Because of my back or leg problem, I go upstairs more slowly than usual.. ☐
24. I stay in bed most of the time because of my back or leg pain…….. ☐
25. Because of my back or leg problem, my sexual activity is decreased…….. ☐
26. I keep rubbing or holding areas of my body that hurt or are uncomfortable. ☐
27. Because of my back or leg problem, I am doing less of the daily work around the house than I would usually do……. ☐
28. I often express concern to other people over what might be happening to my health……. ☐

Appendix B

This question asks about the **back** pain that you may **currently** be experiencing.

29. How would you rate your **back pain** on a 0-10 scale **at the present time**, that is right now, where 0 is 'no pain' and 10 is 'pain as bad as could be'? *(Please cross **one** box)*

No pain — Pain as bad as could be

0	1	2	3	4	5	6	7	8	9	10
☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

This question asks about the **leg** pain that you may **currently** be experiencing.

30. How would you rate your **leg pain** on a 0-10 scale **at the present time**, that is right now, where 0 is 'no pain' and 10 is 'pain as bad as could be'? *(Please cross **one** box)*

No pain — Pain as bad as could be

0	1	2	3	4	5	6	7	8	9	10
☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

The next question is about how long it has been since you were **pain free**.

31. How long is it since you had a **whole month without** any back pain ***or*** leg pain?

Less than 3 months	4 to 6 months	7 to 12 months	1 to 3 years	More than 3 years
☐	☐	☐	☐	☐

Appendix B

Section F

In this section we are asking about your **general health**.

For each of the five sets of statements below, please **cross the one box** that best describes your own health state **today**.

1. **Mobility**

I have no problems in walking about…….. ☐

I have some problems in walking about……. ☐

I am confined to bed……..…… ☐

2. **Self-care**

I have no problems with self-care……..……. ☐

I have some problems washing and dressing myself…….. ☐

I am unable to wash or dress myself……..…… ☐

3. **Usual activities** *(e.g. work, study, housework, family or leisure activities)*

I have no problems with performing my usual activities……..…… ☐

I have some problems with performing my usual activities……. ☐

I am unable to perform my usual activities……..…….. ☐

4. **Pain / discomfort**

I have no pain or discomfort……..…….……. ☐

I have moderate pain or discomfort……. ☐

I have extreme pain or discomfort……..……. ☐

5. **Anxiety / depression**

I am not anxious or depressed……..…….…… ☐

I am moderately anxious or depressed…….…… ☐

I am extremely anxious or depressed……..…….…… ☐

For the following question, please **cross one box** that best describes your **general health** at **present.**

6. In **general**, would you say your health is?

Excellent	Very good	Good	Fair	Poor
☐	☐	☐	☐	☐

Appendix B

Section G

Please rate how **confident** you are that you can do the following things **at present**, **despite the pain**. To indicate your answer, cross the box of **one** of the numbers on the scale under each item, where **0** = not at all confident, and **6** = completely confident.

Remember, these questions are **not** asking whether or not you have been doing these things, but rather **how confident you are that you can do them at present, despite the pain.**

	Not at all confident						Completely confident
1. I can enjoy things, despite the pain……..	0	1	2	3	4	5	6
2. I can do most of the household chores (e.g. tidying-up, washing dishes), despite the pain…..	0	1	2	3	4	5	6
3. I can socialise with my friends or family members as often as I used to do, despite the pain ……	0	1	2	3	4	5	6
4. I can cope with my pain in most situations……..	0	1	2	3	4	5	6
5. I can do some form of work, despite the pain (“work” includes housework, paid and unpaid)…	0	1	2	3	4	5	6
6. I can still do many of the things I enjoy doing, such as hobbies or leisure activity, despite the pain…….	0	1	2	3	4	5	6
7. I can cope with my pain without medication……	0	1	2	3	4	5	6
8. I can still accomplish most of my goals in life, despite the pain…….	0	1	2	3	4	5	6
9. I can live a normal lifestyle, despite the pain……	0	1	2	3	4	5	6
10. I can gradually become more active, despite the pain…….	0	1	2	3	4	5	6

Appendix B

Section H

Listed below are a number of symptoms that you may or may not have experienced **because of your back and / or leg problem**.

1. Please indicate by putting a cross in the box for yes or no, **to tell us whether** you have experienced **any of these symptoms** because of your back and / or leg problem.

	Yes	No
Back pain/ache……	☐	☐
Leg pain/ache……..	☐	☐
Unable to sit comfortably……...	☐	☐
Fatigue…….	☐	☐
Stiff joints…….....	☐	☐
Sleep difficulties…….....	☐	☐
Loss of strength……..	☐	☐

We are interested in your **own personal views** on how you see your current back and / or leg problem. Please indicate how much you agree or disagree with the following statements about your back or leg problem by putting **a cross in one box on each line**.

2. My back and / or leg problem will last for a long time. *(Please cross **one** box)*

Strongly disagree	Disagree	Neither agree or disagree	Agree	Strongly agree
☐	☐	☐	☐	☐

3. My back and / or leg problem has major consequences on my life. *(Please cross **one** box)*

Strongly disagree	Disagree	Neither agree or disagree	Agree	Strongly agree
☐	☐	☐	☐	☐

Appendix B

Remember, we are interested in your **own personal views** on how you see your current back and / or leg problem.

4. There is a lot which I can do to control my back and / or leg symptoms. *(Please cross **one** box)*

Strongly disagree	Disagree	Neither agree or disagree	Agree	Strongly agree
☐	☐	☐	☐	☐

5. What I do can determine whether my back and / or leg problem gets better or worse. *(Please cross **one** box)*

Strongly disagree	Disagree	Neither agree or disagree	Agree	Strongly agree
☐	☐	☐	☐	☐

6. Treatment can control my back and / or leg problem. *(Please cross **one** box)*

Strongly disagree	Disagree	Neither agree or disagree	Agree	Strongly agree
☐	☐	☐	☐	☐

7. I don't understand my back and / or leg problem. *(Please cross **one** box)*

Strongly disagree	Disagree	Neither agree or disagree	Agree	Strongly agree
☐	☐	☐	☐	☐

8. My back and / or leg symptoms come and go in cycles. *(Please cross **one** box)*

Strongly disagree	Disagree	Neither agree or disagree	Agree	Strongly agree
☐	☐	☐	☐	☐

9. My back and / or leg problem affects me emotionally (e.g. it makes me feel frustrated, anxious, angry, afraid, upset or depressed). *(Please cross **one** box)*

Strongly disagree	Disagree	Neither agree or disagree	Agree	Strongly agree
☐	☐	☐	☐	☐

We are **now** interested in what you consider **may have been the cause** of your back and / or leg problem. As people are very different, there is no correct answer for this question. We are **most** interested in **your own views** about the factors that cause your back and/ or leg problem rather than what others (including doctors or family) may have suggested.

10. Below is a list of possible causes for your back and / or leg problem. Please indicate how much you agree or disagree that they were causes for your back and / or leg problem by **putting a cross in one box on each line**.

	Strongly disagree	Disagree	Neither agree or disagree	Agree	Strongly agree
Hereditary, it runs in my family	☐	☐	☐	☐	☐
Ageing	☐	☐	☐	☐	☐
An accident or injury	☐	☐	☐	☐	☐
Chance or bad luck	☐	☐	☐	☐	☐
My own behaviour	☐	☐	☐	☐	☐
Overwork	☐	☐	☐	☐	☐
Lifting and carrying objects	☐	☐	☐	☐	☐
Wear and tear in the spine	☐	☐	☐	☐	☐
My work	☐	☐	☐	☐	☐

Please list in rank-order the three most important factors that you believe caused <u>your</u> back and / or leg problem. ***The most important causes for me:-***

1. ……

2. ……

3. ……

Appendix B

This section contains questions about you and your recent **employment** history.

1. Are you: Female ☐ Male ☐

2. What is your **date of birth**? day ☐☐ month ☐☐ 1 9 year ☐☐

3. What is your current or most recent paid **job title**?

……

4. What does / did the firm / organisation you work for mainly make or do?

……

5. What do / did you mainly do in your job?

……

6. Are you currently in a paid job? *(Please put a cross in **one** box only)*

Yes ☐ *Please continue with **question 7** below*

No ☐ *Please turn to **question 16** on page 17*

7. How would you best describe your **typical working week** in the **last 12 months**? *(Please put a cross in **one** box only)*

☐ Working full time (35 hours or more per week)

OR

☐ Working part time (less than 35 hours per week)

8. Have you **taken time off work** during the **last 12 months** because of your back or leg pain? *(Please put a cross in **one** box only)*

- ☐ Yes, mainly due to my **back pain**
- ☐ Yes, mainly due to my **leg pain**
- ☐ No, not due to my back or leg pain

If yes, please write in the **number** of days, weeks or months you were off work due to your back or leg pain in the **last 12 months**.

- ☐ Days
- ☐ Weeks
- ☐ Months

*Please only enter a number in **one** of these boxes.*

9. Are you **currently**… *(Please put a cross in **one** box only)*

- Doing your usual job ☐ *Please go to **question 12***
- On paid annual leave / holiday ☐
- Working fewer hours ☐
- Doing lighter duties ☐
- On paid sick leave ☐
- On unpaid leave ☐

*Please continue with **question 10***

10. If you are **not** doing your usual job, is this because of your back or leg pain? *(Please put a cross in **one** box only)*

- ☐ Yes, mainly due to my **back pain**
- ☐ Yes, mainly due to my **leg pain**
- ☐ No, not due to my back or leg pain

Appendix B

11. **If you are currently off work**, how long have you been off work for?
*(Please put a cross in **one** box only)*

Less than a month	1 to 6 months	More than 6 months
☐	☐	☐

The next few questions are about your current bout / episode of back or leg pain.

Sometimes when people are off work sick they "**self-certify**", this means they are off work for only a few days and do not need a note from their doctor.

12. Have you **self-certified** time off work because of your **current bout/ episode** of back or leg pain? *(Please put a cross in **one** box only)*

Yes ☐

No ☐

If yes, for **how many days** in total have you self-certified time off work for your current bout/ episode of back or leg pain?

☐ *Please write in the number of **days***
(give your best guess if you don't know exactly)

13. Have you been given any "**Sick Notes"** or **"Fit Notes"** from your doctor because of your **current bout/ episode** of back or leg pain?
*(Please put a cross in **one** box only)*

Yes ☐

No ☐

If yes, for **how many days** in total have your sick or fit note(s) for your current bout/ episode of back or leg pain lasted?

☐ *Please write in the number of **days***
(give your best guess if you don't know exactly)

14. Does your current bout / episode of back or leg pain make it **difficult for you to manage** at work?
*(Please put a cross in **one** box only)*

Not at all	Slightly	Moderately	Very much	Extremely
☐	☐	☐	☐	☐

15. On average, to what extent has your back or leg pain affected your **performance at work** since your back or leg pain started?
Please rate this on a 0-10 scale, where 0 is 'not at all' and 10 is 'the pain is so bad that I am unable to do my job'
*(Please put a cross in **one** box only)*

Not at all — The pain is so bad that I am unable to do my job

0	1	2	3	4	5	6	7	8	9	10
☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

16. **If you are not working**, which of the following best describes your **current** situation? *(Please put a cross in **one** box only)*

Retired	☐
Student	☐
Looking after children / home	☐
Voluntary worker	☐
Unemployed due to leg pain	☐
Unemployed due to other health reasons	☐
Unemployed (not health-related)	☐
Other (please specify)	☐

……

Appendix B

Now please enter today's date:

	day	month	year
Today's Date			2 0

This is the end of the questions.

Thank you for taking the time to fill in this questionnaire.

We assure you that any information will be held in strictest confidence and the information you have given in this questionnaire will not be linked to you.

Please bring this completed questionnaire with you to your clinic appointment.

If you have any further questions about this questionnaire or the study in general, you can telephone Keele University on 01782 733921 during office hours, and ask to speak to Shirley Caldwell, the study nurse.

Thank you very much for your help.

Supported by the National Institute for Health Research (NIHR)

Appendix B

KEELE UNIVERSITY

ATLAS

Assessment and Treatment of Leg pain Associated with the Spine

Clinic Questionnaire

- When completing this questionnaire, please try to be as accurate and honest as you can throughout. There are no 'correct' or 'incorrect' answers. Answer according to your own feelings, rather than how you think most people will answer.
- Try not to let your answer to one question influence your answers to other questions.
- **Please return this questionnaire to the clinic nurse.**
- If you have any further questions about this questionnaire or the study in general, please speak to the clinic nurse.

