# The Effects of Axial Loading on the Disc and Motion Segment Relative to Disc Degeneration and Pain Using Novel MRI Biomarkers

by

Vahid Abdollah

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

In

Rehabilitation Science

Faculty of Rehabilitation Medicine

University of Alberta

## ABSTRACT

The underlying pathology of low back pain (LBP) is poorly understood. Although MRI is often used as a sensitive modality for depicting pathoanatomical abnormalities, it is often clinically inconclusive due to a lack of specificity. While most patients experience symptoms when the spine is loaded, MR imaging is usually conducted in a relaxed supine position. Loading may induce morphological changes in the spine. Therefore, to improve the specificity of MRI, one solution could be to apply loading using axial compression devices.

This PhD investigated the response of the lumbar spine to compression and traction in participants with and without chronic LBP using MRI $T_2$-mapping to identify imaging biomarkers holding promise for further investigation towards meaningful subgrouping of LBP or degeneration. Four studies informed the planning of the loading studies: 1) a systematic review of acute loading responses on imaging measurements of the disc and vertebrae to identify potential biomarkers, 2&3) two reliability studies of the novel candidate biomarkers measured using a novel semi-automated segmentation algorithm, and 4) a comparison of disc fluid content biomarkers' responses to extension exercise. Finally, two pilot experimental studies were conducted to determine: 5) the relation between disc degeneration severity and effects of compression and traction on lumbar discs and vertebrae and 6) the effects of compression and traction on lumbar MRI findings in relation to chronic LBP.

The systematic review illustrated a lack of comparisons of the response to loading between participants with and without pain and no studies on traction. The review identified 14 biomarkers; most did not detect effects of loading. However, limited evidences of forward shift in the nucleus, increased disc diameter, and decreased disc height were observed under compression. Since 50% body weight was commonly used in traditional MRI, it was adopted in this thesis. Twelve

candidate biomarkers were chosen based on having detected changes in the review or representing novel measurements of fluid distribution postulated to be sensitive to loading: the location of the $T_2$-weighted centroid ($T_2$WC: mean position of the points in an ROI, weighted by their $T_2$), and geometric weighted centroid (GWC: mean position of the points in an ROI, weighted based on their location), disc height, disc and nucleus mean $T_2$ time and diameter, and motion segment angle.

Reliability studies indicated that $T_2$WC, GWC, and disc height using different disc width, presented excellent reliability. The novel measurements of fluid distribution showed a better ability to detect changes in response to extension exercises compared to mean signal intensity. Only the horizontal location of the disc $T_2$WC detected significant differences immediately after exercise.

The study of the relation between disc degeneration and response to loading indicated that smaller responses to loading correlated with more severe disc degeneration. Traction compared to compression was most sensitive to capture the load-induced changes in the lumbar spine. Disc height and diameter, disc and nucleus mean $T_2$ time and $T_2$WC, as well as, the nucleus GWC hold promise as sensitive biomarkers to capture loading responses.

The study comparing loading responses between participants with and without LBP indicated that different responses to loading between groups were most often observed at L5-S1. The most sensitive biomarkers for pain were the horizontal-coordinate of the disc and nucleus $T_2$WC (large effect sizes), the horizontal- and vertical-coordinates of the nucleus GWC, and the vertical-coordinates of the discs (small effect) and nucleus $T_2$WC (moderate effect). These biomarkers hold promise for future investigation. In contrast, the nucleus' width, the disc and nucleus mean $T_2$ times, and disc width were least sensitive to capture pain-related differences.

Four biomarkers showed the potential to detect pain-related differences by comparing a single supine unloaded scan. In the pain group, the disc $T_2$WC and nucleus GWC between L3-4 and L5-S1 and nucleus $T_2$WC at L4-5 were moderately to largely more anterior. Likewise, the disc height was taller at L3-4 and L4-5 by a large effect size.

Because the response to loading of our proposed parameters was correlated to with degeneration and some loading responses of parameters detected differences related to pain grouping, our results justify further investigation of the clinical value of the proposed biomarkers. These biomarkers show promise to improve the specificity of MRI for LBP and the ability to detect the response to loading.

## PREFACE

This thesis is an original work by Vahid Abdollah. The research project, of which this thesis is a part, received research ethics approvals from the University of Alberta Research Ethics Board, for the following projects:

1. Project Name "RELIABILITY OF QUANTITATIVE MRI MEASUREMENTS", No. Pro00050135, 8/6/2014 to 8/15/2016.
2. Project Name "EVALUATION OF THE EFFECTS OF COMPRESSION AND TRACTION ON LUMBAR INTERVERTEBRAL DISC FLUID CONTENT AND DISC MORPHOLOGY IN PATIENTS WITH LOW BACK PAIN", No. Pro00052494, 7/21/2015.

The image segmentation program referred to in chapters 4, 5, 6, 7 and 8 was developed by myself using a research grant from Alberta Spine Foundation. All members of my supervisory committee provided invaluable inputs during the development of the software. $T_2$-mapping software implemented in image segmentation program was developed by Dr. Keith Wachowicz.

The MRI compatible loading apparatus referred to in chapters 7 and 8 was designed by myself and manufactured by Professor Robert Lederer Department of Art and Design. The data used in chapter 4, 5, and 6 was collected by Dr. Eric C Parent.

The literature review in chapter 3 and the data analysis in chapters 4, 5, 6, 7 and 8 are my original work.

For chapter 3 of this thesis "The Immediate Effects of Loading on Lumbar Imaging Findings: A Systematic Review", I was responsible for concept formation, designing quality appraisal from, running the search, data collection and analysis as well as the manuscript composition. E.C. Parent assisted for designing quality appraisal from, running the search, data collection and analysis. E.C. Parent was the supervisory author and was involved in concept formation and manuscript composition. S. Adria assisted for data collection. M.C. Battié assisted for abstract screening. M.C. Battié was the supervisory author and was involved in concept formation and manuscript composition.

For chapter 4 of this thesis "Is the Location of the Signal Intensity Weighted Centroid a Reliable Measurement of Fluid Displacement within the Disc?" I was responsible for concept formation, development of the image segmentation software, image and data analysis as well as

the manuscript composition. E.C. Parent assisted for data collection, image segmentation and data analysis. E.C. Parent and M.C. Battié were the supervisory authors and were involved with concept formation and manuscript composition.

For chapter 5 of this thesis "Reliability and Validity of Lumbar Disc Height Quantification Methods on Magnetic Resonance Images", I was responsible for concept formation, development of the image segmentation software, image and data analysis as well as the manuscript composition. E.C. Parent acquired images in a previous project, assisted for data collection, image segmentation and data analysis. E.C. Parent and M.C. Battié were the supervisory authors and were involved with concept formation and manuscript composition.

For chapter 6 of this thesis "Evaluation of the Effects of Extension Exercises on Disc Fluid Using MR Images", I was responsible for concept formation, development of the image segmentation software, image and data analysis as well as the manuscript composition. E.C. Parent acquired images in a previous project, assisted for data collection, image segmentation and data analysis. E.C. Parent and M.C. Battié were the supervisory authors and were involved with concept formation and manuscript composition.

For chapter 7 of this thesis "Evaluation of the Effects of Loading and Disc Degeneration on Lumbar Intervertebral Discs and Motion Segments", I was responsible for concept formation, development of the image segmentation software and loading table, data collection and analysis as well as the manuscript composition. E.C. Parent assisted for data analysis. A. Su assisted for data collection, image analysis and manuscript composition. K. Wachowicz assisted for data collection. E.C. Parent and M.C. Battié were the supervisory authors and were involved with concept formation and manuscript composition.

For chapter 8 of this thesis "The Effects of Compression and Traction on Lumbar MRI Findings in Relation to Chronic Low Back Pain", I was responsible for concept formation, development of the image segmentation software and loading table, data collection and analysis as well as the manuscript composition. E.C. Parent assisted for data analysis. A. Su assisted for data collection, image analysis and manuscript composition. K. Wachowicz assisted for data collection. E.C. Parent and M.C. Battié were the supervisory authors and were involved with concept formation and manuscript composition.

*For Shima, Joubin and my parents;*

*For Ms. Rjabi, my grade I teacher;*

*For Saied my heavenly brother who chose a big school in the Skye to let other children enjoy school;*

*...with all my love*

## ACKNOWLEDGMENTS

I would like to express my most sincere thanks to:

***Shima***, without your support, patience and motivating power I would not have been able to begin and finish this PhD program.

My supervisors **Doctor Eric Parent** and **Professor Michele C Battié**, who provided extensive feedback and support during this PhD project. Thank you, **Eric**, for all our research related and personal discussions, providing constructive critiques, helping me know more people to expand my academic network and your financial support. Thank you, **Michele**, for your constructive critiques and helping me expand my academic network.

I also want to express my gratitude to **Doctor Keith Wachowicz** from my supervisory committee, who provided invaluable extensive feedback and support during the experimental part of this PhD program and facilitated access to MRI services.

I would like to thank **Doctor Alireza Simichi** from the Faculty of Science, who provided statistical consultation for chapters 7 and 8. I also want to express my gratitude to **Robert Lederer** Associate Professor, Faculty of Art and Design, who helped me develop and build the loading apparatus.

I would like to thank **Doctor Jill Urban** (Oxford University) who opened my eyes to the wonderful world of the disc nutrition and biochemistry. I would also like to thank **Doctor Susan Armijo Olivo** who taught me how to write a good paper.

To all the students and volunteers who helped me for this PhD project, **Alex Su**, **Dave Boryski**, **Emily Mostofi**, **John McFarlane**, **Lucy Lou**, **Shaja Chaudhry**, **Jack Underschultz**, **Claire Kee**, **Jia Hueih Ting**, **Sanja Schreiber**, **Alan Richter**, **Warren Matthews, Kathleen Shearer, Linda Slater** and **Gurpreet Chaggar**, without your help I would not have completed this research work. Thank you!

I am also grateful to the following agencies for the excellent financial support I received during my PhD: Fellowship and Grants from the Department of Physical Therapy, Alberta Spine Foundation, Canadian Physiotherapy Association, University of Alberta fellowship and travel grants

**CONTENTS**

## LIST OF TABLES

## LIST OF FIGURES

## LIST OF EQUATIONS

## LIST OF SYMBOLS

CAM: Computer-aided measurement

CT: Computed tomography

DWI: Diffusion weighted imaging

ECM: Extracellular matrix

ES: Effect size

FoV: Field of view

GAGs: Glycosaminoglycans

GRE: Gradient echo

GWC: Geometric weighted centroid

ICC: Intra-rater reliability

IDD: Intervertebral disc degeneration

IL: Interleukin

IVD: Intervertebral disc

LBP: Low back pain

L: Lumbar

µg: Microgram

MHz: Mega Hertz

mm: Millimetre

MPa: Mega Pascal

MRI: Magnetic resonance imaging

MSI: Mean signal intensity

N: Newton

ng: nanogram

ODI: Oswestry Disability Index

PDWI: Proton-density-weighted image

PG: Proteoglycans

RF: Radio frequency

ROI: Region of interest

SD: Standard deviation

SE: Spin echo

SEM: Standard error of measurement

SNR: Signal to noise ratio

T: Tesla

TE: Echo time

TM: Transverse Magnetisation

TNF-α: Tumour necrosis factor alpha

TR: Repetition time

$T_1$RT: $T_1$ relaxation time

$T_2$RT: $T_2$ relaxation time

$T_2$WC: $T_2$-weighted centroid

# CHAPTER 1 OVERVIEW OF THE THESIS

The core interest of this doctoral research was to explore novel MR imaging biomarkers relevant to disc degeneration and common low back pain (LBP) by observing the response of the lumbar spine, and the disc, in particular, using MRI under various loading conditions. LBP is a major cause of discomfort and disability among adults globally [1]. While any spinal structure that is innervated, able to trigger pain, or is susceptible to painful disease or injury could be a possible source for LBP [2, 3], little is known about the pathology underlying LBP. In over 85% of cases, despite using advanced imaging techniques to investigate pathoanatomy, the cause of the pain remains unclear (e.g. non-specific low back pain) [4].

Intervertebral disc degeneration (IDD) has been implicated in the development of LBP, either directly through the infiltration of blood vessels and nerves or indirectly through effects on spinal biomechanics and other structures [5–7]. Disc degeneration has been defined as the abnormal response of the cells of the intervertebral disc (IVD) to progressive structural failure [5]. Vertebral bone, which is densely and richly innervated by both motor and sympathetic nerves, also can be a possible source of spinal pain [8] and a structure affected by degeneration.

Diagnostic imaging modalities, including plain X-ray, computed tomography and magnetic resonance imaging (MRI), are used for diagnosing the underlying pathology of LBP [9]. MRI is preferred due to concerns about radiation exposure and better contrast and visualisation of soft tissue [10]. Conventional MRI is a sensitive modality for depicting structural abnormalities of the lumbar spine, including degenerative changes, but MRI is often clinically inconclusive due to a lack of specificity [11]. Asymptomatic individuals usually demonstrate degenerative features including, disc desiccation, fissures, high-intensity zones, vacuum phenomena, calcification, herniation [9], Schmorl's nodes, erosion, Modic changes, and calcification lesions of the endplates [12–14]. The lack of MRI specificity and the inability to identify conclusively a pathoanatomical cause of LBP may be due to numerous issues with current MRI assessments.

Employing a small number of discrete ordinal scores limits the ability of conventional MR imaging to detect subtle variations in disc degeneration. Further, the subjective nature of assessors' ratings limits the intra- and inter-observer reliability of the scores, with ICCs ranging from 0.33[15] to 0.9 [16]. Reliability is particularly low for higher grades of degeneration, which are typically of the greatest clinical interest [15]. Early disc degenerative changes can be observed as a subtle decrease in the signal intensity on $T_2$-weighted images [17] which reflect the water distribution within the disc. Further, in later stages degeneration is associated with adjacent Modic

type 2 changes (increased signal intensity on $T_1$-weighted images and an iso- or slightly hyperintense signal on $T_2$-weighted images) and high fat content of the paraspinal muscles [18]. Thus, quantifying MR findings such as MR signal intensity may provide a continuous measurement of disc ageing or degeneration, which may allow earlier detection than is possible with ordinal ratings of macroscopic findings. Such continuous measurements may be particularly useful for research involving the intervertebral disc, as they may be more sensitive to detecting variations and changes over time, including small diurnal variations in fluid content and responses of the disc to provocative loading. Quantitative measurements may also assist in tracking response to treatments targeting the degeneration or for defining clear phenotypes in genetic research.

The changes in the disc in response to degeneration [19, 20] and loading [21–27] have been investigated in vivo quantitatively using MR imaging using $T_2$-weighted MR images. $T_2$ relaxation time ($T_2$RT) is the rate at which transverse magnetisation is decreased and disappears [28–31]. Signal intensity on $T_2$-weighted images varies among different MR scanners due to field characteristics, sequence parameters, and patient positioning relative to receiving coils. In contrast, $T_2$ times from $T_2$ maps are independent of these methodological differences and represent a unique tool for comparing repeated scans under different loading conditions and positions or for inter-scanner comparisons. Quantitative $T_2$ measurements are likely helpful to detect subtle fluid and biochemical changes within discs that may not be apparent with qualitative or semi-quantitative measures on other sequences [32]. $T_2$-mapping can quantitatively evaluate the fluid content of the IVDs [27, 33–40].

MR imaging is usually carried out while the subject is lying in a relaxed supine position. The loading of the lumbar spine is minimal in this position as compared to standing or sitting [41, 42], positions in which pain and discomfort are usually worse [43, 44]. Furthermore, changes in the MR signal intensity in response to loading, such as in extension positions, have shown an association with changes in pain symptoms [45]. More specifically, subjects who reported an immediate reduction in pain intensity after a treatment with posterior-to-anterior–directed manual pressures, followed by prone press-up exercises into extension, showed an increase in diffusion of water in the nuclear region of the L5-S1 disc, while those who did not report a pain reduction did not have a change in diffusion [45].

We hypothesised that variations observed in the magnitude and distribution pattern of fluid within the lumbar tissues, as well as laxity and changes in segmental alignment of the lumbar spine

during imaging under various loading conditions, would be associated with degeneration severity and might provide further insight into degenerative conditions and into the pathology underlying persistent back pain. Specifically, refined quantification of vertebra and disc area, height, angulation, alignment, and translational and rotational displacements, as well as MR $T_2$ variations under different loading conditions may lead to the identification of new imaging biomarkers to assist with classification, diagnosis and prognosis of low back pain, and intervention planning or evaluation. The main purpose of this PhD study was to investigate the acute response of the lumbar discs and motion segments to compression and traction loading using continuous, quantitative measurements obtained from MRI, including $T_2$ mapping.

This PhD dissertation consisted of a core clinical study to identify biomarkers for evaluating effects of loading on a set of measures of interest and four preliminary studies. The main clinical study investigated the response of the disc and vertebra to standardised compression and traction loading using $T_2$-weighted and $T_2$ maps magnetic resonance images, with a particular focus on the disc fluid content in response to loading. Thirty-five participants with and without chronic back pain participated in this study. The spine was loaded using a custom-designed lumbar spine MR compatible loading apparatus developed by myself during this PhD and spine images were analysed using a computer aided measurement system also developed by myself during this PhD. The preliminary studies informing the planning of the clinical study were:

- An extensive systematic review of the immediate effects of various loading conditions on the lumbar spine and intervertebral disc, in particular, to guide the identification of potential biomarkers to investigate in the clinical study and assist in interpreting subsequent imaging findings (Chapter 3).
- A reliability study of the location of the signal-intensity-weighted-centroid (SIWC), and the mean signal intensity and area measurements of the whole disc, nucleus and anterior and posterior annulus (Chapter 4).
- A reliability and construct validity study of eight different disc height quantification methods on $T_2$-weighted MR images. (Chapter 5)
- A study to investigate the ability of candidate quantitative measurements to detect the immediate effects of prone press-up extension exercises on the mean signal intensity and location of the SIWC of lumbar discs using $T_2$-weighted MR images (Chapter 6).

- The clinical study is then presented in chapters 7 and 8:
    - An evaluation of the relation between disc degeneration and the effects of loading on lumbar intervertebral discs and vertebrae (Chapter 7).
    - An evaluation of the effects of compression and traction on lumbar MRI findings in relation to chronic low back pain (Chapter 8).
- Chapter 9 is the general discussion and conclusion.

# CHAPTER 2 OVERVIEW OF RELEVANT STRUCTURES AND IMAGING

## Introduction

Low back pain (LBP) is a major cause of discomfort and disability among adults globally [1]. While any spinal structure, which is innervated and susceptible to painful disease or injury could be a possible source of LBP [3], little is known about the specific pathology underlying LBP. In over 85% of LBP cases, despite using advanced imaging techniques to investigate pathoanatomy, the cause of the pain remains unclear (e.g. non-specific LBP) [4].

LBP has been defined as any pain perceived between the ribs and the top of the legs, from any cause [46]. LBP has also been categorised based on the duration of the pain. Acute LBP is used when pain is present for less than four weeks. Sub-acute LBP is used when the pain is present between 4 and 12 weeks. The chronic LBP term is used when the pain is present for more than 12 weeks [47].

## Causes and Epidemiologic Patterns

Non-specific LBP is the most common form of LBP [48]. Experimental studies suggest that pain may originate from any innervated structure, including ligaments, facet joints, the vertebral periosteum, the paravertebral musculature and fascia, blood vessels, the annulus, and spinal nerve roots [3]. LBP is more common in female than male and in those aged 40–69 years than in other age groups [1]. Prevalence is greater in high-income countries than middle- or low-income countries, but there is no difference in prevalence between rural and urban areas [1]. Hoy et al. study reported a positive correlation between a country's human development index and overall mean prevalence of LBP ($r=0{\cdot}088$; $p<0{\cdot}001$) [1].

## Chronic Low Back Pain

The National Institute of Health has defined chronic LBP as a back pain persisting at least three months and resulting in pain on at least half days in the past six months [49]. Intervertebral disc degeneration (IDD) has been implicated as a major factor in the development of chronic LBP [5–7].

***Intervertebral Discs***. IVDs are pads of fibrocartilage lying between vertebral bodies. IVDs comprise a central mass of gelatinous tissue, the nucleus pulposus, a peripheral rim of tough fibrous tissue, the annulus fibrosus, and two cartilaginous endplates at the extremities. Three main functions are considered for IVDs: load transmission, shock absorption and motion control [2].

The average anterior-posterior diameter and cross-sectional area of the lumbar IVDs are 40 mm and 1800 $mm^2$, respectively [50, 51].

Lumbar IVDs have a wide dispersion of cells in a large volume of an extracellular matrix (ECM). The ECM mainly contains water and proteins and organised into two components: inter-fibrillar and fibrillar [52]. The inter-fibrillar component mainly contains water, glycoproteins and proteoglycans (PGs) and fibrillar component mainly contains collagen and elastin [52]. The ECM is continuously degraded and replaced by newly synthesised molecules [53].

Cells make up approximately 1% of the total dry weight of the disc [54]. There are approximately 9000 and 4000 cells per $mm^3$ of the annulus and nucleus, respectively [55]. Cellular activities depend on the preservation of a proper level of nutrients and metabolites inside the disc [53]. Any decline in the glucose, oxygen and pH level can trigger a marked reduction in the cellular activity [53, 56].

In a healthy disc, water makes between 60% and 80% of the disc total weight [2]. Water molecules either bind to the macromolecules to form a hydration layer or move freely [57]. The water content of the disc is variable and is a function of an equilibrium between swelling and mechanical pressure. Swelling pressure is created by the negative charge of the PGs, which tends to absorb $H^+$ of the water molecules. Mechanical pressure, which is due to the weight of the upper structures, the loads induced by physiological motions, muscle and ligament tension, tends to dehydrate the disc [58]. It is estimated that in a healthy disc, almost 25% of the disc fluid content is expressed out when a load larger than the osmotic pressure is placed on the disc and is re-imbibed when the load is reduced or removed [34].

Each collagen is composed numerous tropocollagen molecules, arrayed side-by-side and end-by-end through covalent bonds. Collagens have a long life and their half-life in the IVDs is about 95 years [59]. The annulus and nucleus contain mainly type I and II collagen, respectively. PGs are made up of a protein core attached to at least one GAG [60]. The GAGs are bonded covalently to a central protein core, except hyaluronate, which connects to the protein core using a link protein. The presence of cations, such as carboxyl ($COO^-$) and sulphate ($SO_3^-$) in the GAGs, imparts a net fixed negative charge to the ECM, and a negative osmotic swelling pressure as an outcome [61]. The osmotic pressure helps discs bear spinal compressive loads. The half-life of the PGs in the IVDs is about 5 years [62]. Elastin makes up approximately 10 % of the ECM of mature

human discs [63] and is thought to play an important supporting role for collagen fibres to recover after deformation.

*Blood supply and innervation of the intervertebral discs.* The vertebral arteries terminate in loops at the interface between the bony and cartilaginous endplate making the lumbar IVDs the largest avascular tissue of the human body [53], with cells at the centre of the disc lying almost six to eight mm from the nearest blood supply [61]. The cells of the outer third of the annulus acquire nutrients directly from the local vasculature [64], and the rest of the cells through either passive diffusion and/or water imbibition [65–68]. The rate of solute transport depends on the loading history and shape of the disc [53, 58, 69, 70] whereas the degree of solute penetration depends on both the size and charge of the solute [65]. Passive diffusion is a homeostatic mechanism, which creates a concentration gradient within the disc [67]. This gradient is a function of three variables: the rate of nutrient supply, the rate of nutrient transport and the rate of cellular metabolism [67]. The fluid imbibition mechanism is based on the premise that the disc loses its fluid when it is loaded and retrieves it when the load is reduced or removed [2]. This cycle brings nutrients and waste materials in and out, respectively [2]. Prolonged or abnormal loading is therefore thought to be able to negatively affect the fluid content of the disc and normal imbibition process [53].

The lumbar IVDs is the largest aneural tissue of the human body [53]. The outer annulus is innervated by two nerve plexuses that accompany the anterior and posterior longitudinal ligaments [71]. The anterior plexus, which supplies the anterior and lateral annulus, is derived from the sympathetic trunk and gray rami communicans [71]. The posterior plexus, which supplies the posterior and lateral annulus, is derived from sinuvertebral nerves [71]. The posterior plexus contains both somatic and autonomic roots [71]. Both mechanical and neurogenic factors have been proposed for the neo-innervation in the degenerated IVDs. Intradiscal cracks have a lower pressure and PG content [71]. This creates an ideal environment for both angiogenesis and neo-innervation [71]. Secretion of pro-inflammatory and neurogenic factors such as interleukin (IL)-1, IL-6, IL-8; tumour necrosis factor (TNF)-α, prostaglandin E2, monocyte chemotactic protein-1, and nerve growth factor by disc tissue can trigger an autoimmune response. These chemicals can both sensitise existing nerves and promote new nerve growth in both the disc and the adjacent vertebra. The new nerves are positive for substance-P and have nerve-ending morphologies consistent with nociception [7].

**Intervertebral Disc Degeneration**

IDD has been defined as an abnormal cell-mediated response to the progressive structural failure, associated with morphological, chemical and metabolic changes [5]. A degenerated disc is a desiccated, discoloured and fragmented structure with a lower height and signal intensity on $T_2$-weighted MR images [72]. Degeneration also affects both the magnitude and distribution of fluid in the disc detectable with imaging [5]. The IDD is thus associated with morphological, chemical, metabolic, and biomechanical changes in the disc and affects disc function and biomechanics [5].

The pathophysiology of IDD is not well understood. Various factors, including genetic and mechanical, are thought to contribute to the development of the degeneration. The influence of genetics in the development of IDD is well established. It is estimated that up to one-quarter of the genetic influences on the development of LBP are shared with those affecting disc height narrowing, a degenerative finding associated with LBP symptoms [73]. However, a substantial portion of the genetic influence on pain has been left unexplained [73]. This could be due to other changes affecting pain processing or other underlying pathological conditions [73].

Common minor trauma could trigger minor failures in the endplates. This could trigger a rapid structural failure of the disc [74]. Endplate damage reduces pressure on the adjacent nucleus by 25%-27% [5]. This decreases the role of the nucleus in weight bearing and increases compressive stress in the annulus, particularly in the posterior annulus [5]. The subsequent changes in matrix compressive stress can endanger disc cell metabolism and could lead to the progressive deterioration of the ECM [75]. The results of Carragee et al. study, however, indicated that the development of LBP was not associated with the development of any new structural change in the motion segments [76].

Preserving a proper supply of both nutrients and metabolites is essential for keeping disc's cells viable and active [77]. A larger discal cross-section has been considered the most important morphological factor associated with the development of IDD based on the premise that larger cross-sectional area endangers normal disc nutrition [78]. The shape and size of the endplate are also thought to be associated with the development of IDD. Bigger central region [79] and flat and

irregular shaped endplates[I] are more associated with severe disc degeneration than concave ones [80]. Aortic atherosclerosis[II] is also suggested to be associated with the development IDD [81]. The results of a post-mortem study indicated a positive correlation between aorta atherosclerosis and IDD [81]. A total of 88% of the subjects with positive back pain history had one or more missing arteries, 80% of them had narrow arteries and 72% had developed collateral arteries [81]. The fourth and fifth lumbar arteries are the most common sites to observe stenosis [81, 82].

***Pain generation mechanism.*** The source of the pain in LBP has not been well understood yet. Disc degeneration could elicit pain either directly or indirectly.

***Direct pain generation mechanisms.*** Secretion of proinflammatory cytokines and sensitisation of silent nociceptors are two major direct pain generation hypotheses.

1. Nuclear cells can produce proinflammatory cytokines such as IL, including IL-1, IL-4, IL-6, IL-8, IL-12 and IL-17, interferon-γ and TNF-α [71]. The cells of the inner annulus and to some extent the cells of the outer annulus are also able to produce IL-1 and TNFα [71]. The cytokines, which can reach the outer annulus or endplates through diffusion in a diurnal cycle can trigger nociceptors or simulate noxious effects [71]. Secretion of inflammatory cytokines, in particular, IL-1, may stimulate the expression of angiogenic and neurotrophic factors. The latter can trigger nociceptive nerve ingrowth into the IVD and endplates [71].
2. Trauma and inflammation may trigger the synthesis of factors such as bradykinin and prostanoids, thought to be able to sensitise silent nociceptors, which are usually unresponsive to even maximal mechanical stimulation [6]. The sensitised nociceptors could then respond to changes in intradiscal pressure during movement and trigger pain [6].

***Indirect pain generation mechanism.*** Pain has also been suggested to relate to the changes induced in the spine due to compression loading, including increased intra-discal pressure [41, 42], decreased disc height [83, 84], lumbar instability [85–87], and possibly subsequent stenosis of the intervertebral foramen affecting the neurovascular structures [88], excessive disc or vertebral deformation [89], increased lumbar lordosis [83, 88–91], and subsequent stenosis of the vertebral

---

[I] It is not clear whether irregularity is the cause of the degeneration or degeneration is the cause of the irregularity.
[II] Deposition of plaques of fatty material on the inner walls of the blood vessels.

canal [92]. Disc prolapse has been also suggested to be responsible for the generation of the radicular pain due to mechanical or chemical irritation of a spinal nerve or its root [71]. Significant height loss and listhesis of the vertebra in advanced grades of disc degeneration could cause foraminal or central stenosis, which can also cause mechanical irritation or inflammation of the nerves [71]. Degeneration of the zygapophysial joints either due to overloaded facet joints or inflammatory process of the joint capsules could be another possible source of pain generation [71].

**Imaging of the Lumbar Spine**

Diagnostic imaging modalities, including plain X-ray, computed tomography (CT) and magnetic resonance imaging (MRI) have been widely used in attempts for diagnosing the underlying pathology of pain [9]. The most commonly reported manifestations of the lumbar spine degeneration are disc space narrowing, loss of $T_2$ signal intensity, fissures, high-intensity zones, vacuum changes, calcification, ligamentous changes, vertebra marrow changes, herniation, osteophyte formation, malalignment, and stenosis [2].

***Plain radiology.*** Low cost and availability have made plain radiography the most common spinal imaging method. The anterior-posterior and lateral views can help visualise the lumbar spine alignment, disc space, vertebral body height, and provide a gross assessment of bone density and architecture. Soft tissue structures, however, cannot be evaluated extensively. The oblique view can visualise the pars interarticularis and is useful for diagnosing spondylolysis when clinical evidence exists. Other special views, including flexion-extension and angled views, are usually employed to assess lumbar instability and ankylosing spondylitis of the sacroiliac joint, respectively.

Provocation discography (PD) is done by injecting a contrast agent into the discs of a lightly sedated patient while monitoring the injected volume, pressure, contrast distribution pattern, and patient's response to pain [93]. Results of a systematic review with meta-analysis indicated that PD results can be quite conclusive (specificity of 0.94 and a false-positive rate of 6%) if performed using low-pressure technique [93]. However, this technique is invasive [94, 95].

***Computed tomography.*** Computed tomography (CT) can provide an excellent visualisation of bony structures, particularly in the axial view. CT is, therefore, the best imaging

modality for evaluating the facet joints [96]. However, facet joint arthritis is a relatively common finding in a large percentage of asymptomatic subjects, possibly due to the normal ageing process [96].

***MR imaging.*** MR imaging offers several advantages over conventional radiography and CT for imaging of the lumbar spine. It provides an excellent visualisation of the IVDs, neural elements, paraspinal muscles, facet joints, vertebral marrow and contents of the spinal canal [96]. It is also a highly sensitive and specific modality for diagnosing pathologies such as tumour and infection.

***Basic principles of MRI.*** When a tissue is exposed to a strong external magnetic field ($\vec{B}_0$), the spinning axes of its protons are aligned either parallel or antiparallel to the long axis (z-axis) of the external magnetic field and begin to precess around the long axis of the magnetic field [28]. This creates a net magnetisation vector ($\vec{M}$), which is the sum of all magnetic moments [28, 29]. If a radiofrequency (RF) pulse with the same precession frequency of the protons and strength of $\vec{B}_1$ is applied to the magnetic field, the net magnetisation vector ($\vec{M}$) will be flipped to an angle (α). This will also increase the population of the protons with higher energy state. The new magnetisation vector can be resolved into two components: longitudinal ($M_z$), which is static and cannot induce a current on the receiver coils, and transverse ($M_{xy}$), which is dynamic and can induce a current on the receiver coils [28, 29, 31]. As soon as the RF pulse is switched off the protons begin to fall out of phase and to return to the equilibrium state, that is, the net magnetisation vector thus realigns with the external magnetic field ($\vec{B}_0$) [28].

Sensors (coils) embedded in the MRI machine detect the energy released as the protons realign with the magnetic field. The time it takes for the protons to realign with the magnetic field, as well as the amount of energy released changes depending on the environment and the chemical nature of the molecules. Images are constructed based on these data.

*$T_1$ relaxation time.* $T_1$ relaxation or spin-lattice relaxation is the process whereby protons exchange energy with their surrounding lattice to return to their lower energy equilibrium state and thus restoration of longitudinal magnetisation [30]. $T_1$ relaxation time ($T_1$RT) therefore is the rate at which the longitudinal magnetisation is increased [28–31]. $T_1$RT is a function of the field strength [28–31].

*$T_2$ and $T_2^*$ relaxation time.* Transverse relaxation or spin-spin relaxation describes the process whereby protons fall out of phase in the X–Y plane and net transverse magnetisation decreases and disappears [30]. $T_2$ relaxation time ($T_2$RT) is, therefore, the rate at which TM is decreased and disappears [28–31]. The maximum magnitude of the TM is immediately after its formation when all of the protons are in phase [28–31]. It then begins to decay due to tissue inhomogeneity (small magnetic fields from neighbouring nuclei) and external magnetic field ($\vec{B}_0$) inhomogeneity [30]. $T_2$RT is always shorter than $T_1$RT for a given tissue. Water-based tissues tend to have a longer $T_2$ than fat-based tissues.

$T_2$ relaxation times correlate with hydration (water content) and to a lesser extent with PG content and (negatively) with the collagen content of the IVDs [37, 39, 96, 97]. Thus, variations of the matrix content, including water, PG and collagen or modification of the organisation of the collagen network could probably induce $T_2$ variations [98]. $T_1$ and $T_2$ values of different tissues differ in the human body allowing the production of images contrasting tissues.

*Echo time and repetition time.* The echo time (TE) refers to the time between the application of RF pulse to generate spatially encoded signal and the acquisition of signal [28, 29, 31]. Repetition time (TR) is the time interval between two consecutive RF pulses [28, 29, 31]. These times can be varied to generate images with different contrast between tissues.

***Proton-density-weighted images.*** The contrast of the proton-density-weighted images (PDWI) is mainly due to the density of the spins (Figure 1) [97]. Although PDWIs are theoretically appropriate for the depiction of both the anatomy and pathologic conditions, they have a lower contrast compared to $T_1$- or $T_2$-weighted MR images [29].

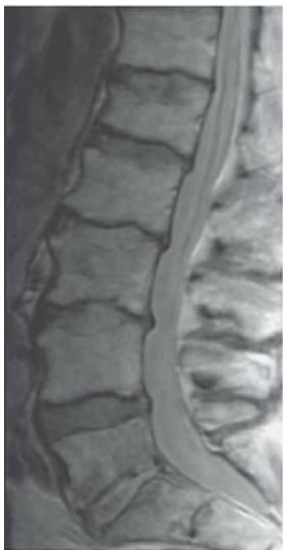

Figure 1 A proton density MR image of the lumbar spine [28]

***$T_1$-weighted images.*** The contrast of $T_1$-weighted MR images ($T_1$-WIs) is mainly due to the $T_1$ effect (Figure 2) [28–30]. Tissues with a short $T_1$RT produce a bright appearance on $T_1$-WIs [29]. $T_1$WIs are sometimes called 'anatomy' scans as tissues with high-fat content appear bright and compartments filled with water appears dark.

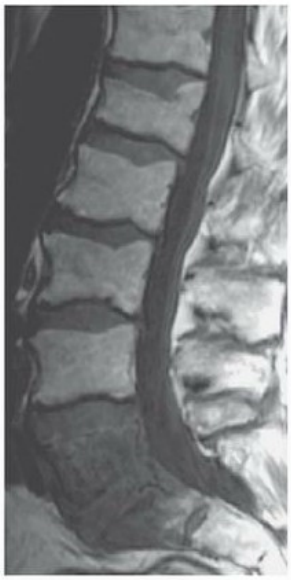

Figure 2 $T_1$-weighted MR image of the lumbar spine [28]

***$T_2$/$T_2$* weighted images.*** The contrast on $T_2/T_2^*$ weighted images ($T_2/T_2^*$-WIs) is mainly due to the $T_2/T_2^*$ effect on SE and GRE sequences, respectively (Figure 3) [28, 29]. Tissues with

a long $T_2/T_2^*$ RT give the highest signal intensities, producing a bright appearance on $T_2$-WIs [28, 29]. Tissues, which contain more water which has a long $T_2$ time, are thus brighter than those that contain more fat [28, 29]. $T_2$-WIs are sometimes called 'pathology' scans as collections of abnormal fluid are bright against the darker normal tissue [29]. $T_2$ based contrasts are particularly useful to study disc degeneration.

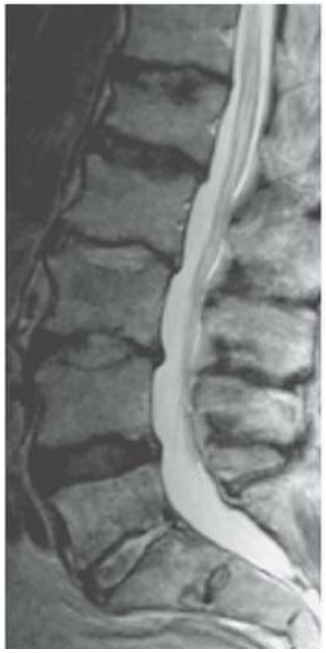

Figure 3 $T_2$-weighted MR image of the lumbar spine [28]

***Quantitative MRI.*** Quantitative MRI (qMRI) includes techniques used for determining the amounts or proportions of the components of a tissue. qMRI, therefore, allows statistical analysis of tissue by providing quantitative scale measurements. In the qMRI, the MRI scanner is a scientific measuring instrument while in contrast the traditional MRI considers scanner a camera for visualising the tissue. It can detect subtle changes that may not be apparent to radiologist' eye. The most common qMRI techniques are relaxometry, diffusion-weighting and MR spectroscopy (MRS).

***Magnetic resonance spectroscopy.*** Magnetic resonance spectroscopy (MRS) is a non-invasive in-vivo MRI-based analytic technique to determine the molecular composition of tissues. It is based on the premise that each element has its unique resonance frequency, making it distinguishable from other elements [28]. Hydrogen ($^1$H), carbon ($^{13}$C), fluorine ($^{19}$F), Sodium ($^{23}$Na) and phosphorus ($^{31}$P) nuclei are usually employed in MR spectrum acquisition [98].

However, MR spectroscopy has a limited application in the lumbar spine due to low signal-to-noise ratio, the presence of water in the disc and adjacent tissues, the presence of lipids in the adjacent bone marrow, and the broad line widths seen in vivo due to bone susceptibility-induced line broadening [99].

***Diffusion-weighted imaging.*** Diffusion-weighted imaging (DWI) is based on the premise that the motion or diffusion of water molecules results in a decrease in the MR signal intensity [100]. Unlike $T_1$- and $T_2$-RTs, diffusion is an intrinsic physical process, which is completely independent of either the magnetic resonance effect or magnetic field [101]. It also reflects the velocity of water molecules, while both $T_1$ and $T_2$ only reflect the concentration of the water molecules [102]. This technique has been used to quantify the effects of loading in the IVDs fluid content [102]. Unfortunately, this technique does not offer detailed anatomical contrast for the study of lumbar morphometry [103].

***MR Relaxometry.*** MR Relaxometry relies on physical characteristics of nuclei relaxation after being excited by an RF pulse. $T_{1\rho}$ and $T_2$ mapping are two commonly used MR relaxometry techniques for studying the lumbar spine.

The $T_{1\rho}$ phenomenon reflects the spin-lattice relaxation time in the rotating frame [104, 105]. Signals are generated when long-duration, low-power RF pulses are applied to lock net magnetisation vector into the transverse plane along one axis [104]. As the strength of the locker pulses ($\vec{B}_{LP}$) are much less than the strength of the local magnetic fields ($\vec{B}_0$), normal $T_1$ and $T_2$ phenomenon will not happen [104]. The spin-lock mechanism excites the spins to couple at a lower signal than Larmor's frequency. Thus the low-frequency regime, such as physicochemical interactions between the ECM molecules and water can be investigated [105–108]. Any ECM variations such as PGs loss can be evaluated using this technique [99, 104, 105, 109, 110]. While this is interesting in the context of studying disc degeneration, the long acquisition times (Sagittal $T_2$-weighted: 285 secs vs. Sagittal $T_{1\rho}$: 400 sec) required for quantitative imaging technique preclude its use in the study of the effect of loading or clinically [111].

While a $T_2$-weighted image shows qualitative intensity variation based on $T_2$RT, $T_2$-relaxometry attempts to quantify the actual $T_2$ relaxation rate at each pixel of the image. $T_2$-maps are constructed by obtaining the $T_2$ constant using a non-linear least squares curve fitting on a pixel-by-pixel basis. The $T_2$ map is then displayed on a pixel-by-pixel basis for the whole image.

At least two consecutive echoes are needed for generating a $T_2$ map. $T_2$ mapping reflects the interaction of water and ECM and is therefore particularly relevant in imaging disc degeneration [112]. This method was therefore adopted in the present study due to its providing detailed anatomical contrast, potential to inform about disc degeneration and to detect the effect of loading on fluid distribution in the lumbar spine.

***MRI-based disc findings.*** Various grading and morphological algorithms, including Pfirrmann method, have been proposed for the MRI-based classification of IVDs degeneration and reporting of degeneration findings [16, 113].

***Pfirrmann's method.*** Pfirrmann's method is a five-grade classification algorithm based on MR signal intensity, disc structure, the distinction between nucleus and annulus and disc height on the sagittal $T_2$-weight images (table 1) [16]. The method mostly relies on the first three characteristics of discs. Disc height plays a major role in making out the difference between grades four and five, but it is not a discriminative feature for grades three and four [16].

**Table 1 Pfirrmann grading method for lumbar intervertebral disc degeneration** [16]

| Grade | Structure | Distinction of Nucleus and Annulus | Signal Intensity | Disc Height |
|---|---|---|---|---|
| I | Homogeneous, Bright White | Clear | Hyperintense, Isointense to Cerebrospinal Fluid | Normal |
| II | Heterogeneous With or Without Horizontal Bands | Clear | Hyperintense, Isointense to Cerebrospinal Fluid | Normal |
| III | Heterogeneous, Gray | Unclear | Intermediate | Normal to Slightly Decreased |
| IV | Heterogeneous, Gray to Black | Lost | Intermediate to Hypointense | Normal to Moderately Decreased |
| V | Heterogeneous, Black | Lost | Hypointense | Collapsed Disc Space |

***Sensitivity and specificity of MR imaging in low back pain.*** MRI is an excellent modality for depicting internal disc morphology as it can detect changes in the fluid content of the lumbar

tissues, and produce highly detailed images. However, MRI is often clinically inconclusive. Abnormalities seen on the lumbar MRI of patients with chronic LBP are often seen in asymptomatic individuals [114, 115]. The results of a recent systematic review of 33 studies indicated that many spine imaging findings have a high prevalence among asymptomatic people [115]. Disc degeneration grading, disc signal loss, disc height loss, disc protrusion, and facet arthropathy are generally part of the normal ageing process rather than pathologic processes requiring intervention (table 2) [115]. The results of another systematic review of 12 studies also indicated no consistent association between LBP and MR imaging findings of Modic changes, disc degeneration and herniation [116].

**Table 2 Age-specific prevalence estimates of degenerative spine imaging findings in asymptomatic participants** [115]

| Imaging Findings | Age (years) | | | | | | |
|---|---|---|---|---|---|---|---|
| | 20 | 30 | 40 | 50 | 60 | 70 | 80 |
| Disc Degeneration (%) | 37 | 52 | 68 | 80 | 88 | 93 | 96 |
| Disc Signal Loss (%) | 17 | 33 | 54 | 73 | 86 | 94 | 97 |
| Disc Height Loss (%) | 24 | 34 | 45 | 56 | 67 | 76 | 84 |
| Disc Bulge (%) | 30 | 40 | 50 | 60 | 69 | 77 | 84 |
| Disc Protrusion (%) | 29 | 31 | 33 | 36 | 38 | 40 | 43 |
| Annular Fissure (%) | 19 | 20 | 22 | 23 | 25 | 27 | 29 |
| Facet Degeneration (%) | 4 | 9 | 18 | 32 | 50 | 69 | 83 |
| Spondylolisthesis (%) | 3 | 5 | 8 | 14 | 23 | 35 | 50 |

The results of a recent meta-analysis indicated that imaging findings with a higher prevalence in people with LBP compared with asymptomatic individuals 50 years of age or younger were disc bulge, spondylolysis, disc extrusion, Modic 1 changes, disc protrusion, and disc degeneration (table) [117]. While any Modic change, central canal stenosis, high-intensity zone, annular fissures, and spondylolisthesis were imaging findings that were not associated with LBP (table 3) [117].

The lack of specificity and inability to conclusively identify the pathoanatomical cause of LBP may be due to conducting MR imaging in a relaxed supine position. How imaging findings are quantified is another consideration possibly responsible for limiting the specificity of MR imaging. Further, frequently, MRI grading algorithms employ a small number of discrete ordinal scores to assess degeneration. Such rating algorithms suffer from limited reliability and ability to distinguish differences between individuals, as well as to detect changes within individuals over

time [118, 119]. Degeneration is a slow process and only a quarter of subjects show an increase in the degeneration grade assessed on MRI after five years of follow-up [120].