Thank you very much for your help with this research study

Appendix C

INSTRUCTIONS FOR THIS QUESTIONNAIRE

Please answer **all** of the questions, even if you feel that they do not apply to you. Some of the questions are arranged in sections according to the period of time that they ask about. Some questions may look like others, but it is important that you fill all of them in.

Some of the questions are about your back and / or leg problem and how you feel about your back or leg problem. Other questions are about you and your general health. Please take the time to read and answer each question carefully.

Most of the questions can be answered by putting a **cross** in a box next to or under your answer. For example, if you wish to answer 'Not at all', **cross** the box like this:

Not at all	Slightly	Moderately	Very much	Extremely
☒	☐	☐	☐	☐

Now please continue and fill in this questionnaire

Appendix C

Section A

This section asks questions about any back and / or leg pain that you may be experiencing.

Think about how your back and leg pain has felt **over the last week.** Please place a cross in the box next to the descriptions that best match your pain. These descriptions may or may not match your pain no matter how severe it feels. Please **only** cross the boxes that describe your pain.

1. In the area where you have pain, do you also have 'pins and needles', tingling or prickling sensations?

*Please cross **one** box*

NO - I don't get these sensations……... ☐

YES - I get these sensations often……. ☐

2. Does the painful area change colour (perhaps look mottled or more red) when the pain is particularly bad?

*Please cross **one** box*

NO - The pain does not affect the colour of my skin…... ☐

YES - I have noticed that the pain does make my skin look different from normal……. ☐

3. Does your pain make the affected skin abnormally sensitive to touch? Getting unpleasant sensations or pain when lightly stroking the skin might describe this.

*Please cross **one** box*

NO - The pain does not make my skin in that area abnormally sensitive to touch…….. ☐

YES - My skin in that area is abnormally sensitive to touch…… ☐

4. Does your pain come on suddenly and in bursts for no apparent reason when you are completely still? Words like 'electric shocks', jumping and bursting might describe this.

*Please cross **one** box*

NO - My pain doesn't really feel like this……... ☐

YES - I get these sensations often……. ☐

Appendix C

Remember these questions are about how your back and leg pain has felt **over the last week.**

5. In the area where you have pain, does your skin feel unusually hot like a burning pain?

*Please cross **one** box*

NO - I don't have burning pain…… ☐

YES - I get burning pain often……. ☐

6. Gently **rub** the painful area with your index finger and then rub a non-painful area (for example, an area of skin further away or on the opposite side from the painful area). How does this rubbing feel in the painful area?

*Please cross **one** box*

The painful area feels no different from the non-painful area…….. ☐

I feel discomfort, like pins and needles, tingling or burning in the painful area that is different from the non-painful area…….. ☐

7. Gently **press** on the painful area with your finger tip then gently press in the same way onto a non-painful area (the same non-painful area you chose in the last question). How does this feel in the painful area?

*Please cross **one** box*

The painful area does not feel different from the non-painful area……. ☐

I feel numbness or tenderness in the painful area that is different from the non-painful area……. ☐

8. If you have back pain and leg pain, which one has been worse for you in the **last week**?

*(Please cross **one** box)*

Back pain is worse	Leg pain is worse	Not applicable (I do not have both back and leg pain)
☐	☐	☐

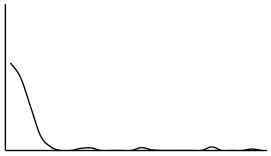
Appendix C

Section B

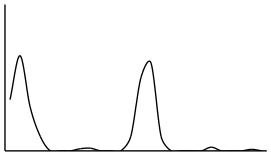
This question is about your back and / or leg pain over the **last year**.

Below are some descriptions of how some people's **back or leg pain can change over time**, with pictures to show how their pain might go up or down.

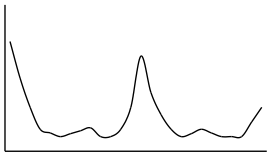
1. Please look at these and cross the box next to the one option that you think comes closest to how your pain has been over the **last year**. Please include pain that you have in your back as well as any pain spreading down your leg(s).

a) 

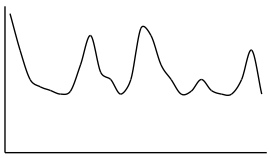
First ever episode of back or leg pain (first time you have had it) ☐

b) 

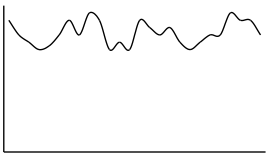
A few episodes of back or leg pain, with mostly pain-free periods in between ☐

c) 

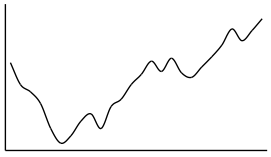
Some back or leg pain most of the time, and a few episodes of severe pain ☐

d) 

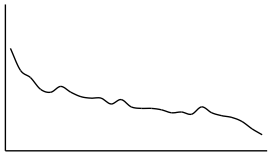
Pain that goes up and down all the time, with episodes of severe back or leg pain ☐

e) 

Severe back or leg pain all or nearly all of the time ☐

f) 

Back or leg pain that has got gradually worse ☐

g) 

Back or leg pain that has improved gradually ☐

Appendix C

Section C

This question is about other health problems.

1. Do you suffer from any of the following?

*(Please put a cross in the box next to **any that apply** to you)*

a) Chest problems ☐

b) Heart problems ☐

c) Raised blood pressure ☐

d) Diabetes ☐

e) Circulation problems in the leg ☐

These next few questions are about smoking.

2. Do you currently smoke cigarettes?

Yes…… ☐ *Please move on to **question 4***

No……. ☐ *Please continue with **question 3***

3. If no, have you ever smoked?

Yes….. ☐ *Please continue with **question 4***

No…… ☐ *Please go to **Section D** on the next page*

4. At what age did you start smoking?

…….years

5. On average, how many cigarettes per day do you smoke / did you smoke?
(Please estimate a number of cigarettes)

……

6. If you do not currently smoke, at what age did you stop smoking?

……years

Appendix C

Section D

The next questions are about how you have been feeling in the **last week.**

For the next questions, please read each item and *cross the box* under the reply that comes closest to how you have been feeling in the **last week.** Don't take too long over your replies; your immediate reaction to each item will usually be more accurate than a long thought out response.

1. I feel tense or wound up:

Most of the time	A lot of the time	From time to time, occasionally	Not at all
☐	☐	☐	☐

2. I still enjoy the things I used to enjoy:

Definitely	Not quite so much	Only a little	Hardly at all
☐	☐	☐	☐

3. I get a sort of frightened feeling as if something awful is about to happen:

Very definitely and quite badly	Yes, but not too badly	A little, but it doesn't worry me	Not at all
☐	☐	☐	☐

4. I can laugh and see the funny side of things:

As much as I always could	Not quite so much now	Definitely not so much now	Not at all
☐	☐	☐	☐

5. Worrying thoughts go through my mind:

A great deal of the time	A lot of the time	From time to time, but not too often	Only occasionally
☐	☐	☐	☐

Appendix C

Please remember to think about the last **week.**

6. I feel cheerful:

Not at all	Not often	Sometimes	Most of the time
☐	☐	☐	☐

7. I can sit at ease and feel relaxed:

Definitely	Usually	Not often	Not at all
☐	☐	☐	☐

8. I feel as if I am slowed down:

Nearly all the time	Very often	Sometimes	Not at all
☐	☐	☐	☐

9. I get a sort of frightened feeling like 'butterflies' in the stomach:

Not at all	Occasionally	Quite often	Very often
☐	☐	☐	☐

10. I have lost interest in my appearance:

Definitely	I don't take as much care as I should	I may not take quite as much care	I take just as much care as ever
☐	☐	☐	☐

11. I feel restless as if I have to be on the move:

Very much indeed	Quite a lot	Not very much	Not at all
☐	☐	☐	☐

12. I look forward with enjoyment to things:

As much as I ever did	Rather less than I used to	Definitely less than I used to	Hardly at all
☐	☐	☐	☐

Appendix C

13. I get sudden feelings of panic:

Very often indeed	Quite often	Not very often	Not at all
☐	☐	☐	☐

14. I can enjoy a good book or radio or TV programme:

Often	Sometimes	Not often	Very seldom
☐	☐	☐	☐

Appendix C

Section E

This section asks questions about your current work or your most recent job.

1. Thinking of your current job, how many hours **per day**, on average, do you spend doing the following **at work**?

   (*Please put a cross in **one** box only **on each line**)*

	Not at all	Less than 2 hours	2 to 4 hours	Over 4 hours
a) Sitting……....	☐	☐	☐	☐
b) Standing……...	☐	☐	☐	☐
c) Operating a motor vehicle…….	☐	☐	☐	☐
d) Working on a vibrating floor or seat.	☐	☐	☐	☐

	Not at all	Less than ½ hour	½ hour to 1 hour	Over 1 hour
e) Kneeling or squatting…….	☐	☐	☐	☐
f) Bending forward (in standing or kneeling position)……	☐	☐	☐	☐

2. If your job involves walking, what distance do you walk **at work each day**, on average?

   *(Please cross **one** box)*

Less than 2 miles	More than 2 miles	Not applicable (I do not walk at work)
☐	☐	☐

Appendix C

Remember these questions are about your current **work** or your most recent job.

3. How many hours **at work**, on average, do you spend manually lifting, carrying or pushing the following loads?

   *(Please put a cross in **one** box only on each line*)

	Not at all	Less than 1 hour	1 to 4 hours	Over 4 hours
a) Less than 10kg (22 lbs) – equivalent to the weight of a portable TV……	☐	☐	☐	☐
b) 10kg (22lbs) or more……	☐	☐	☐	☐

This is the end of the questions.

Thank you for taking the time to fill in this questionnaire.

We assure you that any information will be held in strictest confidence and the information you have given in this questionnaire will not be linked to you.

Now **please return this questionnaire to the clinic nurse.**

If you have any further questions about this questionnaire or the study in general, you can speak to the nurse at the clinic.

Thank you very much for your help.

Appendix C

Section F – Clinic Nurse to Complete

Today's Date

day		month		year			
				2	0		

Nurse to measure and record patient's height and weight:

Height metres

Weight kilograms

Supported by the National Institute for Health Research (NIHR)

Appendix C

ATLAS

Assessment and Treatment of Leg pain Associated with the Spine

Clinical Assessment Form

Date of Assessment | | | | | | |

Assessor's Full Name

Assessor's initials | | Signature

Patient Eligible for Study: Yes / No

Study ID label:

Please state reason if patient non-eligible for study……

……..……

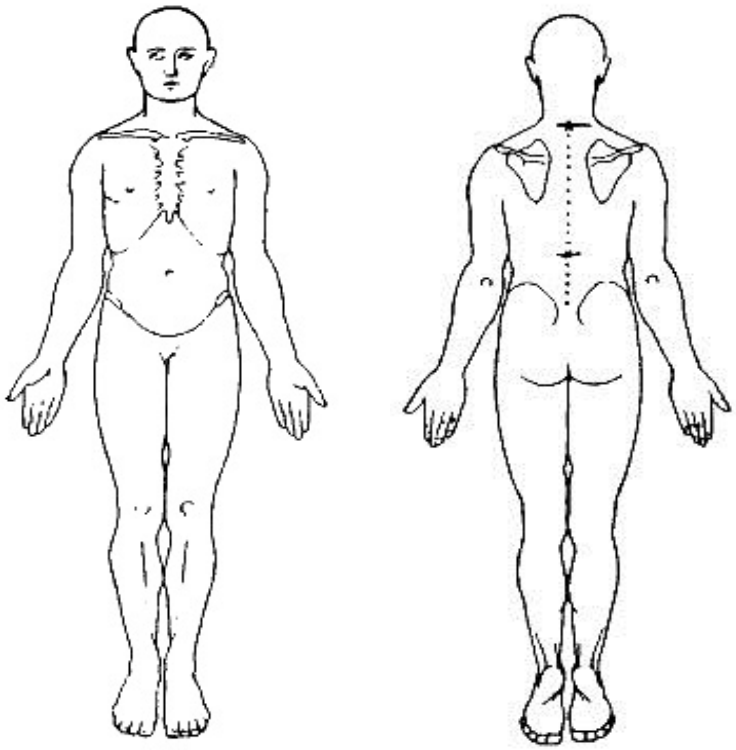
……..

Patient label

Appendix D

Age:

Clinical History



Quality of Pain in Leg

Burning, tingling, sharp, throbbing, toothache, like an electric shock.

Other (please specify)…….

Pain Rating (0 – 10)

LBP: at worst……

at best……

average……

Leg pain:

at worst

Present condition: ……..

……..

……..

HPC……..