Quantitative MRI provides non-invasive measures of the degeneration at the earliest stages of degeneration, which may be essential for efforts towards prevention and early intervention [121]. Symptoms are usually worse in a sitting or standing posture compared to a supine relaxed position [43, 44] suggesting conducting imaging under loading may improve imaging specificity for relevant pathoanatomical findings.

**Table 3 The prevalence of MRI findings of disc degeneration in asymptomatic and symptomatic participants under 50 years** [117]

| Outcome | No. of Studies | OR (95% CI) | Prevalence Asymptomatic | Prevalence Symptomatic | $P_{Value}$ |
|---|---|---|---|---|---|
| **Annular Fissure** | 6 | 1.79 (0.97–3.31) | 11.3% (9.0%–14.2%) | 20.1% (17.7%–22.8%) | 0.06 |
| **High-Intensity Zone** | 4 | 2.10 (0.73–6.02) | 9.5% (6.7%–13.4%) | 10.4% (8.0%–13.4%) | 0.17 |
| **Central Spinal Canal Stenosis** | 2 | 20.58 (0.05–798.77) | 14.0% (10.4%–18.6%) | 59.5% (54.9%–63.9%) | 0.32 |
| **Disc Bulge** | 3 | 7.54 (1.28–44.56) | 5.9% (3.8%–8.9%) | 43.2% (38.2%–48.2%) | ***0.03*** |
| **Disc Degeneration** | 12 | 2.24 (1.21–4.15) | 34.4% (31.5%–37.5%) | 57.4% (54.8%–59.8%) | ***0.01*** |
| **Disc Extrusion** | 4 | 4.38 (1.98–9.68) | 1.8% (0.1%–3.7%) | 7.1% (5.4%–9.4%) | ***<0.01*** |
| **Disc Protrusion** | 9 | 2.65 (1.52–4.62) | 19.1% (16.5%–22.3%) | 42.2% (39.3%–45.1%) | ***0.00*** |
| **Modic Changes** | 5 | 1.62 (0.48–5.41) | 12.1% (9.6%–15.2%) | 23.2% (21.7%–27.3%) | 0.43 |
| **Modic 1 Changes** | 2 | 4.01 (1.10–14.55) | 3.2% (0.7%–9.4%) | 6.7% (4.2%–10.4%) | ***0.04*** |
| **Spondylolisthesis** | 4 | 1.59 (0.78–3.24) | 3.2% (1.8%–5.8%) | 6.2% (4.4%–8.7%) | 0.20 |
| **Spondylolysis** | 2 | 5.06 (1.65–15.53) | 1.8% (0.0%–5.3%) | 9.4% (6.6%–12.4%) | ***<0.01*** |

***Stress imaging.*** The vast majority of low back symptoms, including pain and numbness, are worsened in a loaded position [122]. This is possibly due to changes occurring in the disc space, vertebral foramen, alignment of the spine and segmental stability. The intradiscal pressure increases significantly in an upright standing or sitting posture compared to a relaxed supine position [42, 123]. Results of several studies have indicated that the cross-sectional area of the lumbar central canal and lateral recesses decreases significantly in an extended position, and increases in a flexed position compared to a neutral position [124–126]. Nilsson et al. indicated an increase of mean $T_2$ in the anterior disc with a concomitant decrease in posterior disc mean $T_2$ when the spine was loaded using an axial compression load equal to 50% body weight in the supine position. The lumbar spine was likely extended during loading possibly resulting in a higher compression load on posterior spinal structures [127].

The imaging of the spinal structures in a loaded position is therefore recommended for improving sensitivity and specificity of the imaging modalities. However, this is not always possible with current imaging modalities. Upright MRI is probably the modality of choice for physiologic imaging of the spine. Most current open magnet MR imagers allowing upright imaging are however using either low or moderate strength magnets with limited resolution and signal to noise ratio. This has led to the development of the devices, such as the DynaWell compression device (DynaMed AB, Stockholm, Sweden), to load the lumbar spine in a supine resting position to improve the ability to detect lesions, which were otherwise missed in supine resting position [128–130]. The results to date also appear heterogeneous but suggest there may be potential in improving diagnostic accuracy and the ability of MR imaging to detect clinically relevant imaging findings in patients with LBP. A review of the investigations of the effect of loading on the lumbar spine is needed.

# CHAPTER 3 THE IMMEDIATE EFFECTS OF LOADING ON LUMBAR IMAGING FINDINGS: A SYSTEMATIC REVIEW

Vahid Abdollah; Eric C Parent; Steffen Adria; Michele C Battié

**Abstract**

**Study Design:** Systematic literature review

**Objective:** This review aimed to determine the immediate response of the lumbar spine to different loadings using quantitative assessments from imaging.

**Summary of Background Data:** Imaging for lower back pain is often done while resting supine but lacks diagnostic specificity. However, symptoms are usually triggered in loaded positions. Imaging studies have used various loading protocols and quantitative measurements to better understand lower back pain.

**Methods:** Medline, EMBASE, Scopus, Web of Science, BIOSIS Previews and CINAHL were searched from inception until August 2014. Two of three reviewers independently screened titles with abstracts and then full-text articles focusing on quantifying effects of loading on the lumbar spine using imaging in at least 10 participants with or without back pain. Quality was appraised using a standard checklist. Quantitative outcomes of interest were responses of the disc, spinal canal, lateral foramen, vertebra and endplate to loading reflected on imaging. Levels of evidence summary statements were formulated.

**Results:** Fourteen measurements and 16 loading conditions were identified from 21 included studies**.** No moderate or strong evidence was found. No studies compared statistically the responses to loading between those with and without pain. There was limited evidence of no effect for six measurements for comparison of unloaded spines to any compression loading, device compression or postural loading. There was limited evidence of increases in the anterior distance of the nucleus position and disc diameter combining all loadings or with postural loading, of decreased posterior disc height with devices and decreased cumulative height with postural loading. There was limited evidence of a decrease in anterior disc height going into supine extension and of a contralateral shift in nucleus peak signal intensity with side bending. For four measurements, there was limited evidence of no difference comparing among axial loading conditions. For five measurements, there was limited evidence of no difference between end-range positions. For passive flexion/extension in the supine position, one statement was conflicting, and two had a limited evidence of effect.

**Conclusion:** The results showed that a gap in knowledge about the effects of loading on lumbar imaging still exists, with a lack of head-to-head comparisons of participants with and without pain. Many measurements have not yet been used with some loading conditions.

**Keywords:** imaging, lumbar spine, intervertebral disc, dural sac cross-sectional area, axial loading, unloaded, flexion/extension, systematic review

## Introduction

Low back pain (LBP) is a major cause of discomfort and disability, with a lifetime prevalence of activity-limiting LBP estimated at 39% and healthcare expenditures in the billions of dollars [1, 131]. Despite this high prevalence, little is known about the underlying pathology, with over 85% of complaints remaining undiagnosed and labelled as non-specific LBP [3, 132].

Diagnostic imaging including plain X-ray, computed tomography (CT), and magnetic resonance imaging (MRI), has been widely used for diagnosing the underlying pathology [9]. MRI is generally preferred due to better contrast and visualisation of soft tissue, as well as the absence of radiation [133]. Unfortunately, the results of conventional imaging are often clinically inconclusive despite identifying degenerative changes [11]. Asymptomatic subjects have similar degenerative imaging features among those with LBP for disc degeneration, herniation and Schmorl's nodes, for example, limiting the specificity of diagnostic imaging and promoting guidelines to limit the use of imaging in uncomplicated LBP [9, 134]. However, current imaging protocols, which involve scanning the spine while resting in the supine position and using qualitative assessments, may be missing important clinically relevant pain-related findings.

MRI and CT scans are mostly performed in a relaxed supine position, while back or leg symptoms are often worsened in sitting or standing positions [135], when the lumbar spine is loaded. Symptoms may be related to changes in segmental alignment and biomechanical forces affecting lumbar structures during loading [129]. This has driven researchers to employ axial loading during CT and MRI examinations, either using loading devices [128–130], or kinetic MR systems in which patients can be imaged in sitting, standing, or while performing loading tasks [136–138].

Iwata et al. suggested that imaging under loading provides enhanced information and showed a significant decrease in height, width, and cross-sectional area of the lumbar intervertebral foramina during axial loading at most levels on asymptomatic healthy subjects [128]. Kanno et al.’s findings indicated that changes in the dural sac cross-sectional area (DCSA) following compression with a device significantly correlated with the severity of symptoms in patients with lumbar spinal canal stenosis (LSCS), which conventional MRI could not detect [139]. Lumbar segmental instability (LSI), defined as an excessive intervertebral translational and rotational

motion detected on imaging, has also been considered a possible source of pain in people with LBP [86, 87, 140, 141]. Spinal instability may also predispose to secondary lesions, such as ligamentous sprain and nerve compression [142]. While sagittal and coronal displacements are usually evaluated on radiographs, displacements in all planes in response to loading can be evaluated in response to loading using CT or MR images [143]. Therefore, it appears important to determine if assessing the spine under various loading conditions reveals additional findings relevant to painful spinal conditions.

How imaging findings are quantified is another consideration possibly responsible for limiting the specificity of examinations. Frequently, MRI grading algorithms (e.g. Thompson [113], Pfirrmann [16], and Modic [144]) employ a small number of discrete ordinal scores to assess degeneration. Such ratings are limited, in term of their reliability, their ability to distinguish differences between individuals and detect changes within individuals over time [118, 119]. The use of quantitative continuous measurements of imaging findings may prove more informative about the effect of loading than qualitative assessments.

A review is needed comparing imaging done in a resting supine position or under various loading conditions focused on studies using quantitative assessments to better understand the pathology of LBP. The review will help identify the best measurements and loading protocols to understand which imaging findings in response to loading are most closely related to LBP symptoms or spine degeneration. The main purpose of this systematic review was to determine the immediate (within-day) response of the lumbar spine to different loading conditions using quantitative assessments of the disc's morphology and signal intensity from spine imaging in healthy controls and in those with LBP. This review thereby aimed to identify candidate biomarkers and loading protocols to advance diagnosis or meaningful subgrouping of common spinal disorders.

**Methods**

The methodology for this systematic literature review was guided by the Cochrane Handbook for Systematic Reviews of Interventions [145]. In this review, we considered loading as the intervention in the appraisal and therefore employed Cochrane's review guidelines. The

reporting was planned according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement [146].

***Literature search and study selection.*** Articles were searched from database inception until August 10, 2014, from Medline (Ovid), EMBASE (Ovid), Scopus, Web of Science Core Collection, BIOSIS Previews and CINAHL (EBSCO). The free text and indexed search keywords used were variants related to population defined by back pain, intervertebral disc, lumbar disc, disc, lumbar vertebra, or lumbar spine. These search terms were combined with terms related to measurement methods using AND including water, MRI, magnetic resonance, signal intensity, strain, stress, or finite element. Finally, the search above was combined with terms related to loading using AND: traction, compression, decompression, load, unload, weight-bearing, bending, flexion, extension or rotation. Only English manuscripts were retrieved (Appendix II-Search Strategies).

The following keywords were used to eliminate studies not meeting our inclusion criteria: abstract type (proceeding, symposium), animals (ovine, sheep, dog, canine, rat), cadaveric, surgery (titanium, stent, rod, fixation, screw, cage, implants, microsurgery, laminectomy), and cancer (malignant, tumor). Search terms were selected through brainstorming among three rehabilitation experts, including physical therapists, an orthotist/biomechanist, and a university librarian. The first 500 abstracts in a preliminary search were reviewed to refine search keywords, and the search was run again using the following keywords as new filters to narrow the search: osteoporosis, deformity (scoliosis, kyphosis), failure testing, hematoma, lower extremity, leg, angiography, cervical, implantation, dis* decompression), viscoelastic, veterinary journal, animal models and small animals. References cited in included articles were reviewed to identify additional articles. Results from each database were uploaded to RefWorks® (ProQuest, USA) and duplicates were excluded after review by using the close duplicate function.

***Selection of studies.*** Two of three investigators screened each reference for eligibility. First, reviewers screened the titles and abstracts (EP, MB and VA), and then screened the full text of the articles of retained abstracts (VA, SA). Reviewers had two to 20 years of experience with this literature. Disagreements during titles and abstracts screening that remained after consensus

discussion led to the inclusion in the full-text stage of screening. Disagreements after consensus discussion following full-text screening were resolved by a third reviewer (EP).

***Selection criteria.*** Only peer-reviewed studies reporting in-vivo measurements of the immediate effects of experimental or postural loading on imaging findings on the lumbar spine, in over 10 healthy or symptomatic adults (to ensure stable estimates), were included. Studies were excluded if they investigated specific pathologies (fracture, cancer, congenital deformities and systemic diseases), focused only on spinal levels other than lumbar, on only post-surgery effects, animals, or cadavers. Letters to the editors, publications only of an abstract, or not providing pre- and post-loading information were also excluded.

***Outcomes.*** The quantitative outcomes extracted were lumbar spine responses to loading on imaging, including geometric, texture and signal intensity characteristics of the disc, spinal canal, lateral foramen, vertebra and endplate. No information was extracted about muscles or any qualitative rating.

***Data extraction.*** A standard form (Appendix III) was developed to extract data based on published guidelines [147–149]. Data on participants, loading conditions, imaging methods, imaging measurements, and effects of loading were extracted from full-text articles by two investigators (VA, SA). Disagreements were resolved by a third investigator (EP). Extracted information on participants included: age, gender, height, diagnosis, pain duration, location and intensity. The loading information extracted included: type, magnitude, mode of application, duration and pre-loading conditions. The imaging information extracted included: method, spinal levels, and timing relative to loading and measurements definitions. Finally, we also extracted measures of central tendency and dispersion for each imaging measurement before, after and/or for the change in response to loading. Detailed extraction tables are available as supplementary online material. (http://dx.doi.org/10.7939/DVN/10973)

***Methodological quality.*** The quality of included studies was assessed as recommended by the PRISMA (Appendix V), and the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guidelines [150]. The quality appraisal focused on seven categories: subject recruitment, examiners, methodology, outcomes, handling missing data, statistical analysis and results (Appendix III). Two reviewers (VA, SA) conducted the critical appraisal

independently. Practice appraisals and discussion of five full-text papers occurred as training before the full review. Studies with a minimum score of 70% were considered to be of high quality and those with a lower score to be of low quality [151]. Three questions, one from the subject domain and two from the examiner domain, could be scored "not relevant" if the study was a single group study or only assessed one measurement. If a question on the quality appraisal form was not applicable, it was excluded and the score was calculated as the ratio of the number of relevant criteria met to the total number of relevant criteria for this domain and reported out of the original number of questions.

***Data synthesis and analysis.*** A PRISMA flowchart was completed to summarise article selection (Figure 1) [146]. Agreement between reviewers on article selection at each stage and on the quality appraisal of the included full-text articles was determined using percent agreement and Kappa coefficients [152]. Summary tables were prepared for participants' descriptions (table 4), loading conditions (table 5), imaging methods (table 6), quality appraisal scores (table 7), the level of evidence summary statements and outcomes extracted (table 8).

**Table 4 Levels of evidence for summary statements and description of criteria adopted a priori to determine the level of evidence** [151]

| *Level* | *Description* |
|---|---|
| *Strong* | Consistent results (≥75%) from at least 2 high-quality* studies |
| *Moderate* | 1 high-quality* study and consistent findings (≥75%) in 1 or more low-quality studies |
| *Limited* | Findings in 1 high-quality* study or consistent results (≥75%) among low-quality studies |
| *No* | No study identified |
| *Conflicting* | Inconsistent results irrespective of study quality |

*Studies with quality scores over 70% were deemed high quality.

The level of evidence (strong, moderate, limited, no, and inconclusive evidence) for the effect of each form of loading on each imaging variable was determined according to the consistency of the research findings and the methodological quality of the included studies [151]. If ≥75% of the relevant comparisons reported that a loading condition affected an imaging variable similarly, the evidence was considered consistent [151]. To formulate levels of evidence summary statements, three different loading conditions were considered: postural, axial using any loading device and loads due to physiologic positioning (e.g. flexion, extension…) of the lumbar spine.

The effects of the postural and axial loads were combined and compared, while the loads due to physiological movements were considered a separate category as the spine likely bears different loads and torques during such movements.

**Results**

***Studies included.*** The search identified 4225 references after removing duplicates. After titles and abstracts screening, 4104 were excluded and 151 were included for full-text screening (Figure 4). After screening, 21 papers met selection criteria. The major reasons for exclusion were not focusing on loading (n=71), narrative reviews (n=14) and employing less than 10 subjects (n=13) (Figure 4).

***Demographic information.*** Prospective case series studies were the most frequent study type. One was a prospective case-control study [153], and two were retrospective case-series studies[154, 155] (table 5). Most reported age varying between 20 and 77 years. Ten studies included participants with LBP [83, 87, 124, 130, 140, 154, 156–160], but only five reported on the chronicity [154, 156, 157], or duration of pain[87, 139] (table 5). Only one study reported the intensity of pain [139], and five reported the location of the pain [139, 154–156, 160] on subjects with LBP.

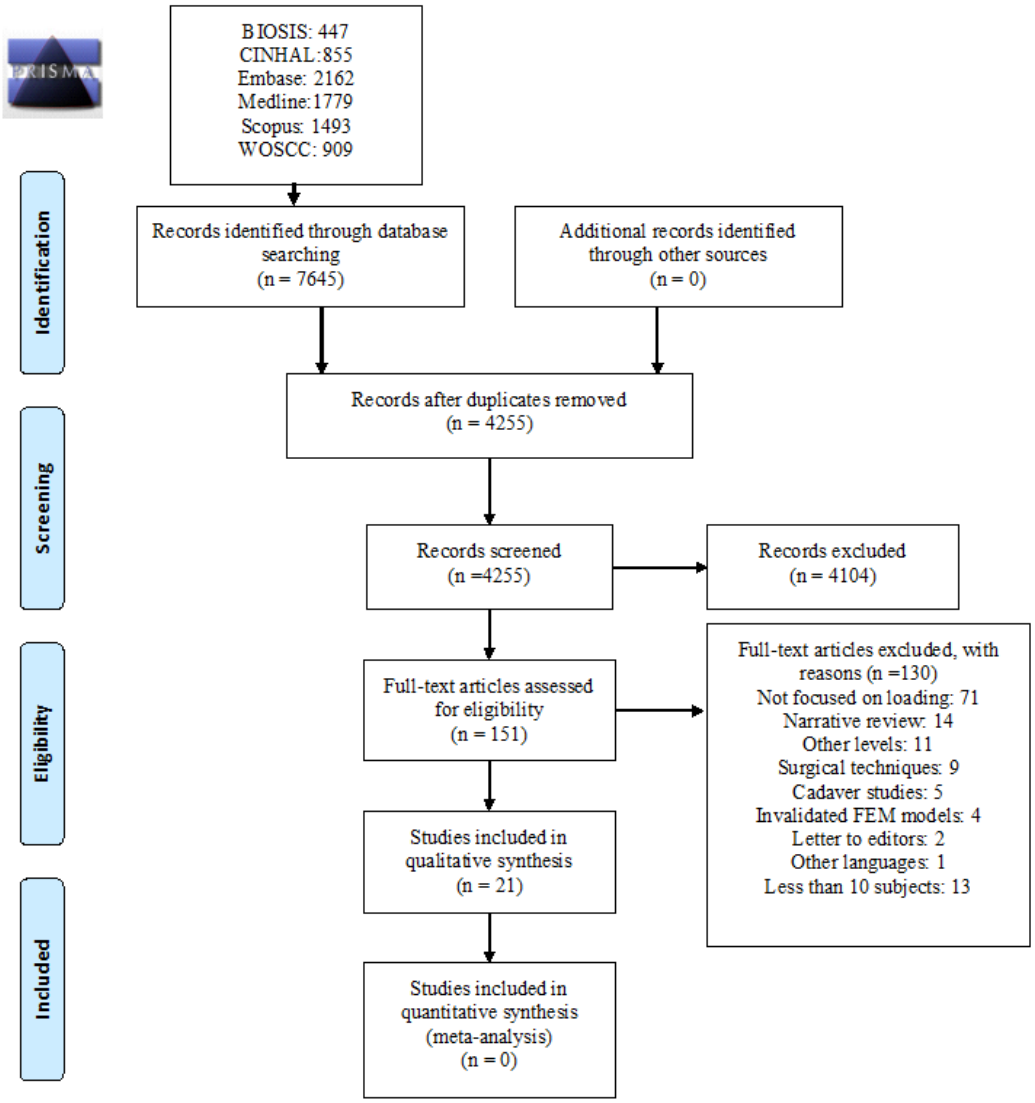


Figure 4 PRISMA flow diagram

**Table 5 Description of study type and study participants in the included studies.**

| Authors | Study Type | Recruitment strategy and Selection Criteria | Number of subjects and groups | Descriptive | | | Diagnosis | Pain (Duration, intensity, location) |
|---|---|---|---|---|---|---|---|---|
| | | | | Age (years) | Height (cm) | Weight (kg) | | |
| Aguilera-Repiso et al. [130] | Prospective Case-series | Consecutive; with pain due to a lumbar hernia and unsuccessful conservative treatment | 15 (12♂, 3♀) | 41.2 (27-57) | | | Lumbar herniation Lumbar herniation | |
| Ahn et al. [154] | Retrospective Case-series | Consecutive with chronic LBP, sciatica, or neurologic intermittent claudication and without lumbar surgery | 51 (24♂, 27♀) | 51 (21-77) | | | Chronic BP | 2wks - 5yrs at L2-S1 |
| Chung et al. [161] | Prospective Case-series | Volunteers, asymptomatic with normal range of motion, without LBP requiring medical attention, connective tissue/ neurologic | 20 (9♂, 11♀) | 26 (22-30) | | | Healthy | |

| Authors | Study Type | Recruitment strategy and Selection Criteria | Number of subjects and groups | Descriptive | | | Diagnosis | Pain (Duration, intensity, location) |
|---|---|---|---|---|---|---|---|---|
| | | | | Age (years) | Height (cm) | Weight (kg) | | |
| | | disease, spinal column fracture and neurologic impairment | | | | | | |
| Eberhardt et al. [153] | Prospective Case-control | Volunteers with suspected lumbar spinal stenosis, sensory disturbance, paresis or claudication and patients prepared for surgery and without any contra-indications for MRI, surgery, lack of spinal stenosis | Healthy: 43 (19♂, 24♀) Symptomatic: 47 (22♂, 25♀) | 38±9.0<br><br>67.7±13.1 | | | Healthy<br><br>Lumbar spinal stenosis | |
| Edmondston et al. [162] | Prospective Case-series | Volunteers without lifetime low back pain, contra- | 10 (4♂, 6♀) | 30±5.8 | 171±6 | 68±32 | Healthy | |

| Authors | Study Type | Recruitment strategy and Selection Criteria | Number of subjects and groups | Descriptive | | | Diagnosis | Pain (Duration, intensity, location) |
|---|---|---|---|---|---|---|---|---|
| | | | | Age (years) | Height (cm) | Weight (kg) | | |
| | | indication for MRI | | | | | | |
| Fazey et al. [163] | Prospective Case-series | Healthy volunteers | 21 (10♂, 11♀) | 24.8 (20-34) | | | Healthy | |
| Fujiwara et al. [164] | Prospective Case-series | Consecutive, LBP, leg symptoms, or both and without traumatic, inflammatory or tumorous disorders, nor prior surgery, spondylolysis or transitional vertebra | 70 (36♂, 34♀) | 45.9±19 | | | LBP or leg symptoms | |
| Hirasawa et al. [165] | Prospective Case-series | Volunteers, males without LBP | 32 (32♂) | 32 (21-61) | | | Healthy | |
| Kanno et al. [139] | Prospective Case-series | Consecutive with typical case of LSCS and without prior lumbar surgery, spinal anomalies, | 88 (54♂, 34♀) | 68±10 | | 64±11 | LSCS | 33±33 mths 39±34/100 Leg |

| Authors | Study Type | Recruitment strategy and Selection Criteria | Number of subjects and groups | Descriptive | | | Diagnosis | Pain (Duration, intensity, location) |
|---|---|---|---|---|---|---|---|---|
| | | | | Age (years) | Height (cm) | Weight (kg) | | |
| | | severe osteoporosis, poly-neuropathy, arterial insufficiency | | | | | | |
| Karadimas et al. [156] | Prospective Case-series | Volunteers with chronic degenerative back pain, conservative treatments failed | 30 (14♂, 16♀) | 44.5 (25-61) | | | Chronic degenerative LBP | Low back |
| Keorochana et al. [155] | Retrospective Case-series | Consecutive, with LBP and without previous history of spinal surgery, inability to participate in weight-bearing MRI, vertebral fractures, tumours, spondylo- | Straight/ kyphosis: 84 (54♂, 30♀)<br>Normal lordosis: 292 (155♂, 139♀)<br>Hyper-lordosis: 52 (32♂, 20♀) | 40.9±11<br>42.9±11.9<br>46.4 ±12.6 | 179.6± 28.4<br>182.2±31.2<br>191.0±34.5 | | LBP<br>LBP<br>LBP | Low back |

| Authors | Study Type | Recruitment strategy and Selection Criteria | Number of subjects and groups | Descriptive | | | Diagnosis | Pain (Duration, intensity, location) |
|---|---|---|---|---|---|---|---|---|
| | | | | Age (years) | Height (cm) | Weight (kg) | | |
| | | listhesis or scoliosis | | | | | | |
| Kim et al. [160] | Prospective Case-series | Volunteers with LBP | 54 (22♂, 32♀) | 61 (35-78) | | | LBP | Low back |
| Kinder et al. [159] | Prospective Case-series | Volunteers, symptomatic with sciatic or neurogenic claudication and without prior spinal surgery | 120 (60♂, 60♀) | 20-92 | | | Sciatic or neurogenic claudication | |
| Kong et al. [140] | Prospective Case-series | Consecutive, with LBP and without coronal or severe sagittal plane deformity, trauma, tumour, infection, lumbar spine surgery | 316 (200♂, 116♀) | 42.1 (16-85) | | | LBP | |
| Kozanek et al. [166] | Prospective Case-series | Consecutive, healthy without lower back pain or | 11 (5♂, 6♀) | 50-60 | | | Healthy | |

| Authors | Study Type | Recruitment strategy and Selection Criteria | Number of subjects and groups | Descriptive | | | Diagnosis | Pain (Duration, intensity, location) |
|---|---|---|---|---|---|---|---|---|
| | | | | Age (years) | Height (cm) | Weight (kg) | | |
| | | other spinal disorders | | | | | | |
| Lee et al. [90] | Prospective Case-series | Volunteers, without back pain, prior back surgery, contra-indications for MRI | 15 (12♂, 3♀) | 24.4±4.8 | | 45-90 | Healthy | |
| Madsen et al. [83] | Prospective Case-series | Consecutive: with neurogenic claudication, suspected discogenic pain, minor degenerative spondylo-listhesis and without severe degenerative changes in lumbar spine, markedly reduced disc height, pronounced osteophytic | Comparing vertical vs. supine unloaded, supine with 40% and 50% body weight compression 16 (?♂, ?♀) Comparing supine psoas relaxed vs. supine extended vs. supine extended with 50% body weight compression | 53 (18 -80)<br>57 (40-70) | | 75 (55-88)<br>75 (48-95) | Neurogenic claudication, Discogenic pain, Degenerative spondylo-listhesis/ Neurogenic claudication | |

| Authors | Study Type | Recruitment strategy and Selection Criteria | Number of subjects and groups | Descriptive | | | Diagnosis | Pain (Duration, intensity, location) |
|---|---|---|---|---|---|---|---|---|
| | | | | Age (years) | Height (cm) | Weight (kg) | | |
| | | enlargement of facet joints, MRI contra-indications, tumours, vertebral trauma, osteoporosis, rheumatic disease, severe cardio-pulmonary disease, athero-sclerosis with peripheral vascular disease | 20 (?♂, ?♀) | | | | | |
| Maigne et al. [87] | Prospective Case-series | LBP worsened by sitting<br>LBP relieved by standing | 42 (6♂, 36♀)<br>32 (4♂, 28♀) | 54.9±9.8<br>57.5±13.4 | | | LBP<br>LBP | 2-120 mths<br>3-150 mths |
| Mauch et al. [88] | Prospective Case-series | Volunteers, without any lumbar spine pain, or lumbar surgical or | 35 (20♂, 15♀) | 28.0±6.5 | | | Healthy | |

| Authors | Study Type | Recruitment strategy and Selection Criteria | Number of subjects and groups | Descriptive | | | Diagnosis | Pain (Duration, intensity, location) |
|---|---|---|---|---|---|---|---|---|
| | | | | Age (years) | Height (cm) | Weight (kg) | | |
| | | conservative treatment | | | | | | |
| Nazari et al. [167] | Prospective Case-series | Volunteers, without LBP requiring treatment or causing absence at work | 25 (25♂) | 20-38 | | | Healthy | |
| Okawa et al. [157] | Prospective Case-series | Volunteers, healthy or with chronic LBP or degenerative spondylo-listhesis | 17 (13♂, 4♀)<br>8 (4♂, 4♀)<br>8 (4♂, 4♀) | 22.3 (21-24)<br>43.5 (31-60)<br>63.1 (46-83) | | | Healthy<br>Chronic LBP<br>Degenerative spondylo-listhesis | |

Abbreviations and symbols: LBP=Low back pain; MRI=Magnetic Resonance Imaging; ♂=males; ♀=females; LSCS= lumbar spinal canal stenosis; wks=weeks; mths= months; yrs= years

**Table 6 Description of the imaging methods and loading strategies tested in the included studies**

| Studies | Imaging Modalities | Type of Loading | Mode of Application | Duration of Loading | Magnitude of Loading | Time of Day of Loading | Precondition of Loading |
|---|---|---|---|---|---|---|---|
| Aguilera-Repiso et al. [130] | MRI | Compression | DynaWell | ≤20 mins during MRI | 50% BW | | |
| Ahn et al. [154] | MRI | Compression | DynaWell | ≥5 mins prior to MRI and during MRI | 50% BW | | |
| Chung et al. [161] | MRI | Passive Flexion/ extension, rotation in supine | Positional | During MRI | | | |
| Eberhardt et al. [153] | MRI | Passive-flexion/ extension in supine Compression | Mechanical bracing devices | 40 min without compression. 60 min with compression, during MRI | 75% BW | | |
| Edmondston et al. [162] | MRI | Passive flexion/ extension, rotation in supine | Rolled towels | During MRI | | | |
| Fazey et al. [163] | MRI | Voluntarily side-bending in supine | Patient active left side flexion, with passive overpressure during MRI | During MRI (6 min 34sec) | | | |

| Studies | Imaging Modalities | Type of Loading | Mode of Application | Duration of Loading | Magnitude of Loading | Time of Day of Loading | Precondition of Loading |
|---|---|---|---|---|---|---|---|
| Fujiwara et al. [164] | X-ray | Voluntarily flexion/ extension in lateral decubitus | Positional | During Imaging | Maximum strength | | |
| Hirasawa et al. [165] | MRI | Standing voluntarily flexion/ extension in sitting | Positional | During MRI | | Morning | |
| Kanno et al. [139] | MRI | Compression | DynaWell | During MRI | 50% BW | | |
| Karadimas et al. [156] | MRI | Postural loading (sitting/ lying) Standing discontinued due to patient difficulty staying still | Postural | During MRI | | | |
| Keorochana et al. [155] | MRI | Standing/ sitting/ lying | Postural | During MRI | | | |
| Kim et al. [160] | MRI | Compression | DynaWell | During MRI | 50% BW | | |
| Kinder et al. [159] | MRI | Compression | Non-magnetic compression device | ≥5 mins during MRI | 50% BW | | |
| Kong et al. [140] | MRI | Flexion/ extension | Positional | During MRI | | | |

| Studies | Imaging Modalities | Type of Loading | Mode of Application | Duration of Loading | Magnitude of Loading | Time of Day of Loading | Precondition of Loading |
|---|---|---|---|---|---|---|---|
| Kozanek et al. [166] | X-ray fluoroscopic | Standing, Active trunk flexion, extension, maximal left-right and left-right twisting | Positional | During Fluoroscopic | | | |
| Lee et al. [90] | MRI | Kneeling<br><br>Compression | Kneeling<br><br>DynaWell | During MRI | 50% BW | Morning | Unloaded: supine 1 hr. Loaded: 50%BW compression or kneeling torso vertical 30 mins each |
| Madsen et al. [83] | MRI | Passive flexion/ extension in supine Compression | Positional DynaWell | During MRI | Postural 40% BW, 50% BW, | | |
| Maigne et al. [87] | X-ray | Standing in an erect posture, standing in maximum extension, sitting in the most painful position, sitting in maximum | Positional | During x-ray | | | |

| Studies | Imaging Modalities | Type of Loading | Mode of Application | Duration of Loading | Magnitude of Loading | Time of Day of Loading | Precondition of Loading |
|---|---|---|---|---|---|---|---|
| | | lumbar flexion | | | | | |
| Mauch et al. [88] | MRI | Standing/ sitting/ lying | Postural | During MRI | | Late afternoon | 20 mins lying before standing segment |
| Nazari et al. [167] | MRI | Standing/ sitting/ lying | Postural | During MRI | | Midday | 15-min. Recumbent between scans |
| Okawa et al. [157] | X-ray video fluoroscopy | Active flexion/ extension in standing | Bending forward and backward | During fluoroscopy | | | |

Abbreviations: MRI= Magnetic resonance imaging, Mins= minutes, hrs=hours

***Imaging and loading modalities used.*** MRI was employed alone in 17 studies. X-ray fluoroscopy was used in two studies [157, 166] and two studies employed plain x-ray imaging [87, 164] (table 6). Eight studies employed positional loading in the form of end range positions of the lumbar spine during imaging including: flexion/extension [83, 87, 124, 140, 157, 161, 165, 166], side-bending [163] or rotation [161, 166]. The DynaWell® compression device was employed to load the lumbar spine with 40 to 50% body weight (BW) in six studies [87, 90, 130, 139, 154, 160], while a mechanical bracing device to load with 75% BW [153], and a nonmagnetic compression device (50% BW) [159] were used in two other studies. Postural loading was used in five studies (standing/sitting/lying [88, 155, 156, 167] or kneeling [90]). In all but 3 of 21 included studies, loading was applied only during the imaging. These 3 studies applied loading before and during imaging with pre-imaging loading durations of >5min [154], and ≤20min [130, 153]. Only three studies described standardised preconditioning before imaging [88, 90, 167] with one more specifying time of examination to control the loading history [165].

***Methodological quality.*** None of the included studies met the 70% high-quality threshold (table 7). Two studies scored over 50% [87, 164], and eight scored between 40% and 50% [87, 88, 130, 139, 140, 155, 156, 165]. Lack of a well-defined subject's recruitment and examiners' training strategy were the most common sources of bias. Only 9/21 studies reported qualification of the examiners [87, 88, 130, 140, 154, 155, 159, 160, 163]. Eight studies employed a consecutive recruitment strategy to recruit subjects with back pain [83, 130, 139, 140, 154, 155, 164], and without back pain [166]. The rest employed volunteers with pain [90, 156, 160], with sciatic or neurogenic claudication [159], or without pain [88, 153, 157, 161–163, 165, 167]. Study hypothesis or objective and validity of the outcome measured were reported in all studies, except two with unclear outcomes reporting strategy [161, 164]. Only one study met the requirements for handling missing data [160]. A possible common source of bias was the majority of studies failing to formulate correlation and mean difference-testing hypotheses a priori. Only one study had a well-defined statistical analysis strategy [165], and four met 80% of the statistical analysis requirements [83, 88, 156, 165]. No study met all the requirements for reporting results because none reported on important adverse events related to loading.

**Table 7 Quality appraisal of the studies included**

| Authors | Subjects Recruitment /7* | Examiners /4* | Methodology /5 | Outcomes /2 | Missing Data /8 | Statistical Analysis /5 | Results /2 | Overall Score /33* | Overall Score (%)* |
|---|---|---|---|---|---|---|---|---|---|
| Aguilera-Repiso et al. [130] | 4 | 1 | 2 | 2 | 6 | 2 | 1 | 14 | 42 |
| Ahn et al. [154] | 3 | 1 | 3 | 2 | 4 | 2 | 1 | 13 | 39 |
| Chung et al. [161] | 2 | 0 | 2 | 0 | 4 | 2 | 1 | 9 | 27 |
| Eberhardt et al. [153] | 3 | 0 | 3 | 2 | 5 | 1 | 1 | 12 | 36 |
| Edmondston et al. [162] | 2 | 0 | 2 | 2 | 4 | 3 | 1 | 12 | 36 |
| Fazey et al. [163] | 2 | 1 | 2 | 2 | 5 | 1 | 1 | 12 | 36 |
| Fujiwara et al. [164] | 5 | 0 | 5 | 1 | 6 | 4 | 1 | 17 | 52 |
| Hirasawa et al. [165] | 2 | 0 | 3 | 2 | 5 | 5 | 1 | 16 | 48 |
| Kanno et al. [139] | 4 | 0 | 3 | 2 | 6 | 3 | 1 | 15 | 45 |
| Karadimas et al. [156] | 4 | 0 | 3 | 2 | 6 | 4 | 1 | 16 | 48 |
| Keorochana et al. [155] | 4 | 1 | 3 | 2 | 6 | 3 | 1 | 16 | 48 |
| Kim et al. [160] | 5 | 1 | 3 | 2 | 8 | 2 | 1 | 17 | 52 |
| Kinder et al. [159] | 4 | 1 | 3 | 2 | 4 | 1 | 1 | 12 | 36 |
| Kong et al. [140] | 5 | 1 | 4 | 2 | 5 | 2 | 1 | 15 | 45 |
| Kozanek et al. [166] | 4 | 0 | 2 | 2 | 3 | 1 | 1 | 9 | 27 |
| Lee et al. [90] | 3 | 0 | 2 | 2 | 6 | 1 | 1 | 12 | 36 |
| Madsen et al. [83] | 2 | 0 | 2 | 2 | 4 | 4 | 1 | 13 | 39 |
| Maigne et al. [87] | 5 | 1 | 3 | 2 | 6 | 3 | 1 | 16 | 48 |
| Mauch et al. [88] | 2 | 1 | 2 | 2 | 4 | 4 | 1 | 14 | 42 |
| Nazari et al. [167] | 3 | 0 | 2 | 2 | 6 | 1 | 1 | 12 | 36 |
| Okawa et al. [157] | 3 | 0 | 2 | 2 | 4 | 2 | 1 | 11 | 33 |

Overall score: sum of the all scores

*One question from the subject domain and two from the examiner domain could be scored “not relevant” and excluded from score calculations. The affected scores were calculated as the ratio of the number of relevant criteria met to the total number of relevant criteria and reported out of the original number of questions for the score.

***Measurement outcomes.*** Fourteen quantitative measurements of the effect of lumbar loading were identified and the level evidence summary statements were formulated for each measure and loading conditions (table 8). As no study reached the high-quality threshold, the highest level of evidence possible was limited (table 4).

***Unloaded vs. any type of compression loading (postural or axial).*** When formulating summary statement combining the evidence of both postural and axial loading, there was limited evidence of a statistically significant increase of the distance between the anterior wall of the nucleus and the anterior wall of the disc [167] and anterior-posterior diameter of the disc [167] in response to compression loading (table 8). There was also limited evidence of no change in posterior disc bulge [90], range of motion [153], and lumbar lordosis angle in response to loading (table 8). The evidence was conflicting on the effects of loading on the cumulative [83], anterior and posterior disc heights [90, 156, 167], anterior-posterior diameter of the nucleus [167], posterior distance of the nucleus position [167], and dural sac cross-sectional area (DCSA) [83, 88, 130, 139, 154, 159, 160, 165] (table 8). No evidence on the effect of compression loading was found for other quantitative imaging measurements (table 8).

***Unloaded vs. axially loaded spine using any compression devices.*** When focusing only on axial loading conditions using compression device, there was limited evidence of statistically significant decrease in the posterior disc height in response to compression loading (table 8**Table 8**) [90]. There was also limited evidence of no change in response to loading for the cumulative disc height [83], posterior bulging of the disc [4], range of motion [153], and lumbar lordosis angle [83] (table 8). The evidence was conflicting on the effects of loading on the dural sac cross-sectional area (DSCA) [83, 130, 154, 159, 160]. No evidence about this type of loading was found about any other quantitative imaging measurements (table 8).

**Table 8 Levels of evidence for summary statements for each quantitative measurement of the effect of lumbar loading.**

| Level of evidence | From (n healthy; n LBP) | Changes | Measurement Construct | Loading conditions compared |
|---|---|---|---|---|
| **Unloaded vs. any type of loading (axial loading/postural loading)** | | | | |
| Limited | 1(1,0)[167] | ↑ | Anterior distance of the nucleus position | Standing; sitting; lying supine |
| Limited | 1(1,0)[167] | ↑ | Anterior-posterior diameter of the disc | Standing; sitting; lying supine |
| Limited | 1(1,0)[90] | No difference | Change in the posterior bulging | Kneeling; 50% BW axial loading |
| Limited | 1(0,1)[153] | No difference | Range of motion | Supine with 75% BW compression |
| Limited | 3(2,1)[83, 88, 165] | No difference | Lumbar lordosis angle | Standing; Supine with 40% and 50% BW on compression; Sitting |
| Conflicting | 1 (0;1)[83] | Any change | Cumulative disc height | Supine with 40% and 50% BW on compression |
| Conflicting | 3 (2;1) [90, 156, 167] | Any change | Anterior disc height | Sitting; Kneeling; supine with 50% BW compression; Standing/sitting |
| Conflicting | 3 (2,1)[90, 156, 167] | Any change | Posterior disc height | Sitting; Kneeling; supine with 50% BW compression; Standing/sitting |
| Conflicting | 1(1,0)[167] | Any change | Anterior-posterior diameter of the nucleus | Standing; sitting; lying supine |
| Conflicting | 1(1,0)[167] | Any change | Posterior distance of the nucleus position | Standing; sitting; lying supine |

| | | | | |
|---|---|---|---|---|
| Conflicting | 8(3,5) [83, 88, 130, 139, 154, 159, 160, 165] | Any change | Dural sac cross-sectional area | 50% BW on compression Standing; Supine with 40% and 50% BW compression Standing/sitting |
| **Unloaded vs Axial Loading using Any Compression Device** | | | | |
| Limited | 1(0,0)[83] | No difference | Cumulative disc height | Supine with 40% and 50% BW on compression |
| Limited | 1 (1,0)[90] | ↓ | Posterior disc height | Supine with 50% BW on compression |
| Limited | 1(1,0)[90] | No difference | Change in the posterior bulging | Supine with 50% BW on compression |
| Limited | 1(0,1) [153] | No difference | Range of motion | Supine with 75% BW on compression |
| Limited | 1(1,0)[83] | No difference | Lumbar lordosis angle | Supine with 40% and 50% BW on compression |
| Limited | 3(2,1) [83, 88, 165] | No difference | Lumbar lordosis angle | Supine with 40% and 50% BW on compression |
| Conflicting | 5(1,4)[83, 130, 154, 159, 160] | Any change | Dural sac cross-sectional area | Supine with 40% and 50% BW on compression |
| **Unloaded vs. Postural Loading** | | | | |
| Limited | 1 (0;1)[83] | ↓ | Cumulative disc height | Standing |
| Limited | 1(1,0)[167] | ↑ | Anterior distance of the nucleus position | Standing; Sitting |
| Limited | 1(1,0)[167] | ↑ | Anterior-posterior diameter of the disc | Standing; Sitting |
| Limited | 2 (2;0)[90, 167] | No difference | Anterior disc height | Kneeling Standing; Sitting |
| Limited | 1(1,0)[90] | No difference | Change in the posterior bulging | Kneeling |
| Limited | 3(2,1)[83, 88, 165] | No difference | Lumbar lordosis angle | Kneeling |

| | | | | |
|---|---|---|---|---|
| | | | | Standing; Sitting |
| Conflicting | 3 (2,1)[90, 156, 167] | Any change | Posterior disc height | Standing; sitting; kneeling |
| Conflicting | 1(1,0) [167] | Any change | Anterior-posterior diameter of the nucleus | Standing; Sitting |
| Conflicting | 1(1,0)[167] | Any change | Posterior distance of the nucleus position | Standing; Sitting |
| Conflicting | 3(2,1)[88, 139, 165] | Any change | Dural sac cross-sectional area | Standing; Sitting |
| **Among postural or axial loading** | | | | |
| Limited | 1 (0;1)[83] | No difference | Cumulative disc height | Supine with 40% and 50% BW compression |
| Limited | 2(2,0)[90, 167] | No difference | Posterior disc height | Kneeling; 50% BW axial loading<br>Standing; Sitting |
| Limited | 1(1,0)[90] | No difference | Change in the posterior bulging | Kneeling; 50% BW compression |
| Limited | 1(0,1)[83] | No difference | Lumbar lordosis angle | Supine with 40% and 50% BW compression |
| **Unloaded vs end range movements** | | | | |
| Limited | 1 (1;0)[162] | ↓ | Anterior disc height | Passive flexion/extension, rotation in supine |
| Limited | 1(1,0)[163] | ↑ Towards contralateral side to SB | Change in the location of peak signal intensity of the nucleus on the mid-coronal plane | Active side-bending in supine |
| Limited | 1(0,1)[164] | No difference | Abnormal tilting movement/ translatory instability/ rotatory instability | Active flexion/extension in lateral decubitus |
| Limited | 2(1,1)[83, 165] | No difference | Lumbar lordosis angle | Voluntarily flexion/extension in sitting |

| | | | | |
|---|---|---|---|---|
| | | | | Passive extension in supine with 50% BW compression |
| **Among end range movements** | | | | |
| Limited | 1 (1;0)[162] | ↓ | Anterior disc height | Passive flexion/extension in supine |
| Limited | 1(1,0)[162] | No difference | Posterior disc height | Passive flexion/extension in supine |
| Limited | 1(1,0)[162] | ↓ | Anterior distance of the nucleus position | Passive flexion/extension in supine |
| Limited | 1(1,00)[161] | No difference | Anterior-posterior diameter of the disc | Passive Flexion/extension in supine |
| Limited | 1 (1;0)[161] | No difference | Left and right parasagittal, mid coronal disc diameters | Passive Flexion/extension in supine |
| Limited | 1(0,1)[164] | No difference | Range of motion | Active flexion/extension in lateral decubitus |
| Limited | 1(1,0)[165] | No difference | Lumbar lordosis angle | Active flexion/extension in sitting |
| Conflicting | 1(1,0)[161] | Any change | Dural sac cross-sectional area | Passive Flexion/extension in supine |

***Unloaded vs. postural loading (sitting/standing).*** When focusing only on axial compression using postural loading, there was limited evidence of a statistically significant decrease in the cumulative disc height in response to postural loading (table 8) [83]. There was also limited evidence of a statistically significant increase in the anterior distance of the nucleus position and anterior-posterior diameter of the disc in response to loading (table 8) [167]. There was limited evidence of no difference in anterior disc height [90, 167], changes in the posterior disc bulging [90] and lumbar lordosis angle [83, 88, 165] (table 8). The evidence was conflicting about the posterior disc height [90, 156, 167], the anterior-posterior diameter of the nucleus [167], the posterior distance of the nucleus position [167], and DSCA [88, 139, 165]. No evidence about this type of loading was found about any other quantitative imaging measurements (table 8).

***Comparison among different postural and axial loading conditions.*** Few studies compared multiple loaded conditions [90]. There was limited evidence of no difference of the cumulative [83] (Passive flexion/extension in supine; 40% BW, 50% BW compression) and posterior disc heights (Kneeling; 50% BW axial loading; standing; sitting; lying supine) [90, 167], change in the posterior bulging (kneeling; 50% BW axial loading) [90] and lumbar lordosis angle (passive flexion/extension in supine; 40% and 50% BW on compression) [83]. No evidence was found about any other quantitative imaging measurements (table 8).

***Unloaded vs. end range movements.*** Only five studies compared unloaded spine images with end-range movement conditions. There was limited evidence of a statistically significant decrease of the anterior disc height in response to going from lumbar flexion to extension in the supine position (table 8) [162]. There was also limited evidence of a statistically significant shift of the location of the nucleus peak signal intensity toward the contralateral side in the mid-coronal plane during side bending [163]. There was limited evidence of no difference in abnormal tilting movement, translatory and rotatory instability measurements [164] as well as in lumbar lordosis angle [83, 165] among supine psoas relaxed, supine unloaded, supine extended, supine with 40% and 50% body weight, supine extended position with additional 50% body weight, sitting, sitting flexed, sitting extended and standing. No evidence about this type of end-range loading was identified for about any other quantitative imaging measurements (table 8).

***Among end range movements.*** Only four studies compared measurements among different end-range movements. There was limited evidence of a statistically significant increase in the anterior disc height and the anterior distance of the nucleus position going from a flexed position

to an extended position [162]. There was also limited evidence of no change for the posterior disc height (supine extension vs. flexion) [162], anterior-posterior, left and right parasagittal, mid-coronal diameter of the disc (among supine neutral, flexion, extension, rotation to the right and left) [161], range of segmental motion (standing flexion vs. extension) [164], and lumbar lordosis angle (among supine unloaded, normal sitting, sitting flexed, sitting extended and standing) [165]. The evidence was conflicting on the DSCA changes on neutral, extension and rotation to the right and left compared to the flexion position [161]. No evidence comparing end-range movements was identified about the any of the other quantitative MRI measurements (table 8).

**Discussion**

This systematic review found no moderate or strong evidence about the effect of loading or of end-range movement positions on lumbar imaging quantitative findings due to the insufficient quality of the included studies. Further, only two studies included both healthy participants and participants with low back pain and these studies did not report the statistical significance of the comparisons of the effects between groups [153, 157]. We can therefore only rely on comparing studies conducted with only one type of participants to understand the difference in response to loading of participants with and without pain.

A number of conflicting (0 to 6 measurements) and limited (4 to 7 measurements) evidence summary statements concerning the effects of loading and movements on the lumbar spine were observed depending on the loading conditions compared. This conflicting evidence could be due to heterogeneity in the studies in terms of the population studied (e.g. with or without pain, different ages) or because of other heterogeneous methodological considerations discussed below. The limited evidence summary statements often showed no effect of the loading conditions compared (2 to 5 measurements depending on the loadings compared). Overall, the observation of conflicting results and the small number of measurement constructs that detected consistent effects of loading may suggest that the types of loading examined to date have a relatively small effect on the quantitative measurements of the lumbar spine studied so far.

Nine measurement constructs were compared among studies where both studies focused only on either healthy or LBP participants were available (table 6). For those comparisons, the results of three measurement constructs were consistent showing no effect of loading in both participants with and without pain from different studies. For the remaining six other instances,

results were conflicting when combining results from participants with and without pain. When trying to formulate levels of evidence summary statements separately for participants with and without pain, the summary statements on two outcomes remained conflicting in both types of patients. For the DSCA, there was limited evidence of a decrease with postural loading in those with low back pain and the results remained conflicting in healthy participants. For DSCA, there was limited evidence of a decrease in healthy with combined loading studies and with device compression while evidence remained conflicting in those with LBP. For posterior disc height, comparing unloaded to any loading there was limited evidence of no difference with loading in LBP and the evidence remained conflicting in healthy participants. Other measurement constructs were evaluated only on participants with back pain or asymptomatic participants, and not both (table 6).

The studies compared were heterogeneous, which may have contributed to the frequent conflicting evidence summary statements and limiting our ability to observe consistent effects of loading. Fourteen different measurement constructs were observed among included studies. The sample size varied from 10 to 316 (64±123) participants, with age ranging from 20 to 92 years. Nine studies only included healthy participants, ten included only LBP participants, and two included both, with and without pain (table 5). More than 16 different loading conditions (table 6) and 3 imaging modalities were inventoried among the included studies. Only three studies employed preloading of the spine before imaging. A few studies conducted manually digitised measurements using OSIRIX [83, 156, 165, 167], while the rest employed other software or did not report software details. Three studies employed automatic measurements [140, 155, 158], one study employed semi-automatic measurements [153], and the rest employed manual measurements.

Difficulty in formulating stronger levels of evidence summary statements is partially due to the lack of quality research on this topic. Most of the studies reviewed did not report the degree to which assessors were trained before conducting measurements or whether the measurements extracted presented adequate reliability. Further, nine of the 21 included studies included samples with fewer than 30 participants suggesting that a limited power may have limited the authors' ability to observe statistically significant results. Of 21 studies included, only two compared subjects with and without pain. Therefore, the clinical value of some of the proposed measures

still remains unclear and more high-quality studies among matched patients, and asymptomatic participants are needed to determine the clinical value of these measurement constructs.

It is unclear if compression devices are limited regarding the maximum load that can be applied to the spine safely. In the present review, only one study applied a compression load equal to 75% body weight [153], others employed compression loads less than or equal to 50% body weight, with only one study controlling for how long before the imaging the loads were applied. The relatively small number of consistent effects of loading observed with supine compression devices (only decreased posterior disc height) versus when using postural loads (increased anterior distance of nucleus, A-P diameter of disc and decreased cumulative disc height) may be due to the authors using less than 80% BW in all the studies where loads were specified. Overall, the results of this review were not in favour of the hypothesis that axial loading using postural or device compression could induce significant morphologic changes in the lumbar spine. Further research using a wide array of quantitative measurement and loading conditions is needed to determine if loading during imaging may increase the specificity for clinically relevant lumbar pathology.

The evidence was conflicting about the effects of loading on the imaging presentations of DSCA. Clinically, imaging is not always correlated with the severity of LSCS symptoms [168, 169]. Subjects with significant narrowing of the spinal canal, may not demonstrate any symptoms, while subjects with severe symptoms may not demonstrate any significant narrowing of the LSCS [139]. Identifying imaging findings correlating with clinical symptoms in LSCS would thus be beneficial to achieve a more accurate diagnosis and help plan a more appropriate treatment. Our review yielded conflicting results on the effects of loading on the LSCS, and only one study was identified evaluating the effects of loading specifically on LSCS symptoms [139]. The results indicated a statistically significant negative correlations between the DCSA in axially loaded MRI with both walking distance and the Japanese Orthopedic Association score [139]. Our review did not yield any evidence on the possible effects of extension motion compared to an unloaded supine position on the LSCS. Although some researchers have suggested that the extension could trigger LSCS symptoms by a reduction of the dural sac size [170, 171]. Further high-quality studies are therefore needed to examine the possible correlation between the change in the DCSA caused by the loading and the severity of the symptoms in patients with LSCS.

Although imaging the lumbar spine while it is loaded using sensitive quantitative measurements may help capture abnormal vertebral motion, and thus increase the sensitivity of

imaging for diagnosis of LSI, there were no studies included in this review examining the effect of compression on LSI measurements. Further, only one study with 70 participants with LBP examined the effect of end-range movements on quantitative LSI measurements. Therefore, there is still insufficient evidence to formulate strong level of evidence summary statements related to the correlation between LSI imaging findings in response to loading and patients' symptoms (table 8).

The present review identified important research gaps. Interestingly, for numerous quantitative measurements of the effect of loading on the lumbar spine, no evidence could be found suggesting further research is needed. Indeed, depending on the loading conditions compared, for 3 to 10 of the 14 measurements used in at least one study, these measurements had not been used to quantify the effect of loading. Further, many measurements have only been used with either healthy participants or participants with low back pain but not both (table 8). Therefore, drawing a solid conclusion regarding any possible effect of loading on the lumbar spine still remains challenging for most measurement constructs of interest. Similarly, drawing conclusions about whether the response to loading differs between participants with and without pain also remains challenging.

Further, despite encountering 16 loading conditions, our review did not yield any evidence on the possible effects of traction loading on the lumbar spine. Traction has been employed traditionally to reduce pain and discomfort among participants with LBP and sciatica [47], affect the fluid content, promote molecular transport in the IVDs [172], and reduce the size of herniated material [173]. Therefore, traction theoretically could be a loading condition useful to identify biomarkers of relevant low back pathology. Chung et al. in a study published after the search period of the present study found a significant elongation of the lumbar spine and a decrease of the size of disc herniation after 30 minutes of traction loading [173]. As no studies about the in-vivo effects of traction on other measurements or comparing participants with and without pain have been conducted, the gap of knowledge still exists.

Although few studies have employed MR imaging to quantify effects of loading on the disc fluid content [22, 127], our review did not yield any results on the effects of loading on the disc fluid content by either measuring mean MR signal intensity or measuring diffusion coefficients. However, recently, Nilsson et al. in a feasibility study on participants with LBP to evaluate effects of compression on disc fluid content using $T_2$-mapping, found a significant

increase in $T_2$ values of the whole disc and subsections [127]. $T_2$-value changes were correlated with degeneration grade, changes in disc angle and lumbar level [127]. As there are no studies about the in-vivo effects of all other forms of loading on the disc fluid content, the gap of knowledge related to disc fluid distribution in response to loading still exists.

***Study limitations.*** This review solely included studies published in English, and no search was conducted of the grey literature. These two factors might have caused a potential bias in selecting relevant papers. As discussed earlier, the results were very heterogeneous which prevented meta-analysis. Unfortunately, the literature is not sufficiently rich to limit the review to studies involving head-to-head comparisons of patients with and without low back pain. This review focused only on quantitative measurements of the effect of loading on the lumbar spine. Studies using discrete subjective ratings have been published to describe the effect of loading on the lumbar spine with some studies suggesting that loading helps identify relevant clinical findings [174]. However, discrete subjective ratings have routinely been criticised for their lack of reliability justifying excluding such measurements from this review.

**Conclusion**

The heterogeneous results highlighted inconsistent evidence regarding the effects of loading on the lumbar spine. The review did not yield any moderate or strong evidence because of the insufficient quality scores of the included studies. For many measurement constructs, no evidence was identified to draw a solid conclusion regarding their possible effects of loading on the lumbar spine. More than half of the limited evidence observed on cumulative, anterior and posterior disc height, the anterior distance of the nucleus position, change in the posterior bulging, abnormal tilting movement, translatory and rotatory instability, range of motion, and lumbar lordosis was of no effect of loading. All included studies that examined the differences among different compression loading conditions provided limited evidence of no difference in cumulative and posterior disc height, posterior bulging of the disc, and lordosis angle. The results highlighted that the gap of knowledge regarding the effects of loading on the imaging presentations of the lumbar spine still exists. Particularly, there is a lack of research on whether the response to loading could help increase the specificity of MRI for low back pain. Therefore, further high-quality studies, comparing responses in subjects with symptoms of LBP and LSCS to matched healthy subjects, are needed to establish a strong correlation between imaging findings and patients' symptoms. Our review also did not yield any results on many quantitative measurements such as

the effects of loading on the disc fluid content by either measuring MR signal intensity or diffusion coefficients. Further, high-quality studies are needed to quantify the effects of loading on additional quantitative MRI measurements and to establish a strong correlation between these measurement constructs and patients' symptoms.

# CHAPTER 4 IS THE LOCATION OF THE SIGNAL INTENSITY WEIGHTED CENTROID A RELIABLE MEASUREMENT OF FLUID DISPLACEMENT WITHIN THE DISC?

Vahid Abdollah; Eric C Parent; Michele C Battié

**Abstract**

Degenerated discs have shorter $T_2$-relaxation time and lower MR signal. The location of the signal-intensity-weighted-centroid reflects the water distribution within a region-of-interest. This study compared the reliability of the location of the signal-intensity-weighted-centroid to mean signal intensity and area measurements. L4-5 and L5-S1 discs were measured on 43 mid-sagittal $T_2$-weighted 3T MRI images in adults with back pain. One rater analysed images twice and another once, blinded to measurements. Discs were semi-automatically segmented into a whole disc, nucleus, and the anterior and posterior annulus. The coordinates of the signal-intensity-weighted-centroid for all regions demonstrated excellent intraclass correlation coefficients for intra- (0.99-1.00) and inter-rater reliability (0.97-1.00). The standard errors of measurement for the vertical-coordinates of the signal-intensity-weighted-centroid for all region of interests were zero at both levels and 0 to 2.2 mm for horizontal-coordinates. The mean signal intensity and area for the whole disc and nucleus presented excellent intra-rater reliability with intraclass correlation coefficients from 0.93 to 1.00, and 0.92 to 1.00 for inter-rater reliability. Mean signal intensity and area had lower reliability for annulus region-of-interests, with intra-rater intraclass-correlation-coefficient from 0.5 to 0.76 and inter-rater from 0.33 to 0.58. The location of the signal-intensity-weighted-centroid is a reliable biomarker for investigating the effects of disc interventions.

**Keywords:** MRI; signal intensity; weighted centroid; segmentation; reliability; intervertebral disc degeneration

## Introduction

Low back pain (LBP) is a major cause of discomfort and disability among adults globally [1]. Intervertebral disc degeneration has been implicated as a major factor in the development of LBP [5–7]. Diagnostic imaging modalities, including plain x-ray and magnetic resonance imaging (MRI), are used for diagnosing the pathology underlying LBP [9], with MRI preferred due to concerns about radiation exposure and better contrast and visualisation of soft tissue [10].

In clinical practice, both $T_1$- and $T_2$-weighted images are routinely acquired. The former is typically used for diagnosis of inflammation and infections, and the latter for evaluating disc morphology [175], and hydration, as well as characterising pathology [16, 113, 176]. Long $T_2$ relaxation times are correlated with higher water content reflected by brighter pixels on T2-weighted images. Regions, such as the nucleus, with higher water content, demonstrate higher $T_2$ values and signal intensity, whereas the annulus, due to a high concentration of fibres, demonstrates lower $T_2$ values and signal intensity on $T_2$-weighted images [20]. $T_2$ values also provide information about the arrangement of collagen fibres and the anisotropy of free water movements within the disc both associated with degeneration status [177, 178].

A degenerated disc is a desiccated, fragmented structure [72], with a lower $T_2$ relaxation time reflected by a darker signal on $T_2$-weighted MR images and showing marked height decrease in later phases of degeneration [16, 72, 113, 176]. Degeneration not only affects the quantity, but also the distribution of fluid in the disc [5]. Various qualitative rating algorithms, including Thompson and Pfirrmann's, based on disc morphology and MR signal intensity characteristics, have been proposed for the MRI assessment of intervertebral disc degeneration [16, 113]. However, the subjective nature of these methods can lead to variations and uncertainty in evaluating the severity of disc degeneration. Moreover, classifying a continuous process, such as degeneration, using only a few discrete grades limits the ability to assess the progression of degenerative changes and response to therapeutic intervention. The ability to detect early degeneration is particularly limited, as at this stage degeneration may not yet be accompanied by any morphological changes. Quantitative assessment of intervertebral disc degeneration based on the mean signal intensity within a region of interest offers an attractive alternative but the mean is dependent on the area and location of the region of interest (ROI) selected.

The location of the MR signal intensity weighted centroid (SIWC) is the location of the arithmetic mean ("average") of the signal intensity of all the pixels in a region of interest. The SIWC, therefore, reflects the distribution of fluid within a structure. The distribution of fluid, as indicated by the location of the SIWC, is likely systematically different in various loading conditions and discs with different degeneration severities. The SIWC could provide a new criterion for quantifying the effects of different loading conditions and intervertebral disc degeneration. To our knowledge, this novel measurement has not been assessed for reliability in both anterior-posterior and cephalo-caudal directions, particularly when used to assess degenerated discs in patients with low back pain. The aims of this study were therefore to determine the intra- and inter-rater reliability of the geometric and signal intensity weighted centroid and compared to the mean signal intensity and area of the whole disc, nucleus and annulus using a semi-automatic signal intensity-based segmentation technique.

**Methods**

***Subjects.*** The baseline mid-sagittal $T_2$-weighted MR images from 43 volunteers with low back pain recruited for a previous study on extension exercise were used in this reliability study [179]. The participants' mean age was 43±13 years, and all had back pain with or without leg pain and a minimum 25% Oswestry Disability Index score. The present study was approved by the University of Alberta Health Research Ethics Review Board.

***MRI image procedures.*** Sagittal $T_2$-weighted images were obtained in the supine position, while knees were supported by a standard imaging pillow, using a 3T Siemens TrioTim MRI scanner (Siemens Healthcare, Erlangen, Germany). Acquisition parameters were TR 1600, TE 254, 2 averages, slice thickness 2 mm, field of view of 320×320 mm; and a matrix size of 640×640 pixels. The mid-sagittal slice was selected by the technician based on the localizer sequence illustrating the presence of the spinous processes and with a clear demarcation of the spinal cord on the anatomical $T_2$-weighted images [180]. In cases where a scoliotic curve was present, the localizer was placed to ensure capturing the L4-5 disc and spinous processes.

***Image analysis and processing.*** To describe the sample, all lumbar discs were classified as proposed by Pfirrmann et al [16] through consensus by two raters. Image post-processing was carried out offline and took approximately 90 seconds for each disc. One rater analysed images twice and another once, while blinded to prior and each other's measurements with a one-week

interval between measurements. We developed a custom MATLAB® algorithm (MathWorks, Natick, MA) using signal variations in the neighbouring structures to semi-automatically segment the L4-5 and L5-S1 discs and then automatically segment each disc into four ROIs: 1) whole disc, 2) nucleus region, 3) anterior and 4) posterior annulus regions.

***Disc segmentation.*** Segmentation began by drawing a line through the vertebral body along the superior endplate of the disc below (Figure 5a). By searching vertically below each pixel along the drawn line the program then determined the pixels with maximum signal intensity difference as the upper boundary of the disc (Figure 5b and c). A similar strategy was used for segmentation of the inferior boundary of the disc, searching vertically above the line drawn by the user (Figure 5d). To locate the anterior and posterior corners of the vertebrae, 20 pixels adjacent to the initial endpoints were scanned further anteriorly or posteriorly. The pixel with maximum signal intensity difference in the anterior-posterior direction was then selected as the vertebra endpoint on each side.

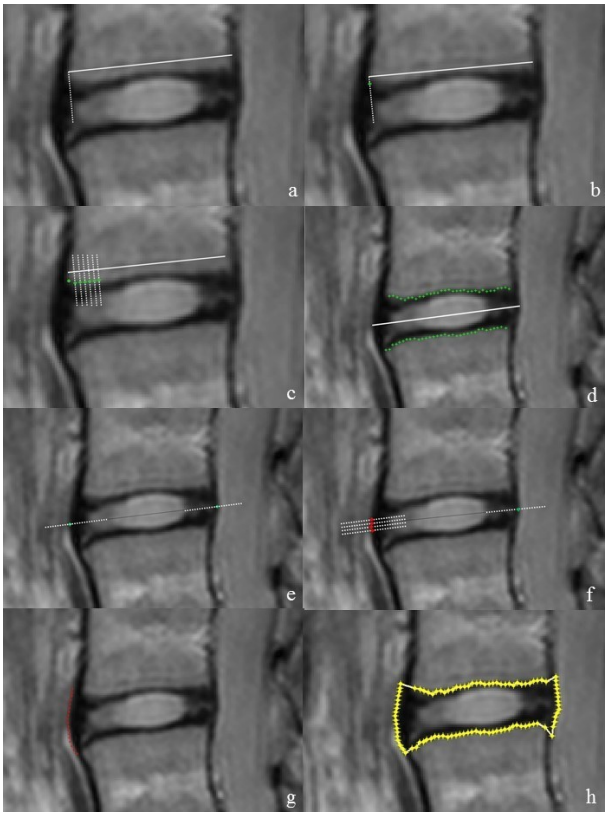


Figure 5. Disc segmentation process, a:) drawing the first tangential line to the upper disc-vertebra boundary of the disc of interest and scanning 20 pixels below the left endpoint of the line, b:) determining the first point of the disc-bone boundary, c:) determining the upper

boundary of disc pixel by pixel along the upper disc-vertebra boundary, d:) generating the disc bisector line between the upper and lower boundary of the disc and determining the real vertebra endpoint, e:) determining anterior and posterior endpoint of disc along the DBL, f: determining the anterior boundary of disc pixel by pixel along the anterior contour, g:) anterior boundary of disc, h:) a fully segmented disc

Users were able to modify the location of individual disc boundary pixels in the case of outlier errors. After determining the slope of the best-fit lines passing through the points of the upper and lower disc-endplate boundaries, a disc bisector line with a slope corresponding to the average of these two lines was then automatically generated dividing the disc into superior and inferior halves (Figure 5e). The anterior and posterior boundary of the disc was then segmented by locating pixels with maximum signal intensity difference along lines drawn parallel to the disc bisector line (Figure 5f, g, and h).

The disc was segmented into further regions of interest: the nucleus and anterior and posterior annular regions as demonstrated in Figure 6. Three parameters were calculated for each ROI: 1) mean signal intensity, 2) area, and 3) X and Y coordinates of the SIWC obtained using Equation 1 [181].

Equation 1 Mathematical description of SIWC

$$X_{SIWC} = \frac{\sum_{i=1}^{n} X_i SI_i}{\sum_{i}^{n} SI_i} \qquad Y_{SIWC} = \frac{\sum_{i=1}^{n} Y_i SI_i}{\sum_{i}^{n} SI_i}$$

Where $X_i$, $Y_i$ and $SI_i$ are the X (horizontal) and Y (vertical) coordinates and the signal intensity of each pixel, and where *n* is the total number of pixels in the ROI.

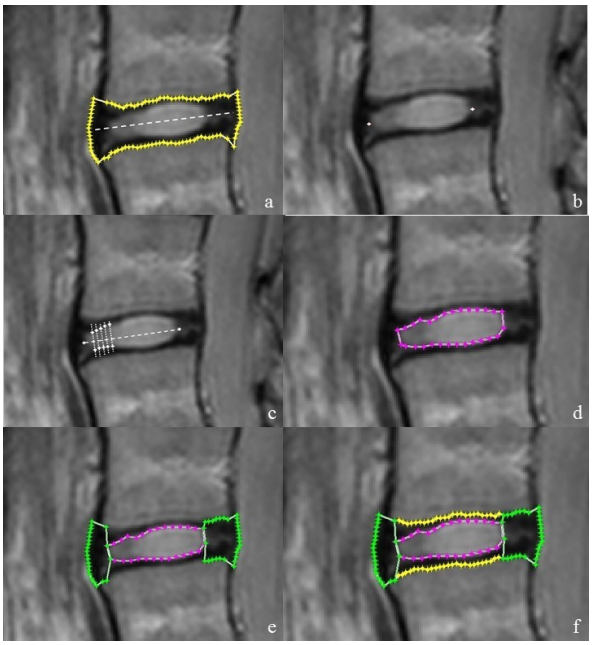


Figure 6 Segmentation of the regions of interest within the disc a:) generating the disc bisector line, b:)determining the anterior and posterior endpoints of the nucleus, c:) determining the upper and lower boundaries of the nucleus, d:) segmented nucleus, e:) segmenting the anterior and posterior annulus f: a fully segmented disc

***Data analysis.*** The mean and standard deviation of each parameter for the first set of measurements by the first rater were reported for the whole sample at each disc level. Intra-rater reliability was assessed for each region of interest and measurement with intra-class correlation coefficients ($ICC_{(3.1)}$) using the data recorded in the two sets of measurements by the first rater. Inter-rater reliability was assessed with an $ICC_{(2.1)}$ model comparing the first measurements by each rater. Standard errors of measurements (SEM) were calculated to estimate measurement error in the measurement unit ($SEM = SD\sqrt{1 - ICC}$, where SD is the standard deviation of the first measurement). ICC estimates higher than 0.75 were considered adequate for research conducted at a group level, and estimates of 0.9 and higher were considered adequate for clinical use at the individual level [152]. All statistical analyses were carried out using SPSS® statistical software, version 22.0 (IBM Corp. Armonk, NY., USA) with the level of statistical significance set at 0.05.

**Results**

A total of 43 L4-5 and 43 L5-S1 discs were analysed. The sample included no discs with Pfirrmann grade I degeneration (normal), and no more than three discs per level were rated as grade V (advanced degeneration with a collapsed disc space) (table 9).

The intra-rater ($ICC_{(3.1)}$) and inter-rater ($ICC_{(2.1)}$) reliability coefficients for the horizontal and vertical coordinates of the SIWC for all ROIs at both disc levels ranged from 0.97 to 1 (table 10). The intra-rater and inter-rater SEMs for the vertical coordinates of the SIWC for all ROIs were zero at both disc levels and varied from zero to 2.2 mm for the horizontal coordinates (table 10). Intra-rater ($ICC_{(3.1)}$) and inter-rater reliability ($ICC_{(2.1)}$) estimates for the average signal intensity and area for the L4-5 and L5-S1 discs ranged from 0.88 to 1.00 for the whole disc and nucleus region, and from 0.27 to 0.78, for the anterior and posterior annulus regions for the two disc levels assessed (table 10). The intra- and inter-rater SEMs of the measurements for the whole disc varied between 0 and 10% of the mean measurements. The intra- and inter-rater SEM of the measurements for the anterior and posterior annulus varied between 0 and 31% of the mean measurements.

**Table 9. Distribution of Pfirrmann's disc degeneration scores at L4-5 and L5-S1.**

| Pfirrmann's Grades | Description | L4-5 | L5-S1 |
|---|---|---|---|
| Grade I | Bright hyperintense homogeneous disc with a normal height | 0[I] | 0 |
| Grade II | Hyperintense inhomogeneous disc, with a clear distinction between nucleus and annulus and a normal height with or without horizontal grey bands. | 8 | 11 |
| Grade III | Inhomogeneous grey disc, with an unclear distinction between nucleus and annulus and a normal or slightly decreased disc height | 17 | 18 |
| Grade IV | Inhomogeneous hypointense dark grey disc, with a normal to moderately decreased disc height | 16 | 11 |
| Grade V | Inhomogeneous hypointense dark disc, with a collapsed disc height | 2 | 3 |

[I] Number of discs

**Table 10. Intra- and inter-rater reliability coefficients (ICC) and standard error of measurement (SEM) for measurements of the different ROIs at both disc levels**

| Regions of interest | Measurements | | Mean | Intra-rater | | | Inter-rater | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | $ICC_{3,1}$ | 95 % CI | SEM | $ICC_{2,1}$ | 95% CI | SEM |
| L4-5 | | | | | | | | | |
| Whole Disc | MSI | | 30.6 (17.5) | 0.99 | 0.98-0.99 | 1.7 | 0.99 | 0.98-0.99 | 1.7 |
| | Area ($mm^2$) | | 178.0 (71.6) | 0.93 | 0.87-0.96 | 18.9 | 0.93 | 0.87-0.96 | 18.9 |
| | SIWC (mm) | X | 150.0 (26.0) | 1.00 | 1.00-1.00 | 0.0 | 1.00 | 1.00-1.00 | 0.0 |
| | | Y | 229.0 (31.0) | 1.00 | 1.00-1.00 | 0.0 | 1.00 | 1.00-1.00 | 0.0 |
| Nucleus | MSI | | 21.7 (7.8) | 0.99 | 0.99-1.00 | 0.8 | 0.97 | 0.94-0.98 | 1.3 |
| | Area ($mm^2$) | | 394.3 (98.3) | 0.99 | 0.98-0.99 | 9.8 | 0.88 | 0.79-0.94 | 34.0 |
| | SIWC (mm) | X | 150.5 (26.0) | 1.00 | 1.00-1.00 | 0.0 | 0.97 | 0.95-0.98 | 2.2 |
| | | Y | 114.5 (31.0) | 1.00 | 1.00-1.00 | 0.0 | 1.00 | 1.00-1.00 | 0.0 |
| Anterior Annulus | MSI | | 9.3 (2.1) | 0.76 | 0.60-0.86 | 1.0 | 0.55 | 0.29-0.73 | 1.4 |
| | Area ($mm^2$) | | 90.2 (34.7) | 0.58 | 0.33-0.75 | 22.5 | 0.33 | 0.03-0.58 | 28.4 |
| | SIWC (mm) | X | 134.0 (27.0) | 0.99 | 0.99-1.00 | 2.7 | 0.99 | 0.98-0.99 | 1.3 |
| | | Y | 233.0 (33.0) | 1.00 | 1.00-1.00 | 0.0 | 1.00 | 1.00-1.00 | 0.0 |
| Posterior Annulus | MSI | | 20.7 (5.7) | 0.65 | 0.43-0.79 | 3.4 | 0.27 | -0.04-0.53 | 4.9 |
| | Area ($mm^2$) | | 53.2 (18.8) | 0.31 | 0.01-0.56 | 15.6 | 0.50 | 0.23-0.70 | 13.3 |
| | SIWC (mm) | X | 165.5 (26.0) | 1.00 | 0.99-1.00 | 0.0 | 0.99 | 0.97-0.99 | 1.3 |
| | | Y | 226.5 (30.0) | 1.00 | 1.00-1.00 | 0.0 | 1.00 | 0.99-1.00 | 0.0 |
| **L5-S1** | | | | | | | | | |
| Whole Disc | MSI | | 24.5 (12.0) | 1.00 | 1.00-1.00 | 0.0 | 0.99 | 0.98-0.99 | 1.2 |
| | Area ($mm^2$) | | 352.8 (100.6) | 0.97 | 0.95-0.98 | 17.4 | 0.94 | 0.90-0.97 | 24.6 |

| Regions of interest | Measurements | | Mean | Intra-rater | | | Inter-rater | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | $ICC_{3,1}$ | 95 % CI | SEM | $ICC_{2,1}$ | 95% CI | SEM |
| | SIWC (mm) | X | 159.5 (28.0) | 1.00 | 1.00-1.00 | 0.0 | 0.99 | 0.99-1.00 | 1.4 |
| | | Y | 261.5 (30.0) | 1.00 | 1.00-1.00 | 0.0 | 1.00 | 1.00-1.00 | 0.0 |
| Nucleus | MSI | | 37.4 (26.0) | 1.00 | 0.99-1.00 | 0.0 | 1.00 | 1.00-1.00 | 0.0 |
| | Area ($mm^2$) | | 151.1 (65.2) | 0.97 | 0.94-0.98 | 11.3 | 0.97 | 0.95-0.99 | 11.3 |
| | SIWC (mm) | X | 160.0 (28.0) | 1.00 | 1.00-1.00 | 0.0 | 1.00 | 1.00-1.00 | 0.0 |
| | | Y | 261.0 (30.0) | 1.00 | 1.00-1.00 | 0.0 | 1.00 | 1.00-1.00 | 0.0 |
| Anterior Annulus | MSI | | 10.1 (2.8) | 0.75 | 0.54-0.84 | 1.4 | 0.78 | 0.63-0.88 | 1.3 |
| | Area ($mm^2$) | | 158.3 (31.2) | 0.72 | 0.54-0.84 | 16.5 | 0.58 | 0.33-0.75 | 20.2 |
| | SIWC (mm) | X | 147.0 (30.0) | 1.00 | 0.99-1.00 | 0.0 | 0.99 | 0.99-1.00 | 1.4 |
| | | Y | 269.0 (31.0) | 1.00 | 1.00-1.00 | 0.0 | 1.00 | 0.99-1.00 | 0.0 |
| Posterior Annulus | MSI | | 18.9 (4.7) | 0.50 | 0.23-0.70 | 3.3 | 0.52 | 0.25-0.71 | 3.3 |
| | Area ($mm^2$) | | 42.1 (14.7) | 0.62 | 0.38-0.78 | 9.1 | 0.32 | 0.03-0.58 | 12.1 |
| | SIWC (mm) | X | 254.0 (29.0) | 1.00 | 0.99-1.00 | 0.0 | 1.00 | 1.00-1.00 | 0.0 |
| | | Y | 173.5 (27.0) | 1.00 | 0.99-1.00 | 0.0 | 1.00 | 0.99-1.00 | 0.0 |

Abbreviations: SIWC: signal intensity weighted centroid; MSI: mean signal intensity; SEM: standard error of measurement; ICC: intraclass correlation coefficient; CI: confidence interval

**Discussion**

Measurements of the whole disc and nucleus area, mean signal intensity and location of the SIWC acquired using our semi-automatic computer-aided segmentation technique, within a sample of patients with chronic low back pain, were found to have high intra- and inter-rater reliability (ICC>0.90) suitable for research and clinical use. However, the reliability estimates of mean signal intensity and area suggested that the measurements of the annular regions are not suitable for clinical use. As hypothesised, the reliability of the location of the SIWC was found excellent for all regions. The location of the SIWC, therefore, showed excellent reliability for all ROIs even when results for the mean signal intensity or ROI area estimates showed poor reliability.

The location of the SIWC of a distribution of signal intensities in an area is the unique point where the weighted relative position of the distributed signal intensities sums to zero. It thus reflects the location where the distribution of signal intensities is balanced around the centroid. In the present study, the SIWC of all regions demonstrated excellent reliability and little if any measurement error for the horizontal and vertical coordinates. The SIWC is, therefore, suitable for research and clinical use. Périé and Curnier also reported good reliability for the nucleus weighted centroid but only reported data related to the horizontal direction [181]. Since loading reportedly affects the amount of fluid and its distribution pattern in many directions within the disc, tracking the weighted centroid pathway may demonstrate the effects of different loading conditions on the location of the highest concentrations of fluid within the whole disc and regionally, and may relate to disc health.

A limitation is that detection of the disc-vertebra boundary and vertebral level using our algorithm relies on user input for drawing lines to begin the segmentation of the upper and lower disc-vertebra boundaries for each disc, which increases the analysis time. Automating disc-vertebra numbering may be a solution but existing algorithms require determining image specific intensity or require manual selection of an ROI to begin segmentation. Therefore, these algorithms also require user input/time which ultimately also affects the segmentation results [182]. To our knowledge, no fully automated algorithms are yet available. Another limitation is that we analysed only the mid-sagittal scan. As a result, lateral portions of the disc were not included in this study. Pathologically relevant disc measurements may be observed in other planes. Our methodology would likely perform well on other slices and could be investigated in the future.

Our reliability estimates for measurements of the whole disc area were slightly lower than those from semi-automatic segmentation techniques proposed by Neubert et al (ICC=0.98-0.99) in seven asymptomatic subjects without a prior history of low back pain [183]. This may be because all subjects in the present study had a history of chronic low back pain, with the majority of discs graded as moderately to severely degenerated using Pfirrmann's classification. Degeneration increases the difficulty is determining the location of boundaries for the regions of interest.

The reliability estimates of the mean signal intensity and area of the annulus regions were lower than the whole disc and nucleus region, possibly due to difficulties in defining a clear boundary between the annulus regions and neighbouring tissues, particularly in severely degenerated discs. These difficulties affected the semi-automated segmentation results and required more manual adjustments of outlier points along the nucleus region boundaries thereby affecting the reliability. Point estimates for all reliability coefficients were lowest for the posterior annulus. In discs with herniation where the posterior annulus displays a higher intensity than normal, such as in the presence of high-intensity zones [184–186], it is challenging to determine a clear boundary between the posterior annulus and the cerebrospinal fluid. Further, in patients with severe degeneration with low annulus intensity and contours occupying the spinal canal and displacing cerebrospinal fluid and nerve roots, it is difficult to contrast the annulus from the nerve tissue or canal after signal intensity in the canal is reduced by displaced cerebrospinal fluid. It is not possible to distinguish the annulus from the adjacent ventral and dorsal ligamentous structures with $T_2$ based imaging techniques [187]. Therefore, portions of the anterior and posterior longitudinal ligaments were included in the disc measurements reported.

The location of the SIWC of different disc ROIs may be employed as a reliable measurement for the investigation of the effects of different loading conditions or therapeutic interventions. The measurement demonstrated high reliability for all regions of interest despite the increased difficulty in reliably segmenting the annulus regions.

**Acknowledgement**

Vahid Abdollah was a recipient of the University of Alberta Doctoral Recruitment Scholarship. All authors were recipients of a research grant provided by the Alberta Spine Foundation for the development of a program for automatic quantitative measurements of the

lumbar disc and vertebra signal intensity and morphology using $T_2$-weighted MR images. Data collection was as part of a study entitled: Creating a Prediction Rule to Identify Patients Likely to Respond to Extension-Oriented Exercises and Understanding the Mechanisms in Persons with Low Back Pain: A Study Using Clinical, MRI and Neuromuscular Assessment. conducted by J. Fritz (PI), A. Cottrell, B. Hayes, E. Parent, K. Sanders funded by a collaborative grant from the Center for Contemporary Rehabilitation Research, Education and Practice (A collaborative Between Rehabilitation Services, PM&R and College of Health) from December 2006 to 2007.

# CHAPTER 5 RELIABILITY AND VALIDITY OF LUMBAR DISC HEIGHT QUANTIFICATION METHODS ON MAGNETIC RESONANCE IMAGES

Vahid Abdollah; Eric C Parent; Michele C Battié

**Abstract**

**Study design:** Reliability and validity study

**Objective:** To evaluate the reliability and construct validity of disc height quantification methods on $T_2$-weighted MR images.

**Summary of background data:** The lumbar intervertebral disc (IVD) is the focus of extensive research, as low back pain (LBP) has often been attributed to IVD degeneration and pathology. Variations in disc height, as an important sign and consequence of disc degeneration, have been of particular interest. However, the choice of the most appropriate method to quantify disc height on spine imaging has remained a subject of controversy.

**Methods:** L4-5 and L5-S1 discs were measured on 43 mid-sagittal 3T MRI images of 22 subjects with back pain (43±13yrs). One rater measured twice and another once, while blinded to measurements. Discs were segmented semi-automatically. Disc heights were calculated with a disc area-based method, using 60%, 80% or 100% of disc width, as well as Hurxthal's, Dabbs' and combining these two methods. Reliability was estimated using intraclass correlation coefficients and standard error of measurement. Construct validity was assessed using correlation coefficients amongst disc height methods.

**Results:** The intra-rater reliability of the area-based disc height measurements ranged from ($ICC_{(3,1)}$) 0.84 to 0.99 with an inter-rater reliability of ($ICC_{(2,1)}$) 0.99. Measurements with the point-based disc height methods had lower intra-rater reliability ranging from 0.76 to 0.96 and inter-rater reliability from 0.84 to 0.98. Inter-rater standard error of measurement varied between 0.2 and 0.3mm for area-based methods and between 0.3 and 0.7mm for point-based methods. Excluding Dabbs' method, high correlation ($r>0.9$) was observed between methods.

**Conclusion:** Area-based height measurements using partial disc width demonstrated excellent reliability and construct validity, and were superior to point-based disc height quantification methods, particularly Dabbs' method.

**Keywords:** intervertebral disc; intervertebral disc degeneration; low back pain; computer aided measurement; intervertebral disc height; magnetic resonance imaging; $T_2$ weighted MR images; reproducibility of results; measurement reliability; validation studies

**Mini abstract**

The reliability of area-based methods of disc height measurement using 80% and 60% of disc width demonstrated excellent reliability, which was superior to using 100% of disc width and point-based methods. High construct validity was supported by correlations exceeding 0.9 for all pairs of measurements, excluding Dabbs' method.

## Introduction

The lumbar intervertebral disc (IVD) is the focus of extensive research, as low back pain (LBP) has often been attributed to IVD degeneration and pathology [188]. Defining degenerative spinal phenotypes on imaging, including disc height, and recommendations for standardised measurement are primary goals of the Degenerative Spinal Phenotypes Focus Group of the International Society for the Study of the Lumbar Spine. Although several stakeholders in spine research have initiated efforts to develop detailed definitions of lumbar disc pathology terms [189], to date there has been little effort to define disc height and standardise its measurement. Recently, the combined task forces of the North American Spine Society, the American Society of Spine Radiology and the American Society of Neuroradiology defined disc height as the distance between the planes of the endplates cranial and caudal to the disc [189]. They recommended that disc height be measured at the centre of the disc, not at the periphery [189], which would minimise the effects of osteophytes. However, the choice of the most appropriate measurement to quantify disc height on spine imaging remains a subject of controversy.

Among the quantitative height measurement methods used, Dabbs' method consists of averaging two lines representing disc height at the anterior and posterior corners of the disc [162, 190–194]. Hurxthal's 2$^{nd}$ method defines height as that measured at the midpoint of the disc [195]. Others have quantified disc height by combining and averaging Dabbs' and Hurxthal's 2$^{nd}$ method [196]. Finally, area-based methods define disc height as a ratio of the disc cross-sectional area in sagittal view over the anterior-posterior diameter of the disc [183], and have been introduced in MR imaging studies [197, 198], as well as earlier investigations using X-ray [191, 199, 200]. Area-based methods are able to capture almost all morphometric features of the disc in the sagittal plane. Traditionally, it had limited clinical application possibly because on x-ray based imaging techniques the anterior and posterior contours of the disc are not visible. Further, until improved access to computer-based measurements, this method may have been too time-consuming to quantify disc height on the MR images compared to point-based methods.

Identifying a reliable, valid method for measuring disc height for standard use would facilitate comparisons between studies and advance knowledge related to the determinants and effects of disc height and disc narrowing, as well as their clinical importance. Yet, to date, reported reliability estimates of various disc height measurements are variable and largely suboptimal.[194] Furthermore, we found no studies of the reliability and validity of the area-based method compared

to other disc height quantification methods. Thus, to guide recommendations for a standard disc height measurement method, the aim of this study was to evaluate the intra- and inter-rater reliability and construct validity of quantitative disc height measurements using various methods, including area-based methods using a semi-automatic signal intensity-based segmentation technique on $T_2$-weighted MR images in adults with chronic LBP.

**Materials and Methods**

***Participants.*** Forty-three mid-sagittal $T_2$-weighted MR images from 22 volunteers with LBP recruited for a previous study on extension exercise were used [179]. Mean age was 43±13 years. All participants had back pain with or without leg pain and a minimum 25% Oswestry Disability Index score. The degree of disc degeneration was assessed for each disc using the 5-point scale proposed by Pfirrmann et al [16] through consensus by two raters. The present study was approved by the Health Research Ethics Board of the University of Alberta.

***MRI image procedures.*** Mid-sagittal $T_2$-weighted MR images were obtained in the supine position before and after a bout prone press-up extension exercises, while knees were supported by a standard imaging pillow, using a 3T Siemens TrioTim MRI scanner (Siemens Healthcare, Erlangen, Germany). Acquisition parameters were TR 1600, TE 254, 2 averages, slice thickness 2 mm, field of view of 320×320 mm; and matrix size of 640×640 pixels. The mid-sagittal slices were selected based on the localizer sequence illustrating the presence of the spinous processes and with a clear demarcation of the spinal cord on the anatomical $T_2$-weighted images [180]. In cases where a scoliotic curve was present, the localizer was placed to ensure capturing the L4-5 disc and spinous processes.

***Image analysis, processing and disc height measurements.*** One evaluator obtained the disc height measurements twice, with a one-week interval between measurements, and another once, while blinded to all prior measurements. The measurements were carried out using a computer-aided measurement (CAM) program that segments discs based on signal variations between the disc and neighbouring structures (see text document, Supplemental Digital Content 1, detailing the process of segmenting the discs and determining the points from which the various disc height measurements were calculated). Then, disc height was measured automatically using Dabbs method [191], Hurxthal's 2$^{nd}$ method [195], a combination of both [196] (Figure 1), and area-based methods [197], using different disc width percentages (Figure 1). For the area-based

methods, disc height measurements were obtained by dividing the disc area centred at the middle of the disc bisector line by the entire length of the disc bisector line (100% of the anterior-posterior disc dimension), as well as using 80% of the length of the disc bisector line and corresponding disc area, and 60% (Figure 1).