……

……

2. Onset: **Acute / Gradual**

3. Symptoms at onset: **Back / Thigh / Lower Leg**

4. LBP since onset: **Better / Worse / Same**

5. Leg Pain since onset: **Better / Worse / Same**

6. Pins & Needles: **Yes / No**

Numbness: **Yes / No**

7. Feeling of weakness in the leg: **Yes / No**

8. Constant symptoms: **Back / Thigh / Lower Leg**

9. Intermittent symptoms: **Back / Thigh / Lower Leg**

10. What is worse: **back / leg** (specify further if necessary)……..

11. Aggravating Factors: **When Still / Sitting /Standing /Walking / Bending / Sit to Stand / Lying Down** / Other (please specify)…….....

Appendix D

12. Easing Factors: **On the Move/ Sitting / Standing/ Walking/ Bending/ Sit to Stand/ Lying Down/** Other (please specify).....

.....

13. Functional Limitations: **Yes / No** (what does it stop you from doing – please specify).....

.....

.....

14. Sleep Disturbances: **Yes / No**

Any comments.....

.....

15. EMS: **Yes/No**

Any comments.....

16. Unremitting Night Pain: **Yes / No**

Any comments .....

17. BB function: Normal / Other - please comment.....

.....

18. SA: **No** / **Yes**-please comment .....

19. Unexplained weight loss: **Yes / No**

Any comments.....

20. General Health: **Good / Fair / Poor**

Any comments.....

21. Any Other Red Flags: **No / Yes** – (please explain).....

.....

22. Cough / Sneeze / Strain: **+ve / -ve** (+ve only if it produces patient's leg symptoms)

23. Gait: steady on feet: **Yes / No**

Any comments.....

24. Previous history of similar LBP: **Yes / No**

Any comments.....

.....

25. Previous history of similar Leg Pain: **Yes / No**

Any comments.....

.....

26. Effect of previous treatment for similar symptoms.....

.....

27. Effect of self-management for similar

symptoms.....

Appendix D

28. Investigations for this problem: **No investigations / x-Rays** / **MRI** / **Bloods**
Any comments……....

……

29. Medical History (Past & Present): Chest /Heart / DM /Epil / BP / Ca / steroids / Anticoag / RA / Fract-osteoporosis / serious illnesses / operations……

……

……

30. Drug History and Effect of Medication on Symptoms:……

……

……

……

……

……

Social History

31.Work: **At work / Off work / Non applicable** (e.g. retired)

(Current details of work, ability to do, effect of symptoms, time off)……

……

……

……

32. Any time off work for previous episodes of back and /or leg pain: **Yes / No**

Any comments……

……

33. Benefits: **No / Yes** (please describe)……

……

……

34. Family: (who is at home with them and family situation) (please describe)……..

……

……

……

35. Physical Activity / Leisure / Sports: (what they do, effect of symptoms on ability to do)……

……

……

……

36. Smoker: **Yes** / **If so, how many a day** ☐ **No / Past Smoker**

37. Alcohol Intake: **None / Occasionally / Regular-under recommended limits / above recommended limits**

Assessment of Psychological Factors (Yellow Flags)

38. Evidence of Fear Avoidance: **Yes / No**

……

……

……

39. Evidence of Distress: **Yes / No**

……

……

……

40. Evidence of Low Mood / Depression: **Yes / No**

……

……

……

41. Coping Strategies: **Active / Passive**

……

……

……

42. Work Issues: **Yes / No / Non applicable**

……

……

……

43. Compensation / Litigation: **Yes / No / Non applicable**

……

……

44. Patient's Future Outlook: **Optimistic / Pessimistic**

……

……

……

Physical Examination

1.Observation……..

2. Obvious Abnormalities: **Yes / No**

Any comments……

……

3. Visible Muscle Wasting: **No / Yes** (if yes please describe)

……

4. Gait: **Normal / Antalgic / Unsteady**

Any comments……

5. Lumbar Shift: **Yes / No**

Any comments……

Lumbar Spine Range of Movement

6. Flexion: normal / limited / hypermobile increase of symptoms: **Yes/No** **LBP / leg pain**

7. Extension: normal / limited / hypermobile increase of symptoms: **Yes/No** **LBP / leg pain**

8. Right SF: normal / limited / hypermobile increase of symptoms: **Yes/No** **LBP / leg pain**

9. Left SF: normal / limited / hypermobile increase of symptoms: **Yes/No** **LBP / leg pain**

Neurological Testing; Lower Limbs

10.	**Myotomes**													
	Toe walking		Heel walking		Single leg squatting		EHL		Eversion		Inversion		Hip Flexion	
	R	L	R	L	R	L	R	L	R	L	R	L	R	L
0/5														
1/5														
2/5														
3/5														
4/5														
5/5														
Comments:														

	Knee jerk		Ankle jerk	
	R	L	R	L
Normal				
Absent				
Slightly reduced				
Signif. reduced				
Brisk				
Comments:				

Clonus: **No / Yes** (describe)……..

Plantars: **downgoing / upgoing / not elicited**

	Right	**Left**
12. Sensation (Pin Prick)		
Reduced/absent-describe areas……..		
……		
……		
……		
……		

Please tick all relevant boxes:	**Normal**	**reduced PP sensation**	**loss of PP sensation**	**total anaesthesia**
	☐	☐	☐	☐

Allodynia / Hyperalgesia-describe areas……..…….

……

……

……

	Right	**Left**
13. Neural tension tests		
SLR……		
……		
Crossover SLR……		
……		
Femoral stretch……		
……		
Slump test……...		
……		

14. Lumbar Spine Palpation Findings (if present, should be patient's own pain)

No pain / Local back pain / Radiating pain

Any comments……

15. Hip Assessment Findings: **Normal / Other** (describe)……...

……

16. Any other findings: (please specify)……

……

……

17. **Clinical Impression** ……...

……

……

18. LBP related leg pain: **Yes / No** **(If No, go to treatment decisions)**

19. LBP with nerve root involvement: **Yes / No**

How confident are you in your clinical impression: [%]

(rate on a 0-100% scale, where 100% means absolutely certain/confident):

If you wish to further qualify your rating please use the space below: ….……

……

(List **up to 4** most relevant items that led you to your clinical impression/diagnosis)……

……

……

……

……

20. Specific Diagnosis: Disc prolapse / Stenosis / Not sure

21. In your normal practice would you have wanted this patient to have an MRI scan? **Yes / No**

Treatment Decisions

1.Discharged 2. SOS 3. Physiotherapy referral and follow-up appointment

4. Back Pain Clinic (**via** Back Pain Clinic: 5.Orthopaedic referral 6. Pain Clinic referral)

6. Other (please give details)……..

……

……

Add any notes/comments you feel necessary

The ATLAS Study

Assessment and **T**reatment of **L**eg pain **A**ssociated with the **S**pine

Examiner Manual

Arthritis Research UK Primary Care Centre
Primary Care Sciences Keele University

Kika Konstantinou PhD, MMACP, MCSP, MSc (Manip Ther)

Spinal Physiotherapy Specialist / NIHR Clinical Lecturer

January 2011

Appendix E

Clinical Assessment
History taking

Body chart	Mark areas of pain and areas of reported numbness / P&N / tingling.
Establish pain location/distribution (possible dermatomal distribution).	Prompts: Where is your pain could you trace the pain slowly with your hand to show me where it is
Establish severity and quality of pain.	
P&N	Ask at first whether they have noticed any other symptoms (other than pain) in their leg(s). If they say no, then clarify by specifically asking them whether they have any numbness, tingling or P&N.
Feeling of weakness in leg.	Prompt: have you noticed any weakness in the leg (s)?
Constant symptoms vs Intermittent	When patient reports constant pain, try to clarify the point. Prompts: Do you feel the pain every moment of every hour of every day? Are you ever without any pain, even if it is for few minutes?
Establish which pain is worse; low back or leg.	Prompts: Between your back and your leg pain which one bothers you the most? If the patient cannot decide, ask them that 'if we could only cure one of the pains today, which do you want to get rid of?'
Aggravating and easing factors.	If they have problems defining any factors, ask them: 'if you had to choose two things that will make your pain far worse what would they be'. Another way to establish some factors it to ask them: 'on balance, are you better if you keep still or if you keep active/moving'.
Functional limitations	Prompts: Is your pain (or your symptoms) stopping you from doing anything? What does it stop you from doing?
Sleep disturbances	Prompts: Is your pain interfering with your sleep? Is the pain keeping you up at night? If they report significant sleep disturbances clarify further: How many hours do you sleep uninterruptedly?

Appendix E

	Do you have to get out of bed and walk about? Is the pain during the night worse than the pain during the day?
Cough/Sneeze/Strain and pain response.	This is recorded positive only if it produces patient's own pain down the leg (or on occasion buttock pain only), otherwise it is negative. Cough/sneeze/strain causing increased LBP does not qualify as a positive sign.
Previous history of similar LBP or leg pain.	If episodic symptoms, clarify how many episodes in a normal year and how long it takes on average for an episode to settle, either with or without treatment.
Previous treatment for similar symptoms and its effect, or treatments for this current problem.	What treatments have you had (including physiotherapy)? Did they make things better/worse or had no effect?
Investigations for this problem.	If they had had imaging tests recently or in the past for their back/leg trouble, ask them if they know about the results and establish how they understand the results, and whether results make sense to them.
Medical history.	Mainly ask about relevant medical history. Ask whether they have any major medical problems or whether they have been to their GP with anything serious. You could also specifically ask for diseases, conditions.
Medication.	Ask in detail about analgesic medication and its effect on symptoms. Need to establish if they are using analgesia effectively.
About work.	Establish what they do and if and in what way is affected by pain. If a job title does not reveal extent of physical activity or postures, ask them what they are doing in their job in terms of physical activities. Establish any time off work for present and past LBP problems or any work modifications due to symptoms. Identify whether their symptoms are contributing to any financial difficulties.
Benefits.	If off work, ask them how they survive financially and record any benefits they may be on.
Family.	Find out a bit about the family situation and whether it suffers in any way by their back/leg problem.

Appendix E

Physical activity.	Ask them whether they do any regular physical activity/exercise. Or what they do to keep themselves fit and physically active
Assessment of psychosocial factors (yellow flags)	
Fear avoidance.	Prompts: If an activity is causing an increase in your symptoms do you stop that activity or carry on? Why? When you are doing something and it causes / increases your pain do you ever worry that you could be causing yourself harm / damage? Do you regulate your activities according to your pain / how you feel / according to a plan? Do you think that pain is always a sign that you are causing yourself harm/damage?
Distress (need to differentiate between back pain related distress and other reasons such as illness in the family, death in the family, relationship problems. Physiotherapists can only consider addressing pain related distress only.)	Prompts: Is there anything upsetting you or worrying you about your pain at the moment? Do you feel stressed or feel things are getting out of control? People often tell us that it is difficult to cope with pain as well as other things – is there anything going on in your life at the moment that is impacting on your ability to cope with your pain? Is there anything else you want to tell me?
Low mood / Depression	Prompts: How do you feel you are coping with your pain? What do you find most upsetting about your current pain? What are you most worried / concerned about? Does your pain get you down? OR Often people who have persistent pain (or pain that has gone on for a long time, or pain that keeps coming back) tell us that at times it affects their mood. Have you found this?
Coping strategies	Prompts: What are you currently doing to help you cope with the pain? When you are in increased pain what do you do? How much are you resting?

Appendix E

	Have you had to lie down during the day because of your pain? How do you think you will be in 6 months time
Work issues	Prompts: Do you enjoy your job? Do you think that you will be able to return to work? When? How has your ability to work being affected by your pain? Have you had time off in the past due to back pain?
Compensation / Litigation	Prompts: Are you involved in claims or litigation because of back pain problems? To do with? Why? Have you had to get involved in claims or litigation because of back pain problems in the past?
Future outlook	Prompts: If you look to the future I am sure that you would hope this pain would have settled but how do you expect to be, let's say in five years time?

Appendix E

Physical Examination

General Observation: with patient undressed down to underwear, have a look for any gross abnormalities. Make note of weight if outside normal, e.g. overweight, obese, too thin. Make note of spinal curves if necessary, e.g. reduced or loss of lordosis, kyphosis, scoliosis, etc. Note LLD if clearly present/obvious. Note any muscle wasting. Record partial weight-bearing if present (normally is due to pain). Record any lumbar shift.

Brief note on gait if necessary or appropriate, e.g. visible limping, unsteadiness.

Lumbar ROM: eyeballing and response of pain.

Neurological examination, lower limbs:
Always perform in the same order: Myotomes – Reflexes – Sensation

Myotomes
Dorsiflexion, plantar flexion, knee extension are checked in fully weight-bearing position (functional tests). Only if patient is unable to weight-bear, check in supine.

S1-S2 (ankle plantarflexor strength-toe walking):
Ask the patient to walk on their tip toes for couple of seconds.
Alternative way: hold their hands for balance and ask them to walk on their tip toes on the spot, again for couple of seconds. Inability to maintain heel off the ground while taking few steps is a positive result.

L4-L5 (ankle dorsiflexor strength-heel walking):
Ask the patient to walk on their heels for couple of seconds.
Alternative way: hold their hands for balance and ask them to walk on their heels on the spot, again for couple of seconds. Inability to maintain the forefoot off the ground is a positive result.

L3-L4 (knee extension):
Ask the patient to stand one leg and to squat slightly once.
Alternative way: hold their hands for balance and ask them to squat slightly once.
Alternative way: single leg sit-to-stand (perform test on good side first); patient is sitting in a chair, examiner in front of patient (holding hands for balance-optional). Ask patient to lift good leg of the ground and rise up using strength of painful leg only.

L5 (big toe extension-EHL):
In supine, ask patient to pull foot and toes towards face/nose and stop you from pulling big toe down.

L2 (L1) (hip flexion):
In supine, ask patient to pull their knee towards chest (or examiner does that), keep it there and not let you push it down.

Reflexes
Knee jerk: in sitting (can be done in supine as well).
Ankle jerk: supine or prone (there are at least 3 different ways to assess ankle jerk).
Use reinforcement techniques if unable to elicit.

If you are unsure whether lower limb reflexes are brisk or reduced, assess reflexes in upper limbs as well to get a feel for overall reflex quality.
Clonus at the ankle: up to 3 beats are normal. Describe any findings.
Clonus at the patella, optional (any clonus at the patella is considered abnormal).
Plantar responses.

Sensation
Pin prick (PP)
Tell patient that you will check the feeling in their legs with a small scratchy pin. Tell them that you want them to tell you whether the scratchy feeling in one side feels the same as the other, as a pin. Start with good side. If patient says it feels different, clarify. The feeling might be reduced but still feels as a pin (or scratchy).
If the feeling in completely lost (they cannot feel you touch them at all), record as anaesthesia.
If the feeling is not scratchy and it feels blunt (doesn't feel like a pin), record as 'loss of PP'.
If the feeling is still scratchy but less so compared to good side, record as 'subjective changes to PP' (or reduced PP sensation).

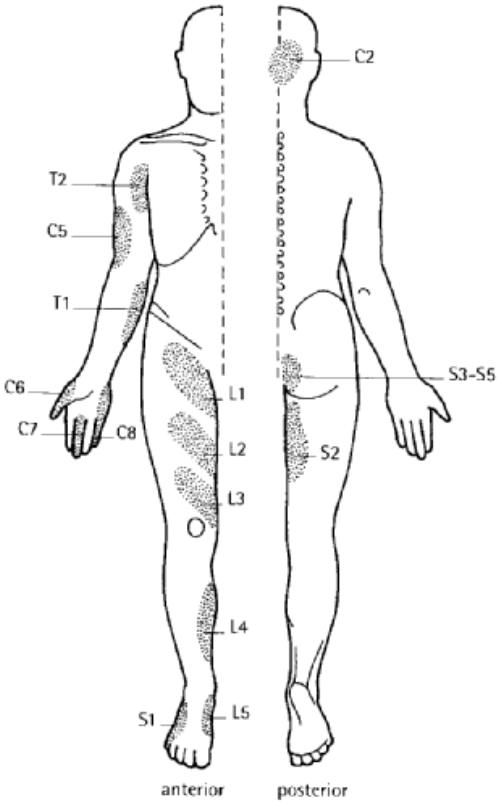
Neural tension tests
SLR (assesses low nerve roots; L5, S1): it is positive for neural tension if patient's own leg pain is reproduced and preferably in a dermatomal distribution.
Should 'sensitise' it with dorsiflexion.
Crossed SLR: patient's own leg pain reproduced upon passive SLR of the asymptomatic leg).
Slump test, optional (note predominately pain response as opposed to tightness).
(Note: in about 20% of patients with radiculopathy due to disc herniation, confirmed on MRI, SLR can be negative.)
Record angle at which symptoms are positive (e.g. right SLR +ve @ 50°)

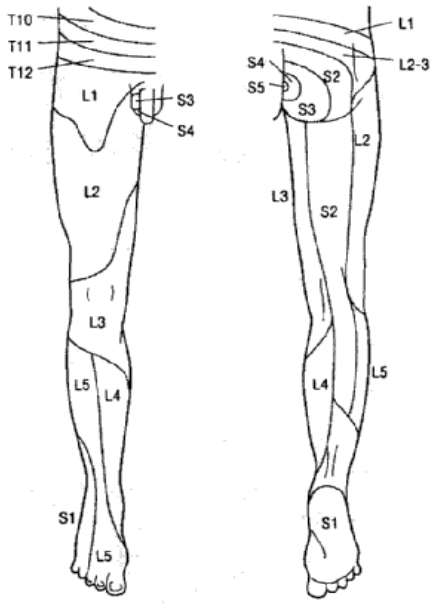
Femoral stretch test (assesses midlumbar nerve roots; L2, L3, L4): Normally performed with the patient in prone position. Reproduction of patient's typical pain in the leg constitutes a positive test.

Palpation findings
Simply record pain response to lumbar spine palpation. Any pain/discomfort should be the one that the patient complains about as opposed to just tenderness in general on pressure which does not resemble patient's own symptoms.
For example, severe pain or tenderness over lower or middle or upper parts of lumbar spine, or no tenderness on palpation.

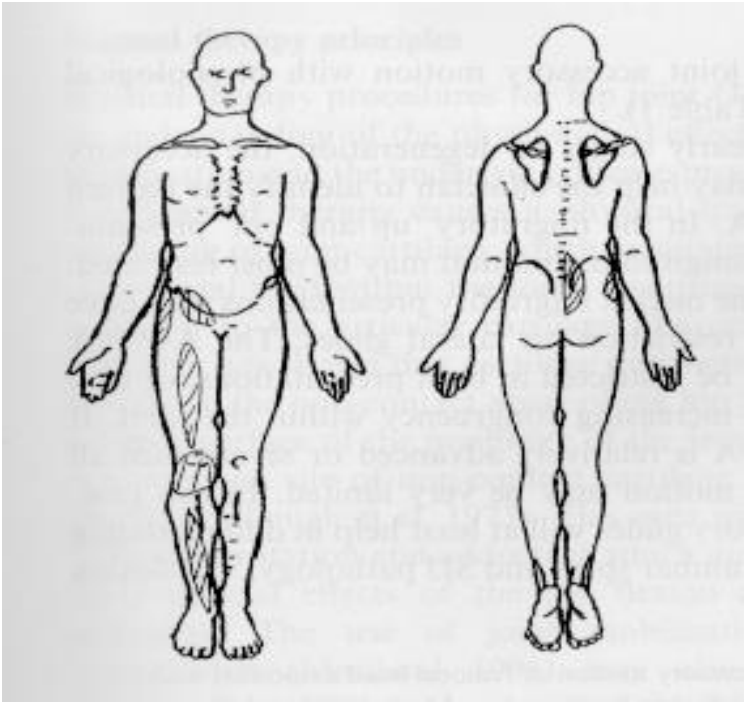
Hip assessment
Should be brief and mainly for excluding the hip as source of symptoms.
Functional test: squatting and gently moving hips from side to side.
In supine: passive testing of flexion, rotations, abduction. 'Leg roll' if necessary.

Pulses in legs
Consider checking pulses (mainly pedis dorsalis) if there is any suspicion of vascular problems presenting with leg pain


C2
T2
C5
T1
C6
C7
C8
L1
L2
L3
L4
S1
L5
S3-S5
S2
anterior
posterior


T10
T11
T12
L1
S3
S4
L2
L3
L5
L4
S1
L5
L1
L2-3
S4
S5
S2
S3
L2
L3
S2
L5
L4
S1

Appendix E



Characteristics attributable to hip involvement

- Pain worsened by weight-bearing, better with rest
- Worse with few first steps after rest
- Pain felt in groin, anterior thigh, knee, anterior shin, lateral thigh (around the greater trochanter), sometimes buttock
- Restricted hip movements (rotations, flexion, abduction, extension), pain on assessing hip movements-usually end range, hip quadrant test (flexion/adduction)

Neurological examination: Head and Arms

Movement	Nerve Root	Reflex
Neck flexion	C1-C2 (CN XI)	
Neck side-flexion	C3	
Shoulder elevation	C4 (CN XI)	
Shoulder abduction (deltoid)	C5 (axillary nerve)	Deltoid reflex
Elbow flexion (biceps) (keep arm supinated)	C5, C6 (musculocutaneous nerve)	Biceps C5 (C6)

Appendix E

Elbow flexion (brachioradialis) (keep arm semi-pronated)	C6 (radial nerve)	Supinator (C5) C6
Elbow extension (triceps)	C6, C7, C8 (radial nerve)	Triceps C7
Finger extension (extensor digitorum)	C7 (C8) (posterior interosseous, branch of radial)	
Finger flexion (flexor digitorum superficialis and profundus)	C8 (median nerve)	Finger flexion C8
Intrinsic hand muscles (fingers abduction, adduction)	T1 (ulnar nerve)	
Thumb abduction (abductor pollicis brevis)	T1 (median nerve)	

Appendix E

Systematic Review search strategies

EMBASE search 25 july 2013

1 Back Pain/ (31253)
2 Spine/ (26311)
3 Back/ (8667)
4 lumbo$.ti,ab. (13822)
5 backache.ti,ab. (2357)
6 back ache.ti,ab. (99)
7 (spinal or spine).ti,ab. (299000)
8 lumbar.ab,ti. (92407)
9 "back pain".ab,ti. (37268)
10 Low Back Pain/ (33604)
11 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 (406573)
12 (leg adj3 pain).ti,ab. (4676)
13 (nerve adj3 pain).ti,ab. (2811)
14 (radi$ adj3 pain).ti,ab. (8467)
15 neuropathic.ti,ab. (22037)
16 (referr$ adj3 pain).ti,ab. (3381)
17 "nerve root$".ti,ab. (10011)
18 Polyradiculopathy/ (6828)
19 Nerve Compression Syndromes/ (11312)
20 radicul$.ti,ab. (13198)
21 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 (68823)
22 11 and 21 (26766)
23 Sciatica/ (55)
24 sciatic$.ti,ab. (25978)
25 ischialgia/ (6145)
26 23 or 24 or 25 (28689)
27 vertebral canal stenosis/ (6622)
28 spinal stenosis.ti,ab. (3688)
29 27 or 28 (7491)
30 intervertebral disk hernia/ (14664)
31 ((disc or discs) adj1 (displacement$ or hernia$ or protru$ or avulsion$)).ti,ab. (7993)
32 ((disk or disks) adj1 (displacement$ or hernia$ or protru$ or avulsion$)).ti,ab. (2859)
33 30 or 31 or 32 (19667)
34 "non specific low back pain".ti,ab. (453)
35 "nonspecific low back pain".ti,ab. (360)
36 "low back-related leg pain".ti,ab. (18)
37 34 or 35 or 36 (808)
38 22 or 26 or 29 or 33 or 37 (71835)
39 Diagnosis/ (925999)
40 Diagnosis, Differential/ (312483)
41 (clinical adj1 predict$).ti,ab. (11616)
42 (clinical adj1 rule$).ti,ab. (238)
43 (predict$ adj3 (model$ or rule$)).ti,ab. (74477)
44 (diagnos$ adj3 (model$ or rule$)).ti,ab. (6038)
45 (classification or classified).ti,ab. (428519)
46 identification.ti,ab. (488098)
47 "subgroup$".ti,ab. (164152)
48 "sub-group$".ti,ab. (9633)
49 39 or 40 or 41 or 42 or 43 or 44 or 45 or 46 or 47 or 48 (2284219)
50 38 and 49 (9757)
51 limit 50 to embase (8049)