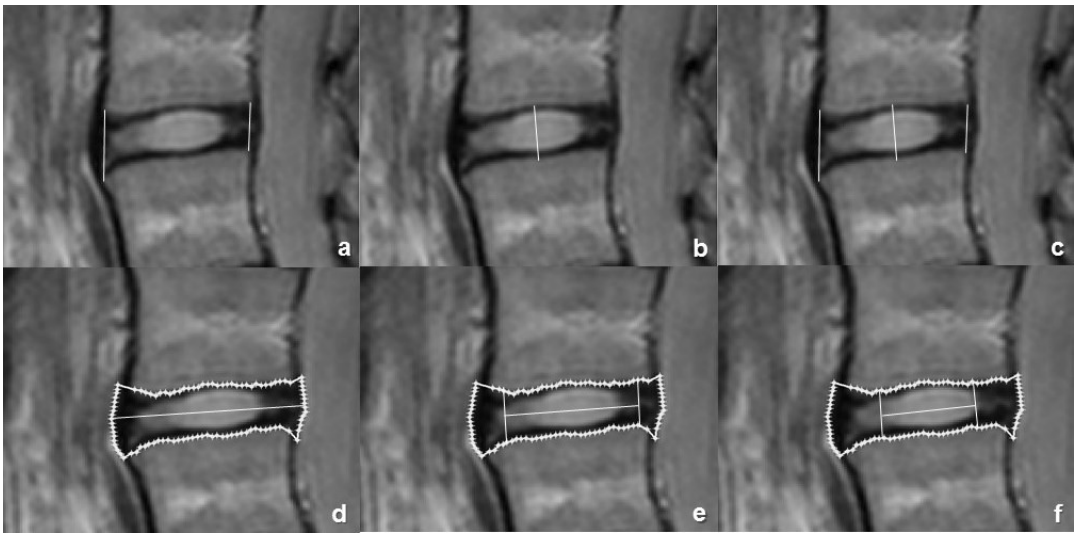


Figure 7. Measuring disc height 1a:) Dabbs' (average of anterior and posterior disc heights), 1b:) Hurxtal's, 1c:) combined method (average of all three), 1d:) area-based method using 100% of the disc width, 1e:) using 80% of the disc width, 1f:) using 60% of the disc width

***Data analysis.*** The mean and standard deviation of each parameter for the second measurements by the first rater were reported for the whole sample at each disc level. Intra-rater reliability was assessed for each height measurement using an intra-class correlation coefficient ($ICC_{(3,1)}$) using the two sets of measurements by the first rater. The inter-rater reliability was assessed with an $ICC_{(2,1)}$ model comparing measurements of each rater. Standard errors of measurements (SEM) were calculated to provide estimates of measurement error in the measurement unit ( $SEM = SD\sqrt{1 - ICC}$ , where SD is the standard deviation of the measurements). ICC estimates higher than 0.75 were considered adequate for research conducted at a group level, and estimates of 0.9 and higher were considered adequate for clinical use at the individual level [152].

Pearson's correlation coefficients were computed for each pair of the disc height measurements to examine the construct validity of the measurements. Correlations exceeding 0.5 have been suggested as a minimum to provide evidence of construct validity (convergent validity)

[201]. Statistical analyses were done using SPSS®, version 22.0 (IBM Corp. Armonk, NY., USA) using a level of statistical significance set at 0.05.

**Results**

A total of 43 discs were analysed at each spinal level (L4-5 and L5-S1). The sample included no discs with Pfirrmann grade I degeneration (normal), and no more than three discs at either level were rated as grade V (advanced degeneration with a collapsed disc space) (table 11).

**Table 11. Distribution of Pfirrmann's disc degeneration scores at L4-5 and L5-S1.**

| Pfirrmann's Grades | Description | L4-5 | L5-S1 |
|---|---|---|---|
| Grade I | Bright hyperintense homogeneous disc with a normal height | 0[I] | 0 |
| Grade II | Hyperintense inhomogeneous disc, with a clear distinction between nucleus and annulus and a normal height with or without horizontal grey bands. | 8 | 11 |
| Grade III | Inhomogeneous grey disc, with an unclear distinction between nucleus and annulus and a normal or slightly decreased disc height | 17 | 18 |
| Grade IV | Inhomogeneous hypointense dark grey disc, with a normal to moderately decreased disc height | 16 | 11 |
| Grade V | Inhomogeneous hypointense dark disc, with a collapsed disc height | 2 | 3 |

The intra-rater reliability ($ICC_{(3,1)}$) of the area-based disc height measurements using the full disc width was 0.84 at the L4-5 level and 0.90 at L5-S1, but was consistently more than 0.91 at both levels when using 60 or 80 percent of the disc width (table 12). Intra-rater reliability of the area-based disc height measurements using 60 or 80% of the disc width was significantly higher at L4-L5 than at L5-S1. At the L4-5 level, the intra-rater reliability of partial disc-width area-based methods was significantly greater than all other methods. The intra-rater reliability estimates of the non-area-based methods ranged from $ICC_{(3,1)}$ 0.86 to 0.96 at L4-5 and from 0.76 to 0.83 at L5-S1.

Inter-rater reliability coefficients ($ICC_{(2,1)}$) of all the area-based methods were 0.99. Inter-rater reliability coefficients were lower for the other height measurement methods, ranging from

[I] Number of discs

0.84 to 0.98 at both spinal levels (table 12). Among the non-area-based measurements, except for the interrater reliability at L5-S1, the point estimates of the intra- and inter-rater reliability for Hurxthal's method were higher than for both Dabbs' and the combined method at both levels (table 12). The area-based disc height measurements using either 80% or 60% of the anterior-posterior disc dimensions also had lower intra- and inter-rater SEM (ranging between 0.2 and 0.7 mm) compared to the other height measurements at both levels (table 12).

Pearson's correlation coefficients of the various disc height measurements with one another ranged from 0.75 to 1.00 (table 13). The highest correlations were observed between the area-based methods using 100% and 80%-disc width, and the lowest correlations were observed between Dabbs' and Hurxthal's 2$^{nd}$ method. All correlations between measurements other than Dabbs' method exceeded 0.90.

**Table 12 Intra- and inter-rater reliability and standard error of measurement for different disc height quantification methods at L4-5 and L5-S1**

| Measurement | | Mean ± SD (mm) | Intra-rater | | | Inter-rater | | |
|---|---|---|---|---|---|---|---|---|
| | | | ICC | 95% CI | SEM (mm) | ICC | 95% CI | SEM (mm) |
| 100% of DWU | L4-5 | 9.8±2.6 | 0.84 | 0.72-0.91 | 1.0 | 0.99 | 0.98-0.99 | 0.3 |
| | L5-S1 | 9.3±2.4 | 0.90 | 0.83-0.94 | 0.8 | 0.99 | 0.99-1.00 | 0.2 |
| 80% of DWU | L4-5 | 10.1±2.3 | 0.99 | 0.99-1.00 | 0.2 | 0.99 | 0.99-1.00 | 0.2 |
| | L5-S1 | 9.5±2.4 | 0.91 | 0.84-0.95 | 0.7 | 0.99 | 0.99-1.00 | 0.2 |
| 60% of DWU | L4-5 | 10.4±2.3 | 0.99 | 0.99-1.00 | 0.2 | 0.99 | 0.99-1.00 | 0.2 |
| | L5-S1 | 9.8±2.5 | 0.91 | 0.83-0.95 | 0.7 | 0.99 | 0.99-1.00 | 0.2 |
| Hurxthal's | L4-5 | 11.0±2.4 | 0.96 | 0.93-0.98 | 0.5 | 0.98 | 0.96-0.99 | 0.3 |
| | L5-S1 | 10.0±2.6 | 0.83 | 0.72-0.91 | 1.1 | 0.93 | 0.88-0.96 | 0.7 |
| Dabbs's | L4-5 | 10.3±1.9 | 0.86 | 0.76-0.92 | 0.7 | 0.84 | 0.71-0.91 | 0.8 |
| | L5-S1 | 9.5±2.4 | 0.76 | 0.60-0.86 | 1.2 | 0.95 | 0.90-0.97 | 0.5 |
| Combined | L4-5 | 10.5±2.0 | 0.92 | 0.86-0.96 | 0.6 | 0.93 | 0.88-0.96 | 0.5 |
| | L5-S1 | 9.8±2.4 | 0.80 | 0.67-0.89 | 1.1 | 0.96 | 0.93-0.98 | 0.5 |

Abbreviations: DWU: disc width used; SEM: standard error of measurement; ICC: intraclass correlation coefficient; CI: confidence interval

**Table 13 Pearson's Correlation coefficients of the construct validity of different disc height quantification methods used at L4-5 and L5-S1**

| Disc Level | Method | 100% of DWU | 80% of DWU | 60% of DWU | Hurxthal | Dabbs | Combined |
|---|---|---|---|---|---|---|---|
| L4-5 | 100% of DW[I]U | 1.00 | | | | | |
| | 80% of DWU | 1.00** | 1.00 | | | | |
| | 60% of DWU | 0.99** | 1.00** | 1.00 | | | |
| | Hurxthal | 0.93** | 0.93** | 0.94** | 1.00 | | |
| | Dabbs | 0.83** | 0.80** | 0.78** | 0.75** | 1.00 | |
| | Combined | 0.93** | 0.91** | 0.90** | 0.90** | 0.96** | 1.00 |
| L5-S1 | 100% of DWU | 1.00 | | | | | |
| | 80% of DWU | 1.00** | 1.00 | | | | |
| | 60% of DWU | 0.98** | 0.99** | 1.00 | | | |
| | Hurxthal | 0.93** | 0.94** | 0.94** | 1.00 | | |
| | Dabbs | 0.91** | 0.89** | 0.86** | 0.80** | 1.00 | |
| | Combined | 0.97** | 0.95** | 0.94** | 0.92** | 0.97** | 1.00 |

Abbreviations: DWU: disc width used

[I] Disc width used

**Discussion**

The various methods of disc height measurements using semi-automatic, computer-aided segmentation included in the current study yielded measurements of high intra- and inter-rater reliability. As expected, based on the rationale for proposing them [199], the area-based methods of disc height measurement demonstrated excellent intra- and inter-rater reliability (ICC >0.90) when using 60% or 80% of disc width. These measurements also demonstrated the lowest measurement error, with SEM estimates ranging from 0.2 to 0.7 mm, showing promise for detecting changes beyond measurement error in longitudinal applications. The small confidence intervals for the area-based methods also indicate that the methods are adequate to demonstrate significant changes even in a small sample size. As highly degenerated discs, as those included in the current sample, are generally more difficult to measure compared to less degenerated discs with clear borders, the present study likely represents a worst-case scenario with respect to reliability estimates.

The area-based method includes almost all morphometric features of the disc in the sagittal plane and thus is less influenced by variations in identifying only a few specific anatomical landmarks. The lower reliability when using 100% of the disc width, as compared to using 80% or 60%, may be due to the relatively indistinct anterior and posterior edges of the vertebra-disc interface where osteophytes are often present. Excluding the disc periphery also eliminates unconfined disc areas located beyond the vertebral margins, which may reflect bulging or herniated material with contours harder to detect reproducibly.

The reliability estimates for Dabbs' method using semi-automated segmentation were comparable to prior estimates for measurements carried out manually [194, 202, 203], and semi-automatically [194], although measurement reliability has varied substantially when considering all studies on this method [162, 183, 193, 194]. In the present study, as the comparison of methods was based on measurements acquired using a CAM, the findings may not be fully representative of how the measures would compare to one another if measurements were obtained manually. Current evidence suggests a tendency for higher reliability when disc height is measured using a CAM than manually [183, 194]. Measurement reliability using Dabbs' method is influenced by the examiner's perception of the location of the vertebral corners at the disc-vertebra interface. The presence of osteophytes or inadequate contrast between hard and soft tissue can make corner detection particularly difficult in moderately to severely degenerated spines. Furthermore, the

presence or absence of endplate irregularities or Schmorl's nodes may affect measurement outcomes. Dabbs' method, which was originally employed for X-ray based imaging techniques, also overlooks the increasingly concave contour of the endplates with degeneration [51, 204]. Drawing the lines at the middle third of the disc has been suggested as a solution for this problem [162]. This may minimise the influence of osteophytes and central concavity, but it still overlooks the overall contour of the disc. It may also yield higher disc height measurements in discs with severely concave endplates. The distortion-corrected radiographic analysis method, which is a modification of Dabbs' method to overcome imaging distortions, suffers from the same limitations [205].

While our results for Dabbs' method were comparable with Pearson et al.'s ICC (0.67-0.84),[194] we found lower reliability for the method than Neubert et al. using a CAM (ICC=0.98-0.99) [183], and Edmondston et al. (ICC=0.94) obtaining measurements manually [162]. This difference may be due to their assessing young subjects with more clearly defined interfaces between vertebrae and discs, whereas in the present study the majority of the discs presented with Pfirrmann grade III or IV disc degeneration. The algorithm proposed by Neubert et al. also employed an active shape model technique, which looks for strong edges or contrasting tissues [183].

There is a limited number of studies investigating CAM for measurements of disc height[183, 190, 194] and we found only one study investigating concurrent validity of computerised height measurements compared to manual measurements [183]. The result indicated significant differences in mean disc height by the method used and by spinal level, however, Pearson's correlation coefficients of all height quantification methods used at both levels were higher than 0.75 [183], indicating a high correlation among methods used and supporting the construct validity of all [152].

Area-based disc height quantification methods demonstrated high reliability and encompass all relevant morphological features of the disc in the sagittal plane, more fully representing overall disc height, particularly when the adjacent endplates are not flat. Our results suggest that the area-based methods provide better reliability than point-based methods, particularly Dabbs' method. Area-based height quantification method using 60% or 80% of the disc width demonstrated excellent reliability, superior to that of traditional point-based methods that rely on precisely identifying the location of the vertebral corners. High construct validity was

supported for all methods, as indicated from high correlations among the methods used, with all correlations exceeding 0.9, except for Dabbs' method.

**Acknowledgement**

Vahid Abdollah was a recipient of the University of Alberta Doctoral Recruitment Scholarship. All authors were recipients of a research grant provided by the Alberta Spine Foundation for the development of a program for automatic quantitative measurements of the lumbar disc and vertebra signal intensity and morphology using $T_2$-weighted MR images. Data collection was as part of a study entitled: Creating a Prediction Rule to Identify Patients Likely to Respond to Extension-Oriented Exercises and Understanding the Mechanisms in Persons with Low Back Pain: A Study Using Clinical, MRI and Neuromuscular Assessment. conducted by J. Fritz (PI), A. Cottrell, B. Hayes, E. Parent, K. Sanders funded by a collaborative grant from the Center for Contemporary Rehabilitation Research, Education and Practice (A collaborative Between Rehabilitation Services, PM&R and College of Health) from December 2006 to 2007.

# CHAPTER 6 EVALUATION OF THE EFFECTS OF EXTENSION EXERCISES ON DISC FLUID USING MR IMAGES

**Vahid Abdollah; Eric C Parent; Michele C Battié**

**Abstract**

**Study Design:** Experimental

**Level of Evidence:** Four

**Background:** It has been hypothesised that McKenzie prone press-up exercises reduce intradiscal pressure, allowing fluid to be reabsorbed into the disc, which could improve the internal stability and local chemical milieu of the disc, potentially reducing symptoms.

**Objectives:** To investigate the immediate effects of prone press-up extension exercises on lumbar disc fluid content and movement.

**Methods:** Twenty-two volunteers with low back pain had mid-sagittal $T_2$-weighted MR images of their lumbar spines obtained before and immediately after performing standard extension exercises. The whole disc and nucleus region of the L4-5 and L5-S1 discs were then segmented from adjacent tissues, and their mean signal intensity (MSI) and signal intensity weighted centroid (SIWC) were measured to capture disc fluid content and displacement. T-tests were used for before and after comparisons.

**Results:** There were no significant differences between the MSI and the vertical position of the SIWC of the whole disc before and after extension at either disc level (ES: -0.23 to 0.09). There was a significant anterior displacement (0.1±5.4 mm) of the location of the SIWC of the disc after extension exercise at the L4-5 level (ES:-0.22), but not at L5-S1 (ES: 0.00) or at either level for the nucleus region (ES: -0.06;0.16).

**Conclusion:** Little evidence was found supporting the hypothesis that prone press-up exercises affect disc fluid content and distribution. Novel parameters reflecting fluid distribution detected similar or larger effects of the extension than MSI. If such exercises are effective in reducing symptoms, it is likely through other mechanisms than by changing fluid content.

**Keywords:** intervertebral disc degeneration; low back pain; $T_2$ weighted MR images; exercise therapy

## Introduction

The lumbar intervertebral disc (IVD) is the largest avascular tissue of the human body and it relies largely on diffusion from adjacent cartilaginous endplates for its nutrition. The IVD's metabolic milieu is linked with this diffusion, as are variations in loading tolerance within the IVD [108, 206–208]. When an external load bigger than the internal osmotic pressure is applied to a disc, the disc begins to lose fluids until an equilibrium is reached between the osmotic pressure and external load [209]. The fluid content of the IVDs is therefore not constant and is a function of the interaction between two pressures: swelling pressure due to the presence of proteoglycans that absorb fluid and mechanical pressure due to body weight, muscle forces, ligament tension and external loads, which tends to expel fluid from the disc.

Disc degeneration likely alters fluid diffusion [5, 108, 210], reducing the internal stability of the IVD and lessening its load-bearing capacity, which may trigger symptoms during activity [207, 211]. Interventions that can increase fluid content of the disc are thought to improve the internal stability and local chemical milieu, potentially leading to a meaningful reduction of symptoms [212]. It has been hypothesised that, during prone press-up exercises, intradiscal pressure is reduced by shifting forces posteriorly to the zygapophyseal joints, which enhances the reabsorption of fluid or promotes fluid redistribution within the disc [135].

Quantification of diffusion of fluid within the disc in vivo has been difficult, mostly due to technical limitations. Therefore, the extent to which various interventions affect fluid content and distribution is unknown. The effects of different loading conditions on the fluid content of the IVD can be studied using mean signal intensity (MSI) on $T_2$-weighted magnetic resonance imaging (MRI). An increase in MSI can be attributed to an increase in the fluid content of the disc [34]. The MSI is, however, a scalar quantity, which does not reflect the direction of the fluid movement.

The MRI signal intensity weighted centroid (SIWC) is the location of the arithmetic mean ("average") of the signal intensity of all pixels in a region of interest. The SIWC, therefore, reflects movement or changes in the distribution of fluid within a structure. Therefore, the location of the SIWC could provide a new criterion for quantifying the effects of different loading conditions on fluid distribution within the disc. The distribution of fluid, as indicated by the location of the SIWC, is likely modified predictably by loading conditions. The change in location of the SIWC, unlike MSI, reflects the direction of changes in the fluid distribution. To our knowledge, this novel

measurement has not been used for evaluating the effects of different loading conditions on the disc. The aims of this study were to investigate the immediate effects of prone press-up extension exercises on the MSI and location of the SIWC of lumbar discs using $T_2$-weighted MR images. We hypothesised that prone press-up extension exercise could change both magnitude and location of the disc fluid content reflected by MSI and SIWC, respectively.

**Methods**

***Subjects.*** The present study was approved by the Health Research Ethics Board of the University of Alberta. The baseline mid-sagittal $T_2$-weighted MR images from 22 volunteers with low back pain (LBP) recruited for a previous study on extension exercise were used in this study [179]. The participants' mean age was 43±13 years, and all had back pain with or without leg pain and a minimum 25% Oswestry Disability Index score. The level of disc degeneration of the lumbar discs analysed in the study, as classified through consensus by two raters using a 5-point Pfirrmann grade [16], included no discs with grade I degeneration (no degeneration), and no more than three discs per level were rated as grade V (advanced degeneration with a collapsed disc space.

***MRI Imaging Procedures.*** Mid-sagittal $T_2$-weighted images of the lumbar spine were obtained in the supine position, before and after a standard bout of extension exercises. The mid-sagittal slices were determined based on the localizer sequence illustrating the presence of the spinous processes and with a clear demarcation of the spinal cord on the $T_2$-weighted images [180]. In cases where a spinal curve was present, the localizer was placed to ensure capturing the L4-5 disc and spinous processes. Participants were scanned at baseline after resting 40 minutes in a supine position for the disc to reach a steady state of hydration [213]. Participants were then asked to perform three sets of 10 prone press-up extension exercises to end-range as tolerated (approximately 1.5 minutes of extension) followed by maintaining a passive extension posture for 15 minutes. A physiotherapist trained in using McKenzie exercises for the lumbar spine instructed participants in performing the exercises. Participants were rescanned immediately afterwards. Mid-sagittal $T_2$-weighted images were obtained in a typical supine position with knees supported by a standard imaging pillow, using a 3T Siemens TrioTim MRI scanner (Siemens Healthcare, Erlangen, Germany). Acquisition parameters were TR 1600, TE 254, 2 averages, slice thickness 2 mm, field of view of 320×320 mm; and matrix size of 640×640 pixels.

***Image Analysis and Processing.*** Image post-processing was carried out offline using a computer aided measurement program developed by our team. The software employs signal variations in the neighbouring structures to semi-automatically segment each disc into different regions of interest (ROIs): the whole disc and nucleus region. One rater analysed images twice and another once, with a one-week interval between measurements, while blinded to prior measurements and each other's. Three parameters were then calculated for each ROI: 1) MSI, 2) X and 3) Y coordinates of the SIWC obtained using Equation 2 [214].

Equation 2. Mathematical description of SIWC

$$X_{SIWC} = \frac{\sum_{i=1}^{n} X_i SI_i}{\sum_{i}^{n} SI_i} \qquad Y_{SIWC} = \frac{\sum_{i=1}^{n} Y_i SI_i}{\sum_{i}^{n} SI_i}$$

Where $X_i$, $Y_i$ and $SI_i$ are the X and Y coordinates and the signal intensity of each pixel, and where *n* is the total number of pixels in the ROI. The centre of the Cartesian coordinate system was placed at the geometric weighted centroid of L4 vertebrae and the displacements of the SIWC were determined relative to that point.

***Data analysis.*** Statistical analyses were carried out using SPSS®, version 23.0 (IBM Corp. Armonk, NY., USA) using a level of statistical significance set at 0.05. The analysis was carried out at each level independently. A paired t-test model was used to compare the MSI and location of the SIWC of the disc and nucleus before and after exercise. The centre of the Cartesian coordinate system (0,0) for each disc was placed at the geometric centre of the same disc, and spatial location of the SIWC was determined relative to that point. Intra-class correlation coefficients were used to estimate intra- and inter-rater reliability ($ICC_{(3,1)}$ and $ICC_{(2,1)}$ respectively) for MSI and location of the coordinates of SIWC. ICC estimates higher than 0.75 are considered adequate for research conducted at a group level, and estimates of 0.9 and higher are considered adequate for clinical use at the individual level [152]. Standard error of measurement (SEM) was calculated to provide an estimate of measurement error in the measurement unit ($SEM = SD\sqrt{1 - ICC}$, where SD is the standard deviation of the first measurement). The effect size of differences between pairs of consecutive loading conditions was estimated using Cohen's d (

Equation 3). The most promising biomarkers were then identified as those with Cohen's D effect size of moderate (0.5) or large (0.8) for each comparison of interest [215].

Equation 3 Mathematical description of Cohen's D

$$d = \frac{(M_B - M_A)}{\sqrt{\frac{SD_B^2 + SD_A^2}{2}}}$$

Where $M_B$ and $M_A$ are the means of before and after the change in loading condition and $SD_B$ and $SD_A$ are standard deviations of before and after the loading change. The mean and SD of change between pairs consecutive of loading conditions, baseline vs. compression and compression vs. traction were also estimated with 95% confidence interval (CI).

**Results**

Measurement data were available for analysis for all L4-5 and L5-S1 discs from the 22 participants on both measurement occasions, except for one participant's MRI, which was excluded due to poor quality at follow-up.

***Reliability of Signal Intensity Weighted Centroid Measurements.*** The intra-rater and inter-rater reliability coefficients for the MSI of the whole disc and nucleus ranged from 0.94 to one, and the corresponding intra- and inter-rater SEMs varied from 0% to 6% of the mean measurements (table 14). The intra- and inter-rater reliability coefficients for the horizontal and vertical coordinates of the SIWC for the whole disc and nucleus ranged from a mean 0.97 to 1.00 (table 14). The intra-rater and inter-rater SEMs for the vertical coordinates of the SIWC for all ROIs were zero at both disc levels and varied from zero to 2.2 mm for the horizontal coordinates (table 14).

***Effects of Exercise on the Mean Signal Intensity and Signal Intensity Weighted Centroid.*** There was no significant difference between the MSI or vertical coordinate of the SIWC of the whole disc and nucleus before and after the extension exercises at either disc level (table 15). Also, with the exception of the whole disc at L4-5 (ES:-0.22), there were no significant differences between the location of the horizontal coordinates before and after the intervention (table 15).

**Table 14 Intra- and inter-rater reliability coefficients (ICC) and standard error of measurement (SEM) for measurements of the whole disc and nucleus at both levels**

| ROI | Measurements | | Mean | Intra-rater | | | Inter-rater | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | $ICC_{3,1}$ | 95 % CI | SEM | $ICC_{2,1}$ | 95% CI | SEM |
| **L4-5** | | | | | | | | | |
| Whole Disc | Signal intensity | | 21.7 (7.8) | 0.99 | 0.98-0.99 | 1.7 | 0.99 | 0.98-0.99 | 1.7 |
| | Weighted Centroid | X | 2.0 (3.4) | 1.00 | 1.00-1.00 | 0.0 | 1.00 | 1.00-1.00 | 0.0 |
| | | Y | 17.0 (1.4) | 1.00 | 1.00-1.00 | 0.0 | 1.00 | 1.00-1.00 | 0.0 |
| Nucleus | Signal intensity | | 30.6 (7.8) | 0.99 | 0.99-1.00 | 0.8 | 0.97 | 0.94-0.98 | 1.3 |
| | Weighted Centroid | X | 2.0 (3.3) | 1.00 | 1.00-1.00 | 0.0 | 0.97 | 0.95-0.98 | 2.2 |
| | | Y | 17.0 (1.5) | 1.00 | 1.00-1.00 | 0.0 | 1.00 | 1.00-1.00 | 0.0 |
| **L5-S1** | | | | | | | | | |
| Whole Disc | Signal intensity | | 24.5 (12.0) | 1.00 | 1.00-1.00 | 0.0 | 0.99 | 0.98-0.99 | 1.2 |
| | Weighted Centroid | X | -7.0 (2.3) | 1.00 | 1.00-1.00 | 0.0 | 0.99 | 0.99-1.00 | 1.4 |
| | | Y | -16.0 (2.0) | 1.00 | 1.00-1.00 | 0.0 | 1.00 | 1.00-1.00 | 0.0 |
| Nucleus | Signal intensity | | 37.4 (26.0) | 1.00 | 0.99-1.00 | 0.0 | 1.00 | 1.00-1.00 | 0.0 |
| | Weighted Centroid | X | -7.0 (2.5) | 1.00 | 1.00-1.00 | 0.0 | 1.00 | 1.00-1.00 | 0.0 |
| | | Y | -15.0 (2.2) | 1.00 | 1.00-1.00 | 0.0 | 1.00 | 1.00-1.00 | 0.0 |

Abbreviations: ROI: region of interest; ICC: intra-class correlation coefficient; SEM (mm): standard error of measurement; CI: confidence interval; X Anterior-posterior coordinate relative to the top left corner of the image; Y top-down distance relative to the top left corner of the image.

**Table 15 Mean Signal Intensity (MSI) and Signal Intensity Weighted Centroid (SIWC) of the whole disc and nucleus at both levels before and after intervention**

| ROI | Measurements | | Before | After | Mean Difference | 95% CI | *P-value* | Effect Size |
|---|---|---|---|---|---|---|---|---|
| L4-5 | | | | | | | | |
| Whole Disc | MSI | | 21.9±7.8 | 21.8±7.9 | -0.1±5.4 | -2.4-2.5 | *0.95* | 0.00 |
| | SIWC (mm) | X | 2.4±3.5 | 1.2±6.7 | -1.2±2.4 | 0.1-2.3 | ***0.03*** | -0.22 |
| | | Y | 17.3±1.5 | 17.4±1.4 | 0.1±0.4 | -0.6-0.2 | *0.30* | 0.07 |
| Nucleus | MSI | | 30.1±17.2 | 31.4±17.8 | 1.3±9.4 | -5.6-3.0 | *0.54* | 0.06 |
| | SIWC (mm) | X | 2.6±3.5 | 2.4±3.1 | -0.2±2.8 | -1.1-1.5 | *0.76* | -0.06 |
| | | Y | 17.4±1.6 | 17.3±1.6 | -0.1±0.4 | 0.0-0.3 | *0.17* | -0.06 |
| L5-S1 | | | | | | | | |
| Whole Disc | MSI | | 23.1±10.1 | 24.0±11.7 | 0.9±3.6 | -2.5-0.8 | *0.28* | 0.09 |
| | SIWC (mm) | X | -7.0±2.7* | -7.0±1.8 | 0.0±2.0 | -1.0-0.8 | *0.86* | 0.00 |
| | | Y | -15.5±2.2** | -15.7±3.3 | -0.2±1.7 | -0.4-1.0 | *0.42* | -0.07 |
| Nucleus | MSI | | 34.2±22.5 | 35.9±25.6 | 1.7±8.2 | -5.4-2.1 | *0.36* | 0.08 |
| | SIWC (mm) | X | -7.2±2.8 | -6.8±2.2 | 0.4±2.0 | -1.2-0.5 | *0.42* | 0.16 |
| | | Y | -15.0±2.4 | -15.5±1.9 | -0.5±2.0 | -0.4-1.4 | *0.29* | -0.23 |

*Posterior (dorsal) to the centre of the Cartesian coordinate system

**Inferior (caudal) to the centre of the Cartesian coordinate system

Abbreviations: ROI: region of interest; CI: Confidence interval.

**Discussion**

The MSI and SIWC of the entire intervertebral disc and nucleus region on midsagittal $T_2$-weighted MRI were found to be highly reliable measurements. Yet, they provided no indication of a change in fluid content or and limited evidence of a change in distribution within the whole disc or nucleus region as a result of extension exercises in a sample of patients with chronic low back pain. Only an anterior shift in the position of the SIWC of the whole L4-5 disc indicated a significant movement or enhanced concentration of fluid anteriorly. The effect sizes also suggest that novel measurements reflecting fluid distribution within the disc space showed equal or larger sensitivity to change compared to the more traditional MSI quantification of fluid content.

The SIWC reflects the location or centroid around which the distribution of signal intensities is balanced. In the present study, the excellent reliability of the SIWC measurements, with little, if any, measurement error for the horizontal and vertical coordinates, suggest the SIWC is suitable for research and clinical use. Périé and Curnier also reported good reliability for the nucleus weighted centroid, but only reported data related to the horizontal direction [214].

Our findings suggesting no significant change in disc fluid content were not consistent with Beattie et al.'s who found a significant increase in the fluid diffusion in the nucleus following a 10-minute session of lumbar pressures in a posterior-to-anterior (PA) direction and prone press-up exercises [45]. In the present study, participants spent 1.5 minutes performing repeated extension to end range followed by 15 minutes in a passive extension position. Results of another study conducted by Beattie et al indicated a significant increase in the disc apparent diffusion coefficient following a 10-minute application of PA directed manual pressures [69]. Thus, the observed changes in both studies could be due to the application of PA directed manual pressure. The anterior shift of the signal intensity observed at the L4-5 level indicates more fluid concentration in the anterior part of the disc. This was consistent with our hypothesis that extension increases the pressure on the posterior part of the disc and drives fluids more anteriorly. The fact that the SIWC did not demonstrate significant movement at L5-S1 makes this finding less persuasive, but the range of motion at L5-S1 is generally less than at L4-5 [216], and may not be enough to reduce intradiscal pressure and significantly drive fluid anteriorly.

Several limitations to the present study should be acknowledged. We did not quantify disc fluid content in the extended position. This would increase the distance between the body and the

spine coil and thus reduce the signal to noise ratio. Our small sample of volunteers presented mainly moderately and highly degenerated discs, which are typically harder to measure, compared to less degenerated discs with highly detectable margins, and may provide a worst-case scenario with respect to measurement reliability and detecting changes in fluid distribution within the disc. However, the clinical significance of changes in SIWC locations of different magnitudes at this point is unclear. Another limitation was the pixel size. We were only able to capture changes equal to or greater than 0.5 mm. An additional study limitation is that we only measured the midsagittal slice of the L4-5 and L5-S1 discs, and it is possible that changes in fluid content or distribution in other regions of the disc may have been missed.

**Conclusion**

The location of the SIWC in different disc ROIs may be employed as a highly reliable measurement for investigating changes in signal distribution, which may be useful in studies of the effects of different loading conditions or therapeutic interventions on fluid movement or distribution. Little evidence was found supporting the hypothesis that prone press-up exercises affect disc fluid content and distribution. There was no significant difference in the whole disc and nucleus MSI or location of the SIWC before and after extension exercise, except for the horizontal coordinate of the SIWC of the whole L4-L5 disc.

**Key points**

FINDINGS: There were no significant differences in the whole disc or nucleus MSI and location of the SIWC before and after extension exercise at either disc level examined, except for an anterior shift of the SIWC of the whole disc at the L4-L5 level.

IMPLICATION: These findings suggest prone press-up extension exercises do not increase disc fluid content, but possibly influence fluid distribution within the discs.

CAUTION: These findings are based on a small sample size and may not be applicable to other groups of patients with less advanced disc degeneration.

# CHAPTER 7 EVALUATION OF THE EFFECTS OF LOADING AND DISC DEGENERATION ON LUMBAR INTERVERTEBRAL DISCS AND MOTION SEGMENTS

Vahid Abdollah; Eric C Parent; Alex Su, Keith Wachowicz, Michele C Battié

**Abstract**

**Background:** Conventional MRI is routinely used to depict intervertebral disc degeneration (IDD) which has become the focus of much research. IDD affects the magnitude and the distribution of fluid in the disc. IDD has been associated with changes in the disc biomechanics, which may be best illustrated by imaging spines under loading. This study aimed to compare the effects of compression and traction loading on geometric and $T_2$ measurements of the lumbar discs and vertebrae using quantitative $T_2$-mapping in relation to disc degeneration.

**Methods:** Thirty-five volunteers (30±11yrs, 51% women) with and without chronic LBP rested in a supine position for 15 minutes before an unloaded MRI scan. The lumbar spine was then loaded with 50% body weight for 20 minutes with imaging occurring in the last 5 minutes and the process was then repeated under traction. Disc, nucleus and vertebra from L1-2 through L5-S1 were semi-automatically segmented on mid-sagittal $T_2$-weighted images and measurements extracted from $T_2$-maps. For each ROI and loading condition, the mean $T_2$ ($MT_2$), geometric weighted centroid (GWC), and $T_2$-weighted centroid ($T_2$WC) were calculated. Disc height, motion segment angle (MSA), disc and nucleus width were also computed. A repeated measure analysis of covariance (ANCOVA) with Sidak's post-hoc comparisons was employed to compare different loading conditions, controlling for age and vertebral dimensions. Cohen's d effect size was calculated for differences between loading conditions. Correlation coefficients were computed between the $MT_2$ of the corresponding disc as a measure of degeneration and the change between each pair of consecutive loading conditions.

**Findings:** From compression to traction, we observed a statistically significant: small decrease for the disc $MT_2$ at L1-2 (ES: -0.35); small increase for the Nucleus $MT_2$ at L3-L4 (ES=0.26); trivial or small inferior and posterior shift of the L4-L5 (ES: 0.4, 0.14) and L5-S1 (ES: 0.25, 0.33) disc $T_2$WC; small inferior and posterior shift of the nucleus $T_2$WC (ES: 0.25, 0.31) and GWC (ES: 0.22,0.31) at L5-S1, small posterior shift of the nucleus $T_2$WC (ES:0.49) and GWC (ES: 0.48) at L4-5. From unloaded to compression, we observed a significant: a small increase in disc width at L5-S1 (ES 0.22); a trivial increase in Nucleus $MT_2$ at L1-2 (ES 0.18); small anterior shift of the disc $T_2$WC at L1-2 (ES 0.39); small posterior shift of the nucleus GWC at L3-4 (ES 0.38); moderate segment extension at L3-4 (MSA 2.1º) and L4-5 (MSA 1.8º). Least severe disc

degeneration correlated with larger changes in the disc and nucleus measurements between pairs of loading conditions with a strongest correlation of -0.51.

**Interpretation:** Effect size estimates suggested that the location of $T_2$WC and GWC, the motion segment angle and disc height were the most promising biomarkers for evaluating the effects of loading on the lumbar spine. The largest responses to loading were observed at the lower lumbar levels. Larger responses to loading were observed with less degeneration.

**Keywords:** magnetic resonance imaging; compression; traction; $T_2$ map; mean $T_2$ time; mean $T_2$ weighted centroid, disc height, degeneration, intervertebral disc, low back pain

## Introduction

Conventional MRI is routinely used to depict structural abnormalities of the lumbar spine and diagnose low back pain (LBP), with a particular interest in the intervertebral disc. There is much research on intervertebral disc degeneration (IDD) in hope to better understand the pathology in the majority of patients with LBP receiving a diagnosis of non-specific LBP [115, 217, 218]. Indeed, IDD has been implicated in the development of LBP, either directly through the infiltration of blood vessels and nerves or indirectly through effects on spinal biomechanics and other structures [5–7]. IDD is associated with morphological, chemical and metabolic changes in the disc [5]. A degenerated disc is a dehydrated and fragmented structure [54], with morphological changes including radial and circumferential annulus tears, disruption of the endplates and loss of demarcation between the nucleus and annulus [5, 51]. With disc degeneration the fragmented proteoglycans begin to diffuse out of the nucleus, reducing the negative charge density of the disc, disc osmotic pressure and thus disc fluid content [61]. Degeneration not only affects the magnitude, but also the distribution of fluid in the disc [5]. In addition to changes within the disc, IDD has also been associated with changes in the biomechanics of the motion segments [5].

Despite the frequency with which diagnostic MRI is used to investigate the disc for LBP, findings are often inconclusive due to a lack of specificity. This may be due, in part, to conducting MR imaging in a relaxed supine position while pain is often experienced when the spine is subjected to loading. Thus, imaging under different loading conditions may provide more informative findings than those revealed at rest.

Loading induces morphologic changes in the lumbar spine [161, 219] and variations in fluid exchange [32], which may provide novel biomarkers of degenerative conditions, with possible clinical relevance. On the $T_2$-weighted MR images, $T_2$ time can be employed as an indirect measure for estimating the fluid content of the tissue [37]. A higher $T_2$ indicates more fluid concentration in the region of interest [37]. The location of the MR $T_2$ weighted centroid ($T_2$WC), which is the location of the arithmetic mean $T_2$ ($MT_2$) of all the pixels in a region of interest weighted based on their intensity, can be used to reflect the distribution of fluid within a structure. $T_2$WC is a unique point location, which only changes in response to variations in fluid distribution as reflected by $T_2$ time of each pixel within the region. Yet, this potentially informative measure

has not been previously utilised in studies of disc degeneration and compression or traction loading.

As degeneration affects morphology, biomechanics and biochemistry of the disc [1], this preliminary study aimed to test the hypothesis that variations observed in the magnitude and distribution pattern of fluid within the lumbar tissues, as well as changes in morphology and segmental alignment of the lumbar spine, under different loading conditions would be associated with the degree of degeneration. The response of these measurements to loading was quantified as it may provide further insight into disc degeneration and function, and potentially the pathology underlying persistent or recurrent back pain. Specifically, refined quantification of changes in $T_2$ profile, location of $T_2$WC, disc height, lumbar motion segment angulation and alignment under different loading conditions may lead to the identification of new imaging biomarkers to assist with classification or diagnosis of lumbar degenerative conditions. We also hypothesised that the response to loading of the quantified biomarkers would decrease with increasing disc degeneration.

**Materials and methods**

This study consisted of a single testing session to compare MRI geometric and $T_2$ measurements in the relaxed supine position, and with the lumbar spine under compression and then traction.

***Participants.*** Thirty-five volunteers (17♂, 18♀), including 20 (11♂, 9♀) with and 15 (6♂, 9♀) without LBP, participated in the study. We employed both volunteers with and without LBP as we were looking for a wide range of degeneration in this preliminary evaluation of novel biomarkers. The participants were recruited using the mass email service to students affiliated with our university. The inclusion criteria for both groups were 18 to 65 years of age and the ability to read and understand English instructions. Participants with claustrophobia, any contraindications for MRI or traction, tumour, uncontrolled hypertension, severe osteoporosis, and previous spine surgery were excluded. Additional inclusion criteria for participants with LBP were an Oswestry Disability Index score of 25% or greater and being in pain for at least half of the days in the past six months before imaging [49]. We excluded participants with confirmed causes of back pain, such as malignancy and fracture; significant compression of the spinal cord/ nerves; on-going workers' compensation or litigation cases; and having received injections within the last 12 weeks.

Additional inclusion criteria for volunteers without LBP were no current back pain and no memorable (severe or disabling) back pain lasting more than one day over the past ten years. The study was approved by the Health Research Ethics Board of the University of Alberta.

G*Power 3.1 software (University of Kiel, Kiel, Germany) was used to calculate the required sample size to secure a power of 0.80 using an α value of 0.05, to detect an effect size corresponding to the mean difference ± standard deviation observed in the disc height in a study of 6 weeks of motorized traction with 50% BW [220]. This study was the only study providing preliminary traction loading response estimates at the time. The required sample size was 40 participants. With 35 participants, an effect size of 0.49 could be detected as statistically significant.

***MRI image acquisition procedures.*** Use of a custom-designed MRI compatible loading table allowed compression and traction to be applied while the participants were in the MRI scanner. Mid-sagittal $T_2$-weighted MR images were acquired during the last five minutes of each loading period that lasted for 20 minutes. The position of mid-sagittal slices were selected by the technician based on the localizer sequence illustrating the presence of the spinous processes and with a clear demarcation of the spinal cord on the $T_2$-weighted images [180]. In cases where a lumbar curve was present, the localizer was placed to ensure capturing the L4-5 disc and spinous processes. The participant spent the first 20 minutes lying in a relaxed unloaded supine position, followed by 20 minutes loaded in compression with 50% body weight (BW), and then 20 minutes in traction with 50% BW applied (Figure 8). A compressive load equal to 50% of the subject's BW was applied to simulate loading in the erect spine [221], followed by a traction force of 50% BW to investigate the effects of a contrasting unloaded condition on imaging measurements [220]. We chose 50% BW as it showed promise to induce changes in disc height [90],while also minimising subject discomfort and possible drop out during scanning [222].

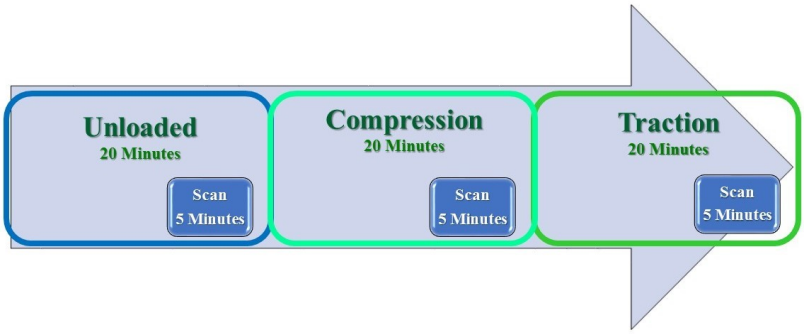


Figure 8 A schematic presentation of loading and imaging workflow

Compression force was applied by a deadweight of 50% BW tied to a rope tensioning a harness strapped around the chest with straps applying load over the shoulders (Figure 9). Loads were left in place during the first 5 minutes of loading with the rope free to move through a pulley at the foot of the device. Then the rope was locked in place by a compression screw and the patient was moved into the MRI while still under loading. For traction, the chest harness was released from the foot of the device and anchored by straps at the head of the device to fix the upper spine. Traction loading was applied via straps pulling on a harness securely tightened around the pelvis by applying a deadweight of 50% BW for 5 minutes and then the tensioned cable was locked in place (Figure 9). To ensure that friction did not limit the application of loading force, the pelvis and lower extremities rested on the lower half of the table which was resting on wheels and free to move in the cephalad-caudal direction.

Images were acquired using a 3T whole-body Philips MR scanner (Philips Healthcare Intera). The $T_2$-weighted image acquisition parameters used to obtain a $T_2$ map were TR 2500, TE (five echoes: 16.9, 44, 71, 98 and 125 ms), 2 averages, a field of view of 500×500 mm, pixel size 0.49×0.49 mm, image thickness 5mm and a matrix size of 1024×1024 pixel. The location of mid-sagittal slices was determined using the localizer sequence to select the location best illustrating the spinous processes and with a clear demarcation of the spinal cord [180]. In cases where a lumbar lateral curve was present, the localizer was placed to ensure capturing the L4-5 disc and spinous processes.

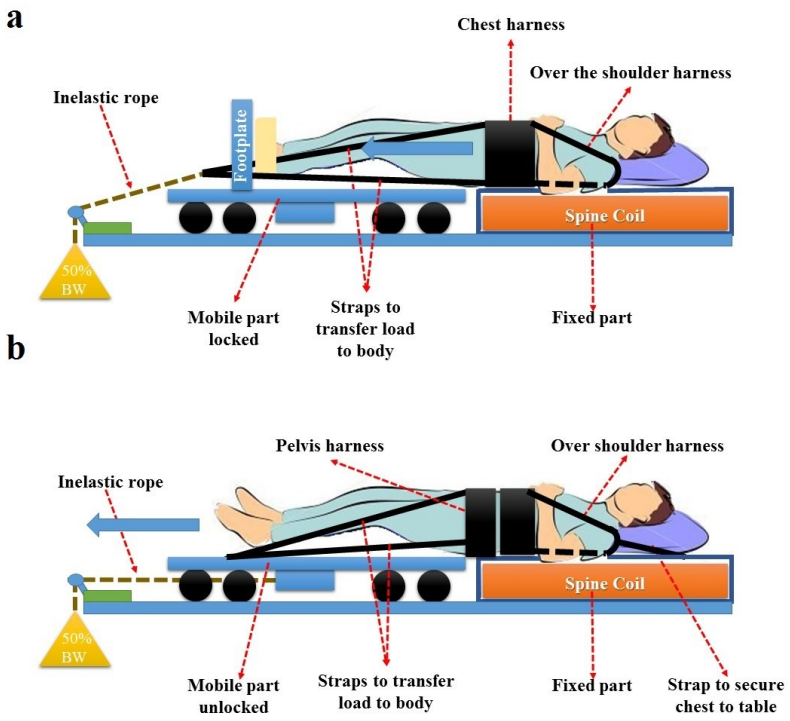


Figure 9 Schematic representation of the loading table a:) compression b:) traction

***Image processing and disc height measurements.*** Image post-processing was carried out offline using a computer aided measurement program developed by our team (Appendix I). A graduate student (AS) completed the semi-automatic image segmentation following training in using the program and practising segmentation on 70 images. $T_2$-maps were constructed automatically by obtaining $T_2$ relaxation time estimates for a given pixel using the Nelder-Mead Simplex method [223].