CINAHL search 25 July 2013

1. CINAHL; BACK PAIN/; 5634 results
2. CINAHL; LOW BACK PAIN/; 8965 results
3. CINAHL; BACK/; 1221 results
4. CINAHL; SPINE/; 4172 results
5. CINAHL; lumbo*.ti,ab; 1012 results
6. CINAHL; lumbar.ti,ab; 8224 results
7. CINAHL; "back ache".ti,ab; 5 results
8. CINAHL; "backache".ti,ab; 156 results
9. CINAHL; ((spinal OR spine)).ti,ab; 29021 results
10. CINAHL; 1 OR 2 OR 3 OR 4 OR 5 OR 6 OR 7 OR 8 OR 9; 45516 results
11. CINAHL; ((leg adj3 pain)).ti,ab; 918 results
12. CINAHL; ((nerve adj3 pain)).ti,ab; 481 results
13. CINAHL; ((radi* adj3 pain)).ti,ab; 1250 results
14. CINAHL; neuropathic.ti,ab; 2740 results
15. CINAHL; ((referr* adj3 pain)).ti,ab; 646 results
16. CINAHL; POLYRADICULOPATHY/; 128 results
17. CINAHL; "nerve root*".ti,ab; 720 results
18. CINAHL; radicul*.ti,ab; 1273 results
19. CINAHL; 11 OR 12 OR 13 OR 14 OR 15 OR 16 OR 17 OR 18; 6970 results
20. CINAHL; 10 AND 19; 2549 results
21. CINAHL; SCIATICA/; 604 results
22. CINAHL; sciatic*.ti,ab; 1137 results
23. CINAHL; 21 OR 22; 1384 results
24. CINAHL; "spinal stenosis".ti,ab; 593 results
25. CINAHL; SPINAL STENOSIS/; 806 results
26. CINAHL; 24 OR 25; 968 results
27. CINAHL; INTERVERTEBRAL DISK DISPLACEMENT/; 1646 results
28. CINAHL; ((disc OR discs) adj1 (displacement* OR hernia* OR protru* OR avulsion*)).ti,ab; 960 results
29. CINAHL; ((disk OR disks) adj1 (displacement* OR hernia* OR protru* OR avulsion*)).ti,ab; 253 results
30. CINAHL; 27 OR 28 OR 29; 2201 results
31. CINAHL; DIAGNOSIS/; 2759 results
32. CINAHL; DIAGNOSIS, DIFFERENTIAL/; 26024 results
33. CINAHL; ((clinical adj1 predict*)).ti,ab; 2000 results
34. CINAHL; ((clinical adj1 rule*)).ti,ab; 497 results
35. CINAHL; ((predict* adj3 (model* OR rule*))).ti,ab; 6227 results
36. CINAHL; ((diagnos* adj3 (model* OR rule*))).ti,ab; 810 results
37. CINAHL; ((classification OR classified)).ti,ab; 29740 results
38. CINAHL; identification.ti,ab; 23163 results
39. CINAHL; "subgroup*".ti,ab; 13665 results
40. CINAHL; "sub-group*".ti,ab; 733 results
41. CINAHL; 31 OR 32 OR 33 OR 34 OR 35 OR 36 OR 37 OR 38 OR 39 OR 40; 99779 results
42. CINAHL; "non specific low back pain".ti,ab; 178 results
43. CINAHL; "nonspecific low back pain".ti,ab; 203 results
44. CINAHL; "low back-related leg pain".ti,ab; 11 results
45. CINAHL; 42 OR 43 OR 44; 388 results
46. CINAHL; 20 OR 23 OR 26 OR 30 OR 45; 6366 results
47. CINAHL; 41 AND 46; 722 results

Cochrane Library search31 July 2013

#1 MeSH descriptor: [Back Pain] explode all trees 2632
#2 lumbo* 800
#3 backache 534
#4 "backache" 532
#5 (spinal or spine) 14971
#6 lumbar 6332
#7 "back pain" 5214
#8 MeSH descriptor: [Spine] explode all trees 3365
#9 {or #1-#8} 20511
#10 (leg adj3 pain) 148
#11 (nerve adj3 pain) 226
#12 (radi* adj3 pain) 403
#13 "neuropathic" 1225
#14 (referr* adj3 pain) 404
#15 (nerve root*) 703
#16 MeSH descriptor: [Polyradiculopathy] explode all trees 10
#17 MeSH descriptor: [Nerve Compression Syndromes] explode all trees 472
#18 radicul* 761
#19 {or #10-#18} 3448
#20 #9 and #19 1331
#21 MeSH descriptor: [Sciatica] explode all trees 220
#22 "sciatic*" 852
#23 MeSH descriptor: [Spinal Stenosis] explode all trees 155
#24 "spinal stenosis" 284
#25 MeSH descriptor: [Intervertebral Disc Displacement] explode all trees 574
#26 ((disc or discs) adj1 (displacement* or hernia* or protru* or avulsion*)) 9
#27 ((disk or disks) adj1 (displacement* or hernia* or protru* or avulsion*)) 9
#28 "non specific low back pain" 174
#29 "nonspecific low back pain" 105
#30 "low back-related leg pain" 1
#31 {or #20-#30} 2749
#32 MeSH descriptor: [Diagnosis] explode all trees 232499
#33 (clinical adj 1 predict*) 722
#34 (clinical adj 1 rule*) 531
#35 (predict* adj3 (model* or rule*)) 657
#36 (diagnos* adj3 (model* or rule*)) 1249
#37 (classification or classified) 21347
#38 identification 14958
#39 "subgroup*" 23193
#40 "sub-group*" 2255
#41 {or #32-#40} 260240
#42 #41 and #31 1703

AMED search 25 July 2013

Back

1. AMED; lumbo*.ti,ab; 501 results.
2. AMED; backache.ti,ab; 69 results.
3. AMED; "back ache".ti,ab; 7 results.
4. AMED; ((spinal OR spine)).ti,ab; 10540 results.
5. AMED; lumbar.ti,ab; 3305 results.
7. AMED; "back pain".ti,ab; 5024 results.
8. AMED; LOW BACK PAIN/; 3694 results.
9. AMED; SPINE/; 1513 results.
10. AMED; BACK/; 403 results.
11. AMED; 1 OR 2 OR 3 OR 4 OR 5 OR 7 OR 8 OR 9 OR 10; 16717 results.

Leg

12. AMED; (leg ADJ3 pain).ti,ab; 300 results.
13. AMED; (nerve adj3 pain).ti,ab; 160 results.
14. AMED; (radi* adj3 pain).ti,ab; 479 results.
15. AMED; neuropathic.ti,ab; 595 results.
16. AMED; (referr* adj3 pain).ti,ab; 275 results.
17. AMED; "nerve root*".ti,ab; 202 results.
18. AMED; radicul*.ti,ab; 381 results.
19. AMED; 12 OR 13 OR 14 OR 15 OR 16 OR 17 OR 18; 2022 results.

Back and Leg

20. AMED; 11 AND 19; 861 results.

Sciatica

21. AMED; SCIATICA/; 135 results.
22. AMED; sciatic*.ti,ab; 338 results.
23. AMED; 21 OR 22; 360 results.

Stenosis

24. AMED; SPINAL STENOSIS/; 113 results.
25. AMED; "spinal stenosis".ti,ab; 157 results.
26. AMED; 24 OR 25; 184 results.

Disc

27. AMED; INTERVERTEBRAL DISK DISPLACEMENT/; 291 results.
28. AMED; ((disc OR discs) adj1 (displacement* OR hernia* OR protru* OR avulsion*)).ti,ab; 263 results.
29. AMED; ((disk OR disks) adj1 (displacement* OR hernia* OR protru* OR avulsion*)).ti,ab; 76 results.
30. AMED; 27 OR 28 OR 29; 495 results.

NSLBP

31. AMED; "non specific low back pain".ti,ab; 115 results.
32. AMED; "nonspecific low back pain".ti,ab; 98 results.
33. AMED; "low back-related leg pain".ti,ab; 3 results.
34. AMED; 31 OR 32 OR 33; 212 results.

All back and leg pain

35. AMED; 20 OR 23 OR 26 OR 30 OR 34; 1862 results.

Diagnosis/Classification

36. AMED; DIAGNOSIS/; 12509 results.
38. AMED; DIAGNOSIS DIFFERENTIAL/; 316 results.
39. AMED; ((clinical adj1 predict*)).ti,ab; 184 results.
40. AMED; ((clinical adj1 rule*)).ti,ab; 6 results.
41. AMED; ((predict* adj3 (model* OR rule*))).ti,ab; 958 results.
42. AMED; ((diagnos* adj3 (model* OR rule*))).ti,ab; 106 results.
43. AMED; ((classification OR classified)).ti,ab; 3989 results.
44. AMED; identification.ti,ab; 2464 results.
45. AMED; "subgroup*".ti,ab; 1314 results.
46. AMED; "sub-group*".ti,ab; 124 results.
47. AMED; 36 OR 38 OR 39 OR 40 OR 41 OR 42 OR 43 OR 44 OR 45 OR 46; 20331 results.

Diagnosis/Classification LBLP

48. AMED; 35 AND 47; 425 results.

Appraisal scores for individual studies in systematic review (Y= yes; N= no; P=partial; DK=don't know; NA= non applicable)(Chapter four)

	Albert 2012	Barker 1990	Ben Debba 2000	Bernard 1985	Cassisi 1995	Delitto 2012	Fritz 2007	Glassman 2011	Hahne 2011	Hall 1994	Mckenzie 1981	Nachemson 1982	Nijs 2015	Paatelma 2009	Petersen 2003	Roach 1997	Schafer 2009	Scholz 2009	Spitzer 1987	Smart 2011	Sweetman 1992	Vining 2013
Purpose Is purpose, population and setting clearly specified?	Y	Y	Y	Y	Y	Y	Y	P	Y	P	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Content validity (i) Is the domain and all specific exclusions from the domain clearly specified?	Y	P	Y	P	Y	Y	Y	P	Y	P	Y	Y	Y	Y	Y	P	Y	Y	Y	Y	P	Y
(ii) Are all relevant categories included?	P	N	N	P	N	Y	N	N	N	Y	N	P	Y	Y	Y	P	Y	Y	Y	Y	P	Y
(iii) Is the breakdown of categories appropriate, considering the purpose?	Y	N	Y	Y	Y	Y	Y	P	Y	Y	Y	N	Y	Y	Y	DK	Y	Y	P	Y	P	Y
(iv) Are the categories mutually exclusive?	Y	N	Y	N	N	N	P	N	Y	Y	N	N	N	Y	N	N	P	Y	N	Y	N	N
(v) Was the method of development appropriate?	P	N	P	N	N	Y	Y	N	P	P	N	DK	P	N	P	N	P	Y	Y	Y	P	Y
(vi) If multiaxial, are criteria of content validity satisfied for each additional axis?	NA	NA	NA	NA	Y	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	P	NA	NA	NA
Face Validity (i) Is the nomenclature used to label the categories satisfactory?	Y	N	Y	Y	P	P	N	Y	Y	Y	Y	P	Y	Y	Y	Y	Y	Y	Y	Y	N	Y
(ii) Are the terms used based upon empirical (directly observable) evidence?	Y	N	Y	N	P	Y	P	Y	Y	Y	Y	Y	Y	P	Y	N	Y	Y	Y	Y	Y	Y
(iii) Are the criteria for determining inclusion into each category clearly specified?	Y	Y	Y	N	Y	P	Y	P	Y	Y	Y	Y	N	N	Y	P	Y	Y	Y	Y	Y	Y
(iv) If yes do these criteria appear reasonable?	Y	N	P	P	P	Y	Y	P	Y	P	Y	N	P	P	Y	P	Y	P	Y	P	N	Y
(v) Have the criteria been demonstrated to have reliability or validity?	Y	N	P	N	P	P	P	N	P	P	P	DK	P	DK	P	P	Y	P	P	P	N	P
(vi) Are the definitions of criteria clearly specified?	Y	P	Y	N	Y	P	Y	P	Y	Y	Y	Y	P	Y	P	P	Y	P	N	Y	N	Y
(vi) If multiaxial are criteria of face validity satisfied for each additional axis?	NA	NA	NA	NA	Y	NA	NA	N	NA	NA	NA	DK	NA	NA	NA	NA	NA	NA	N	NA	NA	NA

Appraisal scores for individual studies in systematic review (Y= yes; N= no; P=partial; DK=don't know; NA= non-applicable)

	Albert 2012	Barker 1990	Ben Debba 2000	Bernard 1985	Cassisi 1995	Delitto 2012	Fritz 2007	Glassman 2011	Hahne 2011	Hall 1994	Mckenzie 1981	Nachemson 1982	Nijs 2015	Paatelma 2009	Petersen 2003	Roach 1997	Schafer 2009	Scholz 2009	Spitzer 1987	Smart 2011	Sweetman 1992	Vining 2013
Feasibility (i) Is the classification simple to understand?	Y	P	Y	N	Y	N	Y	P	P	Y	N	Y	Y	Y	Y	P	Y	Y	Y	Y	N	Y
(ii) Is the classification easy to perform?	Y	DK	Y	N	P	N	DK	Y	P	Y	N	Y	DK	Y	N	DK	Y	Y	P	Y	N	Y
(iv) Any special skills/tools or training required?	Y	N	P	Y	DK	N	P	DK	Y	Y	Y	N	Y	Y	Y	DK	Y	P	Y	Y	Y	P
(v) How long does it take to perform?	DK	DK	DK	DK	DK	DK	DK	DK	DK	DK	DK	DK	DK	30m	1hr	DK	DK	DK	DK	DK	DK	DK
Construct Validity (i) Does it discriminate between entities thought to be different in a way appropriate for the purpose?	P	DK	P	DK	Y	Y	P	DK	N	Y	P	P	DK	DK	DK	DK	Y	Y	Y	Y	DK	DK
(ii) Does it perform satisfactorily compared to other systems classifying the same domain?	P	DK	DK	DK	DK	DK	DK	DK	N	DK	P	DK	DK	DK	P	DK	DK	Y	DK	DK	DK	DK
Reliability (i) Does the system provide consistent results when classifying the same conditions?	DK	DK	DK	N	DK	P	DK	Y	DK	Y	Y	DK	DK	Y	DK	P	Y	DK	DK	Y	P	DK
(ii) Is the intraobserver and interobserver reliability satisfactory?	DK	DK	DK	N	DK	P	DK	Y	DK	Y	Y	DK	DK	Y	DK	P	P	DK	DK	Y	P	DK
Generalisability (i) Has it been used in other studies &/or settings?	P	N	N	N	N	Y	N	N	N	Y	Y	Y	N	N	Y	N	Y	N	Y	N	N	P
TOTAL OVERALL SCORE	4	2	3.5	2	3	3.5	3	2.5	3	5	5.5	3.5	2.5	3.5	4	3	5	4	4	5	2.5	3.5

Appendix G

Reasons for diagnosis among raters A, B and C in cases of disagreement (Y yes; N no) (chapter five)

Case		Rater A	Rater B	Rater C
1	**Y/N sciatica**	**No**	**Yes**	**Yes**
	Confidence (%)	85%	85%	75%
	Reason 1	Reduced sensation objectively -non dermatomal	Distribution and quality description pain	Description of broadly dermatomal pain from thigh and calf
	Reason 2	Negative SLR	Numbness	Dulling *pin prick (difficult to read)* sensation
	Reason 3		Decreased sensation objectively	Description of heaviness to leg
	Reason 4		positive SLR	
2	**Y/N sciatica**	**No**	**Yes**	**Yes**
	Confidence (%)	70%	70%	80%
	Reason 1	Back worse than leg	Numbness	Description of numbness to lateral thigh
	Reason 2	Normal neurology	Decreased pin prick sensation	Sign of dulling pinprick to lateral thigh
				Leg pain aggravated by motion, eased by rest
3	**Y/N sciatica**	**Yes**	**No**	**No**
	Confidence (%)	75%	85%	70%
	Reason 1	Positive cough sneeze	Back more than leg pain	Pain on examination all in back not leg
	Reason 2	Sensitivity (objective allodynia/hyperalgesia)	No leg pain below knee	Normal neurology, few elements of nerve /radicular pain but majority point to referred pain.
4	**Y/N sciatica**	**No**	**Yes**	**Yes**
	Confidence (%)	75%	70%	55%
	Reason 1	Back worse than leg	Leg pain below knee	no typical root symptoms other than pain below the knee
	Reason 2		Positive SLR	All provoked pain in back other than possible positive sciatic stretch
	Reason 3		Reduced sensation objectively	

Case		Rater A	Rater B	Rater C
5	**Y/N sciatica**	**Yes**	**No**	**No**
	Confidence (%)	80%	80%	80%
	Reason 1	Worse with weight bearing, better sitting	Back more than leg pain	Pain to knee
	Reason 2	Strange sensation in legs (subjectively)	No leg pain below knee	Recurrent
	Reason 3		No neuro deficit or neural stretch	Other aches and pains ?osteoarthritis
	Reason 4		No subjective pins and needles/numbness	
6	**Y/N sciatica**	**No**	**Yes**	**Yes**
	Confidence (%)	85%	60%	60%
	Reason 1	Above knee symptoms	Reduced sensation objectively	Apparently absent left knee reflex
	Reason 2			Sensation in left foot diminished
7	**Y/N sciatica**	**Yes**	**No**	**No**
	Confidence (%)	90%	80%	95%
	Reason 1	Reduced strength Extensor Hallucis Longus	Back worse than leg	Central pain > peripheral pain
	Reason 2		Description of leg pain - dull and deep	Lack of dermatomal specificity
	Reason 3		Cough/ sneeze negative	No radiation/restriction on ROM testing-
	Reason 4			Lack of neurology/ positive SLR findings
8	**Y/N sciatica**	**Yes**	**No**	**No**
	Confidence (%)	70%	80%	70%
	Reason 1	Pins and needles right foot subjectively	Back worse than leg	Central pain>peripheral pain
	Reason 2	Hyperalgesia right big toe, shin ,calf objectively	SLR/femoral stretch/ slump negative	Lack of peripheralisation with ROM
	Reason 3		Hyperalgesia general areas	Lack of neurological findings
	Reason 4		Poor general well being/ fibromyalgia/ overweight	Pins and needles ? red herring related to fibromyalgia: warrants further investigation
9	**Y/N sciatica**	**Yes**	**No**	**No**
	Confidence (%)	75%	80%	70%
	Reason 1	Positive femoral stretch	No real neurology	Negative neural tension tests
	Reason 2	Reduced sensation medial thigh objectively	Negative femoral nerve stretch	Site of pain

Case		Rater A	Rater B	Rater C
	Reason 3		LBP more problematic	Back>leg pain
10	**Y/N sciatica**	**No**	**Yes**	**No**
	Confidence (%)	80%	75%	70%
	Reason 1	No dermatomal pattern	Below knee symptoms	Pain extends above knee
	Reason 2	Sensory changes non dermatomal	Sensory changes	Character of pain is throbbing
	Reason 3	No pattern to reflex changes	Positive SLR	Inconsistent neurological signs- Hyperaesthesia and non-dermatomal
	Reason 4			Long history and previous MRI scan – degenerative changes
11	**Y/N sciatica**	**Yes**	**Yes**	**No**
	Confidence (%)	75%	85%	60%
	Reason 1	Positive neural tension tests	Positive slump and femoral stretch	Back pain > leg pain
	Reason 2	Altered myotomes	Quads and hip flexor weakness	History of hip joint pain & hip flexor weakness
	Reason 3			Only transient pain on femoral nerve stretch
	Reason 4			Lack of sensory abnormality, historical leg weakness

2x2 Table of descriptive cross-tabulated frequency data of categorical predictors (<u>before re-categorisation)</u> for diagnosis of sciatica in patients with LBLP reference standard one: high confidence (≥ 80%) clinical diagnosis and reference standard two: high confidence (≥ 80%) clinical diagnosis plus confirmatory MRI findings (clear or possible nerve root compression).

		Model one Reference standard		Model two Reference standard	
Self report items		Sciatica (*n*=295)	Referred (*n*=100)	Sciatica (*n*=200)	Referred (*n*=195)
Sensory changes	Yes	161 (54.6%)	12 (12.0%)	110 (55.0%)	63 (32.3%)
	No	134 (45.4%)	88 (88.0%)	90 (45.0%)	132 (67.7%)
Below knee pain	Yes	250 (84.7%)	28 (28.0%)	173 (86.5%)	105 (53.8%)
	No	45 (15.3%)	72 (72.0%)	27 (13.5%)	90 (46.2%)
Leg pain worse than back pain	Yes	183 (62.0%)	15 (15.0%)	142 (71.0%)	56 (28.7%)
	No	112 (38.0%)	85 (85.0%)	58 (29.0%)	139 (71.3%)
Cough/Sneeze	Yes	95 (32.2%0	7 (7.0%)	79 (39.5%)	23 (11.8%)
	No	200 (67.8%)	93 (93.0%)	121 (60.5%)	172 (88.2%)
Clinical Examination items					
Straight leg raise (SLR)	Yes	211 (71.5%)	10 (10.0%)	146 (73.0%)	75 (38.5%)
	No	84 (28.5%)	90 (90.0%)	54 (27.0%)	120 (61.5%)
Crossed SLR	Yes	21 (7.1%)	0 (0.0%)	19 (9.5%)	2 (1.0%)
	No	274 (92.9%)	100 (100%)	181 (90.5%)	193 (99.0%)
Slump	Yes	62 (21.0%)	2 (2.0%)	40 (20%)	24 (12.3%)
	No	233 (79.0%)	98 (98.0%)	160 (80%)	171 (87.7%)
Femoral nerve stretch	Yes	27 (9.2%)	0 (0.0%)	14 (7.0%)	13 (6.7%)
	No	268 (90.8%)	100 (100.0%)	186 (93.0%)	182 (93.3%)
Myotomes (n=394)	Yes	81 (27.5%)	0 (0.0%)	59 (29.5%)	22 (11.3%)
	No	214 (72.5%)	99 (100.0%)	141 (70.5%)	172 (88.7%)
Sensation	Yes	161 (54.6%)	12 (12.0%)	110 (55.0%)	63 (32.3%)
	No	134 (45.4%)	88 (88.0%)	90 (45.0%)	132 (67.7%)
Reflexes	Yes	86 (29.2%)	5 (5.0%)	75 (37.5%)	16 (8.2%)
	No	209 (70.8%)	95 (95.0%)	125 (62.5%)	179(91.8%)

Table comparing key characteristics for latent class 3 and high confidence clinical diagnosis referred leg pain group (chapter seven).

	Latent class 3	**Referred pain**
	n=83	**n=100**
Age (years) mean (SD)	48.7 (14.7)	48.6 (13.8)
Gender, Female	55 (66.3)	67 (67)
Current smoker	19 (22.9)	25 (25)
BMI mean (SD)	29.0 (5.6)	29.0 (5.7)
Self-certified time off work or given sick note due to current episode (100)	22 (26.5)	24 (24.0)
Back/leg interference with work performance, mean (SD)	4.7 (3.0)	4.9 (2.8)
RMDQ disability score (0-23) mean (SD)	10.9 (5.7)	11.3 (5.8)
Back pain intensity,(0-10) mean (SD)	4.9 (2.0)	5.2 (2.0)
Leg pain intensity,(0-10) mean (SD)	3.3 (1.9)	3.7 (2.1)
Sciatica Bothersomeness Index (SBI)(0-24), mean (SD)	10.8 (5.3)	11.5 (5.2)
S-LANSS, neuropathic pain score (≥12) predominantly neuropathic	26 (31.7)	35 (35.4)
STarTBack risk score		
Low	18 (22.8)	22 (23.2)
Medium	35 (44.3)	37 (38.9)
High	26 (32.9)	36 (37.9)
HADS Anxiety		
Mild/possible	16 (19.3)	24 (24.0)
Probable/moderate/severe	25 (30.1)	30 (30.0)
HADS Depression		
Mild/possible	16 (19.3)	19 (19.0)
Probable/moderate/severe	12 (14.5)	14 (14.0)
EQ—5D summary index	0.5 (0.3)	0.5 (0.3)
Sleep Disturbance due to back/leg pain	50 (60.2)	61 (20.7)
MRI findings		
Clear nerve root compression	16 (19.3)	18 (18.0)
Possible nerve root compression	10 (12.0)	13 (13.0)
Clinicians confidence in diagnosis, mean (SD)	87.6 (7.3)	87.2 (6.9)

SD, standard deviation; BMI, body mass index; RMDQ, Roland Morris Disability Questionnaire; s-LANSS, self-report Leeds Assessment of Neuropathic Symptoms and Signs; HADS, Hospital Anxiety and Depression Scale.
All figures are frequencies (percentages) unless otherwise stated as mean (SD).