The software was used to semi-automatically segment each lumbar motion segment by identifying boundaries displaying maximal $T_2$ signal intensity differences into the following regions of interest (ROIs): the whole disc, the hyperintense nucleus region (hereafter referred to as the nucleus) and vertebra (Appendix I). The semi-automated segmentation procedures have been reported previously [224]. The $T_2$-weighted image corresponding to the third echo (TE=71ms) was used for image segmentation and each ROI was then reconstructed on the corresponding $T_2$-map to extract measurements of interest. Three parameters were calculated for each vertebra and nucleus ROI: 1) mean $T_2$ time ($MT_2$), 2) horitzontal (X) and and vertical (Y)

coordinates of the geometric weighted centroid (GWC), and 3) of $T_2$WC obtained using *Equation 4* and Equation *5*, respectively [214]. The $MT_2$ and location of $T_2$WC were calculated for IVDs.

Equation 4. Mathematical description of $T_2$WC

$$T_2WC = \begin{bmatrix} X = \frac{\sum_{i=1}^{n} X_i T2_i}{\sum_i^n T2_i} \\ Y = \frac{\sum_{i=1}^{n} Y_i T2}{\sum_i^n T2_i} \end{bmatrix}$$

Equation 5. Mathematical description of the GWC

$$GWC = \begin{bmatrix} X = \frac{\sum_{i=1}^{n} X_i A_i}{\sum_i^n A_i} \\ Y = \frac{\sum_{i=1}^{n} Y_i A_i}{\sum_i^n A_i} \end{bmatrix}$$

Where $X_i$ and $Y_i$, are the horizontal and vertical coordinates of each pixel within an ROI, and $A_i$ and $T_{2i}$ are the area of each pixel and the $T_2$ time of each pixel, respectively. As the location of the lumbar spine varied relative to the centre point of the image in different loading conditions, the centre of the Cartesian coordinate system was placed at the GWC of the superior vertebra for each motion segment. All GWC and $T_2$WC measurements for the motion segment were determined relative to that reference point. The negative horizontal and vertical coordinate values indicate more anterior and superior position relative to the reference vertebra's GWC, respectively.

Disc height was also measured automatically using an area-based method [197] by dividing the area of the disc by the corresponding diameter representing 80% of the disc width centred at the mid-point of the whole disc. The angle between the best-fit lines passing through the segmented contour of the superior and inferior border of the disc was computed as the motion segment angle. The width of the disc and nucleus along the disc bisector line was also computed automatically. The $ICC_{3,1}$ intra-rater reliability estimates of the motion segment angle and disc and nucleus width ranged from 0.75 to 0.95. The reliability estimate of the motion segment angle and disc and nucleus width ranged from 0.75 to 0.95. All other measurement constructs were found to have good to excellent intra- and inter-rater reliability, with reliability coefficients (ICC) ranging from 0.88 to 1.00 [224].

**Disc degeneration grading.** The degree of disc degeneration was assessed for each disc using the 5-point scale proposed by Pfirrmann et al [16] through consensus by two raters (VA and EP) with more than five years of experience in spine imaging. Degeneration was graded using the $T_2$-weighted image with a TE of 71 ms.

***Data analysis.*** The analysis was carried out at each level independently. The normality of the data was examined using Shapiro-Wilk test for each measure. A repeated model of analysis of covariance (ANCOVA) was employed to compare different loading conditions [152]. Age and the area of the second lumbar vertebra (to adjust for variations in lumbar spine size among participants), were included as covariates. Sidak's post-hoc comparisons were used to determine the significance of differences between the baseline unloaded and compression measurements and between compression and traction [152]. The effect size of differences between pairs of consecutive loading conditions was estimated using Cohen's d (Equation 6). The most promising biomarkers were then identified as those with Cohen's D effect size of moderate (0.5) or large (0.8) for each comparison of interest [215].

Equation 6 Mathematical description of Cohen's D

$$d = \frac{(M_B - M_A)}{\sqrt{\frac{SD_B^2 + SD_A^2}{2}}}$$

Where $M_B$ and $M_A$ are the means of before and after the change in loading condition and $SD_B$ and $SD_A$ are standard deviations of before and after the loading change. The mean and SD of change between pairs consecutive of loading conditions, unloaded vs. compression and compression vs. traction were also estimated with 95% confidence interval (CI).

Pearson's correlation coefficients were computed for the disc $MT_2$ while unloaded as a measure of disc degeneration and the change in each quantitative MRI measurement between consecutive loading conditions to test the hypothesis that the response to loading decreases with increasing disc degeneration. Finally, a post-hoc Pearson's correlation coefficient estimation was conducted among the change variables of all candidate biomarkers when going from compression to traction at L5-S1 (the level, which demonstrated the most response to loading). This was to explore which variables shared large amounts of variance. We considered the variables with r-

square exceeding 0.50, which corresponds to sharing over 50% of common variance as providing possibly redundant information. Correlation estimates larger than r=0.71 were used to identify variables possibly providing redundant information as the corresponding coefficient of determination ($r^2$) indicates more than 50% of common variance [225]. All statistical analyses were carried out using SPSS® statistical software, version 23.0 (IBM Corp. Armonk, NY., USA) with the level of statistical significance set at 0.05.

**Results**

The mean age of the 35 participants was 30.4±11.2 [18-53] years, and the mean weight was 75.9±13.5 [56-96] kg for females and 69.9±12.8 [55-94] kg for males. Thirty-five discs and vertebrae were analysed at each level from L1-2 to L5-S1 for each loading condition. Of the total 175 discs, the sample included only three discs (1.7%) with Pfirrmann grade I (normal), and two (1.1%) grade V (advanced degeneration with a collapsed disc space), both at L5-S1 (table 16).

**Table 16 Distribution of Pfirrmann scores of IDD at each level from L1-2 through L5-S1**

| Level | Grade I | Grade II | Grade III | Grade IV | Grade V |
|---|---|---|---|---|---|
| L1-2 | 0 | 29 | 4 | 2 | 0 |
| L2-3 | 1 | 32 | 2 | 0 | 0 |
| L3-4 | 1 | 28 | 5 | 1 | 0 |
| L4-5 | 1 | 14 | 13 | 7 | 0 |
| L5-S1 | 0 | 12 | 11 | 10 | 2 |
| Total | 3 | 115 | 35 | 20 | 2 |
| Percentage | 1.7 | 65.7 | 20.0 | 11.4 | 1.1 |

**Differences in response to loading.** Most differences in response to loading were observed going from compression to traction (table 17). Differences were most often observed at L4-5 and L5-S1 and most often from measurements reflecting fluid distribution rather than geometric measurements.

**Going from compression to traction.** Statistically significant trivial or small inferior ($ES_{L4-5}$:0.14; $ES_{L5-S1}$:0.33) and posterior shifts ($ES_{L4-5}$:0.40; $ES_{L5-S1}$:0.30) of the disc $T_2$WC (Figure 10 a) were observed at the two lowest lumbar levels (table 17). A statistically significant small posterior shift was also observed for the nucleus $T_2$WC and GWC at L4-5 ($ES_{T2WC}$:0.49; $ES_{GWC}$: 0.48) and L5-S1 ($ES_{T2WC}$:0.25; $ES_{GWC}$:0.22) (Figure 10b). There was also a statistically significant small inferior shift of the nucleus $T_2$WC (ES: 0.31) and GWC (ES:0.31) at L5-S1 (Figure 10 b).

Statistically significant increases of the disc height were observed at L1-2 (trivial ES:0.08) and L5-S1 (small ES:0.32) were observed (Figure 10 c). There was also a statistically significant small decrease in the disc $MT_2$ at L1-2 (ES:-0.35) and increase in the nucleus $MT_2$ at L3-4 (Figure 10 d) (ES:0.26).

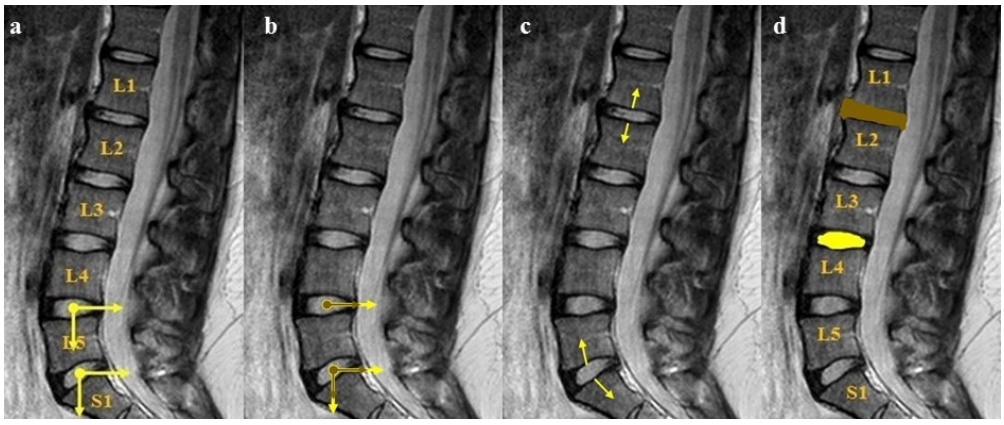


Figure 10 Significant differences from compression to traction for a:) disc $T_2$WC b:) nucleus $T_2$WC (yellow) and GWC (brown) and c:) disc height d:) disc and nucleus $MT_2$ (images illustrate direction and location of the effects but are not scaled)

Moreover, a small effect size but not statistically significant was observed for increased disc $MT_2$ at L4-5 (ES:0.22) as well as decreased nucleus $MT_2$ at L1-2 (ES:-0.26). Likewise, a small effect size but not statistically significant was also observed for decreased disc width (ES:-0.22) at L5-S1 and increased nucleus width at L1-2 (ES:0.23) and L3-4 (ES:0.20). A small effect size but not statistically significant was observed for a decrease in the motion segment angle at L3-4 (ES:-0.27) and L4-5 (ES:-0.27), and an increase at L5-S1 (ES:0.25).

***Going from unloaded to compression*.** Statistically significant small posterior shifts were observed for the disc $T_2$WC (ES:0.39) at the L1-2 level and for the nucleus GWC at L3-4 (ES:0.38) (Figure 11 a) (table 17). The nucleus $MT_2$ increased by a trivial effect size (ES:0.18) at L1-2 (Figure 11 b). The motion segment angle showed statistically significant increases with moderate effect sizes by an average 2.1$^o$ at L3-4 (ES:0.56) and 1.8$^o$ at L4-5 (ES:0.53) (Figure 11 c). A significant small increase of the disc width was observed at L5-S1 (ES:0.22) (Figure 11 d).

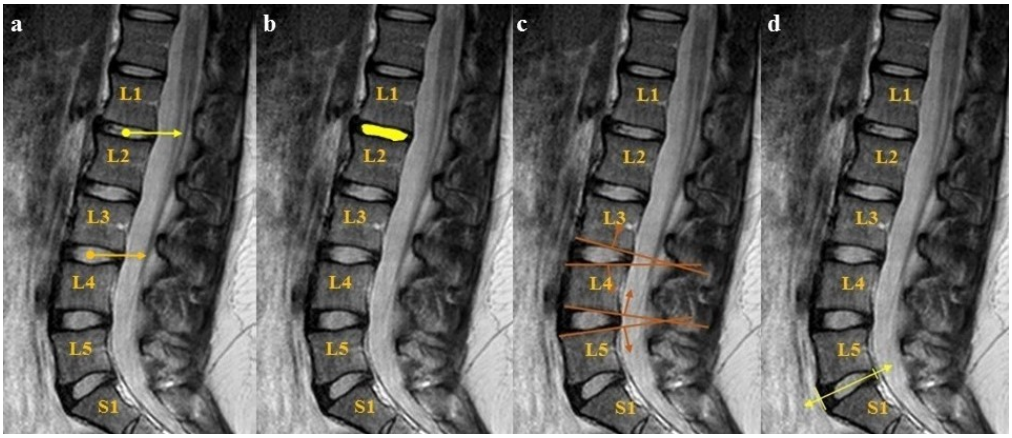


Figure 11 Significant differences from unloaded to compression a:) disc $T_2$WC (illustrated in yellow) and nucleus GWC (illustrated in brown) b:) nucleus $MT_2$ and c:) motion segment angle d:) disc width (images illustrate direction and location of the effects but are not scaled)

A small effect size but not statistically significant was observed for the posterior shift in the disc and nucleus $T_2$WC at L2-3 ($ES_{DISC}$:0.44; $ES_{NUCLEUS}$:0.39), and only for the nucleus at L3-4 (ES:0.35). Likewise, a small effect size but not statistically significant was also observed for the posterior shift in the nucleus GWC at L2-3 (ES:0.41). In contrast, a small effect size but not statistically significant was observed for an anterior shift of the nucleus T2WC at L5-S1 (ES: -0.20) and the Nucleus GWC at L4-5 (ES: -0.23). A small effect size but not statistically significant was observed for a superior shift in the nucleus $T_2$WC and GWC at L5-S1 (ES: -0.21) as well as for a reduction in the disc $MT_2$ at L3-4 (ES: -0.26) and in the nucleus width at L1-2 (ES: -0.34), L3-4 (ES: -0.27) and L5-S1 (ES: -0.26).

**Table 17 Differences between loading conditions in mean $T_2$, location of the $T_2$ weighted and geometric centroids of the disc and nucleus ROIs for L1-2 through L5-S1**

| Measure | Level | Unloaded | 50% BW C | 50% BW T | Effect Size C vs. U | Effect Size T vs. C |
|---|---|---|---|---|---|---|
| ***Disc*** | | | | | | |
| Horizontal coordinate of $T_2$WC (mm) | L1-2 | 7.3±0.5 | 8.6±0.6[a] | 8.6±0.4 | 0.39 | 0.00 |
| | L2-3 | 6.7±1.3 | 9.2±0.5 | 9.4±0.4 | 0.44 | 0.07 |
| | L3-4 | 6.5±0.5 | 7.1±0.6 | 7.5±0.5 | 0.17 | 0.12 |
| | L4-5 | 0.2±0.8 | 0.1±0.7 | 1.8±0.7[b] | -0.02 | 0.40 |
| | L5-S1 | -11.7±0.6 | -12.3±0.7 | -10.8±0.8[b] | -0.13 | 0.30 |
| Vertical coordinate of $T_2$WC (mm) | L1-2 | 33.6±0.2 | 33.4±0.2 | 33.5±0.2 | -0.08 | 0.04 |
| | L2-3 | 35.3±0.6 | 35.7±0.2 | 35.8±0.2 | 0.16 | 0.08 |
| | L3-4 | 37.2±.20 | 37.4±0.2 | 37.4±0.2 | 0.09 | 0.00 |
| | L4-5 | 38.3±0.3 | 38.3±0.3 | 38.6±0.3[b] | 0.00 | 0.14 |
| | L5-S1 | 35.4±0.4 | 34.9±0.5 | 36.1±0.5[b] | -0.15 | 0.33 |
| Mean $T_2$ (ms) | L1-2 | 133±3 | 136±3 | 130±2[b] | 0.17 | -0.35 |
| | L2-3 | 137±3 | 140±3 | 140±3 | 0.17 | 0.00 |
| | L3-4 | 142±3 | 137±3 | 143±3 | -0.26 | 0.31 |
| | L4-5 | 132±3 | 130±3 | 135±4 | -0.09 | 0.22 |
| | L5-S1 | 120±3 | 118±3 | 120±3 | -0.10 | 0.10 |
| Height 80% Disc Width Used | L1-2 | 8.7±0.1 | 8.9±0.1 | 9.0±0.1[b] | 0.16 | 0.08 |
| | L2-3 | 9.9±0.2 | 10.0±0.2 | 10.1±0.2 | 0.07 | 0.07 |
| | L3-4 | 10.5±0.2 | 10.6±0.2 | 10.7±0.2 | 0.07 | 0.07 |
| | L4-5 | 11.1±0.2 | 11.2±0.2 | 11.4±0.2 | 0.08 | 0.17 |
| | L5-S1 | 10.5±0.2 | 10.3±0.2 | 10.8±0.3[b] | -0.13 | 0.32 |
| Disc Width (mm) | L1-2 | 71.8±0.6 | 72.0±0.8 | 71.3±0.6 | 0.03 | -0.10 |
| | L2-3 | 74.1±0.7 | 74.7±0.7 | 73.8±0.7 | 0.10 | -0.15 |
| | L3-4 | 75.8±0.8 | 75.3±0.8 | 74.7±0.7 | -0.07 | -0.09 |
| | L4-5 | 78.4±0.9 | 78.6±0.9 | 78.5±0.9 | 0.02 | -0.01 |
| | L5-S1 | 75.1±0.9 | 77.0±1.1[a] | 75.1±1.0 | 0.22 | -0.22 |
| ***Nucleus*** | | | | | | |
| Horizontal coordinate of $T_2$WC (mm) | L1-2 | 6.2±0.5 | 7.0±0.5 | 7.2±0.5 | 0.23 | 0.07 |
| | L2-3 | 5.8±1.2 | 8.0±0.4 | 8.2±0.5 | 0.39 | 0.07 |
| | L3-4 | 5.5±0.5 | 6.7±0.6 | 6.9±0.4 | 0.35 | 0.06 |
| | L4-5 | 0.0±0.7 | -0.8±0.8 | 1.2±0.6[b] | -0.18 | 0.49 |
| | L5-S1 | -10.9±0.7 | -11.9±0.8 | -10.6±0.8[b] | -0.20 | 0.25 |
| Vertical-coordinate of $T_2$WC (mm) | L1-2 | 33.6±0.2 | 33.5±0.2 | 33.6±0.2 | -0.04 | 0.04 |
| | L2-3 | 35.1±0.6 | 35.6±0.2 | 35.7±0.2 | 0.12 | 0.08 |
| | L3-4 | 37.0±0.2 | 37.2±0.2 | 37.1±0.2 | 0.09 | -0.04 |
| | L4-5 | 38.1±0.3 | 37.9±0.3 | 38.2±0.3 | -0.09 | 0.13 |
| | L5-S1 | 35.1±0.5 | 34.3±0.6 | 35.5±0.5[b] | -0.21 | 0.31 |

| Measure | Level | Unloaded | 50% BW C | 50% BW T | Effect Size C vs. U | Effect Size T vs. C |
|---|---|---|---|---|---|---|
| Horizontal coordinate of GWC (mm) | L1-2 | 6.7±0.5 | 7.1±0.6 | 7.2±0.5 | 0.13 | 0.03 |
| | L2-3 | 5.6±1.2 | 7.8±0.4 | 7.8±0.4 | 0.41 | 0.00 |
| | L3-4 | 4.8±0.6 | 6.2±0.6[a] | 6.1±0.5 | 0.38 | -0.03 |
| | L4-5 | -0.8±0.7 | -1.8±0.8 | 0.2±0.7[b] | -0.23 | 0.48 |
| | L5-S1 | -11.3±0.7 | -12.2±0.8 | -11.1±0.7[b] | -0.18 | 0.22 |
| Vertical coordinate of GWC (mm) | L1-2 | 33.5±0.2 | 33.4±0.2 | 33.5±0.2 | -0.04 | 0.04 |
| | L2-3 | 35.0±0.6 | 35.5±0.2 | 35.6±0.2 | 0.16 | 0.08 |
| | L3-4 | 37.0±0.2 | 37.2±0.2 | 37.1±0.2 | 0.09 | -0.05 |
| | L4-5 | 38.0±0.3 | 37.7±0.3 | 38.1±0.3 | -0.13 | 0.18 |
| | L5-S1 | 35.0±0.4 | 34.3±0.5 | 35.4±0.5[b] | -0.21 | 0.31 |
| Mean $T_2$ (ms) | L1-2 | 161±4 | 166±5[a] | 159±4 | 0.18 | -0.26 |
| | L2-3 | 167±4 | 172±4 | 170±3 | 0.19 | -0.08 |
| | L3-4 | 171±4 | 165±5 | 174±5[b] | -0.15 | 0.26 |
| | L4-5 | 152±5 | 151±5 | 155±5 | -0.03 | 0.11 |
| | L5-S1 | 141±5 | 139±6 | 142±5 | -0.06 | 0.08 |
| Nucleus Width (mm) | L1-2 | 24.5±0.4 | 23.1±0.5 | 24.1±0.4 | -0.34 | 0.23 |
| | L2-3 | 25.2±0.5 | 25.3±0.5 | 25.0±0.5 | 0.02 | -0.07 |
| | L3-4 | 26.3±0.6 | 25.1±0.8 | 26.0±0.6 | -0.27 | 0.20 |
| | L4-5 | 26.1±0.5 | 25.1±0.7 | 24.9±0.6 | -0.26 | -0.05 |
| | L5-S1 | 25.3±0.5 | 24.9±0.7 | 24.7±0.6 | -0.10 | -0.05 |
| ***Motion Segment*** | | | | | | |
| Angle° | L1-2 | 2.5±0.5 | 3.1±0.6 | 3.1±0.5 | 0.19 | 0.00 |
| | L2-3 | 4.6±0.5 | 5.1±0.4 | 4.8±0.6 | 0.19 | -0.10 |
| | L3-4 | 7.1±0.5 | 9.1±0.6[a] | 8.1±0.6 | 0.56* | -0.27 |
| | L4-5 | 10.5±0.6 | 12.3±0.6[a] | 11.4±0.5 | 0.53* | -0.27 |
| | L5-S1 | 15.2±0.6 | 15.2±0.9 | 16.4±0.7 | 0.00 | 0.25 |

[a] *Significant difference of compression vs unloaded. (Sidak $p<0.05$)*

[b] *Significant difference of traction vs compression (Sidak $p<0.05$)*

Negative horizontal means moving anteriorly relative to geometric centroid of the vertebra above.

Positive vertical means moving caudally relative to geometric centroid of the vertebra above.

* Medium effect size

**Bold** values represent $P_{Value}<0.05$

Abbreviations: BW: body weight; GWC: geometric weighted centroid; $T_2$WC: $T_2$weighted centroid; mm: millimetre; U: unloaded; C: compression; T: traction

***Correlations between degeneration and response to loading.*** The largest negative and positive Pearson's correlation coefficients for disc degeneration estimated using $MT_2$ and the differences in the disc and nucleus measurements from unloaded to compression or from

compression to traction were -0.51 and 0.49, respectively (table 18). For changes from compression to traction, less degeneration was associated with more superior movement of the disc $T_2$WC ($r$=-0.45), the nucleus $T_2$WC ($r$=-0.51) and the GWC ($r$=-0.48) at L2-3 (table 18). Likewise, lower degeneration associated with larger posterior movement of the nucleus $T_2$WC at L1-2 ($r$=0.42), L3-4 ($r$=0.39) and L5-S1 ($r$=0.36), and the nucleus GWC at L5-S1 ($r$=0.39). Conversely lower degeneration associated with larger anterior movement of the disc $T_2$WC ($r$=-0.46) and the nucleus GWC ($r$=-0.51) at L1-2. Less degeneration was also associated with more decreased disc width at L4-5 ($r$=-0.36).

For changes from unloaded to compression, less disc degeneration (higher $MT_2$) was associated with a larger decrease in $MT_2$ of the disc at L2-3 ($r$=-0.45), L3-4 ($r$=-0.51) and L4-5 ($r$=-0.42), and of the nucleus at L3-4 ($r$=-0.36). Likewise, less degeneration was associated with larger anterior movement of the disc $T_2$WC at L1-2 ($r$=-0.45) and L5-S1 ($r$=-0.36) and downward movement of the disc $T_2$WC at L1-2 ($r$=0.49). Less degeneration was also associated with more decrease in disc height at L1-2 ($r$=-0.33)

**Table 18 Pearson's correlation coefficients between baseline mean $T_2$ as a marker of IDD and changes in the quantitative measurements from unloaded to compression or from compression to traction at each lumbar level**

| Measure | Level | Mean ± SD difference U to C | Correlation (r) of difference from U to C with IDD | Mean ± SD difference C to T | Correlation (r) of difference from C to T with IDD |
|---|---|---|---|---|---|
| **Disc** | | | | | |
| Mean $T_2$ (ms) | L1-2 | 3±11 | -0.32 | -5±11 | -0.07 |
| | L2-3 | 3±15 | **-0.45**** | 0±19 | -0.08 |
| | L3-4 | -4±16 | **-0.51**** | 6±15 | 0.25 |
| | L4-5 | -2±16 | **-0.42*** | 5±16 | 0.28 |
| | L5-S1 | -2±14 | -0.28 | 2±13 | -0.01 |
| Horizontal coordinate of $T_2$WC (mm) | L1-2 | 1.3±2.6 | **-0.45**** | 0.0±2.9 | **-0.46**** |
| | L2-3 | 2.4±6.9 | -0.05 | 0.2±2.7 | 0.22 |
| | L3-4 | 0.6±2.1 | -0.05 | 0.5±2.6 | 0.30 |
| | L4-5 | -0.2±3.7 | 0.07 | 1.7±3.7 | -0.33 |
| | L5-S1 | -0.6±2.4 | **-0.36*** | 1.5±2.4 | 0.19 |
| Vertical coordinate of $T_2$WC (mm) | L1-2 | -0.2±0.6 | **0.49**** | 0.1±0.6 | -0.10 |
| | L2-3 | 0.4±3.5 | -0.04 | 0.2±0.8 | **-0.45**** |
| | L3-4 | 0.2±0.9 | 0.10 | 0.0±0.9 | 0.01 |
| | L4-5 | 0.0±0.7 | 0.20 | 0.3±0.7 | -0.25 |
| | L5-S1 | -0.5±1.4 | 0.08 | 1.2±1.5 | -0.12 |
| Height (mm) | L1-2 | 0.1±0.3 | **-0.33*** | 0.2±0.3 | 0.14 |
| | L2-3 | 0.0±0.4 | -0.19 | 0.1±0.3 | -0.04 |
| | L3-4 | 0.1±0.6 | 0.09 | 0.0±0.7 | -0.13 |
| | L4-5 | 0.1±0.5 | 0.02 | 0.2±0.6 | -0.09 |
| | L5-S1 | -0.2±0.7 | 0.33 | 0.5±0.8 | -0.29 |
| Width (mm) | L1-2 | 0.2±3.2 | -0.05 | -1.0±2.5 | 0.15 |
| | L2-3 | 0.6±3.2 | 0.04 | -1.0±2.5 | -0.04 |
| | L3-4 | -0.5±2.6 | 0.03 | -0.7±2.7 | -0.01 |
| | L4-5 | 0.2±3.9 | 0.31 | 0.0±3.9 | **-0.36*** |
| | L5-S1 | 1.9±4.1 | 0.30 | -1.9±3.1 | -0.09 |
| Nucleus | | | | | |
| Mean $T_2$ (ms) | L1-2 | 5±11 | -0.03 | -7±15 | -0.22 |
| | L2-3 | 4±16 | -0.32 | -2±19 | 0.04 |
| | L3-4 | -5±19 | **-0.36*** | 8±19 | 0.22 |
| | L4-5 | -1±15 | -0.28 | 4±17 | 0.14 |
| | L5-S1 | -1±19 | -0.18 | 3±16 | 0.12 |
| Horizontal | L1-2 | 0.7±1.8 | -0.18 | 0.2±1.8 | **0.42*** |
| | L2-3 | 2.2±7.0 | -0.07 | 0.1±1.8 | 0.14 |

| Measure | Level | Mean ± SD difference U to C | Correlation (r) of difference from U to C with IDD | Mean ± SD difference C to T | Correlation (r) of difference from C to T with IDD |
|---|---|---|---|---|---|
| coordinate of $T_2$WC (mm) | L3-4 | 1.1±2.1 | -0.11 | 0.2±2.3 | **0.39*** |
| | L4-5 | -0.8±3.3 | 0.08 | 2.0±3.4 | -0.21 |
| | L5-S1 | -1.0±3.9 | -0.21 | 1.3±2.5 | **0.36*** |
| Vertical coordinate of $T_2$WC (mm) | L1-2 | -0.1±0.5 | 0.24 | 0.1±0.6 | -0.15 |
| | L2-3 | 0.5±3.5 | -0.03 | 0.2±0.7 | **-0.51**** |
| | L3-4 | 0.2±0.8 | 0.12 | -0.1±0.7 | 0.01 |
| | L4-5 | -0.2±1.0 | 0.05 | 0.4±1.1 | -0.16 |
| | L5-S1 | -0.8±2.4 | 0.01 | 1.1±1.9 | 0.07 |
| Horizontal coordinate of GWC (mm) | L1-2 | 0.4±2.3 | -0.14 | 0.2±2.1 | **0.40*** |
| | L2-3 | 2.2±7.2 | -0.07 | 0.0±2.2 | 0.02 |
| | L3-4 | 1.3±2.2 | -0.02 | -0.1±2.7 | 0.23 |
| | L4-5 | -1.0±3.0 | 0.04 | 2.0±3.2 | -0.19 |
| | L5-S1 | -0.9±3.8 | -0.21 | 1.2±2.3 | **0.39*** |
| Vertical-coordinate of GWC (mm) | L1-2 | 0.0±0.6 | 0.13 | 0.1±0.7 | -0.19 |
| | L2-3 | 0.5±3.5 | -0.04 | 0.2±0.8 | **-0.48**** |
| | L3-4 | 0.2±0.8 | 0.12 | -0.1±0.8 | 0.02 |
| | L4-5 | -0.2±1.0 | 0.02 | 0.4±1.1 | -0.13 |
| | L5-S1 | -0.7±2.2 | -0.01 | 1.1±1.8 | 0.12 |
| Width (mm) | L1-2 | -6.6±25.7 | -0.20 | -0.4±2.9 | -0.11 |
| | L2-3 | 0.1±4.3 | 0.01 | -0.4±2.9 | -0.25 |
| | L3-4 | -1.2±5.2 | 0.05 | 0.8±4.8 | -0.14 |
| | L4-5 | -0.9±4.4 | -0.11 | -0.3±5.8 | 0.11 |
| | L5-S1 | -0.4±4.1 | 0.12 | -0.2±4.6 | -0.03 |
| Motion Segment | | | | | |
| Angle° | L1-2 | 1±4 | -0.07 | 0±3 | 0.11 |
| | L2-3 | 0±3 | -0.12 | -3±3 | 0.17 |
| | L3-4 | 2±4 | -0.19 | -1±4 | 0.09 |
| | L4-5 | 2±4 | 0.32 | -1±3 | -0.24 |
| | L5-S1 | 0±4 | -0.05 | 1±3 | 0.04 |

* $p<0.05$, ** $p<0.01$

Abbreviations: BW: Body weight, GWC: geometric weighted centroid. $T_2$WC: $T_2$weighted centroid, mm: Millimetre

***Correlation among changes in the quantitative measurements from compression to traction at L5-S1***. A correlation corresponding to over 50% shared variance was observed between the horizontal coordinates of the nucleus GWC and $T_2$WC ($r_X$=0.93), and with the vertical coordinates of the GWC ($r_X$=0.95) (table 19). There was also a high correlation between horizontal

and vertical coordinates of the nucleus $T_2$WC (r=0.78) and between the horizontal and vertical coordinates of the GWC (*r*=0.75), as well as horizontal coordinates of the nucleus $T_2$WC and vertical coordinates GWC (*r*=0.71). The vertical coordinates of the nucleus $T_2$WC and horizontal coordinates of the GWC were also highly correlated (*r*=0.73).

**Table 19 Pearson's correlation coefficients among changes in the quantitative measurements from compression to traction at L5-S1**

| | | | Mean $T_2$ U | Disc changes in | | | | | Nucleus changes in | | | | | | MSA changes |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | Mean | Height | Width | $T_2$WC | | Mean | Width | $T_2$WC | | GWC | | |
| | | | | | | | H | V | | | V | H | H | V | |
| Mean $T_2$ U | | | 1 | -0.01 | -0.29 | -0.09 | 0.19 | -0.12 | 0.12 | -0.03 | 0.07 | 0.36* | 0.39* | 0.12 | 0.04 |
| Disc Changes in | Mean | | | 1 | -0.07 | -0.50** | 0.1 | -0.03 | 0.73** | -0.04 | 0.3 | 0.36* | 0.35* | 0.26 | 0.04 |
| | Height | | | | 1 | -0.16 | -0.03 | 0.53** | -0.11 | 0.09 | 0.37* | 0.04 | 0.12 | 0.41* | 0.1 |
| | Width | | | | | 1 | -0.34* | -0.24 | -0.54** | 0.19 | -0.51** | -0.61** | -0.54** | -0.44** | 0.19 |
| | $T_2$WC | H | | | | | 1 | 0.67** | 0.31 | 0.12 | 0.51** | 0.70** | 0.52** | 0.39* | 0.06 |
| | | V | | | | | | 1 | 0.1 | 0.08 | 0.71** | 0.52** | 0.41* | 0.61** | 0.29 |
| Nucleus changes in | Mean | | | | | | | | 1 | -0.12 | 0.25 | 0.41* | 0.35* | 0.18 | -0.05 |
| | Width | | | | | | | | | 1 | -0.04 | -0.09 | -0.01 | 0.01 | 0.12 |
| | $T_2$WC | V | | | | | | | | | 1 | 0.78** | 0.73** | 0.95** | 0.11 |
| | | H | | | | | | | | | | 1 | 0.93** | 0.71** | 0.1 |
| | GWC | H | | | | | | | | | | | 1 | 0.75** | 0.12 |
| | | V | | | | | | | | | | | | 1 | 0.07 |
| MSA changes | | | | | | | | | | | | | | | 1 |

* $p<0.05$, ** $p<0.01$

Abbreviations: GWC: geometric weighted centroid. $T_2$WC: $T_2$weighted centroid, mm: Millimeter; V: vertical; H: Horizontal; MSA: motion segment angle; U:unloaded

**Discussion**

This clinical study examined the response of lumbar discs and motion segments to identify biomarkers which may help define subgroups with different degeneration phenotypes by using loading during MR imaging. The significant responses of lumbar motion segments to loading conditions were most evident at the lower lumbar motion segments where the spine was more degenerated. It was also more consistent with the hypothesised response of the lumbar spine based on a theoretical model of compression and traction, i.e. compression reduced or did not change the disc height and expelled fluid out from the discs while traction increased disc height and brought fluid into the discs. The behaviour of the upper lumbar motion segments (L1-2 and L2-3) was not consistent with that theoretical model.

As hypothesised, the pairwise comparisons between compression and traction demonstrated more significant findings than between unloaded and compression suggesting that this loading contrast may be more effective to detect variations in response to loading. The location of the disc and nucleus $T_2$WC and GWC, the motion segment angle and disc height were the most promising biomarkers for evaluating the effects of loading on the lumbar spine. Response to loading was further shown to relate to degeneration severity suggesting many of the investigated measurements show promise as biomarkers for degeneration.

Results support our hypothesis that the fluid distribution related measurements are more sensitive to capture the effects of degeneration and loading as the majority of the statistically significant changes and largest effect sizes were observed for $T_2$WC parameters of both the disc and nucleus. The $T_2$WC is a unique point corresponding to the mean position of all pixels of an ROI, weighted based on their $T_2$ and changes in its position reflect the distribution of fluids in the ROI. Likewise, a significant posterior and inferior shift of the nucleus GWC was also observed at lower motion segments when the lumbar spine underwent traction from compression. Any variation in the location of the GWC reflects changes in the geometry and/or location of the ROI.

However, when the lumbar spine underwent compression from unloaded, the fluid shift ($T_2$WC) in the disc and nucleus was small and not statistically significant at L5-S1. Further, only, at L3-4, a significant posterior fluid shift was observed in the nucleus. This suggests that

comparing pairs of extreme loading conditions (compression vs traction) was best at detecting the response to loading.

The $T_2$ value of a tissue reflects the fluid content of the tissue [37]. However, the mean signal intensity in $T_2$-weighted MR images which ware widely used has a limitation because it depends on the distance of the tissue from receiving coils [226], the magnetic field inhomogeneities and scanner characteristics [37]. In contrast, the mean $T_2$ measured in $T_2$-maps acquired in this study is independent of these factors and is, therefore, potentially more useful for inter-scanner, between-subjects and within-subject repeated image comparisons.

To the best of our knowledge, the in-vivo effect of traction on the IVD and nucleus fluid content had not been investigated using MRI $T_2$ measurements. Traction can decrease intradiscal pressure below -100 mmHg [227]. This negative pressure would bring fluid within the disc and nucleus. Consistent with this, the calculated effect sizes indicated that the fluid content of the IVDs and nucleuses of the lower lumbar levels increased during traction, however, differences did not reach significance. This may be due to small sample size, or insufficient duration and magnitude of loading. Although, previous studies suggested that 10 minutes of unloading was enough for the spine to recover its height [228–230], a longer loading time may result in larger fluid content change. However, increasing loading duration may limit the clinical applicability. Further, clinical traction recommendations typically recommend using 50%BW to ensure good tolerance [220, 231]. It is unclear whether higher loads would be tolerable.

Surprisingly, at L1-2 the disc $MT_2$ decreased significantly in traction and the nucleus $MT_2$ increased significantly in compression. The IVD and nucleus are expected to lose or not change fluid content when undergoing compression and to regain it when loads are reduced [32]. Although these findings were not consistent with this hypothesis, this observation was consistent with the results of Nilsson et al. indicated an increase in the disc $MT_2$ after lumbar loading with 50% BW [127]. Nillsson et al. results were however, conflicting with ours and others' for the lower lumbar levels [22, 26]. The intended uniaxial compression load may have resulted in a complex loading condition throughout the spine.

A complex mechanical interaction between the nucleus and the surrounding annulus with loading has been suggested to induce tensile hoop stresses within the annulus, causing axial bulging [232–236]. Our results were consistent with this suggestion; the disc width increased in

compression at most levels and decreased at all levels in traction. Although the directions of the changes observed for disc height and for disc and nucleus width at L5-S1 were consistent with previous studies comparing compression to supine unloaded [44, 167], most were not statistically significant. This suggests that these morphologic measurements are not as sensitive to detect response to loading as the fluid distribution parameters. Madsen et al.'s indicated a significant decrease of the cumulative disc height (mean 1.4 mm) using compression in a supine extended position similar to the present study and compared to unloaded supine [83]. Nazari et al., in 25 homogeneous participants without pain, indicated a significant increase in the anterior-posterior diameter of the disc and nucleus when going from supine unloaded to sitting or standing [167]. This may be due to different loading conditions, 50% body weight vs. standing which is almost 5 times bigger than lying unloaded [237], or more heterogeneity among participants in the present study. Spinal loads are also higher during sitting and standing compared to axial loading in a supine position [41, 42].

Our findings for the angles of lumbar motion segments were consistent with the theoretical models suggesting that compression increases the motion segment angle producing an increased lordosis, and traction decreases it by flattening the lumbar spine. However, it was only statistically significant at L4-5 from unloaded to compression. We could not identify other studies on the effect of axial loading on segment angles. The angular changes demonstrate that our uniaxial application of loading to the torso resulted in complex lumbar load transfers.

The significant correlations between disc degeneration and other changes in measurements observed from unloaded to compression and from compression to traction for the disc height, $MT_2$, width, and $T_2WC$, nucleus $MT_2$, $T_2WC$ and GWC, and motion segment angle indicated that changes in response to the loading likely hold promise and are sensitive enough for studying the prognosis of disc degeneration or the response of the degenerated discs to treatments. The response to loading decreased with increasing disc degeneration, which was consistent with our a priori hypothesis. However, the clinical significance of changes of the GWC and $T_2WC$ remains unclear as such changes have not been studied for their impact on symptoms or related to progression in degeneration or disc pathology measurements.

The high significant correlation that was observed between the coordinates of the nucleus $T_2$WC and GWC at L5-S1 indicates that these measures may be used interchangeably and possibly provide redundant information. Given that $T_2$WC provides information on the spatial orientation of the ROI and the fluid distribution of each pixel, it may be more informative than the GWC and more sensitive to minor changes. Our correlation estimates suggest that both disc and nucleus $T_2$WC provide complementary information for evaluating the effect of degeneration as shared variance corresponded to less than 50% for the correlation between the corresponding horizontal coordinates of each ROI.

A key strength of this investigation was the use of $T_2$ maps, which can provide an estimate of the fluid content of each voxel. We used a semi-automated segmentation method leading to high measurement reliability as well as a standardised loading protocol applied during imaging with controlled pre-load. Nevertheless, some limitations should be acknowledged. We employed a relatively small sample, which presented heterogeneity by combining participants with and without back pain with a range of ages. However, we accounted for age and body size differences in the analyses by using age and vertebra size as covariates. Our volunteer sample included a limited number of discs with severe degeneration, which contains less fluid and which, based on our correlation estimates, would reduce the disc response to loading. Many variables were examined in this preliminary study and some statistically significant results may be due to chance. A replication study, justified by the promising results, is needed to confirm which biomarkers are truly useful to detect degeneration difference. With our pixel size, we could capture changes within the sagittal plane of 0.49 mm of larger. However, a smaller pixel would limit the signal to noise ratio (SNR). To maximize the SNR, we also chose sagittal slices with five mm thickness limiting the ability to capture changes laterally less than five millimetres in magnitude. We only measured the midsagittal slice as each parasagittal $T_2$-map requires a similar scanning time, which would increase the total scanning time. Changes in other regions of the disc may have been missed.

**Conclusion**

To our knowledge, this was the first study evaluating the effects of traction and compression taking advantage of quantitative $T_2$-mapping. We examined the effects of loading on geometric and fluid content and distribution measurements, including the recently introduced GWC and $T_2$WC. These biomarkers hold promise for future investigation as objective

measurements of degeneration. Disc's height, width, $MT_2$ and $T_2WC$ as well as nucleus's $MT_2$, and $T_2WC$ hold promise as sensitive biomarkers to capture the response of the lumbar spine to different loading conditions and possibly also to monitor degenerative and regenerative changes of the lumbar motion segments. Traction compared to compression was more sensitive to capture the load-induced changes. However, the clinical value of the proposed measurement constructs still remains unclear and future studies are needed to determine if the responses to loading define degeneration subgroups with different pathoanatomy, prognosis or response to treatments.

**Acknowledgement**

Vahid Abdollah was a recipient of the University of Alberta Doctoral Recruitment Scholarship, and of the Doctoral Fellowship from the Department of Physical Therapy of the University of Alberta. All authors were recipients of a research grant provided by the Alberta Spine Foundation for the development of a program for automatic quantitative measurements of lumbar discs and vertebrae. Data collection was conducted using grants from the Canadian Physiotherapy Association and the Department of Physical Therapy at the University of Alberta.

# CHAPTER 8 THE EFFECTS OF COMPRESSION AND TRACTION ON LUMBAR MRI FINDINGS IN RELATION TO CHRONIC LOW BACK PAIN

**Vahid Abdollah; Eric C Parent; Alex Su, Keith Wachowicz, Michele C Battié**

**Abstract**

**Background:** Low back pain (LBP) is a major cause of adult disability globally. Conventional MRI, which is a sensitive modality to depict lumbar pathoanatomy, is often clinically inconclusive due to lack of specificity. This is likely due to imaging in a relaxed supine position while pain is often experienced during spine loading. This study aimed to investigate the response of the lumbar spine to compression and traction in participants with and without chronic LBP using MRI $T_2$-mapping. Results would identify imaging biomarkers that hold promise to advance diagnosis, prognosis and meaningful subgrouping of LBP.

**Methods:** Fifteen patients with chronic LBP (30.2±10.6yrs, 71.4±13.5kg) were matched for age, weight and gender with 15 healthy volunteers (29.7±10.6 yrs., 72.9±12.7l kg). The participant spent the first 20 minutes lying in a relaxed unloaded supine position, followed by 20 minutes loaded in compression with 50% body weight, and then 20 minutes in traction with 50% body weight applied. Mid-sagittal $T_2$-weighted MR images were acquired during the last five minutes of each loading period. The whole disc, nucleus, and vertebrae from L1 to L5 were semi-automatically segmented on mid-sagittal $T_2$-weighted MR images and measurements extracted from corresponding $T_2$-maps. For each segmented region, the mean $T_2$ time, geometric weighted centroid (GWC), and $T_2$ weighted centroid ($T_2$WC) were calculated. Disc height and width, nucleus width, and motion segment angle were also computed. A two-way repeated model of analysis of variance with Sidak's post-hoc comparisons was conducted for comparing groups and loading conditions. Cohen's d effect sizes were also calculated for differences between loading conditions.

**Findings:** The interaction between loading and pain were most often observed at L5-S1. The interaction between pain and loading was significant for the horizontal coordinate of the disc $T_2$WC at L5-S1, the vertical coordinate of the disc $T_2$WC at L3-4 and L5-S1 and of the nucleus at L5-S1 as well as for disc height at L1-2. The main effect of pain was significant on the horizontal coordinate of the disc $T_2$WC at L3-4 and L4-5, horizontal coordinate of the nucleus $T_2$WC at L4-5, the horizontal coordinate of the nucleus GWC at L3-4 and L4-5 and L5-S1 and disc height at L3-4, and L4-5. The widths and mean $T_2$ for both disc and nucleus ROIs presented the smallest effects for loading and pain.