Table reproduced (in part) from baseline ATLAS results paper (Konstantinou et al. 2015). Comparison of key characteristics for patients diagnosed with or without sciatica. (Chapter seven)

Characteristics	Sciatica pain *n* = 452 (74.2%)	Referred pain *n* = 157 (25.8%)	Sig.*
Socio-demographics			
Age (years), mean (SD)	50.4 (14.0)	49.4 (13.7)	0.451
Gender, Female	276 (61.1)	105 (66.9)	0.194
BMI	29.9 (6.3)	30.0 (8.7)	0.906
Current smoker	151 (33.4)	43 (27.4)	0.163
Back/leg pain interference with work performance, mean (SD) †	6.0 (2.9)	5.4 (3.0)	0.073
Self-certified time off work or given sick note due to current episode (365)†	111 (40.8)	33 (36.3)	0.443
Pain and function			
RMDQ disability score (0-23), mean (SD)	12.9 (5.7)	12.0 (5.7)	0.093
Back pain intensity, mean (SD)	5.6 (2.2)	5.4 (2.1)	0.413
Leg pain intensity, mean (SD)	5.6 (2.3)	4.1 (2.3)	<0.001
Pain below knee	333 (76.6)	61 (40.9)	<0.001
Leg pain is worse	252 (55.8)	28 (17.8)	<0.001
Sleep disturbance due to back/leg pain	325 (72)	103 (66)	0.447
EQ—5D summary index	0.4 (0.3)	0.5 (0.3)	0.391
Start Back risk score			0.086
Low	53 (12.1)	29 (19.1)	
Medium	212 (48.5)	64 (42.1)	
High	172 (39.4)	59 (38.8)	
Sciatica Bothersomeness Index mean (SD)			
Leg pain	4.5 (1.4)	3.6 (1.6)	<0.001
Numbness or tingling in leg, foot or groin	3.5 (2.0)	2.4 (2.1)	<0.001
Weakness in leg or foot	2.8 (2.0)	2.0 (2.1)	<0.001
Back or leg pain while sitting	4.1 (1.6)	4.1 (1.7)	0.084
Composite score	14.9 (5.1)	12.2 (5.4)	<0.001
S-LANSS; neuropathic pain score (≥12)	232 (51.6)	61 (39.0)	<0.007
Psychological measures			
HADs anxiety subscale			0.023
Mild/possible case	86 (19.1)	34 (21.8)	
Probable/moderate/severe case	116 (25.7)	55 (35.3)	
HADs depression subscale			0.325
Mild/possible case	82 (18.1)	37 (23.4)	
Probable/moderate/severe case	75 (16.6)	23 (14.7)	
MRI findings(553)			<0.001
Nerve root compression	252 (60.7)	45 (32.4)	
Normal	163 (39.3)	94 (67.6)	

All figures are frequencies (percentages) unless stated otherwise as mean (SD).

*Significance p-value for the difference between participants diagnosed as having sciatica symptoms and those diagnosed as having referred leg pain based on Chi-squared test for categorical variables and 2-sample t-test for continuous variables.

†Applicable to those currently in paid job

Email to NHS clinical leads for workshop

Dear Manager

I am a musculoskeletal physiotherapist, and currently in the final year of my PhD studies on low back pain, at Keele University. I would like to share my research results with clinicians and hear their feedback regarding the clinical relevance and applicability of my findings.

I would be very grateful if you could identify any clinicians from your service, involved in the assessment and treatment of spinal pain patients, which would potentially be able to participate in a workshop to evaluate a diagnostic tool and classification system for low back-related leg pain.

Title of workshop: A diagnostic tool and classification system for low back-related leg pain- would you use this in your practice?

Objective of workshop

To evaluate the clinical relevance, feasibility and usefulness of:

- A clinical diagnostic tool to identify sciatica in patients with low back-related leg pain
- A classification system for primary care consulters with low back related leg pain

The target audience for the workshop are physiotherapist (band five up to ESP) who assess and/or treat low back pain patients as part of their case workload.

The workshop will be run at the Research Institute for Primary Care Sciences, Keele University, over 3 hours which will include lunch and refreshments. A certificate of attendance will be issued to all participants. The following dates/ times are available and the most convenient time for the majority of participants will be chosen.

Wednesday 29 June (am or pm)
Thursday 30 June (am)
Monday 4 July (am or pm)
Wednesday 6 July (am)
Thursday 14 July (am)

Yours sincerely

Siobhán Stynes

NIHR Clinical Doctoral Research Fellow/Research Physiotherapist
Arthritis Research UK Primary Care Centre
Keele University, Staffordshire
ST5 5BG.

Workshop Participant Details

Name	
Place of work	
Clinical Job Title(s)/role	
Years Qualified	
Years working primarily in MSK physio	
What % of your weekly caseload are spinal patients?	
Do you agree to content from <u>discussions</u> from this workshop to be used in future papers and presentations? You will not be identified by name, but referred to as a participant. Do you give your permission to be included in photographs of the workshop which may be used in future presentations	<u>Signature</u> Yes / No ---------- Yes / No ----------

Many thanks for your time and participation!

Appendix M

Workshop Presentation slides June 2016

Background

Low back-related leg pain (LBLP)

Nerve root involvement (NRI)
Radicular pain
Sciatica

No NRI
Referred pain
Non specific leg pain

LBLP is associated with:
- Increased levels of disability
- Greater work loss
- Higher health care costs

Why differentiate?

Clinical management

Research purposes

Current models

Individual items -poor diagnostic performance..

...."better performance may be obtained when tests are combined" Van der Windt et al 2010

Multivariable models: Variations in predictors, setting and reference standard.

Reference standard

Aims

Part 1

1. Determine clinicians agreement on the diagnosis of LBLP

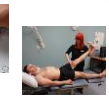
Part 2

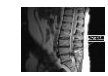
1. Develop a diagnostic tool to identify patients with leg pain due to NRI.

Methods

Sample
LBLP primary care consulters

ATLAS
Assessment and Treatment of Leg pain Associated with the Spine





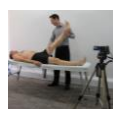
Part one: Reliability study

Aim: Examine the agreement and inter-rater reliability of diagnosing LBLP among clinicians

Method

Sample: 36 adult LBLP consulters

Standardized assessment – Video recorded



Method continued

Assessors:
Part I: MSK physiotherapists (six)
Part II: External Health professionals (six)

All clinicians asked

- Is the leg pain due to NRI? Yes/No
- Rate confidence in their diagnosis (%)

Results of Reliability study

Raters in Group B	Raters in Group A: NRI	Referred	Total
NRI	20	5	25
Referred	5	6	11
Total	25	11	36

Raters in Group C	Raters in Group A: NRI	Referred	Total
NRI	19	4	23
Referred	6	6	12
Total	25	10	35

Group A = Physiotherapists (involved in ATLAS study) assessing the patient
Group B = Physiotherapists (involved in ATLAS study) watching assessments on video
Group C = External health professionals watching assessments on video

Appendix N

Workshop Presentation slides June 2016

Results continued

	Part 1	Part 2
Sample size	N= 36	N=35
Agreement	72%	71%
Reliability Kappa (k)	Fair k=0.34	Fair k=0.34
95% confidence intervals	0.02, 0.68	0.02, 0.69

Conclusions for Reliability study

- Reliability is fair when diagnosing LBLP
- When confidence is higher, agreement and reliability improve substantially

ORIGINAL ARTICLE

Reliability among clinicians diagnosing low back-related leg pain

Shihbala Styans[2] · Kika Konstantinou[2] · Kate M. Dunn[2] · Martyn Lewis[1] · Elaine M. Hay[3]

Received: 9 June 2015 / Revised: 7 December 2015 / Accepted: 7 December 2015
© The Author(s) 2015. This article is published with open access at Springerlink.com

Abstract

Introduction

Part two: Clinical Diagnostic Model

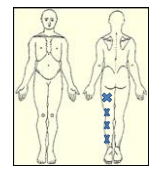
Aim: Identify items from clinical assessment that identify NRI

Multivariate diagnostic prediction model using logistic regression

Methods

Source of Data
Cross sectional cohort study
Participants
LBLP primary care consulters
Outcome
Nerve root involvement
Predictors
Self report and physical examination items

Self report variables



1. **Intensity of leg pain**
2. **Sensory changes**
3. **Pain below knee**
4. **Leg pain worse than back pain**
5. **Positive cough/sneeze**

Physical Examination variables

6. Neural Tension Tests

i. Straight leg raise 

ii. Slump 

iii. Femoral nerve stretch 

7. Neurological deficits:

i. Muscle strength 

ii. Sensation 

iii. Reflexes 

Reference standard

Reference standard

- Clinical Diagnosis

Clinical Diagnosis (when confidence ≥ 80%)

- Incorporation bias issues.....

Clinical Diagnosis (when confidence ≥ 80%) PLUS confirmatory MRI

Appendix N

Workshop Presentation slides June 2016

Statistical Analysis

- Multivariable logistic regression
  - stepwise backwards selection
- Model performance
  - calibration
  - discrimination
- Internal Validation
  - bootstrapping

Results for Diagnostic model

Demographics	N=395	
Female	236 (60%)	
Age (years) mean (standard deviation)	48.8 (13.9)	
Pain below knee	278 (70%)	
Duration of pain	Back	Leg
0-6 weeks	131 (33%)	165 (42%)
6 – 12 weeks	94 (24%)	85 (22%)
> 3 months	167 (42%)	125 (32%)
Intensity pain (0-10) mean (sd)	Back 5.5 (2.2)	Leg 5.3 (2.4)

Participants and outcome events

Clinical Diagnosis 295/394 (75%)

Clinical Diagnosis plus confirmatory MRI 200/394 (51%)

Odds ratios for multivariable model

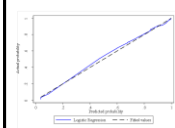
Reference Standard:	Clinical diagnosis NRI confidence ≥80%	Clinical diagnosis confidence ≥80% plus confirmatory MRI
Item	Odds Ratio (95% CI)	Odds Ratio (95% CI)
Subjective sensory changes	2.66 (1.20, 5.90)	--------
Below knee pain	6.25 (2.80, 13.94)	2.13 (1.19, 3.83)
Leg pain worse	4.55 (1.89, 10.99)	2.94 (1.77, 4.89)
Intensity leg pain	--------	1.15 (1.03, 1.29)
Positive cough / sneeze	--------	2.50 (1.34, 4.65)
Neural Tension tests	21.63 (9.00, 51.97)	1.76 (1.03, 3.00)
Neurological deficit	8.50 (3.80, 19.01)	2.81 (1.69, 4.69)

Odds ratios for multivariable model

Reference Standard:	Clinical diagnosis NRI confidence ≥80%	Clinical diagnosis confidence ≥80% plus confirmatory MRI
Item	Odds Ratio (95% CI)	Odds Ratio (95% CI)
Subjective sensory changes	2.66 (1.20, 5.90)	--------
Below knee pain	**6.25 (2.80, 13.94)**	**2.13 (1.19, 3.83)**
Leg pain worse	**4.55 (1.89, 10.99)**	**2.94 (1.77, 4.89)**
Intensity leg pain	--------	1.15 (1.03, 1.29)
Positive cough / sneeze	--------	2.50 (1.34, 4.65)
Neural Tension tests	**21.63 (9.00, 51.97)**	**1.76 (1.03, 3.00)**
Neurological deficit	**8.50 (3.80, 19.01)**	**2.81 (1.69, 4.69)**

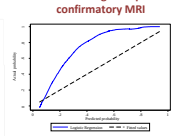
Calibration and discrimination

Clinical Diagnosis NRI



Bootstrapped model
AUC 0.95 (CI 0.93, 0.98)

Clinical Diagnosis plus confirmatory MRI



Bootstrapped model
AUC 0.82 (CI 0.78, 0.86)

Scoring Tool

Predictors	Beta	Beta/ 0.98	Rounded score
Subjective sensory changes	0.98	1	1
Below knee pain	1.83	1.9	2
Leg pain worse than back pain	1.52	1.6	2
Neural Tension tests	3.07	3.1	3
Neurological deficit	2.14	2.2	2
Total			10

Sum Score	Probability of Sciatica
0	4%
1	9%
2	19%
3	42%
4	63%
5	83%
6	93%
7	96%
8	99%
9	100%
10	100%

Conclusions

Interpretation

- Diagnosis of NRI is challenging

Strengths and Limitations

- Reference standard
- Predictor selection

Implications

Potential use of tool in research and clinical settings

Next step

s.stynes@keele.ac.uk

@SiobhanStynes

Appendix N

Workshop Presentation slides June 2016

Acknowledgements

Arthritis Research UK | primary care centre

This research is supported by the National Institute for Health Research .
Siobhán Stynes is supported by an NIHR/CNO Clinical Doctoral Research Fellowship.
Dr. Kika Konstantinou is supported by an HEFCE/NIHR Senior Clinical Lectureship.
Professor Hay is a NIHR Senior Investigator.
The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health.

NHS National Institute for Health Research

Arthritis Research UK | primary care centre

Keele University

Classification of low back-related leg pain

Siobhán Stynes

NHS National Institute for Health Research

Keele University

Agenda for primary care research on LBP 2009

CAN CLINICALLY RELEVANT SUBGROUPS OF LBP BE IDENTIFIED?

Keele University

Aim of this study

- Identify and describe classes of LBLP patients using Latent Class (LC) modelling
- Compare these classes to the clinically defined groups of LBLP patients with and without a diagnosis of sciatica.

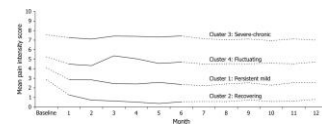
Keele University

Latent Class modelling

- Relates the observed patterns of test results to unobserved (latent) categories of patients
- Creates classes with minimal within-class variation and maximum between-class variation (Kongsted et al 2015).
- Circumvents the use of a reference standard

Keele University

Course of pain over 12 months among primary care low back pain consulters in the United Kingdom, 2001–2003.



Kate M. Dunn et al. Am. J. Epidemiol. 2006;163:754-761

American Journal of EPIDEMIOLOGY

Keele University

Methods

Sample

LBLP primary care consulters

ATLAS Assessment and Treatment of Leg pain Associated with the Spine



Variables

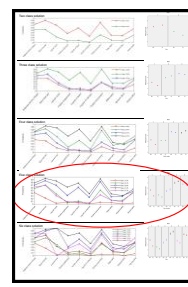
Self report and physical examination items

Keele University

Variables for LC modelling	
Self report items:	1) Leg pain intensity 2) Back pain intensity
History items from clinical examination:	3) Subjective sensory changes 4) Pain below knee 5) Leg pain worse than back pain 6) Positive cough/sneeze/strain
Physical examination items:	7) Leg pain on active flexion 8) Leg pain on active extension 9) Positive neural tension tests 10) Reflex deficit 11) Myotome deficit 12) Sensory deficit

---

Keele University

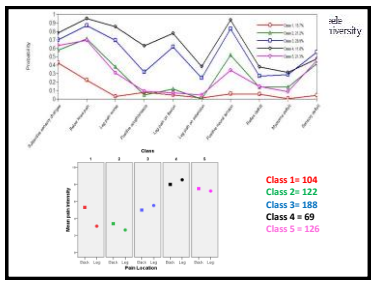


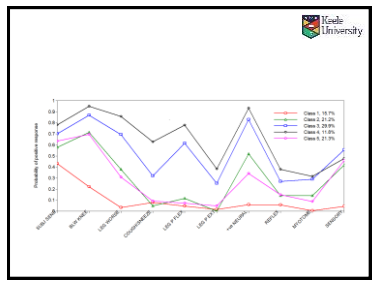
Results

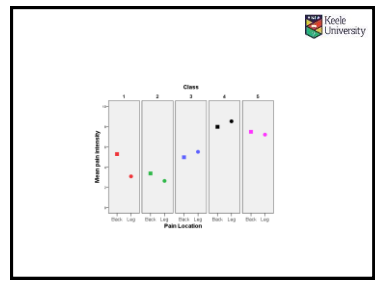
Five class solution was optimal

Appendix N

Workshop Presentation slides June 2016







Class Summary

1. Low intensity leg pain, some subjective sensory changes. Back pain > leg pain.
2. Low intensity back and leg pain. Back > leg pain. High probability of below knee pain; moderate probability of subjective and objective sensory deficits and positive SLR.
3. Moderate leg pain, Leg > back pain; Very high likelihood of below knee pain and positive neural tension. Moderate probability of objective sensory deficit.
4. High intensity back & leg pain. Leg > back pain. High probability neural tension, leg pain worse than back pain and below knee pain. Highest probability of neurological deficit and positive cough/sneeze.
5. High intensity back & leg pain, back pain > leg pain. Strong profile of subjective and objective sensory changes compared to other groups. Not likely to have positive neural tension. 70% probability of below knee pain.

Class Summary

1. Low intensity leg pain, some subjective sensory changes. Back pain > leg pain.
   81% with Clinical diagnosis of referred pain.
2. Low intensity back and leg pain. Back > leg pain. High probability of below knee pain; moderate probability of subjective and objective sensory deficits and positive SLR.
   81% with clinical diagnosis of sciatica.
3. Moderate leg pain, Leg > back pain; Very high likelihood of below knee pain and positive neural tension. Mod prob sensory deficit.
   93% with clinical diagnosis of sciatica.
4. High intensity back & leg pain. Leg > back pain. High probability neural tension, leg pain worse than back pain and below knee pain. Highest probability of neurological deficit and positive cough/sneeze.
   100% with clinical diagnosis of sciatica.
5. High intensity back & leg pain, back pain > leg pain. Strong profile of subjective and objective sensory changes compared to other groups. Not likely to have positive neural tension. 70% probability of below knee pain.
   71.0% with clinical diagnosis of sciatica.

MRI and clinical diagnosis characteristics

Diagnosis n	Class 1 104	Class 2 122	Class 3 188	Class 4 69	Class 5 126	p value*
Clinical diagnosis sciatica (%)	20 (19.2)	99 (81.1)	175 (93.1)	69 (100.0)	89 (70.6)	<0.001
MRI nerve root compression (%)	25 (26.3)	56 (50.5)	106 (63.1)	57 (89.1)	53 (45.7)	<0.001
Clinician confidence in diagnosis ≥80% (%)	72 (69.2)	75 (61.4)	156 (83.0)	63 (91.3)	70 (55.6)	<0.001

Class Summary

1. Low intensity leg pain, some subjective sensory changes. Back pain > leg pain.
   81% with Clinical diagnosis of referred pain.
2. Low intensity back and leg pain. Back > leg pain. High probability of below knee pain; moderate probability of subjective and objective sensory deficits and positive SLR.
   81% with clinical diagnosis of sciatica.
3. Moderate leg pain, Leg > back pain; Very high likelihood of below knee pain and positive neural tension. Mod prob sensory deficit.
   93% with clinical diagnosis of sciatica.
4. High intensity back & leg pain. Leg > back pain. High probability neural tension, leg pain worse than back pain and below knee pain. Highest probability of neurological deficit and positive cough/sneeze.
   100% with clinical diagnosis of sciatica.
5. High intensity back & leg pain, back pain > leg pain. Strong profile of subjective and objective sensory changes compared to other groups. Not likely to have positive neural tension. 70% probability of below knee pain.
   71.0% with clinical diagnosis of sciatica.

Socio-demographic and self report characteristics

	Class 1 104	Class 2 122	Class 3 188	Class 4 69	Class 5 126	p value
Socio-demographics						
Age (mean (sd))	47.2 (13.8)	50.4 (13.3)	50.9 (14.4)	49.2 (12.7)	51.9 (14.1)	0.111
Age > 65 (%)	13 (12.5)	17 (13.9)	33 (17.6)	7 (10.1)	22 (17.5)	0.238
Gender (female)	76 (73.1)	72 (59.0)	113 (60.1)	42 (60.9)	80 (63.5)	0.234
Current **Smoker** (%)	27 (26.0)	29 (23.8)	52 (27.7)	30 (43.5)	56 (44.4)	<0.001
BMI Obese	31 (29.8)	49 (40.5)	78 (41.5)	36 (52.2)	54 (43.2)	0.353
Socioeconomic status Routine work (%)	41 (39.4)	43 (36.1)	85 (46.4)	36 (55.4)	78 (63.9)	<0.001
Self-certified **time off work** (363) Sick note (365) (%)	25 (35.7) 22 (31.4)	20 (25.6) 16 (20.3)	42 (35.0) 34(28.3)	11 (29.7) 14 (37.8)	8 (13.8) 14 (16.2)	0.032 0.279

Health characteristics

	Class 1 104	Class 2 122	Class 3 188	Class 4 69	Class 5 126	p value [a]
Health						
General Health Fair/poor	38 (36.5)	31 (25.5)	59 (31.4)	32 (47.1)	62 (49.2)	<0.001
Co-morbidities Two or more other health problems	16 (15.4)	15 (12.3)	21 (11.2)	5 (7.2)	23 (18.3)	0.139
Sleep Disturbance (yes)	69 (66.3)	73 (59.8)	133 (70.7)	61 (88.4)	92 (73.0)	0.001

Duration						
Back pain > 6 weeks	64 (61.5)	72 (59.0)	117 (62.6)	47 (68.1)	89 (71.2)	0.279
Leg pain > 6 weeks	59 (60.5)	52 (45.2)	105 (57.7)	38 (57.6)	87 (71.9)	0.001
>12 months	15 (15.2)	10 (8.7)	24 (13.2)	3 (4.5)	29 (24.0)	<0.001

Appendix N

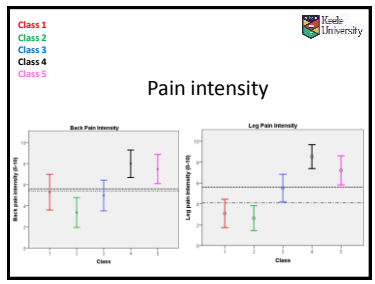
Workshop Presentation slides June 2016

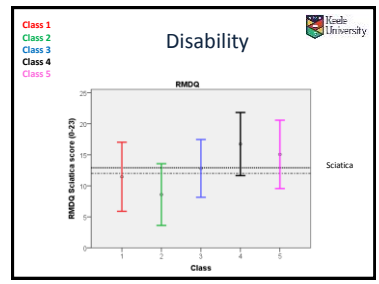
Sciatica v referred leg pain

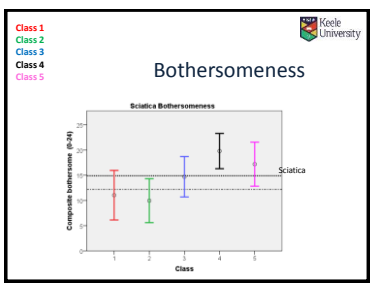
back pain intensity
disability
proportion of smokers
time off work
sleep disturbance
quality of life
Depression
Future risk of poor outcome

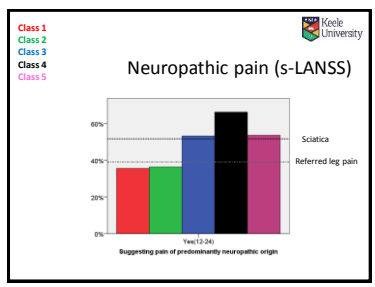
NO DIFFERENCE!

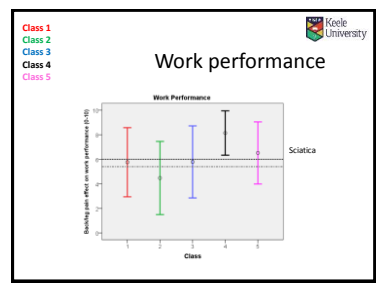
Konstantinou et al 2015

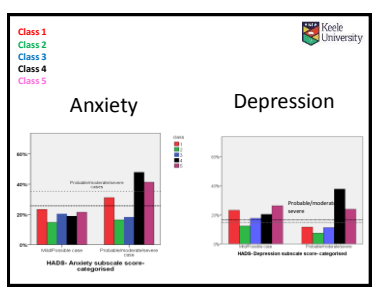


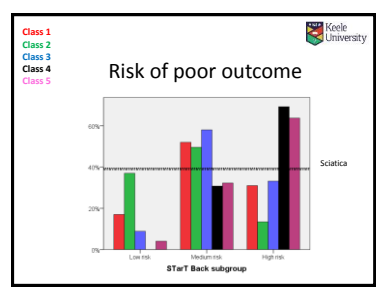








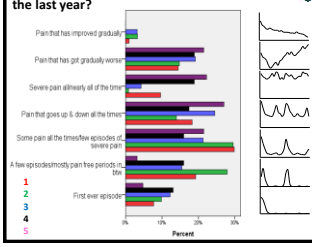




Appendix N

Workshop Presentation slides June 2016

How has the back/leg pain been over the last year?



Summary of classes

- Class 4: clinically easy to recognize
  Classic sciatica signs; High pain and disability; anxious, depressed, smokers; routine workers; high risk poor outcome; worsening, severe pain
- Class 1: clinically easy to recognize
  Back > leg pain, no neurology; anxiety; similar profile to moderate sciatica
- Class 3: typically reflect overall sciatica group, less severe profile
- Class 2: Below knee pain and SLR; least severe profile
- Class 5: more difficult to diagnose, severe profile like severe sciatica

Strengths

- LBLP primary care cohort
- No reference standard for diagnosis
- Highlighted spectrum of sciatica presentation

Limitations

- Judgement for interpretation of classes
- Clinical relevance?

Do you think these subgroups could add useful information to your clinical practice?

Questions

- Name of classes?
- Do you recognise these groups?
- Clinical relevance of these groups?
- Do you perceive that these groups could add useful information to your clinical practice
- Think about targeted treatment

Proposed names

1= REFERRED PAIN
2= MILD SCIATICA
3= MODERATE SCIATICA
4= SEVERE SCIATICA
5= ATYPICAL SCIATICA

Appendix N