**Interpretation:** The response of lumbar motion segments to loading conditions detected more differences between participants with and without LBP than comparisons only of measurement from unloaded images. Biomarkers reflecting an anterior-posterior shift in fluid distribution within discs and nucleus and disc height and their response to loading showed promise to improve the specificity of MRI for LBP.

**Keywords:** magnetic resonance imaging; low back pain; compression; traction; $T_2$ map; $T_2$-weighted centroid; disc height; degeneration; intervertebral disc

## Introduction

Low back pain (LBP) is a major cause of discomfort and disability among adults globally [1]. While any spinal structure that is innervated and susceptible to painful disease or injury could be a possible source of LBP [3], little is known about the specific pathology underlying LBP. In over 85% of LBP cases, despite using advanced imaging techniques to investigate pathoanatomy, the cause of the pain remains unclear (e.g. non-specific LBP) [4].

Intervertebral disc degeneration (IDD) has been implicated as a major factor in the development of LBP [5–7]. A degenerated disc is a dehydrated, discoloured and fragmented structure [54]. The IDD is thus associated with morphological, chemical, metabolic, and biomechanical changes which affect disc function and biomechanics [5]. With disc degeneration the fragmented proteoglycans begin to diffuse out of the nucleus, reducing the negative charge density of the disc, disc osmotic pressure and thus disc fluid content [61] and the normal loading transfer in the motion segment [5]. Degeneration not only affects the magnitude, but also the distribution of fluid in the disc [5]. In addition to changes within the disc, IDD has also been associated with changes in the biomechanics of the motion segments [5].

Diagnostic imaging has been used for diagnosing the underlying pathology of pain [9]. Although conventional MRI is a sensitive modality for depicting structural abnormalities of the lumbar spine, it is often clinically inconclusive due to a lack of specificity with many pathoanatomical findings being nearly as prevalent in those with and without LBP [11]. The lack of specificity and inability to conclusively identify a pathoanatomical cause of LBP may be due to conducting MR imaging in a relaxed supine position. Symptoms are usually worse in a sitting or standing posture compared to a supine relaxed position [43, 44].

Loading affects the morphology of the lumbar spine [161, 219], but our systematic review has shown that there is a gap in understanding if the spine of participants with and without pain responds differently to loading [222]. The review also showed that traction loading has not been investigated quantitatively to date. The fluid content and swelling pressure of the intervertebral discs (IVD) are a function of the interplay between external loads, tension in the collagen fibril networks, non-fibrillar matrix composition and the osmolarity of surrounding media [32]. The IVD is a viscoelastic tissue, therefore, the morphological characteristics of the IVD change in response to loading depending mainly on its fluid content and flow [32].

The changes in the disc in response to degeneration [19, 20], and loading [21–27] have been investigated in vivo quantitatively using magnetic resonance (MR) imaging. $T_2$ relaxation time ($T_2$RT) is the rate at which transverse magnetisation is decreased and disappears [28–31]. Signal intensity on $T_2$-weighted images varies among different MR scanners due to field characteristics, sequence parameters, and patient positioning relative to receiving coils. In contrast, $T_2$ times from $T_2$ maps are independent of these methodological differences. $T_2$ times correlate well with fluid and proteoglycan content of the IVDs [37]. The mean $T_2$ ($MT_2$) time from $T_2$-maps represent a unique tool for comparing repeated scans under different loading conditions and positions. Quantitative $T_2$ measurements are likely helpful to detect subtle fluid and biochemical changes within discs that may not be apparent with qualitative or semi-quantitative measures [32]. $T_2$-mapping can quantitatively evaluate the fluid content of the IVDs [27, 33–40]. The effects of diurnal loading on the IVD fluid contents using $T_2$-mapping has been quantitatively investigated [21–27]. The results indicated a decline in the IVD fluid content after loading the spine in the upright posture during the day [26, 27, 29, 31]. Chokan et al. also indicated a significant decrease of $T_2$ values in the nucleus after exercise and a significant increase after rest [27].

In-vivo measurements of the IVD and vertebra morphology or IVD fluid content under different loading conditions have not been investigated extensively despite the fact that changes in the fluid distribution within the lumbar spine in response to extension loading have shown an association with changes in pain symptoms [45]. Only, participants who reported an immediate reduction in LBP intensity after posterior-to-anterior–directed pressures, and prone extension exercises, showed an increase in diffusion of fluid in the nuclear region of L5-S1 [45, 238].

Pain has also been suggested to relate to the changes induced in the spine due to compression loading, including increased intra-discal pressure [41, 42], decreased disc height [83, 84], lumbar instability [85–87], and possibly subsequent stenosis of the intervertebral foramen affecting the neurovascular structures [88], excessive disc or vertebral deformation [89], increased lumbar lordosis [83, 88–91], and subsequent stenosis of the vertebral canal [92].

For this study, we hypothesized that variations observed in the magnitude and distribution pattern of fluid within the IVDs, nucleus and vertebra (measured by $T_2$-weighted centroid ($T_2$WC), geometric weighted centroid (GWC) and $MT_2$), as well as changes in segmental alignment or morphology of the lumbar spine structures (measured by height, width, and segment angle) during

imaging under compression and traction loading, would differ between participants with and without persistent LBP.

The main purpose of the present study was to compare the response of the lumbar IVDs and vertebrae, to compression and traction loading in participants with and without chronic LBP using MRI $T_2$-mapping. We hypothesised that the response of the candidate biomarkers to compression and traction loading would be systematically different in participants with and without chronic LBP. This study, thereby, aimed to identify imaging biomarkers that hold promise to increase the specificity for LBP for further investigation to advance diagnosis, prognosis and meaningful subgrouping of LBP.

**Materials and methods**

This case-control study consisted of a single MRI testing session to compare the effects of unloaded, compression and traction loading conditions on the lumbar spine.

**Participants.** Fifteen consecutive volunteer patients with chronic LBP (6 females, 9 males) were matched for age (±5 yrs.), weight (±5 kg) and gender with 15 healthy volunteers. The participants were recruited using the mass email service to students of our university. The inclusion criteria for both groups were 18 to 65 years of age and the ability to read and understand English. Participants with claustrophobia, any contraindications for MRI or traction, tumour, uncontrolled hypertension, severe osteoporosis, and previous spine surgery were excluded. The inclusion criteria for participants with LBP were an Oswestry Disability Index score of 25% or greater and being in pain for at least half of the days in the past six months before imaging [49]. We excluded participants with confirmed causes of back pain (malignancy, fracture, etc.) and referred pain from other sources; significant compression of the spinal cord/ nerves; on-going workers' compensation or litigation cases and having received injections within the last 12 weeks. The inclusion criteria for volunteers without LBP were no current back pain and no memorable (severe or disabling) back pain lasting more than one day over the past 10 years. Participants with chronic LBP were asked to indicate the intensity of current pain level on a scale of 0 (no pain) to 10 (worst pain imaginable) using a numeric pain rating scale (NPRS). This study was approved by the Health Research Ethics Board of the University of Alberta.

G*Power 3.1 software (University of Kiel, Kiel, Germany) was used to calculate the required sample size to secure a power of 0.80 using an $\alpha$ value of 0.02 (corrected for comparing

three loading conditions for a unilateral test), to detect an effect size corresponding to the mean difference ± standard deviation observed in the disc height following 6 weeks of motorized traction with 50% BW [220]. This study was the only study providing preliminary traction loading response estimates at the time. The required sample size was 20 participants in each group. With 15 participants in each group, an effect size of 0.78 could be detected as statistically significant. This number of subjects was deemed sufficient for this preliminary study to provide stable effect size estimates to identify promising biomarkers for further investigation in future studies.

**MRI image acquisition procedures.** Mid-sagittal $T_2$-weighted MR images were acquired during the last five minutes of each loading period applied in the following order for all subjects: 20 minutes lying in a relaxed unloaded supine position, 20 minutes loaded in compression with 50% body weight (BW), and 20 minutes in traction with 50% BW (Figure 12). Mid-sagittal slices positioned by the technician based on where the localizer sequence illustrated the presence of the spinous processes and where a clear demarcation of the spinal cord could be observed on $T_2$-weighted images [180]. In cases where a lumbar curve was present, the localizer was placed to ensure capturing the L4-5 disc and spinous processes. Loadings were applied using a custom designed MRI compatible loading table. A compressive load equal to 50% of the subject's BW was applied to partially simulate loading documented in the erect spine [221]. A traction force of 50% BW was applied to investigate the effects of unloading (traction) on study outcomes [220].

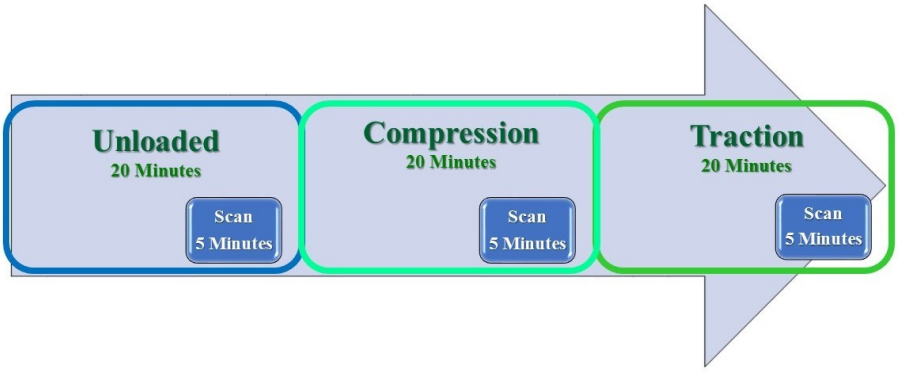


Figure 12 A schematic presentation of loading and imaging workflow

Compression force was applied by a deadweight of 50% BW tied to a rope tensioning a harness strapped around the chest with straps applying load over the shoulders. Loads were left in place during the first 5 minutes of loading with the rope free to move through a pulley at the foot of the device (Figure 13). Then, the rope was locked in place by a compression screw and the

patient was moved into the MRI while still under loading. For traction, the chest harness was released from the foot of the device and anchored by straps at the head of the device to fix the upper spine. Traction loading was applied via straps pulling on a harness securely tightened around the pelvis by applying a deadweight of 50% BW for 5 minutes and then the tensioned cable was locked in place (Figure 13). To ensure that friction did not limit the application of loading force, the pelvis and lower extremities rested on the lower half of the table which was resting on wheels and free to move in the cephalad-caudal direction.

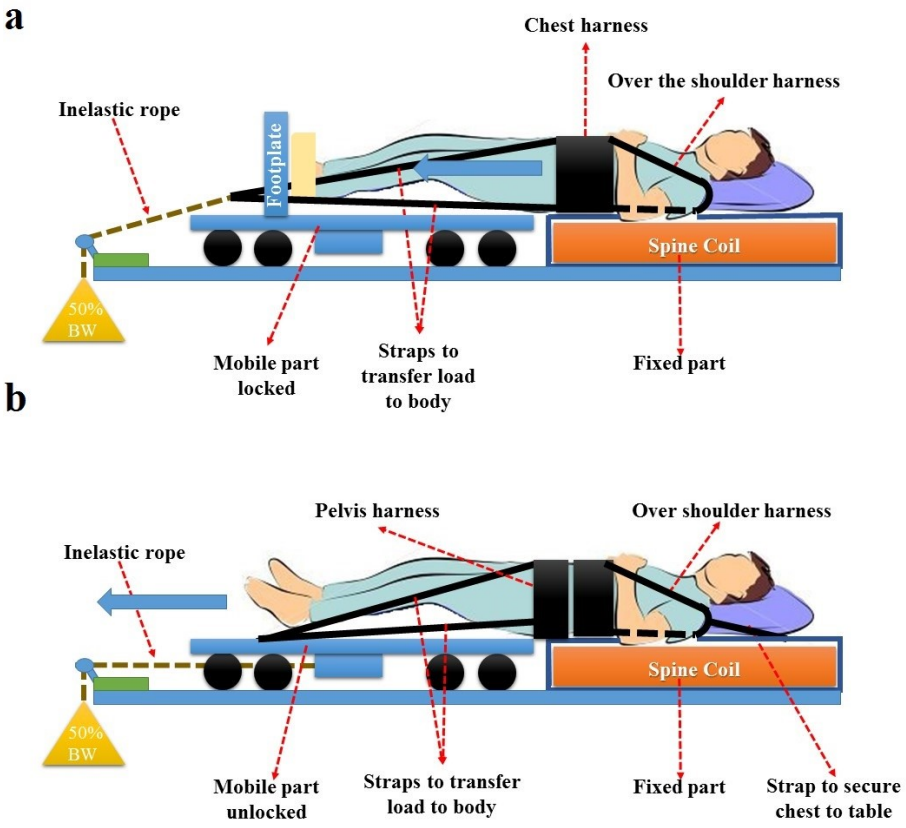


Figure 13 Schematic representation of the loading table a:) compression b:) traction

Images were acquired using a 3T MR scanner (Philips Healthcare Intera whole body MRI scanner). The $T_2$-weighted image acquisition parameters used to obtain a $T_2$ map were TR 2500, TE (five echoes: 16.9, 44, 71, 98 and 125 ms), two averages, field of view of 500×500 mm, pixel size 0.49×0.49 mm, image thickness 5mm and a matrix size of 1024×1024 pixel. The location of mid-sagittal slices was determined using the localizer sequence to select the location best illustrating the spinous processes and with a clear demarcation of the spinal cord [180]. If a lumbar

lateral curve was present, the localizer was placed to ensure capturing the L4-5 disc and spinous processes.

***Image processing and disc height measurements.*** Image post-processing was carried out offline using a computer aided measurement program developed by our team (Appendix I). A graduate student (AS) completed the semi-automatic image segmentation following program demonstration by a team member and practising segmentation of 60 sets of images. $T_2$-maps were constructed automatically by obtaining $T_2$ relaxation time estimates for a given pixel using the Nelder-Mead Simplex method [223].

The software was used to semi-automatically segment each lumbar motion segment by identifying boundaries displaying maximal $T_2$ signal intensity differences into the following regions of interest (ROIs): the whole disc, the hyperintense nucleus region (hereafter referred to as the nucleus) and vertebra (Appendix I). The semi-automated segmentation procedures have been reported previously [224]. The $T_2$-weighted image corresponding to the third echo (TE=71ms) was used for image segmentation and each ROI was then reconstructed on the corresponding T2-map to extract measurements of interest. Three parameters were calculated for each vertebra and nucleus ROI: 1) mean $T_2$ time ($MT_2$), 2) horizontal (X) and vertical (Y) coordinates of the geometric weighted centroid (GWC), and 3) of $T_2$WC obtained using Equation 11 and Equation 12, respectively [214]. The $MT_2$ and location of $T_2$WC were calculated for IVDs.

Equation 7. Mathematical description of $T_2$WC

$$T_2WC = \begin{bmatrix} X = \frac{\sum_{i=1}^{n} X_i T_{2i}}{\sum_{i}^{n} T_{2i}} \\ Y = \frac{\sum_{i=1}^{n} Y_i T_{2i}}{\sum_{i}^{n} T_{2i}} \end{bmatrix}$$

Equation 8. Mathematical description of GWC

$$GWC = \begin{bmatrix} X = \frac{\sum_{i=1}^{n} X_i A_i}{\sum_{i}^{n} A_i} \\ Y = \frac{\sum_{i=1}^{n} Y_i A_i}{\sum_{i}^{n} A_i} \end{bmatrix}$$

Where $X_i$ and $Y_i$, are the horizontal and vertical coordinates of each pixel within an ROI, and $A_i$ and $SI_i$ are the area of each ROI and $T_2$ of each pixel, respectively. As the location of the lumbar spine varied relative to the centre point of the image in different loading conditions, the

centre of the Cartesian coordinate system was placed at the GWC of the superior vertebra for each motion segment. All GWC and $T_2$WC measurements for the motion segment were determined relative to that point. The negative horizontal and vertical coordinate values indicate more anterior and superior position relative to the reference vertebra's GWC, respectively.

Disc height was also measured automatically using an area-based method [197] by dividing the area of the disc by the corresponding diameter representing 80% of the disc width centred at the mid-point of the whole disc. The angle between the best-fit lines passing through the segmented contour of the superior and inferior border of the disc was computed as the motion segment angle. The width of the disc and nucleus along the disc bisector line was also computed automatically. The intra and inter – rater (ICC) reliability estimate for the motion segment angle as well as for the disc and nucleus width ranged from 0.75 to 0.95. All other measurement constructs were found to have good to excellent intra- and inter-rater reliability, with reliability coefficients (ICC) ranging from 0.88 to 1.00 [224].

***Disc degeneration grading.*** The degree of disc degeneration was assessed for each disc using the 5-point scale proposed by Pfirrmann et al [16] through consensus by two raters (VA and EP) with more than five years of experience in spine imaging. Degeneration was graded using the $T_2$-weighted image with a TE of 71 ms.

***Data analysis.*** The analysis was carried out at each level independently because disc size and biomechanical response are likely different at each level. The normality of the data was then examined using Shapiro-Wilk test for each measure. A mixed model analysis of variance (ANOVA) was employed to compare groups (pain vs no pain) and the different loading conditions (unloaded, compression, traction) [152]. There was no adjustment for covariates as participants were matched for gender, age and weight. Sidak's post-hoc comparisons were used to determine the significance of differences between baseline and compression and between compression and traction [152]. Statistical analyses were carried out using SPSS®, version 23.0 (IBM Corp. Armonk, NY., USA) using a level of statistical significance set at 0.05.

The effect size (ES) of differences between pairs of consecutive loading conditions was estimated using Cohen's d (Equation 6) for each pain subgroup.

Equation 9. Mathematical description of Cohen's D

$$d = \frac{(M_B - M_A)}{\sqrt{\frac{SD_B^2 + SD_A^2}{2}}}$$

Where $M_B$ and $M_A$ are the means before and after the change in loading condition and $SD_B$ and $SD_A$ are standard deviations of the measurement before and after the loading change. The Cohen's d effect size was also calculated for the magnitude of differences between groups observed for the unloaded image. A Cohen's d between 0.2 and 0.5 was considered a small effect, a Cohen's between 0.5 and 0.8 considered a medium effect, and a Cohen's d over 0.8 considered a large effect size [215]. Effect sizes smaller than 0.2 were referred to as trivial. The mean and SD of change between pairs consecutive of loading conditions, baseline vs. compression and compression vs. traction were also estimated with 95% confidence interval (CI).

**Results**

The mean age of the 15 participants with LBP was 30.2±10.6 [18 to 53] years, and their mean weight was 75.6±14.3 [55-96] kg for females and 69.4±12.4 [55-96] kg) for males. The average numeric pain rating scores (out of 10 points) before imaging was 3±3 in the LBP group and 0±1 in controls. In the supine unloaded position, mean pain was 3±2 points in the LBP group and 0±1 in controls. After 20 minutes under compression, pain was 4±2 points in the LBP group and 1±1 in controls. Finally, after 20 minutes under traction pain was 4±2 points in the LBP group and 1±1 in controls. The mean age of the 15 participants without LBP was 29.7±10.6 [18 to 53] years and their mean weight was 77.3±13.5 [55-96] kg for females and 71.0±10.7 [55-96] kg) for males. Thirty discs and vertebrae were analysed at each level from L1-2 to L5-S1 for each loading condition.

In participants with LBP, no discs (0.0%) were rated as grade V (advanced degeneration with a collapsed disc space) (table 20). Pfirrmann grade II (52 discs, 69.3%), III (14 discs, 17.3%) and IV (eight discs, 10.7%) presented the highest frequency, respectively (table 20). For the control group, of the total of 75 discs, the sample included two discs (2.7%, at L2-3 and L3-4) with Pfirrmann grade I, and no disc (0.0%) rated as grade V (table 20). Pfirrmann grade II (49 discs, 65.3%), III (14 discs, 18.7%) and IV (10 discs, 13.3%) had the highest frequency, respectively (table 20). In both groups, the lower levels presented the most severe degeneration. The distribution of the degeneration grades was similar at all levels (table 20).

**Table 20 Distribution of Pfirrmann's disc degeneration classification scores at each level for both groups**

| Level | Group | Grade I | Grade II | Grade III | Grade IV | Grade V |
|---|---|---|---|---|---|---|
| L1-2 | Pain | 0 | 14 | 1 | 0 | 0 |
| | Control | 0 | 11 | 2 | 2 | 0 |
| L2-3 | Pain | 1 | 14 | 0 | 0 | 0 |
| | Control | 1 | 14 | 0 | 0 | 0 |
| L3-4 | Pain | 1 | 12 | 2 | 0 | 0 |
| | Control | 1 | 12 | 2 | 0 | 0 |
| L4-5 | Pain | 0 | 7 | 6 | 2 | 0 |
| | Control | 0 | 7 | 6 | 2 | 0 |
| L5-S1 | Pain | 0 | 5 | 4 | 6 | 0 |
| | Control | 0 | 5 | 4 | 6 | 0 |
| Total | Pain | 2 | 52 | 13 | 8 | 0 |
| | Control | 2 | 49 | 14 | 10 | 0 |
| Percentage | Pain | 2.7% | 69.3% | 17.3% | 10.7% | 0.0% |
| | Control | 2.7% | 65.3% | 18.7% | 13.3% | 0.0% |

***Interaction between pain and loading.*** Four biomarkers presented significant interactions between pain and loading showing the potential to detect differences between pain groups by comparing responses to loading: disc height (at L1-2), the horizontal (at L5-S1) and vertical (at L3-4, L5-S1) coordinates of the disc $T_2$WC, and the vertical coordinates of the Nucleus $T_2$WC (at L5-S1) (table 21). Differences in effect sizes for the response to a pair of loading condition between groups were more often larger for the compression to traction conditions. Overall, differences between compression and traction showed the largest effect sizes compared to differences between unloaded and compression (13 vs. 8 significant).

More specifically, at L5-S1 the disc $T_2$WC in the pain group showed a moderate anterior shift (ES:-0.50) when going from unloaded to compression, while it showed a small posterior shift (ES:0.20) in the no-pain group (Figure 14 a) (table 21). When going from compression to traction, the disc $T_2$WC in the pain group showed a moderate posterior shift (ES:0.66), while its location did not significantly change (ES:0.00) in the no-pain group (Figure 14 b).

In the pain group, from unloaded to compression, the disc $T_2$WC presented a moderate anterior shift at L5-S1 (ES:-0.50), while it showed a trivial posterior shift in the no-pain group (ES:0.20) (Figure 14 a). In the pain group, from compression to traction, the disc $T_2$WC presented a moderate posterior shift at L5-S1 (ES:0.66), while going from compression to traction had no effect on the horizontal location of the $T_2$WC in the no-pain group (trivial ES:0.20) (Figure 14 b).

In the pain group, from unloaded to compression, the disc $T_2$WC presented a moderate inferior shift (ES:0.50) at L3-4 and a small superior shift (ES:-0.33) at L5-S1 (Figure 14 c). In the pain group, when going from compression to traction, the disc $T_2$WC showed a moderate superior shift (ES:-0.50) at L3-4 and a small inferior shift (ES:0.33) at L5-S1 (Figure 14 d). In the pain group, from unloaded to compression, the nucleus $T_2$WC presented a moderate superior shift (ES: -0.57) (*Figure 14* e) and from compression to traction, showed a moderate inferior shift (ES: 0.50) at L5-S1 (Figure 14). In contrast, going from unloaded to compression (Figure 14) or from compression to traction (Figure 14) had no effect on the location of the vertical location of the disc $T_2$WC in the no-pain group at L3-4 and for the disc and nucleus at L5-S1 (ES:0.00).

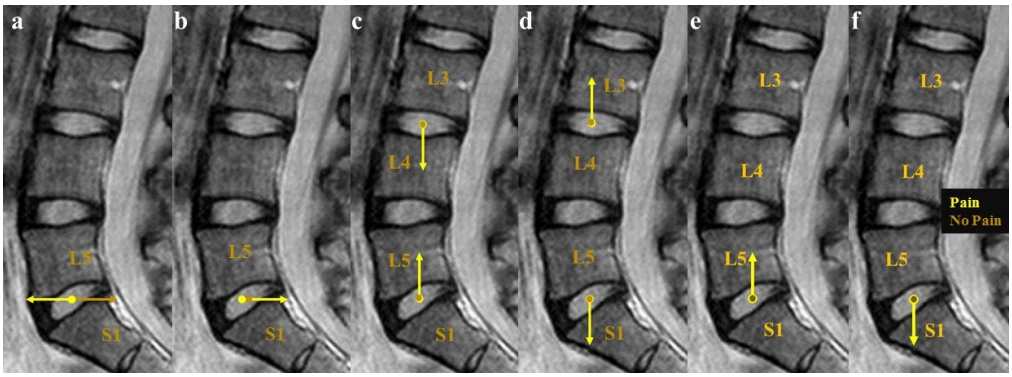


Figure 14 Significant interaction effects between pain and loading for: a:) horizontal coordinate of the disc $T_2$WC when going from unloaded to compression b:) horizontal coordinate of the disc $T_2$WC when going from compression to traction c:) vertical coordinate of the disc $T_2$WC when going from unloaded to compression d:) vertical coordinate of the disc $T_2$WC when going from compression to traction e:) vertical coordinate of the nucleus $T_2$WC when going from unloaded to compression f:) vertical coordinate of the disc $T_2$WC when going from compression to traction (images only illustrate direction and location of the effects and are not scaled)

At L1-2, in both group, disc height presented a statistically significant increase but a trivial effect size when going from unloaded to compression ($ES_{LBP}$:0.10; $ES_C$:0.07) (Figure 15 a). The disc height change was not statistically significant in the pain group (ES:0.00) when going from compression to traction, while a statistically significant increase was observed in those with no pain (trivial ES:0.13) (Figure 15 b).

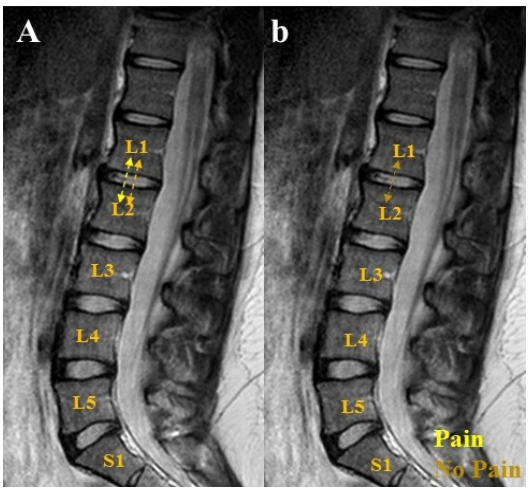


Figure 15 Significant interaction effects between pain and loading for the disc height a:) when going from unloaded to compression b:) when going from compression to traction (images only illustrate direction and location of the effects and are not scaled)

***Main effect of pain.*** Four biomarkers showed a main effect of pain demonstrating the potential to detect differences between groups by comparing a single image: disc height, horizontal coordinates of the disc $T_2$WC and of the nucleus GWC (table 22). The disc $T_2$WC at L3-4 (ES:0.73) and L4-5 (ES: 0.59) and the nucleus $T_2$WC at L4-5 (ES: 0.80) of the pain group were moderately to largely more anterior than in the control group (Figure 16; table 21). Likewise, the nucleus GWC of the pain group was moderately to largely more anterior (ES:0.62:0.83) than for the control group at L3-4, L4-5 and L5-S1 (Figure 16; table 21). The disc height of the control group was smaller than the pain group at L3-4 and L4-5 (Figure 16; table 21) by a large effect size (ES:-0.86:-1.01). The main effect of pain observed for the other levels and variables in table 21 did not reach statistical significance.

***The main effect of loading.*** The following variables presented only a main effect of loading showing an inability to detect differences between those with and without symptoms: disc and nucleus mean $T_2$ at L1-2, vertical coordinates of the nucleus GWC at L5-S1, motion segment angles at L3-4 and L4-5. Further, for the following variables which at some levels detected differences between pain groups, the main effect of loading was also significant: the horizontal coordinate of the disc $T_2$WC at L1-2, L4-5 and L5-S1, vertical coordinate of the disc $T_2$WC at L5-S1, disc height at L1-2, disc height and width at L5-S1, horizontal coordinate of the nucleus $T_2$WC and GWC at L2-3 and L4-5 and the nucleus GWC at L3-4, as well as vertical coordinate of the nucleus $T_2$WC at L5-S1. Overall, the number of significant pairwise comparisons when going from

compression to traction was larger than when going from unloaded to traction (8 vs. 3 significant) (table 21). The main effect of loading observed for the other levels and variables in table 21 did not reach statistical significance.

**Table 21 Difference between loading conditions and between two groups in morphological measurements, mean $T_2$, location of the coordinates of the geometric and $T_2$ weighted centroids.**

| Measure | Level | Group | Unloaded | 50% BW Comp. | 50% BW Trac. | Effect size Comp vs. Unloaded | Effect size Trac vs. Comp | *ANOVA p-value Pain* | *ANOVA p-value Loading* | *ANOVA p-value Interaction* |
|---|---|---|---|---|---|---|---|---|---|---|
| **Disc** | | | | | | | | | | |
| Horizontal coordinate $T_2$WC (mm) | L1-2 | Pain | 8.0±0.7 | 9.0±1.0 | 8.7±0.5 | 0.28 | 0.00 | *0.22* | ***0.04*** | *0.66* |
| | | Control | 6.4±0.7 | 8.0±0.9 | 8.0±0.7 | 0.28 | 0.00 | | | |
| | L2-3 | Pain | 8.2±0.9 | 9.1±0.6 | 9.3±0.7 | 0.32 | 0.00 | *0.31* | *0.06* | *0.23* |
| | | Control | 4.8±2.8 | 9.1±0.8 | 9.5±0.7 | 0.50* | 0.00 | | | |
| | L3-4 | Pain | 5.3±0.8 | 5.4±1.1 | 6.6±0.7 | 0.00 | 0.57* | ***0.03*** | *0.10* | *0.18* |
| | | Control | 7.5±0.8 | 8.4±0.7 | 7.9±0.7 | 0.33 | 0.00 | | | |
| | L4-5 | Pain | -1.6±1.1 | -2.1±1.0 | 0.4±1.0 | 0.00 | 0.50* | ***0.04*** | ***0.00*** | *0.30* |
| | | Control | 0.9±1.0 | 1.7±1.1 | 2.6±1.0 | 0.25 | 0.25 | | | |
| | L5-S1 | Pain | -12.8±1.0 | -14.6±1.0 | -12.3±1.2[b] | -0.50* | 0.66* | *0.07* | ***0.02*** | ***0.00*** |
| | | Control | -10.9±1.2 | -10.1±1.2 | -9.6±1.2[b] | 0.20 | 0.00 | | | |
| Vertical coordinate $T_2$WC (mm) | L1-2 | Pain | 34.4±0.5 | 33.3±0.5 | 33.3±0.5 | -0.07 | -0.01 | *0.82* | *0.15* | *0.33* |
| | | Control | 33.7±0.9 | 33.4±0.9 | 33.6±0.9 | -0.33 | 0.33 | | | |
| | L2-3 | Pain | 35.6±0.5 | 35.7±0.5 | 35.6±0.5 | 0.00 | 0.00 | *0.64* | *0.39* | *0.39* |
| | | Control | 34.4±1.4 | 35.5±0.8 | 35.8±0.8 | 0.49 | 0.00 | | | |
| | L3-4 | Pain | 37.1±0.5 | 37.5±0.4 | 37.3±0.5 | 0.50* | -0.50* | *0.65* | *0.34* | ***0.03*** |
| | | Control | 37.0±0.7 | 36.8±0.7 | 37.3±0.7 | 0.00 | 0.00 | | | |
| | L4-5 | Pain | 38.1±0.5 | 38.0±0.4 | 38.4±0.5 | 0.00 | 0.00 | *0.72* | *0.07* | *0.85* |
| | | Control | 38.3±0.6 | 38.3±0.6 | 38.5±0.6 | 0.00 | 0.00 | | | |
| | L5-S1 | Pain | 34.5±0.7 | 33.5±0.9 | 35.1±0.8[b] | -0.33 | 0.33 | *0.18* | ***0.00*** | ***0.00*** |
| | | Control | 35.6±0.8 | 35.7±0.9 | 36.3±0.9[b] | 0.00 | 0.00 | | | |
| Mean $T_2$ | L1-2 | Pain | 131±3 | 136±3[a] | 130±3[b] | 0.38 | -0.48 | *0.62* | ***0.00*** | *0.99* |

| Measure | Level | Group | Unloaded | 50% BW Comp. | 50% BW Trac. | Effect size Comp vs. Unloaded | Effect size Trac vs. Comp | *ANOVA p-value Pain* | *ANOVA p-value Loading* | *ANOVA p-value Interaction* |
|---|---|---|---|---|---|---|---|---|---|---|
| (ms) | | Control | 134±5 | 139±5[a] | 132±4[b] | 0.24 | -0.31 | | | |
| | L2-3 | Pain | 141±4 | 140±4 | 145±4 | -0.07 | 0.34 | *0.47* | *0.49* | *0.50* |
| | | Control | 136±4 | 141±5 | 139±4 | 0.28 | -0.12 | | | |
| | L3-4 | Pain | 144±5 | 140±5 | 146±5 | -0.21 | 0.31 | *0.60* | *0.18* | *0.98* |
| | | Control | 142±4 | 137±5 | 142±5 | -0.28 | 0.26 | | | |
| | L4-5 | Pain | 131±6 | 129±5 | 137±7 | -0.09 | 0.34 | *0.58* | *0.17* | *0.47* |
| | | Control | 134±6 | 136±5 | 137±5 | 0.10 | 0.05 | | | |
| | L5-S1 | Pain | 125±6 | 121±6 | 126±5 | -0.19 | 0.24 | *0.32* | *0.49* | *0.56* |
| | | Control | 118±5 | 117±5 | 118±5 | -0.05 | 0.05 | | | |
| 80% Disc width height (mm) | L1-2 | Pain | 9.0±0.3 | 9.1±0.3 | 9.1±0.3 | 0.10 | 0.00 | *0.24* | ***0.00*** | ***0.03*** |
| | | Control | 8.5±0.4 | 8.6±0.4 | 8.8±0.4 | 0.07 | 0.13 | | | |
| | L2-3 | Pain | 10.3±0.3 | 10.4±0.3 | 10.4±0.3 | 0.10 | 0.00 | *0.17* | *0.10* | *0.28* |
| | | Control | 9.6±0.4 | 9.7±0.5 | 9.8±0.5 | 0.06 | 0.05 | | | |
| | L3-4 | Pain | 11.1±0.3 | 11.4±0.3 | 11.2±0.3 | 0.23 | -0.17 | ***0.02*** | *0.13* | *0.08* |
| | | Control | 9.9±0.4 | 9.9±0.4 | 10.2±0.4 | 0.00 | 0.19 | | | |
| | L4-5 | Pain | 11.7±0.2 | 11.8±0.2 | 11.8±0.2 | 0.00 | 0.00 | ***0.03*** | *0.09* | *0.28* |
| | | Control | 10.6±0.3 | 10.7±0.3 | 10.9±0.3 | 0.07 | 0.15 | | | |
| | L5-S1 | Pain | 10.1±0.3 | 10.8±0.3 | 11.2±0.3[b] | -0.10 | 0.40 | *0.09* | ***0.00*** | *0.90* |
| | | Control | 10.0±0.5 | 9.9±0.5 | 10.2±0.5[b] | 0.00 | 0.10 | | | |
| Disc Width (mm) | L1-2 | Pain | 71.7±1.7 | 72.4±1.7 | 71.4±1.7 | 0.09 | -0.13 | *0.95* | *0.28* | *0.72* |
| | | Control | 72.1±1.7 | 71.8±2.0 | 71.2±1.8 | -0.04 | -0.08 | | | |
| | L2-3 | Pain | 73.9±1.6 | 75.2±1.6 | 73.3±1.4 | 0.21 | -0.34 | *0.78* | *0.25* | *0.15* |
| | | Control | 74.5±1.6 | 74.5±1.6 | 74.7±1.7 | 0.00 | 0.03 | | | |
| | L3-4 | Pain | 75.2±1.5 | 74.6±1.4 | 74.5±1.5 | -0.09 | -0.04 | *0.85* | *0.14* | *0.79* |
| | | Control | 75.7±1.9 | 75.1±1.8 | 74.4±1.9 | -0.08 | -0.10 | | | |

| Measure | Level | Group | Unloaded | 50% BW Comp. | 50% BW Trac. | Effect size Comp vs. Unloaded | Effect size Trac vs. Comp | *ANOVA p-value Pain* | *ANOVA p-value Loading* | *ANOVA p-value Interaction* |
|---|---|---|---|---|---|---|---|---|---|---|
| | L4-5 | Pain | 77.1±1.9 | 78.1±2.0 | 77.5±2.2 | 0.13 | -0.07 | *0.73* | *0.87* | *0.35* |
| | | Control | 76.9±2.2 | 76.4±1.8 | 77.2±2.1 | -0.06 | 0.11 | | | |
| | L5-S1 | Pain | 73.9±1.6 | 75.8±1.7 | 74.2±1.7 | 0.30 | -0.25 | *0.58* | ***0.04*** | *0.86* |
| | | Control | 75.4±2.5 | 77.0±2.6 | 75.4±2.6 | 0.16 | -0.16 | | | |
| **Nucleus** | | | | | | | | | | |
| Horizontal coordinate $T_2$WC (mm) | L1-2 | Pain | 6.7±0.7 | 7.3±0.8 | 7.5±0.7 | 0.00 | 0.33 | *0.38* | *0.05* | *0.65* |
| | | Control | 5.8±0.7 | 6.7±0.9 | 6.4±0.7 | 0.28 | -0.28 | | | |
| | L2-3 | Pain | 7.2±0.8 | 8.2±0.7 | 8.4±0.7 | 0.33 | 0.00 | *0.19* | ***0.04*** | *0.36* |
| | | Control | 3.8±2.8 | 7.8±0.7 | 7.8±0.7 | 0.54* | 0.00 | | | |
| | L3-4 | Pain | 4.4±0.9 | 5.2±1.1 | 5.9±0.7 | 0.28 | 0.28 | *0.30* | *0.34* | *0.10* |
| | | Control | 6.3±0.7 | 8.1±0.7 | 7.7±0.6 | 0.67* | 0.00 | | | |
| | L4-5 | Pain | -1.9±1.0 | -3.2±1.0 | -0.4±0.9[b] | -0.25 | 0.85** | ***0.02*** | ***0.00*** | *0.13* |
| | | Control | 1.2±1.0 | 0.7±1.1 | 1.6±0.8[b] | 0.00 | 0.28 | | | |
| | L5-S1 | Pain | -12.4±1.1 | -14.6±1.4 | -12.5±1.2 | -0.66* | 0.40 | *0.05* | *0.18* | *0.12* |
| | | Control | -9.9±1.1 | -9.6±1.1 | -9.4±1.1 | 0.00 | 0.25 | | | |
| Vertical coordinate $T_2$WC (mm) | L1-2 | Pain | 33.3±0.4 | 33.2±0.4 | 33.3±0.4 | 0.00 | 0.00 | *0.62* | *0.08* | *0.37* |
| | | Control | 33.7±0.8 | 33.5±0.8 | 33.8±0.8 | -0.33 | 0.33 | | | |
| | L2-3 | Pain | 35.3±0.5 | 35.4±0.5 | 35.3±0.5 | 0.50* | -0.50* | *0.84* | *0.41* | *0.38* |
| | | Control | 34.3±1.4 | 35.4±0.8 | 35.8±0.7 | 0.24 | 0.33 | | | |
| | L3-4 | Pain | 36.9±0.5 | 37.2±0.4 | 37.0±0.4 | 0.00 | 0.00 | *0.80* | *0.47* | *0.28* |
| | | Control | 36.8±0.7 | 36.8±0.7 | 37.0±0.8 | 0.00 | 0.00 | | | |
| | L4-5 | Pain | 37.5±0.5 | 37.2±0.5 | 37.8±0.5 | -0.50* | 0.50* | *0.27* | *0.11* | *0.34* |
| | | Control | 38.2±0.6 | 38.0±0.6 | 38.1±0.6 | 0.00 | 0.00 | | | |
| | L5-S1 | Pain | 33.7±0.7 | 32.3±1.1 | 34.2±0.9[b] | -0.57* | 0.50* | *0.08* | ***0.03*** | ***0.03*** |

| Measure | Level | Group | Unloaded | 50% BW Comp. | 50% BW Trac. | Effect size Comp vs. Unloaded | Effect size Trac vs. Comp | *ANOVA p-value Pain* | *ANOVA p-value Loading* | *ANOVA p-value Interaction* |
|---|---|---|---|---|---|---|---|---|---|---|
| | | Control | 35.6±0.8 | 35.6±0.9 | 35.7±0.9[b] | 0.00 | 0.00 | | | |
| Horizontal coordinate GWC (mm) | L1-2 | Pain | 6.8±0.7 | 7.2±0.9 | 7.5±0.8 | 0.00 | 0.28 | *0.68* | *0.33* | *0.55* |
| | | Control | 6.4±0.8 | 7.2±0.9 | 6.6±0.7 | 0.28 | 0.00 | | | |
| | L2-3 | Pain | 6.6±0.8 | 7.9±0.8 | 7.9±0.7 | 0.33 | 0.00 | *0.36* | ***0.04*** | *0.50* |
| | | Control | 4.0±2.8 | 7.8±0.7 | 7.7±0.7 | 0.41 | 0.00 | | | |
| | L3-4 | Pain | 3.3±0.9 | 4.5±1.1[a] | 4.9±0.9 | 0.25 | 0.00 | ***0.01*** | ***0.01*** | *0.26* |
| | | Control | 6.0±0.7 | 7.8±0.8[a] | 7.0±0.6 | 0.67* | -0.39 | | | |
| | L4-5 | Pain | -2.7±1.0 | -4.1±1.1 | -1.5±0.9 | -0.28 | 0.57* | ***0.02*** | ***0.01*** | *0.09* |
| | | Control | 0.3±1.1 | -0.1±1.0 | 0.7±0.9 | 0.00 | 0.25 | | | |
| | L5-S1 | Pain | -12.8±1.1 | -14.7±1.4 | -12.8±1.1 | -0.44 | 0.40 | ***0.03*** | *0.39* | *0.11* |
| | | Control | -10.2±1.1 | -9.6±1.1 | -9.8±0.9 | 0.00 | 0.00 | | | |
| Vertical coordinate GWC (mm) | L1-2 | Pain | 33.3±0.4 | 33.1±0.4 | 33.2±0.4 | 0.00 | 0.00 | *0.61* | *0.36* | *0.53* |
| | | Control | 33.6±0.8 | 33.4±0.8 | 33.8±0.8 | -0.33 | 0.33 | | | |
| | L2-3 | Pain | 35.3±0.5 | 35.4±0.5 | 35.2±0.5 | 0.00 | 0.00 | *0.90* | *0.36* | *0.30* |
| | | Control | 34.2±1.4 | 35.4±0.8 | 35.8±0.7 | 0.24 | 0.33 | | | |
| | L3-4 | Pain | 36.8±0.5 | 37.1±0.4 | 36.9±0.4 | 0.00 | 0.00 | *0.94* | *0.77* | *0.29* |
| | | Control | 36.9±0.7 | 36.8±0.7 | 36.9±0.7 | 0.00 | 0.00 | | | |
| | L4-5 | Pain | 37.4±0.5 | 37.0±0.5 | 37.7±0.5 | 0.00 | 0.50* | *0.19* | *0.15* | *0.20* |
| | | Control | 38.1±0.6 | 38.0±0.5 | 38.0±0.6 | 0.00 | 0.00 | | | |
| | L5-S1 | Pain | 33.6±0.7 | 32.4±1.0 | 34.2±0.8[b] | -0.67* | 0.67* | *0.06* | ***0.04*** | *0.06* |
| | | Control | 35.5±0.7 | 35.5±0.8 | 35.7±0.8[b] | 0.00 | 0.00 | | | |
| Mean $T_2$ (ms) | L1-2 | Pain | 164±5 | 170±7 | 163±5[b] | 0.26 | -0.31 | *0.67* | ***0.02*** | *0.86* |
| | | Control | 162±9 | 166±8 | 159±7[b] | 0.12 | -0.24 | | | |
| | L2-3 | Pain | 175±6 | 179±7 | 180±5 | 0.17 | 0.04 | *0.13* | *0.38* | *0.75* |
| | | Control | 166±6 | 171±6 | 168±6 | 0.21 | -0.16 | | | |

| Measure | Level | Group | Unloaded | 50% BW Comp. | 50% BW Trac. | Effect size Comp vs. Unloaded | Effect size Trac vs. Comp | *ANOVA p-value Pain* | *ANOVA p-value Loading* | *ANOVA p-value Interaction* |
|---|---|---|---|---|---|---|---|---|---|---|
| | L3-4 | Pain | 180±9 | 176±9 | 182±10 | -0.11 | 0.16 | *0.20* | *0.05* | *0.67* |
| | | Control | 167±6 | 161±7 | 173±9 | -0.23 | 0.39 | | | |
| | L4-5 | Pain | 156±11 | 155±10 | 158±10 | -0.02 | 0.08 | *0.90* | *0.64* | *0.77* |
| | | Control | 156±8 | 158±8 | 159±8 | 0.06 | 0.03 | | | |
| | L5-S1 | Pain | 151±11 | 145±10 | 152±11 | -0.15 | 0.17 | *0.29* | *0.60* | *0.39* |
| | | Control | 137±8 | 139±10 | 139±8 | 0.06 | 0.00 | | | |
| Width (mm) | L1-2 | Pain | 23.9±0.9 | 22.4±1.0 | 23.3±1.2 | -0.42 | 0.21 | *0.25* | *0.19* | *0.94* |
| | | Control | 24.8±1.1 | 22.4±1.2 | 24.6±1.1 | -0.31 | 0.28 | | | |
| | L2-3 | Pain | 25.3±0.9 | 25.7±0.9 | 25.3±1.2 | 0.11 | -0.09 | *0.77* | *0.92* | *0.80* |
| | | Control | 25.6±1.1 | 25.0±0.9 | 25.0±0.8 | -0.15 | 0.00 | | | |
| | L3-4 | Pain | 25.7±0.9 | 26.3±1.0 | 26.1±1.1 | 0.16 | -0.05 | *0.78* | *0.34* | *0.36* |
| | | Control | 26.0±1.1 | 24.2±1.3 | 25.0±1.0 | -0.37 | 0.17 | | | |
| | L4-5 | Pain | 26.1±1.1 | 24.6±1.4 | 25.8±1.2 | -0.32 | 0.23 | *0.40* | *0.48* | *0.11* |
| | | Control | 25.6±0.8 | 25.7±0.9 | 23.3±0.7 | 0.03 | -0.76* | | | |
| | L5-S1 | Pain | 25.3±1.1 | 24.6±1.3 | 24.7±1.1 | -0.15 | 0.02 | *0.84* | *0.36* | *0.66* |
| | | Control | 25.2±0.7 | 24.9±1.1 | 23.8±0.7 | -0.08 | -0.31 | | | |
| **Motion Segment** | | | | | | | | | | |
| Angle$^{o}$ | L1-2 | Pain | 3.4±0.9 | 4.1±0.8 | 3.6±0.6 | 0.23 | -0.18 | *0.13* | *0.62* | *0.08* |
| | | Control | 2.3±0.7 | 3.0±0.9 | 2.9±0.9 | 0.24 | -0.03 | | | |
| | L2-3 | Pain | 4.9±0.6 | 4.9±0.6 | 5.2±0.8 | -0.05 | 0.15 | *0.80* | *0.62* | *0.32* |
| | | Control | 4.8±0.9 | 5.9±0.5 | 4.9±0.8 | 0.38 | -0.38 | | | |
| | L3-4 | Pain | 7.3±0.7 | 8.9±1.2[a] | 7.6±0.8 | 0.42 | -0.33 | *0.47* | ***0.02*** | *0.63* |
| | | Control | 7.4±0.9 | 9.5±0.8[a] | 8.9±1.1 | 0.63* | -0.21 | | | |
| | L4-5 | Pain | 10.1±0.8 | 13.3±0.8 | 11.8±0.8 | 0.74* | -0.48 | *0.74* | ***0.01*** | *0.60* |
| | | Control | 11.0±0.8 | 12.3±0.9 | 11.8±0.8 | 0.39 | -0.16 | | | |

| Measure | Level | Group | Unloaded | 50% BW Comp. | 50% BW Trac. | Effect size Comp vs. Unloaded | Effect size Trac vs. Comp | *ANOVA p-value Pain* | *ANOVA p-value Loading* | *ANOVA p-value Interaction* |
|---|---|---|---|---|---|---|---|---|---|---|
| | L5-S1 | Pain | 15.3±1.0 | 14.2±1.5 | 16.7±1.0 | -0.23 | 0.53 | *0.76* | *0.09* | *0.17* |
| | | Control | 15.0±1.0 | 16.2±1.4 | 16.5±1.2 | 0.26 | 0.06 | | | |

Negative horizontal means located anteriorly relative to geometric centroid of the vertebra above.
Positive change in the vertical direction means moved inferiorly relative to geometric centroid of the vertebra above.
* Medium effect size
** Large effect size
[a] significant difference between baseline and compression
[b] significant difference between compression and traction
Bold values represent $P_{value}<0.05$
Abbreviations: BW: Body weight; GWC: geometric weighted centroid; $T_2$WC: $T_2$weighted centroid; mm: Millimetre; Comp: Compression; Trac: Traction;

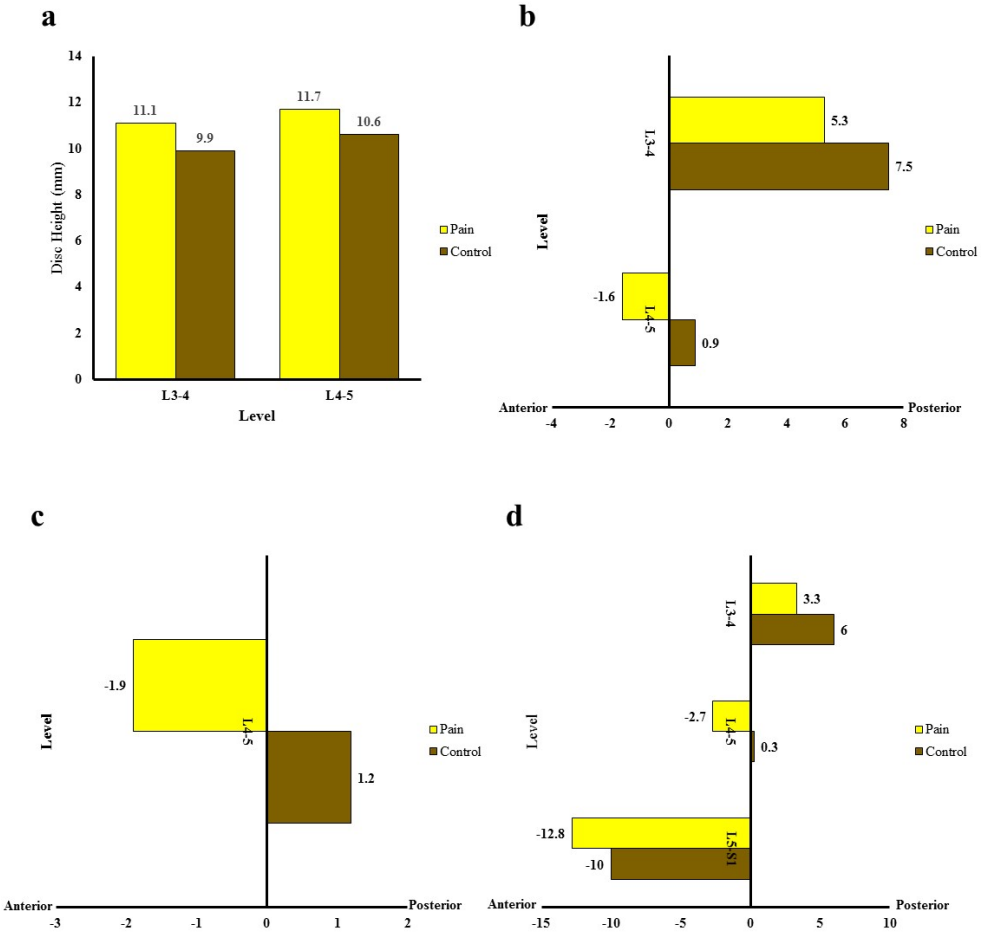


Figure 16 The statistically significant main effects of pain a:) horizontal location of the disc $T_2$WC b:) horizontal location of the nucleus $T_2$WC c:) horizontal location of the nucleus GWC d:) disc height using 80% disc width

**Table 22 Effect sizes for the difference between two groups for the unloaded condition**

| Measure | Level | Pain | Control | Effect size Pain vs. Control |
|---|---|---|---|---|
| Disc | | | | |
| Horizontal coordinate of $T_2$WC (mm) | L1-2 | 8.0±0.7 | 6.4±0.7 | -0.58* |
| | L2-3 | 8.2±0.9 | 4.8±2.8 | -0.43 |
| | L3-4 | 5.3±0.8 | 7.5±0.8 | 0.73* |
| | L4-5 | -1.6±1.1 | 0.9±1.0 | 0.59* |
| | L5-S1 | -12.8±1.0 | -10.9±1.2 | 0.47 |
| Vertical coordinate of $T_2$WC (mm) | L1-2 | 34.4±0.5 | 33.7±0.9 | 0.08 |
| | L2-3 | 35.6±0.5 | 34.4±1.4 | -0.30 |
| | L3-4 | 37.1±0.5 | 37.0±0.7 | -0.04 |
| | L4-5 | 38.1±0.5 | 38.3±0.6 | 0.09 |
| | L5-S1 | 34.5±0.7 | 35.6±0.8 | 0.36 |
| Mean $T_2$ (ms) | L1-2 | 131±3 | 134±5 | 0.17 |
| | L2-3 | 141±4 | 136±4 | -0.31 |
| | L3-4 | 144±5 | 142±4 | -0.14 |
| | L4-5 | 131±6 | 134±6 | 0.09 |
| | L5-S1 | 125±6 | 118±5 | -0.33 |
| Height (mm) | L1-2 | 9.0±0.3 | 8.5±0.4 | -0.42 |
| | L2-3 | 10.3±0.3 | 9.6±0.4 | -0.52* |
| | L3-4 | 11.1±0.3 | 9.9±0.4 | -0.86** |
| | L4-5 | 11.7±0.2 | 10.6±0.3 | -1.01*** |
| | L5-S1 | 10.1±0.3 | 10.0±0.5 | -0.57* |
| Width (mm) | L1-2 | 71.7±1.7 | 72.1±1.7 | 0.06 |
| | L2-3 | 73.9±1.6 | 74.5±1.6 | 0.09 |
| | L3-4 | 75.2±1.5 | 75.7±1.9 | 0.07 |
| | L4-5 | 77.1±1.9 | 76.9±2.2 | -0.02 |
| | L5-S1 | 73.9±1.6 | 75.4±2.5 | 0.19 |
| Nucleus | | | | |
| Horizontal coordinate of $T_2$WC (mm) | L1-2 | 6.7±0.7 | 5.8±0.7 | -0.33 |
| | L2-3 | 7.2±0.8 | 3.8±2.8 | -0.43 |
| | L3-4 | 4.4±0.9 | 6.3±0.7 | 0.62* |
| | L4-5 | -1.9±1.0 | 1.2±1.0 | 0.80** |
| | L5-S1 | -12.4±1.1 | -9.9±1.1 | 0.56* |
| Vertical coordinate of $T_2$WC (mm) | L1-2 | 33.3±0.4 | 33.7±0.8 | 0.14 |
| | L2-3 | 35.3±0.5 | 34.3±1.4 | -0.27 |
| | L3-4 | 36.9±0.5 | 36.8±0.7 | -0.03 |
| | L4-5 | 37.5±0.5 | 38.2±0.6 | 0.31 |
| | L5-S1 | 33.7±0.7 | 35.6±0.8 | 0.61* |
| Horizontal coordinate of GWC | L1-2 | 6.8±0.7 | 6.4±0.8 | -0.14 |
| | L2-3 | 6.6±0.8 | 4.0±2.8 | -0.32 |
| | L3-4 | 3.3±0.9 | 6.0±0.7 | 0.83** |

| Measure | Level | Pain | Control | Effect size Pain vs. Control |
|---|---|---|---|---|
| (mm) | L4-5 | -2.7±1.0 | 0.3±1.1 | 0.75* |
| | L5-S1 | -12.8±1.1 | -10.2±1.1 | 0.62* |
| Vertical coordinate of GWC (mm) | L1-2 | 33.3±0.4 | 33.6±0.8 | 0.13 |
| | L2-3 | 35.3±0.5 | 34.2±1.4 | -0.26 |
| | L3-4 | 36.8±0.5 | 36.9±0.7 | 0.04 |
| | L4-5 | 37.4±0.5 | 38.1±0.6 | 0.37 |
| | L5-S1 | 33.6±0.7 | 35.5±0.7 | 0.69 |
| Mean $T_2$ (ms) | L1-2 | 164±5 | 162±9 | -0.10 |
| | L2-3 | 175±6 | 166±6 | -0.44 |
| | L3-4 | 180±9 | 167±6 | -0.53* |
| | L4-5 | 156±11 | 156±8 | -0.02 |
| | L5-S1 | 151±11 | 137±8 | -0.37 |
| Nucleus Width (mm) | L1-2 | 23.9±0.9 | 24.8±1.1 | 0.23 |
| | L2-3 | 25.3±0.9 | 25.6±1.1 | 0.06 |
| | L3-4 | 25.7±0.9 | 26.0±1.1 | 0.07 |
| | L4-5 | 26.1±1.1 | 25.6±0.8 | -0.16 |
| | L5-S1 | 25.3±1.1 | 25.2±0.7 | -0.01 |
| Motion Segment | | | | |
| Angle (º) | L1-2 | 3.4±0.9 | 2.3±0.7 | -0.39 |
| | L2-3 | 4.9±0.6 | 4.8±0.9 | -0.04 |
| | L3-4 | 7.3±0.7 | 7.4±0.9 | 0.02 |
| | L4-5 | 10.1±0.8 | 11.0±0.8 | 0.02 |
| | L5-S1 | 15.3±1.0 | 15.0±1.0 | -0.07 |

* Medium effect size
** Large effect size
*** Effect size more than one standard deviation

**Discussion**

The present study was innovative in developing new measurements for evaluating the effects of loading in participants with and without chronic low back pain using real-time quantitative $T_2$-mapping. The disc height, horizontal and vertical coordinates of the disc $T_2$WC and the vertical coordinates of the nucleus $T_2$WC were four biomarkers that indicated significant interactions between pain and loading. The horizontal coordinates of the disc and nucleus $T_2$WC, the horizontal coordinates of the nucleus GWC and disc height were four biomarkers which demonstrated a main effect of pain suggesting that they may detect differences between those with and without pain using traditional supine unloaded MRI scan. The response of the lumbar motion segments to three loading conditions (unloaded, compression and traction) was different between participants with and without pain. Specifically, significant interactions between pain and loading

were most often observed at L5-S1 consistent with the beliefs that the lower lumbar levels are most affected by degeneration and implicated in causing LBP [239].

Both $T_2$WC and GWC are highly reliable biomarkers, which suggest they are not affected by subtle inter-user segmentation differences. As their measurement error is low, these biomarkers show increased sensitivity to smaller differences. $T_2$ relaxation time is a scalar quantity and only indicates change in fluid content; in contrast, $T_2$WC is a vector able to capture the direction of change in the fluid distribution independent of changes in fluid content. $T_2$WC is thus more sensitive to changes in response to loading and likely to the differences between subjects with and without pain. The $T_2$WC location may also provide information about the role of the vascular supply adjacent to the disc on the nutrition or of intradiscal fissures, as any deviation in $T_2$WC in response to loading indicates changes corresponding to fluid redistribution. The disc $T_2$WC of the pain group was located more anteriorly than in the controls at L3-4 and L4-5. Changes in hydration, as well as collagen anisotropy, have been reported to be early indicators of cartilage deterioration [112]. This likely indicates that discs are more anisotropic in the people with LBP. This may affect fluid distribution and thus hydrostatic pressure of the disc and cause uneven load distribution on the pain-sensitive endplates or annulus fibres. Trauma and inflammation may trigger synthesis of factors such as bradykinin and prostanoids within the disc [6]. These factors are thought to be able to sensitise silent nociceptors, which are usually unresponsive to even maximal mechanical stimulation and thus minimise their activation threshold [6]. The sensitised nociceptors could then respond to changes in intradiscal pressure during movement and trigger pain [6].

Any displacement of the nucleus $T_2$WC and GWC with loading likely indicate a deformable nucleus hyper-intense area contour. The nucleus GWC of the pain group was located more anteriorly than in the control group at L3-4, L4-5 and L5-S1. The coordinates of the GWC of an ROI reflect the position of a unique point representing the mean position of all the points in the ROI, and any variation in the location of this point reflects changes in the geometry and/or location of the ROI. The GWC is thus able to capture movements of the nucleus relative to the vertebra above. In contrast, traditional measurements such as the distance between the boundary of the nucleus and annulus can only capture changes in one direction and only between two reference points on the contour of each structure. Measurements relying on the selection of two points on the boundaries used for measuring disc width can be subjective. Disc fissures may provide a passage allowing fluid to escape the nucleus toward the posterior corner of the vertebra,

which could explain the lack of differences horizontally in disc width but changes captured by the GWC and $T_2$WC.

Possibly, for the reason outlined above, the $T_2$WC and GWC parameters show promise to improve the specificity of MRI for LBP worth exploring to better understand the pathophysiology of LBP. A difference in the magnitude of responses to traction was observed between groups with the LBP group showing a posterior and inferior shift of the disc $T_2$WC at L5-S1 and the controls showing no response to traction. This shift likely indicates more fluid is brought in the posterior-inferior region of the nucleus with traction in those with pain. The response of the nucleus to loading varied depending on the level, which is also consistent with being a biomarker of pain, as we would not expect all levels to show pain-relevant pathology at the same time.

Among traditional morphologic measurements, disc height was the only measurement able to detect differences between pain groupings either by detecting differences supine unloaded on by illustrating differences in response to loading between groups. The disc height of the pain group was also consistently larger than the controls at L3-4 and L4-5. Further, an interaction between pain and loading was observed at L1-2, where both groups demonstrated a similar response to compression, but the controls showed larger increases in traction. As we measured disc height by dividing disc area over the disc width, a larger height indicated a larger disc cross-sectional area in the pain group. A larger discal cross-section has previously been associated with the development of intervertebral disc degeneration [78]. It is possible that larger cross-sectional area limits normal disc nutrition processes [53].

The directions of the changes observed for disc height were consistent with Madsen et al.'s report of a non-significant increase in the cumulative disc height under axial compression loading using 50% BW. Madsen et al. found a significant decrease of the cumulative disc height (mean difference 2.5 mm) in standing compared to the application of 50% BW suggesting that supine axial compression may lead to different responses than what may be anticipated based on diurnal variation studies [83]. The results of our systematic review indicated limited evidence of no change for the posterior disc height when the spine was loaded comparing to unloaded in LBP participants, while the evidence remained conflicting in healthy participants [222]. The review indicated conflicting evidence for the anterior disc height under compression in both healthy and LBP participants [222]. However, to our knowledge no other study used a similar disc area-based height measurement to monitor the effects of loading.

From examining the effect sizes observed in each group for variables with significant interactions, the pair of loading conditions best illustrating differences related to pain was changing from compression to traction. It would be interesting to conduct a study comparing unloaded to traction to determine if these two conditions would illustrate pain-relevant differences as clearly. If the same information was provided, this may increase the practicality of testing loading effects by requiring only one application of loading and allow obtaining the traditional supine lying images.

The disc and nucleus width and mean $T_2$ time as well as motion segment angles did not capture the difference between pain groups, suggesting these measurements may not relate to pain relevant pathology and may not help improve the specificity of MRI for back pain. The inability to detect significant differences between pain groups for these biomarkers may reflect an insufficient magnitude and/or duration of loading or a low statistical power. Nevertheless, the present study identified better biomarkers for pain groupings. Further, most compression studies to date also employed 50% BW [83, 90, 127, 130, 139, 154, 159, 160] for loading. In our review [222], we identified only one study using more loading (75% BW) on both healthy and participants with symptomatic lumbar spinal stenosis [153]. Although 50% BW is less than what lumbar discs normally bear during the upright postural activities [41, 240], the rationale for choosing a lower load was to minimise discomfort during imaging. Our observing no dropouts and complications suggest that this was carried out successfully, and leaves the possibility that higher load could be tolerable open. Increasing load duration, however, would likely be impractical.

The promising ability of multiple $T_2$WC and GWC biomarkers obtained for both the disc and nucleus ROI raises the question of whether some of these related parameters are redundant and whether all are needed to detect differences between pain groups. For the nucleus $T_2$WC and GWC, we observed a similar number of significant interactions and pain main effects. Given that $T_2$WC captures fluid content information about each pixel in addition to spatial information, this parameter may be preferred to reduce the measurement burden in the future. The disc $T_2$WC demonstrated more significant interactions and main effects of pain compared to the nucleus $T_2$WC indicating that the disc $T_2$WC is more sensitive to capture differences in the participants with pain than the nucleus. This, as well as the fact that disc segmentation is likely easier than the hyper-intense nucleus region, could justify more focus on the disc $T_2$WC for future studies. Nevertheless, future investigations are needed to determine which of the differences detected between pain

groups by $T_2$WC and GWC and for the nucleus compared to the disc; ROIs are most informative for determining meaningful prognosis subgroups or informing management decisions.

Some limitations to the present study should be acknowledged. We employed a small sample size for this preliminary study, which likely affects our power to capture changes between groups and among loading conditions. The present study was explanatory and many variables were examined, there is a chance that some of the statistically significant results were observed by chance. A replication study, justified by the promising results, is needed to confirm which biomarkers are truly useful to detect differences between pain groups. Our small sample included both volunteers with and without LBP with mainly moderately degenerated discs with a limited number of discs with severe degeneration (Pfirrmann's grade four and five). In the present study, we did not control for potential confounder since we matched participants for age, weight and gender. We trusted our matching could rule out the possible effect(s) of confounders, however investigating the role of possible confounders may change the results. Based on our correlation estimates between degeneration and response to loading observed in our previous study (Chapter 6), our relatively unaffected disc would likely present larger responses to loading.

Another limitation was the image pixel size. By limiting acquisition time to limit the risk of motion artefact, we could only capture changes equal to or greater than 0.4 mm within 5mm thick images. A smaller pixel could decrease the signal to noise ratio (SNR) but the increased image resolution could possibly allow detecting smaller responses to loading. Our 5mm slice thickness may allow volume averaging and responses to loading occurring only on a portion of this thickness may be missed. This study also only measured the midsagittal slice because each additional parasagittal $T_2$-map would have required another five minutes of scanning time, which would lead to impractical total scanning time and scanning different regions at different time points. Changes in fluid content or distribution in other regions of the disc may have been missed. Nevertheless, our sample comprised patients with chronic LBP and we were able to identify promising biomarkers detecting differences between pain groups and loading conditions.

**Conclusions**

The present study identified some novel and existing biomarkers detecting differences between those with and without low back pain from unloaded supine MRI. However, as expected the number of biomarkers detecting differences between those with and without pain was larger

when examining the responses to compression and traction loading. Different responses to loading between pain groups were most often observed at L5-S1. Of the 12 candidate biomarkers, the most sensitive biomarkers for pain were: the horizontal coordinate of the disc and nucleus $T_2$WC presenting large effect sizes, the horizontal and vertical coordinates of the disc and nucleus GWC, and the vertical coordinates of the discs and nucleus $T_2$WC presenting moderate effect sizes while disc height presenting a small effect size. These biomarkers hold promise for future investigation. In contrast, the disc and nucleus width and mean $T_2$ times as well as motion segment angle were the biomarkers least sensitive to capture difference between groups. Larger studies are needed to further investigate the clinical value of the proposed biomarkers but these show promise to improve the specificity of MRI for LBP and the ability to detect the response to loading.

**Acknowledgement**

Vahid Abdollah was a recipient of the University of Alberta Doctoral Recruitment Scholarship and of the Department of Physical Therapy of University of Alberta Doctoral Fellowship. All authors were recipients of a research grant provided by the Alberta Spine Foundation for the development of a program for automatic quantitative measurements of the lumbar disc and vertebra mean $T_2$ and morphology using $T_2$-weighted MR images. Data collection was conducted using grants from the Canadian Physiotherapy Association and the Department of Physical Therapy of the University of Alberta.

# CHAPTER 9 GENERAL DISCUSSION AND CONCLUSIONS

The main purpose of this thesis was to compare the response of the lumbar spine's discs and motion segments to loading between participants with and without chronic low back pain or in relation to degeneration to identify new imaging biomarkers for assisting with classification, and diagnosis of LBP and degenerative conditions. First, promising measurements and loading protocols were identified from a systematic review of the literature. Novel measurements (GWC, $T_2$WC) and image segmentation strategies were then proposed to address the gaps identified in the review to capture responses of the spine to loading. Indeed, a novel semi-automated segmentation strategy was proposed achieving high reliability for measuring the candidate biomarkers. Quantitative biomarkers were selected to address sensitivity to change and reliability issues known to affect qualitative discrete grading of MRI findings, which would limit their use to document the effect of loading. Further, measurements were extracted from MRI $T_2$ maps, thereby addressing the limitation of relying only on $T_2$-weighted images for which signal intensity cannot be interpreted quantitatively and because the difference between repeated signal intensity measurements from such images is difficult to interpret.

As a result, the following hypotheses behind this PhD work were partially supported by this preliminary investigation. First, some of the variations observed in the magnitude and distribution pattern of fluid within the lumbar tissues, as well as changes in segmental alignment of the lumbar spine during imaging under different loading conditions, were associated with the degree of degeneration and may provide further insight into degenerative conditions and the pathology underlying persistent back pain. Specifically, refined quantification of the disc fluid distribution using the location of the disc and nucleus $T_2$-weighted centroid ($T_2$WC) as well as the more traditional methods, including disc height and motion segment angle under different loading conditions showed promise as MR imaging biomarkers for disc degeneration. Second, also as postulated, the novel quantitative measurements of the fluid distribution in the lumbar discs showed superior ability to detect differences between those with and without pain compared to measurements that are more traditional. Finally, because symptoms are often experienced when the spine is loaded, as hypothesised, we observed that responses of the candidate biomarkers to loading did provide an additional tool for detection of differences between those with and without pain beyond the information provided by traditional unloaded supine imaging. The evidence related to the hypotheses above produced in the different phases of this PhD is summarised in details hereafter.

**Summary of the main findings**

***Systematic literature review of the effect of loading on quantitative measurements of the lumbar spine.***

***Characteristics of the Included Studies.***

- Of 21 studies included, only ten included participants with LBP [83, 87, 124, 130, 140, 154, 156–160] and only two reported data on participants with and without LBP.
- No study met the high-quality threshold with lack of a well-defined subject's recruitment (selection bias) and examiners' training strategy being the most common sources of bias. Therefore, no moderate or strong summary statements could be formulated.
- Fourteen quantitative measurements of the effect of lumbar loading were identified, but no study had focused on the mean signal intensity or $T_2$ times nor the centroid of the GWC and $T_2$WC as biomarkers of the fluid distribution.

While a high heterogeneity was observed among studies in terms of sample characteristics, loading conditions, and biomarkers used, summary statements could be formulated within five categories of relatively homogeneous of loading conditions comparisons. These statements are as follows:

***Comparing unloaded vs. axially loaded spine using any compression device***

- There was *limited* evidence (1 high-quality* or consistent results among low-quality studies) of:
    - a decrease in posterior disc height in response to compression loading [90].
    - no change in the cumulative disc height (40% and 50% BW) [83], posterior bulging of the disc (50% BW) [4], range of motion (75% BW) [153] and lumbar lordosis angle (40% and 50% BW) [83].
- The evidence was *conflicting* on the effects of loading on the DSCA [83, 130, 154, 159, 160]. The evidence remained conflicting when we tried to determine the effect of loading separately for subjects with and without pain.

***Comparing unloaded vs. postural loading (such as sitting/standing).***

- There was *limited* evidence of:
    - a decrease in cumulative disc height (standing) [83], an increase in the anterior distance of the nucleus position (standing and sitting) and anterior-posterior diameter of the disc (standing and sitting) [167].

  - no difference in anterior disc height (kneeling, standing and sitting) [90, 167], posterior disc bulging (kneeling) [90] and lumbar lordosis angle (kneeling, standing and sitting) [83, 88, 165].

- The evidence was *conflicting* about posterior disc height (kneeling, standing and sitting) [90, 156, 167], anterior-posterior diameter of the nucleus (standing and sitting) [167], posterior distance of the nucleus position (standing and sitting) [167], and DSCA (standing and sitting) [88, 139, 165].
  - When trying to resolve conflicting evidence by examining differences in healthy separately from participants with LBP, limited evidence of a decrease in the DSCA was observed in participants with LBP (standing) [139] and the evidence remained conflicting in healthy subjects [88, 165].

***Comparing among different postural and axial loading conditions.***

- There was evidence of:
  - no difference in cumulative (40% and 50% BW axial compression) [83], and posterior disc heights (kneeling, standing, sitting and with 50% BW axial compression) [90, 167], posterior bulging (kneeling and 50% BW axial compression) [90], and lumbar lordosis angle (40% and 50% BW axial compression) [83] among loading conditions.

***Comparing unloaded vs. end range movements***

- There was limited evidence of:
  - a decrease in anterior disc height going from lumbar flexion to extension in the supine position [162], and of a contralateral shift of the location of the nucleus peak signal intensity in the mid-coronal plane in side bending [163].
  - no difference in abnormal tilting movement, translatory and rotatory instability [164] and lumbar lordosis angle among supine psoas relaxed, supine unloaded [83, 165], supine extended, supine with 40% and 50% body weight, supine extended with 50% body weight compression, sitting, sitting flexed, sitting extended and standing conditions.

***Among end range movements***

- There was *limited* evidence of:
  - a decrease in anterior disc height and anterior distance of the nucleus position going from a flexed to an extended position [162].

  - no change in posterior disc height (supine extension vs. flexion) [162], anterior-posterior, left and right parasagittal, mid-coronal diameter of the disc (supine neutral vs. supine flexion, extension, rotation to right and left) [161], range of motion (standing flexion vs. extension) [164], and lumbar lordosis angle (supine unloaded vs. sitting, sitting flexed, sitting extended and standing) [165].
- The evidence was *conflicting* on the DSCA changes [161].

***Implications of the systematic review on the PhD experiment.*** The review helped choose axial loading with 50% BW as a loading condition easily implemented in conventional MR systems. This was justified because no difference was observed between loading devices and postural loading, between 50% and 75% BW, and between movements at end range and some device loading. The magnitude of loading most commonly used was 50%BW with no adverse effects. The review demonstrated that a knowledge gap still exists in comparing the effects of loading between participants with and without chronic LBP. Thus, we chose a matched sample of both groups. As we could not identify any research about the effects of traction, we chose to evaluate its effects on the disc as it offered an important loading contrast with compression. We postulated that larger loading effects would be created going from compression to traction. This would help detect relevant associations between changes in imaging findings and lumbar disc degeneration and differences between patients with and without pain.

Among the biomarkers identified in the review, we chose to examine disc height, disc and nucleus width, which showed some promise in responding to loading. However, multiple disc height definitions have shown conflicting effects suggesting further investigation of the reliability and validity of the disc height measurements was needed. As no evidence or evidence of no effect was found in some traditional measurements, including disc bulge, canal and foramen space, these measurements were not included. Nevertheless, nerve spaces may be worth investigating in the future as related evidence was conflicting or lacking. We also chose novel biomarkers, which had not yet been investigated as some related parameters such as distance from the anterior and posterior limits of the nucleus to the edges of the disc, had shown some promise to detect the effect of loading. These included the location of GWC, $T_2$WC and motion segment angle.

***Reliability study of the signal intensity weighted centroid***

A novel semi-automated procedure was implemented in Matlab-based software to extract the following novel parameters while minimising user input. The SIWC was previously proposed for the nucleus horizontal coordinates in scoliosis and spondylolisthesis research [214], we expanded that for the vertical coordinates as well. The SIWC was proposed as a biomarker to monitor fluid distribution within lumbar structures. The GWC was developed for use as a reference point in the vertebra to quantify all measurements relative to it and for quantifying nucleus movements within the disc.

- The intra-rater ($ICC_{(3.1)}$) and inter-rater ($ICC_{(2.1)}$) reliability coefficients for the horizontal and vertical coordinates of the SIWC for all regions of interest (ROIs) at L4-5 and L5-S1 disc levels was excellent ranging from 0.97 to 1.
- Intra-rater ($ICC_{(3.1)}$) and inter-rater reliability ($ICC_{(2.1)}$) estimates for the average signal intensity and area for the whole disc and nucleus region at L4-5 and L5-S1 ranged from 0.88 to 1.00, but only from 0.27 to 0.78, for the anterior and posterior annulus regions.

***Implications of the reliability SIWC study on the PhD experiment.*** This study confirmed our hypothesis that semi-automated measurements would offer adequate reliability for use in our loading study and that the novel SIWC offered excellent reliability. However, the limited ability to define annulus regions of interest led to the exclusion of this ROI from further investigations in our loading study.

***Reliability and validity of lumbar disc height quantification methods***

The results of a systematic review conducted by our group indicated that the proposed methods for measuring disc height are numerous and yield heterogeneous reliability and validity results [241]. Thus, seven different disc height definitions were chosen and compared using our semi-automated image analysis strategy.

- The intra-rater reliability $ICC_{(3,1)}$ of the area-based disc height measurements using the full disc width was 0.84 at the L4-5 level and 0.90 at L5-S1, but consistently more than 0.91 at both levels when using 60 or 80 percent of disc width.
- The intra-rater reliability estimates of the non-area-based methods were lower and ranged from $ICC_{(3,1)}$ 0.86 to 0.96 at L4-5 and from 0.76 to 0.83 at L5-S1.

- Similarly, the inter-rater reliability estimates ICC $_{(2,1)}$ of the area-based methods was higher (0.99) than for non-area-based methods for which ICC $_{(2,1)}$ ranged from 0.84 to 0.98 at both spinal levels tested.
- Among the non-area-based measurements, the point estimates of the intra- and inter-rater reliability of Hurxthal's method were higher than for both Dabbs' and the combined method at both levels.
- Regarding the construct validity, the highest Pearson's correlations between different methods were observed between the area-based methods using 100% and 80%-disc width, and the lowest correlations were observed between Dabbs' and Hurxthal's second method.
- All correlations between measurements other than Dabbs' method exceeded 0.90.

***Implications of the height reliability study on the PhD experiment.*** We chose area-based disc height measurement methods as they could capture more disc contour information compared to the point-based methods. Among different width used, we chose 80% as it indicated a higher reliability compared to full width (100%) and captured more disc boundaries relative to lower widths (40% and 60%). This width also presented high construct validity estimates with traditional measurement methods.

***Effects of extension exercise on disc fluid content***

The hypothesis that novel measurements of the SIWC could capture more changes in response to loading than traditional measurements such as the mean signal intensity was tested in patients with back pain in a historical dataset where extension loading had been applied based on a McKenzie internal disruption model to influence symptom [135]. This model suggests that extension loading may have an effect on the fluid distribution and fluid quantity within the disc.

- Results partially supported our hypothesis. While there was no significant difference in the mean signal intensity at the levels tested, a significant anterior displacement in the location of the horizontal coordinates of SIWC for the whole disc was observed from before to after the extension intervention at L4-5. However, the horizontal coordinates of SIWC at other levels and the vertical coordinate of the SIWC of the whole disc and nucleus at either disc levels did not demonstrate any significant differences from before to after the extension exercises. Nevertheless, the effect sizes of our novel fluid distribution indicators were similar or larger compared to traditional fluid content measurements using signal intensity.

***Implications of the extension exercise study on the PhD experiment.*** The results confirmed our primary hypothesis that the location of the SIWC shows promising sensitivity to capture fluid movements within the discs compared to the traditional signal intensity changes.

***Evaluation of the effects of loading on lumbar intervertebral discs and vertebrae in relation to disc degeneration***

The hypothesis that the effect of compression and traction loading on the lumbar spine would represent biomarkers of lumbar disc degeneration was examined in this preliminary study. A preliminary study was deemed necessary before justifying a larger study due to the novelty of many aspects of the methodology. The study used candidate measurements identified in the review and novel measurements (Mean $T_2$ times, GWC, $T_2$WC) developed in this PhD, which were extracted using the developed semi-automated measurement procedures. Compression loading was applied using a custom-designed MRI-compatible apparatus as per loading protocols encountered commonly in the review. In addition, a novel traction loading condition was also investigated to examine response between opposite extremes of axial loading. It was postulated that larger loading effects would be created going from compression to traction, which would help detect relevant associations between changes in imaging findings and lumbar disc degeneration.

Results showed that the effect of loading was detectable especially using the novel GWC, and $T_2$WC measurements and, as expected, more so between compression and traction loading conditions.

- Significant small differences were observed between ***compression and traction***:
    - a trivial or small significant inferior ($ES_{L4\text{-}5}$:0.14; $ES_{L5\text{-}S1}$:0.33) and posterior shift ($ES_{L4\text{-}5}$:0.40; $ES_{L5\text{-}S1}$:0.30) of the horizontal and vertical coordinates of the disc $T_2$WC was observed at the two lowest lumbar levels.
    - a small significant posterior shift was also observed for the horizontal coordinate of the nucleus $T_2$WC and GWC at L4-5 ($ES_{T2WC}$:0.49; $ES_{GWC}$:0.48) and L5-S1 ($ES_{T2WC}$:0.25; $ES_{GWC}$:0.22).
    - a small significant inferior shift for the vertical coordinates of the nucleus $T_2$WC (ES: 0.31) and GWC (ES: 0.31) at L5-S1.
    - a trivial significant increase of the disc height at L1-2 (ES:0.08) and a small significant increase at L5-S1 (ES:0.32).

- a small significant increase in the nucleus $MT_2$ was observed at L3-4 (ES: 0.26).

- Significant small and moderate difference were also observed between ***unloaded and compression*** for:
  - a small significant (ES:0.39) posterior shift of the horizontal coordinate of the disc $T_2$WC at L1-2 level.
  - a small significant posterior shift (ES: 0.38) at the L3-4 level for the horizontal coordinate of the nucleus GWC.
  - a small significant increase (ES:0.22) of the disc width at L5-S1.
  - a moderate increase by an average 2.1$^o$ at L3-4 (ES:0.56) and 1.8$^o$ at L4-5 (ES:0.53) of the segment extension angle.
  - a trivial increase of the nucleus $MT_2$ (ES:0.18) of 5 ms at L1-2.
- As hypothesised, the correlation between disc degeneration (estimated by mean $T_2$ time) and the changes in the disc and nucleus measurements in response to loading reached significance indicating that more degeneration led to smaller responses to loading (Pearson's r=-0.51 to 0.49).
- For the difference between unloaded and compression, significant correlations between degeneration and the changes in quantitative MRI measurements were observed for the following measurements. Less severe disc degeneration (higher mean $T_2$) was associated with:
  - larger anterior shift in the coordinates of the disc $T_2$WC at L1-2 (-0.45) and L5-S1 (-0.36).
  - larger inferior shift in the coordinates of the disc $T_2$WC at L1-2 (0.49).
  - larger reduction in mean $T_2$ at the L2-3 (-0.45), L3-4 (-0.51) and L4-5 (-0.42) discs.
  - larger reduction in disc height at L1-2 (-0.33).
  - larger reduction in nucleus mean $T_2$ at L3-4 (-0.36).
- For the difference between compression and traction, significant correlations between degeneration and the changes in quantitative MRI measurements were observed for the following measurements. Less severe disc degeneration (higher mean $T_2$) was associated with:
  - larger anterior shift in the coordinates of the disc $T_2$WC at L1-2 (-0.46).
  - larger superior shift in the coordinates of the disc $T_2$WC at L2-3 (-0.45).
  - larger reductions in disc width at L4-5 (-0.36).

- larger posterior shift in the coordinates of the nucleus $T_2$WC at L1-2 (0.42), L3-4 (0.39) and L5-S1 (0.36).
- larger superior shift in the coordinates of the nucleus $T_2$WC at L2-3 (-0.51).
- larger posterior shift in the coordinates of the nucleus GWC at L1-2 (0.40) and L5-S1 (0.39).
- larger superior shift in the coordinates of the nucleus GWC at L2-3 (-0.48).

***Implications on the PhD experiment.*** The lower lumbar levels behaviour was consistent with the theoretical model of compression and traction. We postulated that this inter-level inconsistency was likely due to the heterogeneity of the sample related to back pain. Therefore, as planned, in the following chapter, we examined the effects of loading on each group separately. Our result also indicated that disc height, width, $MT_2$ and $T_2$WC as well as nucleus $MT_2$, GWC and $T_2$WC hold promise as sensitive biomarkers to capture the response of the lumbar spine to loading and likely to monitor degenerative and regenerative changes in the lumbar spine. This, as well as the fact that disc segmentation is likely easier than for the hyper-intense nucleus region, could justify focusing on the disc $T_2$WC for future studies. However, the clinical value of the proposed measurements remained unclear. Unlike ordinal scaling methods combining multiple characteristics to assess disc degeneration, the proposed biomarkers and measurement software are sensitive to both subtle changes and local variations in the disc morphology. The results also indicated that changes from compression to traction were larger compared to changes from unloaded to compression, thus to decrease scanning time, one solution could be focusing only on compression and traction. Healthy discs present the largest responses to loading, and therefore loading helps illustrate the severity of degeneration affecting a disc.

### *The effects of compression and traction on lumbar MRI findings in relation to chronic low back pain*

Similar to the previous study, it was postulated that: larger loading effects would be created going from compression to traction and that changes between loading conditions would help detect relevant differences comparing changes in imaging findings between participants with and without low back pain matched for age, weight and gender. Thus, this study would help identify biomarkers of chronic low back pain. This study also took advantage of the novel measurements; semi-automated measurements and loading protocol including a traction condition because traction represents an extreme contrasting with compression.

A small number of biomarkers for pain were identified from variables which demonstrated significant differences between pain groups (main effects for pain) suggesting that these variables may help distinguish patients with a without pain using only unloaded images.

- Coordinates of $T_2$WC more anterior in the pain group for the whole disc at L3-4, and L4-5 (ES: 0.59 to 0.78) and for the nucleus at L4-5 (ES: 0.94).
- Coordinates of the nucleus GWC more anterior in the pain group at L3-4, L4-5, and L5-S1 (ES: 0.62 to 1.03).
- Disc height was larger in the pain group at L3-4, and L4-5 (ES: 0.57 to 1.02).

Biomarkers for pain were also identified from variables with significant interactions between pain and loading suggesting that these variables may help distinguish between patients with and without pain by comparing their responses to compression and traction loading.

- Pairwise comparisons from unloaded to compression revealed the following:
    - a moderate anterior shift in LBP and a small posterior shift in controls for the disc $T_2$WC at L5-S1.
    - For the vertical location of the disc $T_2$WC:
        - a moderate inferior shift at L3-4 in LBP and no effect in controls.
        - a small superior shift at L5-S1 in LBP and no effect in controls.
    - a moderate superior shift in LBP for the nucleus $T_2$WC at L5-S1 and no effect in controls.
- Pairwise comparisons from compression to traction revealed the following:
    - a moderate to large posterior shift in LBP and no effects in controls for the disc $T_2$WC at L5-S1.
    - For the vertical location of the disc $T_2$WC:
        - a moderate superior shift at L3-4 in LBP and no effect in controls.
        - a small inferior shift at L5-S1 in LBP and no effect in controls.
    - a moderate inferior shift in LBP and no effect in controls for the nucleus $T_2$WC at L5-S1.
    - no effect in LBP and trivial increase (ES 0.13) in controls for disc height at L1-2.

- No pain main effect nor interactions between pain and loading were detected for the following measurements suggesting that these do not show promise as biomarkers for low back pain:
    - disc mean $T_2$ time, nucleus mean $T_2$ time, disc width, nucleus width, motion segment angle.

Results indicated that the proposed candidate biomarkers hold promise to help improve the specificity of MRI for low back pain by detecting differences between images of patients with and without pain while unloaded. Further, as hypothesised, differences in responses to loading can also help detect differences between pain groups. Differences were more numerous when comparing changes from compression to traction to the more commonly used compression loading added after unloaded images. Some measurements were only detecting responses to loading and may represent good measurements for biomechanics studies not focused on diagnostic. Further, the preliminary analysis of this relatively small sample yielded a number of small, moderate and large effect sizes justifying further investigations of the proposed biomarkers to improve the specificity to low back pain.

**General discussion**

To the best of our knowledge, the present study was the first comparing the immediate effects of traction and compression between participants with and without chronic low back pain (LBP) and in relation to degeneration levels using quantitative $T_2$-mapping and a highly reliable image segmentation software. The results indicated that the response to three loading conditions was different across lumbar levels with different degeneration and between participants with and without pain. Specifically, significant interactions between loading and pain were most often observed at L5-S1 consistent with beliefs that the lower part of the lumbar spine is most affected by degeneration and implicated in causing LBP [239].

Amongst candidate biomarkers evaluated in this dissertation, coordinates of the disc and nucleus $T_2$WC and GWC were the most promising biomarkers to detect pain as determined by the number of significant main effects of pain or interactions as well as by the magnitude of the effect sizes compared between two groups. Amongst candidate biomarkers for evaluating the effects of loading in relation to degeneration of the lumbar spine, the location of $T_2$WC and GWC, the motion

segment angle and disc height were the most promising ones. Correlation estimates indicated that larger responses to loading were observed with less degeneration.

LBP is a major cause of discomfort and disability among adults globally [1]. Although it has affected humans since at least the Bronze Age, in over 85% of LBP cases, the cause of the pain remains unclear (e.g. non-specific LBP) [4]. Imaging modalities are used for diagnosing the underlying pathology of LBP [9], with MRI preferred due to better visualisation of soft tissue and concerns about radiation exposure with the other methods [10]. MRI is a sensitive modality for depicting anatomical abnormalities of the lumbar spine. However, it is often clinically inconclusive due to a lack of specificity [11]. MR imaging is usually conducted in a relaxed supine position while most patients experience symptoms when the spine is loaded [43, 44]. Compression loading is thought to induce morphological changes, including reducing disc height, increasing disc bulging, thickening the ligamentum flavum, deforming the dural sac and narrowing the spinal canal [242–244]. Therefore, to increase the specificity of MR imaging, one solution could be employing upright MRI scanners or trying to simulate physiologic loading using axial compression devices if tolerable by patients.

***Systematic literature review of the effect of loading on quantitative measurements of the lumbar spine***

A review was needed to identify the best measurements and loading protocols to understand which imaging findings in response to loading were most closely related to LBP symptoms or spine degeneration. The review did not yield any moderate or strong evidence due to the low quality of the included studies but identified the main sources of bias affecting these studies including inadequate information about requirement strategy, examiners training, and reliability of the measurements. We tried to avoid these limitations in our studies.

The review revealed a lack of comparison of the effects of loading between participants with and without LBP. Many measurements have only been used with either healthy participants or participants with LBP but not both. When formulating levels of evidence summary statements separately for participants with and without pain, the statements on two outcomes remained conflicting in both types of participants. For 3 to 10 of the 14 measurements used in at least one study, the measurements had not been used to quantify the effect of loading depending on the

homogeneous set of loading conditions compared. A clear knowledge gap about comparing the effect of loading between groups existed.

The heterogeneity in the review results highlighted many areas with conflicting evidence regarding the effects of loading on the lumbar spine. For many measurement constructs, no evidence was identified depending on the loading conditions considered. More than half of the limited evidence summarised in the opening of this chapter indicated no effect of loading was observed in multiple measurements. For supine axial loading, only limited evidence of decreased posterior disc height in response to loading was observed showing a lack of sensitive MRI biomarkers of the response to loading. For the postural loading, limited evidence of increased anterior distance of the nucleus from the annulus boundary and of increased anterior-posterior diameter of the disc as well as of decreased cumulative disc height was identified in response to loading. There again a limited number of sensitive biomarkers showed a need for further investigation.

Our review yielded conflicting evidence on the effects of loading on the lumbar spinal canal stenosis (LSCS). Only one study was identified that evaluated the effects of loading specifically on the LSCS symptoms [139]. The results indicated negative correlations between the dural sac cross-sectional area (DCSA) in axially loaded MRI and both walking distance and Japanese Orthopedic Association Score [139]. Although some researchers have suggested that extension could trigger LSCS symptoms by a reducing the dural sac size [170, 171], our review did not yield any evidence on the effects of extension compared to an unloaded supine position on the LSCS. Only one study with participants with LBP was identified that examined the effect of end-range movements on quantitative lumbar segmental instability (LSI) measurements, and no evidence was identified about the possible effect of compression or traction on the LSI. These findings supported not focussing the present thesis of candidate biomarkers related to these MR findings.

The results, therefore, highlighted that the knowledge gap regarding the effects of loading on the imaging presentations of the lumbar spine still exists, particularly, a lack of research on whether the response to loading, including traction, could help increase the specificity of MR imaging on LBP. Based on the evidence from the systematic literature review, a few measurement constructs deemed to be relevant to the subjects' symptoms and the state of degeneration were chosen for further investigation on both participants with and without chronic LBP. We could not

identify any evidence about the effect of loading on the disc fluid content. As the degeneration affects disc fluid content and implicated as a cause of LBP [5], we postulated both the fluid content and distribution of the discs would be systematically different in participants with and without pain. The candidate biomarkers selected for investigation in the present thesis were the disc's fluid content (measured using mean $T_2$ time), height, the location of $T_2$WC, and width as well as the nucleus's fluid content, GWC, the location of $T_2$WC and width as well as motion segment angle.

- An updated version of this literature review is available as an addendum in the appendix VI: The immediate effects of loading on lumbar imaging findings: Update of a systematic review.

***Reliability study of the signal intensity weighted centroid***

As manual segmentation is tedious, time-consuming and prone to user error [245, 246], we developed computer aided measurement software and examined the reliability for all candidate biomarkers extracted using the software. However, the similar signal intensity of the lumbar tissues, ambiguous disc boundaries, particularly in degenerated discs, and the inter-individual variability in the shapes and signal intensities of the discs make developing computer-aided segmentation techniques extremely challenging [187]. Thus, only a few groups have developed CAM to speed up and improve the reliability of lumbar segmentation [182, 214, 245–250], and even fewer did so to improve the reliability for detecting possible abnormalities [182].

$T_2$WC/SIWC provides information about the distribution of the fluid within the structure and thus provides some information about fluid movements within the structure. In contrast, many traditional methods [167], measuring the distance between two tissue boundaries can only capture movement in one direction and only using two points rather than all the pixels in the ROI. Moreover, the selection of the two points on the boundaries is subjective and possibly more prone to error than using all points from an area semi-automatically segmented using our procedure, which showed excellent reliability. The results of this study indicated a high intra- and inter-rater reliability for the SIWC, while the reliability estimates of the mean signal intensity and area suggested that the measurements of the annular regions were not suitable for clinical use.

***Reliability and validity of lumbar disc height quantification methods***

As many disc height measurement methods exist, methods were compared head-to-head to determine which would be most appropriate for use with studying loading. Our results indicated that the area-based methods provide better reliability than point-based methods, particularly

Dabbs' method, which has been widely used by many researchers for the quantification of the disc height [190, 191, 202]. The reliability of point-based methods is influenced by the examiner's perception of the location of the vertebral margins at the disc-vertebra boundary, and the presence or absence of endplate irregularities or Schmorl's nodes may also affect measurement outcomes.

The area-based disc height quantification methods demonstrated excellent intra- and inter-rater reliability when using 60% or 80% of disc width for the estimation. These measurements also demonstrated the lowest measurement error, suggesting promise for detecting changes beyond measurement error in longitudinal applications. The lower reliability when using 100% of the disc width as compared to using 80% and 60% was likely due to the relatively indistinct anterior and posterior edges of the vertebra-disc interface. Limiting disc height measurements to a fraction of the disc diameter also reduce the imprecision due to the presence of osteophytes near the corners, which may bias some of the point-based references using the vertebral corners, such as Dabbs' or the combined method. The area-based disc height quantification method using 80% of disc with was therefore adopted in subsequent studies as it also presented high construct validity correlations with traditional measurements. To our knowledge area-based measurements of disc height had not yet been used to quantify the effect of loading on the lumbar spine.

***Effects of extension exercise on disc fluid content***

Loading not only affects the quantity of fluid within discs but also may affect its distribution. The mean $T_2$ relaxation time, however, is a scalar quantity, which does not reflect the direction of the fluid movement. Therefore, another variable is needed to capture changes in the fluid distribution. $T_2$WC reflects movement or changes in the distribution pattern of fluid within a structure. As degeneration does not affect the whole disc at the same time, $T_2$WC may also provide information about the region, which is more affected by degeneration. $T_2$WC was also chosen because it may provide some information about the role of adjacent vascular regions on the disc nutrition, as any deviation in the location of $T_2$WC in response to acute loading may indicate more signal concentration and thus more fluid having moved to that region.

The location of the SIWC could be employed as a highly reliable measurement for investigating changes in signal distribution in ROIs, which may be useful in studies of the effects of different loading conditions or therapeutic interventions on targeting fluid redistribution. Little evidence was found supporting the hypothesis that prone press-up exercises affect disc fluid

content and distribution. There was no significant difference in the whole disc and nucleus MSI or location of the SIWC before and after extension exercise, except for the horizontal coordinate of the SIWC of the whole L4-L5 disc. The significant anterior shift of the signal intensity observed at the L4-5 level indicates more fluid concentration on the anterior part of the disc. This was consistent with theoretical model of extension exercise where extension increases the pressure on the posterior part of the disc and drives fluids more anteriorly. The fact that the SIWC did not demonstrate movement in the same direction at L5-S1 makes this finding less persuasive, but the range of motion at L5-S1 is generally less than at L4-5 [216], and may not be enough to reduce intra-discal pressure and drive fluid anteriorly. Nevertheless, the novel parameters reflecting fluid distribution detected similar or larger effects of extension than MSI suggesting promise in the study of the effect of lumbar loading conditions. Interestingly, our findings suggesting no significant change in disc fluid content (MSI) were not consistent with Beattie et al.'s findings, indicating a significant increase in the fluid diffusion in the nucleus following 10 minutes of lumbar posterior-anterior pressures and prone press-up exercises [45]. These larger observed changes could be due to the application of PA directed manual pressure. Still the novel measurements were adopted to study the effect of loading in the subsequent studies from this PhD.

***Evaluation of the effects of loading and disc degeneration grade on lumbar intervertebral discs and vertebrae in participants with and without back pain.***

In the loading studies described in chapters 7 and 8, when the spine was loaded using compression and traction loads equal to 50% BW, the axial loading appeared to result in a complex load transmission throughout the lumbar spine. The behaviour of the middle and lower lumbar motion segments was more consistent with the hypothesised response based on a theoretical model of compression and traction, where compression is expected to decrease disc and nucleus fluid content and traction to do the opposite. This difference between levels could be due to bearing different loading magnitudes and moment at the upper levels compared to the lower ones.

***Evaluation of the effects of loading and disc degeneration grade on lumbar intervertebral discs and vertebrae.*** The results from our loading studies also indicated as we had hypothesised, that the response of the lumbar motion segments to three loading conditions differed across lumbar levels between participants in relation to levels of disc degeneration. Thereby, our results demonstrate the potential for these candidate biomarkers to improve the specificity of MRI for disc degeneration. Specifically, significant differences in response to loading were most often

observed at L5-S1 consistent with beliefs that the lower part of the lumbar spine is most affected by degeneration and implicated in causing LBP [239]. The magnitude of the loads and moments acting on the lower levels are larger than for the upper levels [251].

The location of $T_2$WC of the disc and nucleus, the GWC of the nucleus, and the motion segment angle were found to be the most promising biomarkers for evaluating the effects of loading on the lumbar motion segments with different degeneration grades (detected significant changes). In addition, the mean $T_2$ time of an ROI was correlated with anterior and inferior shifts of $T_2$WC in response to compression compared to the unloaded condition. Similarly, a significant posterior and superior shift of the disc and nucleus's $T_2$WC and of the nucleus GWC was observed in the lower lumbar motion segments in response to traction compared to compression. Consistent with degeneration causing disc desiccation, narrowing and a reduction in the hyper intense nucleus area [16], discs with more degeneration presented smaller fluid redistribution shifts in response to loading possibly increasing the specificity of MRI for relevant degeneration. Larger changes were associated with the healthiest discs.

The fluid content of the discs and nucleus of the lower lumbar motion segments decreased during compression and increased during traction; however, in our sample differences did not reach significance. This likely indicates that loading affects the distribution of the fluids within the disc more than the amount of fluid. A significant negative correlation was however observed between compression loading and degeneration at these levels, indicating the disc fluid content in the more degenerated discs is less than normal discs, which was consistent with other studies [5]. In contrast, the upper motion segments demonstrated a decrease in fluid content from compression to traction, reaching significance only at L1-2. The behaviour of the upper discs indicated that the axial compression loading likely did not affect these levels and the discs continued to imbibe fluid.

Although the directions of the changes observed for disc height and disc and nucleus width changes at the L5-S1 were consistent with previous studies [44, 167], most were not statistically significant except for the increased disc width from unloaded to compression at L5-S1. The heterogeneity of our participants may have limited our ability to capture more changes, or it may be because of our using different loading conditions than in some of the reviewed studies. Nevertheless, in our study, both the disc and nucleus $T_2$WC, nucleus GWC, disc height and motion segment angle hold more promise for capturing changes in response to loading.

Our findings for the lumbar motion segments angles were consistent with the theoretical models of compression and traction (i.e. compression increases the motion segment angle producing an increased lordosis, and traction decreases the angle by flattening the lumbar spine). However, in this study, differences only reached statistical significance at L4-5 when going from unloaded to compression. The loading table was theoretically designed to apply uniaxial compression and traction load, however, due to its effect on lumbar lordosis; the line of action of the force could be different at different levels. This observation may explain some of the differences in fluid distribution variables between levels.

Significant correlations were observed between degeneration and changes in quantitative MRI measurements from unloaded to compression and from compression to traction for disc's height, mean $T_2$, width and $T_2$WC and nucleus's $T_2$. Of those measurements, the disc height, mean $T_2$ and $T_2$WC as well as nucleus mean $T_2$, $T_2$WC and GWC hold more promise for further investigation, as our results indicated that they are also more sensitive to loading as well. These measurements may be sensitive enough for studying the natural history of degeneration as in genetic studies or the response of degenerated discs to regeneration treatments.

***The effects of compression and traction on lumbar MRI findings in relation to chronic low back pain.*** Amongst candidate biomarkers, the most promising (sensitive) biomarkers to pain were the novel measurements of the coordinates of the disc and nucleus $T_2$WC and GWC determined by the number of significant main effects of pain or interactions as well as by the magnitude of the effect sizes compared between groups. The horizontal coordinates of these novel biomarkers at different levels were also one of only a few parameters detecting differences between patients and controls in the commonly used supine unloaded position. $T_2$WC of the pain group was located more anteriorly and superiorly compared to the controls especially at the lower motion segments and was more sensitive to loading. Compression shifted $T_2$WC more anteriorly and superiorly, while traction did the opposite. The extension moment created by compression might increase pressure on the posterior part of the disc and therefore push the fluid more anteriorly. The unloading effect of traction may decrease that pressure and thus help disc fluid return to its equilibrium state.

$T_2$WC coordinates were more sensitive to loading in the pain group compared to the controls. $T_2$WC of the nucleus was slightly more sensitive to the loading as it contains a higher concentration of fluid compared to the fibrous part of the disc and thus may provide information

about the response of the disc to acute loading. Changes in hydration, as well as collagen anisotropy, have been reported to be early indicators of cartilage deterioration [112]. This likely indicates that discs were more anisotropic in the pain group. These changes may also affect the fluid distribution and thus hydrostatic pressure of the disc and lead to uneven load distribution on the pain-sensitive endplates or annulus fibres. Trauma and inflammation may trigger the synthesis of factors such as bradykinin and prostanoids within the disc [6]. With fluid redistribution under loading, these factors may be able to sensitise silent nociceptors, which are usually unresponsive to even maximal mechanical stimulation [6]. The sensitised nociceptors and usually active nociceptive affected by the fluid redistribution could then respond to changes in intra-discal pressure during movement and trigger pain [6]. This hypothesis would be consistent with our results indicating that the fluid distribution responses to loading in the participants with LBP were different from those without pain.

The GWC also demonstrated promise to improve the specificity of MRI for LBP. The horizontal coordinate at all levels, except L1-2, was sensitive enough to capture a differential response between pain groups or interactions between pain and response to loading. The response to traction loading was different between two groups. While the pain group demonstrated a posterior shift, controls demonstrated no change. This likely indicates more fluid was brought back in the posterior region of the nucleus with traction in those with pain. The response of the nucleus to loading varied depending on the level, which is also consistent with being a biomarker of pain, as we would not expect all levels to show pain-relevant pathology at the same time. As the GWC parameter reflects the spatial area occupied by the hyper-intense nucleus region, it also reflects fluid distribution. For upper levels in the pain group and all levels in controls, the GWC either moved anteriorly or did not change on compression compared to an unloaded condition. This might be a result of the changes in pressure distribution as the motion segment angle of the relatively healthy disc becomes more extended.

The disc height was the only traditional morphologic measurements that demonstrated potential to detect differences between pain groupings using only the supine unloaded images as we defined disc height as the disc area over the disc width, participants in the pain group might had more concave endplates resulting in higher disc heights compared to healthy ones, which would be consistent with Videman et al.'s findings with increasing degeneration [153]. The inability to detect significant differences at most levels may reflect, experiencing different loading

conditions at each level, and insufficient magnitude and/or duration of loading or a low statistical power in this preliminary study. Most compression studies to date employed 50% BW [83, 90, 127, 130, 139, 154, 159, 160] for loading. In our review [222], we identified only one study that applied 75% BW on both healthy participants and participants with symptomatic lumbar spinal stenosis [153].

Both mean $T_2$ times and width of the nucleus were not sensitive enough to capture a statistically significant change in response to loading or difference between pain groups (except at L1-2 for the mean $T_2$ in traction) which might be due to limited power in this preliminary investigation.

**Clinical importance**

In contrast to Pfirrmann and other ordinal rating algorithms that pool different degenerative features to grade the degree of disc degeneration, and which are, therefore, inherently unable to capture subtle and/or local changes, the proposed new quantitative method is sensitive to both subtle and local changes in the disc morphology. The results also indicated that compression and traction demonstrate more changes compared to the unloaded condition, thus to decrease scanning time, one solution could be ignoring the unloaded stage. As in the present study, we did not examine the effects of traction on an unloaded spine, examining this could help determine whether the observed significant differences were due to using only traction or comparing two extreme loading conditions. This would help evaluate the cost-effectiveness of compression-traction loading scenario compared to the unloaded-traction scenario, which requires less time and effort, by clinicians.

Results also indicated that both disc and nucleus should be considered for evaluation as the response to loading at some levels was only significant for the disc and at other levels was only significant for the nucleus, indicating that these measurements may be complementary. A future diagnostic accuracy study could help determine whether the additional time and segmentation efforts required for extracting measurements related to the hyper-intense nucleus area in addition to the whole disc measurements are needed to optimise MRI specificity for LBP.

The results of the present dissertation indicated that the disc and nucleus $T_2$WC, nucleus GWC and disc height hold promise to improve the specificity of MR imaging for diagnosing the possible underlying pathology of pain in patients with chronic LBP. Among traditional clinical

biomarkers employed to date, the disc height was the only biomarker that was different between participants with and without persistent LBP on the unloaded supine scans. However, the systematic differences in the position of the nucleus $T_2$WC and GWC as well as the disc $T_2$WC on the supine unloaded scans between two groups suggest that MRI specificity for LBP may be improved when the novel parameters are employed in addition to disc height. However, further studies are needed to determine a cut-off point for normal and a symptomatic disc on the promising measurements and if diagnostic accuracy would be sufficient.

The high significant correlations that were observed between horizontal and vertical coordinates of the nucleus $T_2$WC and GWC at L5-S1 indicated that these measures could be used interchangeably and possibly provide redundant information. Given that $T_2$WC provides information on the spatial orientation of the ROI and the fluid distribution of each pixel, it may be more informative than the GWC and likely more sensitive to minor changes. In particular, in discs with severe degeneration grades, the differences between $T_2$WC and geometric centroid may be the largest, and $T_2$WC would likely capture the most changes in response to loading as it uses fluid concentration in addition to geometric information. Our correlation estimates did not suggest that both disc and nucleus $T_2$WC provide the same information as the coefficient of determination ($R^2$) was less than 50% for horizontal coordinates. This likely indicates that both points are needed for evaluating the effect of degeneration.

Nevertheless, examining the response to loading would likely also offer improved diagnostic accuracy provided the novel fluid distribution parameters were considered. However, examining the response to loading requires a special loading apparatus, additional image acquisition time, use of an image segmentation software (which will be made freely available soon) and measurement extraction time. Given the resources involved and the worldwide efforts to limit imaging cost for LBP [252], on the basis of our results, we only recommend the application of stress imaging (compression/loading) for those people with chronic LBP once the benefits associated with a meaningful classification for research and clinical application outweigh the burden of additional time and cost. The higher cost of MR imaging under loading may be justified if the uncertainty about MRI specificity is reduced, and if it helps inform better management decisions. Larger studies, including replications, diagnostic accuracy, prognostic, treatment studies shown effective for certain subgroups, defined based on our biomarkers response to loading, are

however needed to investigate further the clinical value of the proposed biomarkers showing promise to improve the specificity of MRI for LBP and the ability to detect the response to loading.

**Study strengths**

***Systematic literature review.*** We conducted an extensive systematic literature review on the effects of loading to identify candidate biomarkers and loading protocols. Our review followed PRISMA reporting methodology. A thorough literature search informed by content experts and a medical librarian helped find more than 4200 abstracts and 105 papers, which were fully reviewed by at least two reviewers each. Twenty-one paper then underwent an extensive standard data extraction. Quality appraisal combined criteria from various tools as no tools yet gather consensus from the scientific community for this type of review. Reviewers received standard training and achieved adequate reliability. Despite the literature presenting too much heterogeneity for meta-analysis, a rigorous level of evidence summary statement formulation strategy was used to draw conclusions. A gap in knowledge regarding the effects of loading on the imaging presentations of the lumbar spine was detected, particularly, a lack of any high-quality research and specifically of research on whether the response to loading could help increase the specificity of MRI for LBP. This gap suggested that further high-quality studies, comparing responses in subjects with symptoms of LBP to matched healthy subjects, are needed to establish a strong correlation between imaging findings and patients' symptoms. Our review also illustrated a gap about many quantitative measurements. For example, no evidence was found on the effects of loading on the disc fluid content by either measuring MR signal intensity, mean $T_2$ times or diffusion coefficients. Although we identified works by Beattie et al. [102, 212, 253] on the disc diffusion, they were not captured by our mesh terms, likely due to different keywords they chose. Further high-quality studies are needed to quantify the effects of loading, identify additional quantitative MRI measurements, to address the heterogeneity issues encountered in this review and to establish a strong correlation between these measurement constructs and patients' symptoms.

***Computer-aided measurement.*** Most quantitative MRI studies of the lumbar spine have used manual segmentation techniques. Manual segmentation is usually a tedious and time-consuming process [245, 246], with limited reliability [247, 248]. We tested the solution of developing a CAM to improve reliability in patients' representative of our targeted population. We developed an accurate and highly reliable CAM for quick measurements of discs and vertebrae for extracting continuous quantitative measures of signal intensity changes, morphology, angulation

and translation of the motion segments. Our reliability experiment followed standard design and reporting guidelines, and our evaluators were blinded to participants information and their prior measurements [254]. Of the few segmentation techniques employed for normal and scoliotic discs [187, 245, 247, 250, 255], most have focused on the accuracy of the segmentation, with limited consideration for specifically the related measures of clinical interest [187, 214, 245, 250]. This gap was addressed by quantifying the reliability for all the following measurements later used in the effect of loading studies: disc mean intensity, height, area and width, nucleus mean intensity, area and width as well as motion segment angle. We also implemented new biomarkers in our software, including $T_2$WC and GWC. By comparing the reliability of a wide array of disc height measurement strategies in the same sample, we obtained objective comparison allowing an informed choice on which was the most reliable approach.

***Effects of loading.*** Novel image acquisition methods and quantitative biomarkers were proposed and shown to have promise to advance meaningful classification of LBP and disc degeneration. Our studies on the effect of loading addressed: 1) the need for measurements to detect responses to loading which may help define clear phenotypes in genetic studies and carefully monitor response to spine degeneration treatments, 2) the need for more specificity when using MRI to identify pain relevant findings, 3) the need for evaluating the clinical value of the traditional biomarkers, by matching participants with and without pain for characteristics known to affect MRI findings and 4) the need to focus on participants with chronic low back pain for which MR imaging may be more indicated than for acute back pain. Further, refining the quantification of the vertebra and disc morphology, including area, height and fluid distribution indicators under different loading conditions was shown to have value.

To our knowledge, this dissertation was the first study on the effects of compression and traction on the motion segment fluid using quantitative $T_2$-mapping, a strategy that addresses the limitation of measuring signal intensity on $T_2$-weighted images. We employed a 3T MRI system that compared to kinetic MRI system had a better signal to noise ratio (SNR) and thus resolution to minimise pixel size and thus to minimise partial volume averaging artefacts. An MRI loading apparatus was developed to load the lumbar spine during MR imaging within the confine of existing systems normally used to image patient in relaxed supine position. Loading applied with the device was well tolerated with no participants requesting interruption in the loading application. We standardised the loading parameters used for all participants including loading

order, pre-imaging loading durations and relative loading magnitude. We also proposed highly reliable imaging biomarkers to capture fluid content, distribution and instability characteristics with $T_2$WC and GWC of different ROIs. The traditional measurements, including disc height, width and nucleus width, were selected based on the presence of a knowledge gap about their clinical value to determine whether they hold promise for larger scale studies.

**Study limitations**

Nevertheless, limitations to the present dissertation should also be acknowledged.

***Systematic literature review.*** Our systematic review solely included quantitative studies published in English. We did not try to capture those studies that employed qualitative ordinal ratings to evaluate the effects of loading. Although two people were bilingual in our team, they could not act as a second reviewer for each other to include additional languages, which may introduce a language bias [145]. We did not register our study in the PROSPERO, as we did not conduct an intervention review. The heterogeneity of the methodological approaches adopted by different authors made it difficult to draw meaningful conclusions and prevented conducting a meta-analysis. We rely on authors' reported results, which possibly introduces an outcome reporting bias [145]. Indeed, outcomes reported by different studies differed significantly and possibly not all measurements recorded by different authors were published. While some recommend that systematic reviews be backed up with correspondence with the authors for replication and/or reproduction of their results [256], this was not feasible due to resource constraints.

***Reliability studies.*** Although our CAM was created to segment annulus and cartilaginous endplates from adjacent tissues, our reliability estimates for these areas were suboptimal, possibly due to difficulties in defining a clear boundary between neighbouring tissues using our largest signal intensity difference segmentation strategy, particularly in severely degenerated discs. We conducted the reliability study on $T_2$-weighted images with a different scanner and a lower resolution than the one used in later chapters. Better resolution and image quality could improve our reliability estimates. Reliability was the lowest for the posterior annulus. In discs with high-intensity zones [184–186], it is challenging to determine a clear boundary between the posterior annulus and the cerebrospinal fluid. Further, in patients with severe degeneration with low annulus intensity, but with contours occupying the spinal canal and displacing cerebrospinal fluid and

nerve roots, it is difficult to contrast the annulus from the nerve tissue or canal because the signal intensity in the canal is reduced. Employment of parabolic curve fitting techniques using more neighbouring points to inform segmentation to detect annulus contours may help increase the reliability of the segmentation. However, displaced materials (e.g. herniation) may be more likely to be missed. Other segmenting solutions could include employing atlas-based segmentation techniques or artificial intelligence tools [246].

Another limitation was that the semi-automated segmentation algorithm still relied partially on user input. First, input was needed for drawing lines to begin the segmentation of the upper and lower disc-vertebra boundaries for each disc, which increased the analysis time. Although the program lets rater to amend outliers, this could introduce another source of inter-rater variability, as the location of the corrected points would be subjective. However, our high inter-rater reliability results indicated that this did not play a big role. Automating disc-vertebra numbering may be a solution, but existing algorithms require determining image specific intensity or still require manual selection of at least one ROI to begin segmentation. Therefore, these algorithms also require user input/time which ultimately may also affect the segmentation results [182]. To our knowledge, no fully automated algorithms are yet available.

***MR imaging and image segmentation.*** In selecting loading conditions to compare, we could not quantify the disc fluid content in the extended position. In addition, in comparing the ability of our measurements to detect changes in the extension study we could not compare to the supine relaxed position to data obtained in the extended position, as extension increases the distance between the body and the spine coil and thus reduces the SNR. An additional limitation in all studies was that we only measured the midsagittal slice of the IVDs, and it is possible that changes in fluid content or distribution in other regions of the disc may have been missed or that reliability would differ if assessed on other slices. Pathologically relevant disc measurements may be observed in other planes. Nevertheless, our analysis of the midsagittal plane found pain relevant findings and our measurements were able to detect changes in response to loading in this plane. Unfortunately, it was not possible to distinguish the annulus from the adjacent ventral and dorsal ligamentous structures with $T_2$ based imaging techniques. Therefore, portions of the anterior and posterior longitudinal ligaments were included in the disc measurements reported. Another limitation was the pixel size. We were only able to capture changes within the midsagittal plane equal to or greater than 0.4 mm. Because slice thickness was 5mm, the ability to detect changes

occurring laterally within the disc may have been limited and partial volume averaging issues may have affected the quantitative $T_2$ times estimated for each pixel. Our samples included a very small number of discs with Pfirrmann grade V degeneration where it may be difficult or impossible to segment the nucleus hyper-intense areas. Generalizability should be investigated in highly degenerated discs. The present study was explanatory and many variables were examined, there is a chance that some of the statistically significant results were observed by chance. A replication study, justified by the promising results, is needed to confirm which biomarkers are truly useful to detect differences in degeneration levels and between subjects with and without pain.

***Sample size.*** Two sets of images were used in the present dissertation. For the reliability and extension loading study, the baseline mid-sagittal $T_2$-weighted MR images from 22 volunteers with LBP recruited for another study on extension exercise were used [179]. Our small sample consisted of consecutive volunteers with low back pain and included mainly Pfirrmann's grade three and four degenerated discs with a limited number of normal or slightly degenerated discs (Pfirrmann's grade one and two) or severe degeneration (Pfirrmann's grade five), which could likely affect the disc response to loading. With a sample of 22 subjects, with repeated measurements, an effect size of 0.49 in $T_2$WC location could be detected with statistical significance.

For the experimental studies, the mid-sagittal $T_2$-weighted MR images from 35 volunteers with and without chronic LBP were analysed. Our small sample consisted of consecutive volunteers responding to a mass email invitation and included mainly slightly degenerated discs (Pfirrmann's grade two (65.7%) with limited number of discs with normal (Pfirrmann's grade one =1.7%) or severe degeneration (Pfirrmann's grade four =11.4% and five =1.1%), which could likely affect the disc response to loading. We did not have enough statistical power to evaluate the effects of loading on discs within each of the different degeneration grades. In the pain chapter, recruitment of participants continues to match all people with chronic LBP with asymptomatic people. Therefore, we were not able to compare our candidate biomarkers among all people with and without LBP with our targeted statistical power. Nevertheless, moderate and large effects sizes were reported and statistical significance observed justifying further investigations of the proposed biomarkers and loading conditions.

**Further Development Work Suggested**

***Systematic literature review.*** Systematic reviews are strong tools for identifying knowledge gaps, highlighting methodological inconsistencies and weaknesses, enhancing and promoting evidence-informed diagnosis and classification of the low back disorders, particularly in areas with a strong and well-developed evidence base. Our systematic review can be expanded by including more languages to minimise selection bias. A hand search of the references and relevant publications may also help capture more studies. Finally, in the future, conducting a meta-analysis of the data within the few categories where multiple studies become available may help identify the most promising candidate biomarkers for future studies.

***Reliability studies.*** Although our CAM was designed to segment annulus and cartilaginous endplates from adjacent tissues, the reliability for these regions was suboptimal. There is mounting evidence about the possible role of endplates in those with back pain, and also the posterior annulus is subject to volume changes under loading due to herniation, which is common in people with back pain. After refining segmentation strategies by using more robust edge detection algorithms, conducting a new inter- and intra-rater reliability study on a sufficient sample size within each category of degenerated discs may help estimate the reliability of these regions. Such reliability studies would help understand if measurements present adequate reliability also for the extremes of degeneration, which were underrepresented in the present study. Further, the reliability generalizability could be tested by using new raters, likely from clinical settings. Another proposed development is to minimise user input for drawing lines to begin the segmentation by fully automating level identification, which could decrease the analysis time and likely further improve reliability. Developing methods to avoid having to rely on user input to fix outlier contour points would also be a good avenue for future research.

***Loading studies.*** To address the limitation of using discrete scores to detect responses to loading, developing additional quantitative imaging biomarkers beyond the promising ones identified in the present study may be needed for meaningful classification of the disc and motion segment degeneration. As the evidence is mounting about the possible role of cartilaginous endplates on the development of disc degeneration and LBP [257], evaluating the effects of loading on $T_2$WC and GWC of the cartilaginous endplate may help determine the value of these biomarkers for detecting clinically relevant endplate abnormalities. Likewise, as the posterior annulus can be affected by herniation, loading could affect the location of GWC and $T_2$WC of an annulus ROI,

exploring these novel biomarkers may help identify the pathway of fluid redistribution with loading non-invasively.

We did not try to capture para-sagittal images and other standard views as it could increase our image acquisition time. However, patho-anatomically relevant information may be found in other images. Parasagittal views can visualise the intervertebral foramen, and axial view can visualise DCSA. Evaluating the effects of loading on the aforementioned planes may help evaluate the clinical value of the suggested biomarkers and capture new pathological biomarkers. Evaluating traction when going from unloaded could also help whether keep all loading conditions or drop some, which are not very informative. Evaluating other loading magnitude or duration could help determine the most informative loading magnitude and duration.

The promising results from this preliminary study suggesting that the candidate biomarkers can detect response to loading and distinguish participants with and without low back pain justify conducting a study with a bigger sample of both participants with and without back pain to help determine the clinical value of the proposed biomarkers. A long-term follow-up study in both participants with and without back pain could also help track changes in the candidate biomarkers to investigate their prognostic value.

**Conclusion**

From the literature review, 12 candidate biomarkers were investigated to improve the specificity of MRI for degeneration and low back pain. All candidate biomarkers were shown to provide reliable measurements. The novel biomarker ($T_2$WC) reflecting disc fluid distribution had equal of better ability to detect significant response from before and after extension exercise. The experimental study indicated that disc height, width, mean $T_2$ time and $T_2$WC, as well as, nucleus mean $T_2$ time, GWC and $T_2$-WC hold promise as sensitive biomarkers to capture the response of the lumbar spine to different loading conditions and likely to monitor degenerative and regenerative changes of the lumbar motion segments. Traction compared to compression was more sensitive to capture the load-induced changes of the lumbar motion segments.

The most sensitive *biomarkers for pain* were the horizontal (large effect sizes) and vertical coordinates (medium effect size) of the disc and nucleus $T_2$WC, the horizontal and vertical coordinates of the nucleus GWC (both medium effect size), and disc height (small effect size). Different responses to loading between pain groups were most often observed at L5-S1. Among these biomarkers, three (horizontal coordinates of the disc $T_2$WC and nucleus GWC and disc height) showed potentials to detect differences between patients with and without pain using traditional unloaded MR imaging, the rest. However, more studies are needed to quantify the cut-off point between normal and problematic biomarker measurements and investigate the clinical value of the proposed biomarkers showing promise to improve the specificity of MRI for LBP and the ability to detect the response to loading.

Since collecting biomarkers requires a special loading apparatus and image segmentation software, we recommend the application of stress imaging (compression/loading) for those people with idiopathic chronic LBP only once the benefits associated with a meaningful classification for research and clinical application outweigh the burden of additional time and cost.

## REFERENCES


1. Hoy D, Bain C, Williams G, et al (2012) A systematic review of the global prevalence of low back pain. Arthritis Rheum 64:2028–2037.
2. Bogduk N (2012) Clinical and radiological anatomy of the lumbar spine, 5th ed. Churchill Livingstone, Edinburgh ; New York
3. Deyo RA, Weinstein JN (2001) Low back pain. N Engl J Med 344:363–370.
4. Chou R, Qaseem A, Snow V, et al (2007) Diagnosis and treatment of low back pain: a joint clinical practice guideline from the American College of Physicians and the American Pain Society. Ann Intern Med 147:478–491.
5. Adams MA, Roughley PJ (2006) What is intervertebral disc degeneration, and what causes it? Spine (Phila Pa 1976) 31:2151–2161.
6. Brown MF, Hukkanen M V, McCarthy ID, et al (1997) Sensory and sympathetic innervation of the vertebral endplate in patients with degenerative disc disease. J bone Jt surgeryBritish Vol 79:147–153.
7. Freemont AJ, Peacock TE, Goupille P, et al (1997) Nerve ingrowth into diseased intervertebral disc in chronic back pain. Lancet 350:178–181.
8. Kauppila LI (2009) Atherosclerosis and disc degeneration/low-back pain--a systematic review. Eur J Vasc Endovasc Surg 37:661–670.
9. Jarvik JG, Deyo RA (2002) Diagnostic evaluation of low back pain with emphasis on imaging. Ann Intern Med 137:586–597.
10. Chou R, Qaseem A, Owens DK, et al (2011) Diagnostic imaging for low back pain: advice for high-value health care from the American College of Physicians. Ann Intern Med 154:181–9.
11. Weiner DK, Kim YS, Bonino P, Wang T (2006) Low back pain in older adults: are we utilizing healthcare resources wisely? Pain Med 7:143–150.
12. Hilton RC, Ball J (1984) Vertebral rim lesions in the dorsolumbar spine. Ann Rheum Dis 43:302–307.
13. Wang Y, Battie MC, Videman T (2012) A morphological study of lumbar vertebral endplates: radiographic, visual and digital measurements. Eur Spine J 21:2316–2323.
14. Vernon-Roberts B, Pirie CJ (1977) Degenerative changes in the intervertebral discs of the lumbar spine and their sequelae. Rheumatol Rehabil 16:13–21.
15. Nagy SA, Juhasz I, Komaromy H, et al (2014) A statistical model for intervertebral disc degeneration: determination of the optimal T2 cut-off values. Clin Neuroradiol 24:355–363.
16. Pfirrmann CW, Metzdorf A, Zanetti M, et al (2001) Magnetic resonance classification of lumbar intervertebral disc degeneration. Spine (Phila Pa 1976) 26:1873–1878.
17. Kornberg M (1988) MRI diagnosis of traumatic Schmorl's node. A case report. Spine (Phila Pa 1976) 13:934–5.
18. Teichtahl AJ, Urquhart DM, Wang Y, et al (2016) Lumbar disc degeneration is associated with modic change and high paraspinal fat content - a 3.0T magnetic resonance imaging study. BMC Musculoskelet Disord 17:439.

19. Takashima H, Takebayashi T, Yoshimoto M, et al (2012) Correlation between T2 relaxation time and intervertebral disk degeneration. Skeletal Radiol 41:163–167.
20. Perry J, Haughton V, Anderson PA, et al (2006) The value of T2 relaxation times to characterize lumbar intervertebral disks: preliminary results. AJNRAmerican J Neuroradiol 27:337–342.
21. Botsford DJ, Esses SI, Ogilvie-Harris DJ (1994) In vivo diurnal variation in intervertebral disc volume and morphology. Spine (Phila Pa 1976) 19:935–40.
22. Lee, S; Fredericson, M; Lang P (2005) The effect of axial loading and spine position on intervertebral disc hydration: an in vivo pilot study. J Back Musculoskelet Rehabil 74:512–516.
23. Paajanen H, Lehto I, Alanen A, et al (1994) Diurnal fluid changes of lumbar discs measured indirectly by magnetic resonance imaging. J Orthop Res 12:509–514.
24. Roberts N, Hogg D, Whitehouse GH, Dangerfield P (1998) Quantitative analysis of diurnal variation in volume and water content of lumbar intervertebral discs. Clin Anat 11:1–8.
25. Malko JA, Hutton WC, Fajman WA (1999) An in vivo magnetic resonance imaging study of changes in the volume (and fluid content) of the lumbar intervertebral discs during a simulated diurnal load cycle. Spine (Phila Pa 1976) 24:1015–22.
26. Ludescher B, Effelsberg J, Martirosian P, et al (2008) T2- and diffusion-maps reveal diurnal changes of intervertebral disc composition: An in vivo MRI study at 1.5 Tesla. J Magn Reson Imaging 28:252–257.
27. Chokan K, Murakami H, Endo H, et al (2016) Evaluation of Water Retention in Lumbar Intervertebral Disks Before and After Exercise Stress With T2 Mapping. Spine (Phila Pa 1976) 41:E430–E436.
28. Hashemi RH, Bradley WG, Lisanti CJ (2010) MRI: the basics. Lippincott Williams & Wilkins, Philadelphia, PA
29. McRobbie DW (2007) MRI from picture to proton. Cambridge University Press, Cambridge, UK ; New York
30. Currie S, Hoggard N, Craven IJ, et al (2013) Understanding MRI: basic MR physics for physicians. Postgrad Med J 89:209–223.
31. Westbrook C, Roth CK, Talbot J (2011) MRI in practice. Wiley-Blackwell, Chichester, West Sussex ; Malden, MA
32. Schmidt H, Reitmaier S, Graichen F, Shirazi-Adl A (2016) Review of the fluid flow within intervertebral discs - How could in vitro measurements replicate in vivo? J Biomech 49:3133–3146.
33. Antoniou J, Pike GB, Steffen T, et al (1998) Quantitative magnetic resonance imaging in the assessment of degenerative disc disease. Magn Reson Med 40:900–7.
34. Boos N, Wallin A, Gbedegbegnon T, et al (1993) Quantitative MR imaging of lumbar intervertebral disks and vertebral bodies: influence of diurnal water content variations. Radiology 188:351–354.
35. Chatani K, Kusaka Y, Mifune T, Nishikawa H (1993) Topographic differences of 1H-NMR relaxation times (T1, T2) in the normal intervertebral disc and its relationship to water content. Spine (Phila Pa 1976) 18:2271–5.

36. Kim DJ, Suh JS, Jeong EK, et al (1999) Correlation of laminated MR appearance of articular cartilage with histology, ascertained by artificial landmarks on the cartilage. J Magn Reson Imaging 10:57–64.

37. Marinelli NL, Haughton VM, Munoz A, Anderson PA (2009) T2 relaxation times of intervertebral disc tissue correlated with water content and proteoglycan content. Spine (Phila Pa 1976) 34:520–524.

38. Mlynárik V, Degrassi A, Toffanin R, et al (1996) Investigation of laminar appearance of articular cartilage by means of magnetic resonance microscopy. Magn Reson Imaging 14:435–42.

39. Tertti M, Paajanen H, Laato M, et al (1991) Disc degeneration in magnetic resonance imaging. A comparative biochemical, histologic, and radiologic study in cadaver spines. Spine (Phila Pa 1976) 16:629–634.

40. Watanabe A, Benneker LM, Boesch C, et al (2007) Classification of Intervertebral Disk Degeneration with Axial T2 Mapping. Am J Roentgenol 189:936–942.

41. Wilke HJ, Neef P, Caimi M, et al (1999) New in vivo measurements of pressures in the intervertebral disc in daily life. Spine (Phila Pa 1976) 24:755–762.

42. Nachemson A, Morris JM (1964) In Vivo Measurements of Intradiscal Pressure. Discometry, a Method for the Determination of Pressure in the Lower Lumbar Discs. J bone Jt surgeryAmerican Vol 46:1077–1092.

43. Claus A, Hides J, Moseley GL, Hodges P (2008) Sitting versus standing: does the intradiscal pressure cause disc degeneration or low back pain? J Electromyogr Kinesiol 18:550–8.

44. Tarantino U, Fanucci E, Iundusi R, et al (2013) Lumbar spine MRI in upright position for diagnosing acute and chronic low back pain: statistical analysis of morphological changes. J Orthop Traumatol 14:15–22.

45. Beattie PF, Arnot CF, Donley JW, et al (2010) The immediate reduction in low back pain intensity following lumbar joint mobilization and prone press-ups is associated with increased diffusion of water in the L5-S1 intervertebral disc. J Orthop Sports Phys Ther 40:256–64.

46. Frank JW, Kerr MS, Brooker AS, et al (1996) Disability resulting from occupational low back pain. Part I: What do we know about primary prevention? A review of the scientific evidence on prevention before disability begins. Spine (Phila Pa 1976) 21:2908–2917.

47. Wegner I, Widyahening IS, van Tulder MW, et al (2013) Traction for low-back pain with or without sciatica. Cochrane database Syst Rev 8:CD003010.

48. Maher C, Underwood M, Buchbinder R (2016) Non-specific low back pain. Lancet. doi: 10.1016/S0140-6736(16)30970-9

49. Deyo RA, Dworkin SF, Amtmann D, et al (2014) Report of the NIH Task Force on Research Standards for Chronic Low Back Pain. Pain Med 15:1249–1267.

50. Roberts S, Menage J, Urban JP (1989) Biochemical and structural properties of the cartilage end-plate and its relation to the intervertebral disc. Spine (Phila Pa 1976) 14:166–174.

51. Twomey LT, Taylor JR (1987) Age changes in lumbar vertebrae and intervertebral discs. Clin Orthop Relat Res (224):97–104.

52. Levangie PK, Norkin CC, Levangie PK (2011) Joint structure and function: a

comprehensive analysis. F.A. Davis Co., Philadelphia

53. Grunhagen T, Shirazi-Adl A, Fairbank JC, Urban JP (2011) Intervertebral disk nutrition: a review of factors influencing concentrations of nutrients and metabolites. Orthop Clin North Am 42:465–77, vii.
54. Urban JP, Roberts S (1995) Development and degeneration of the intervertebral discs. Mol Med Today 1:329–335.
55. Colombini A, Lombardi G, Corsi MM, Banfi G (2008) Pathophysiology of the human intervertebral disc. Int J Biochem Cell Biol 40:837–842.
56. Stairmand JW, Holm S, Urban JP (1991) Factors influencing oxygen concentration gradients in the intervertebral disc. A theoretical analysis. Spine (Phila Pa 1976) 16:444–449.
57. Shah NJ, Ermer V, Oros-Peusquens A-M (2011) Measuring the Absolute Water Content of the Brain Using Quantitative MRI. In: Modo M, Bulte JWM (eds) Magn. Reson. Neuroimaging Methods Protoc. Humana Press, New York, pp 29–64
58. Adams MA, Hutton WC (1983) The effect of posture on the fluid content of lumbar intervertebral discs. Spine (Phila Pa 1976) 8:665–671.
59. Sivan SS, Wachtel E, Tsitron E, et al (2008) Collagen turnover in normal and degenerate human intervertebral discs as determined by the racemization of aspartic acid. J Biol Chem 283:8796–8801.
60. Cassinelli EH, Hall RA, Kang JD (2001) Biochemistry of intervertebral disc degeneration and the potential for gene therapy applications. Spine J 1:205–214.
61. Urban JP, Winlove CP (2007) Pathophysiology of the intervertebral disc and the challenges for MRI. J Magn Reson Imaging 25:419–432.
62. Sivan SS, Tsitron E, Wachtel E, et al (2006) Aggrecan turnover in human intervertebral disc as determined by the racemization of aspartic acid. J Biol Chem 281:13009–13014.
63. Johnson EF, Chetty K, Moore IM, et al (1982) The distribution and arrangement of elastic fibres in the intervertebral disc of the adult human. J Anat 135:301–309.
64. Freemont AJ (2009) The cellular pathobiology of the degenerate intervertebral disc and discogenic back pain. Rheumatology (Oxford) 48:5–10.
65. Urban JP, Smith S, Fairbank JC (2004) Nutrition of the intervertebral disc. Spine (Phila Pa 1976) 29:2700–2709.
66. Healey EL, Burden AM, McEwan IM, Fowler NE (2011) Diurnal variation in stature: do those with chronic low-back pain differ from asymptomatic controls? Clin Biomech (Bristol, Avon) 26:331–6.
67. Jackson AR, Huang C-YC, Brown MD, Gu WY (2011) 3D finite element analysis of nutrient distributions and cell viability in the intervertebral disc: effects of deformation and degeneration. J Biomech Eng 133:91006.
68. Jünger S, Gantenbein-Ritter B, Lezuo P, et al (2009) Effect of limited nutrition on in situ intervertebral disc cells under simulated-physiological loading. Spine (Phila Pa 1976) 34:1264–71.
69. Beattie PF, Donley JW, Arnot CF, Miller R (2009) The change in the diffusion of water in normal and degenerative lumbar intervertebral discs following joint mobilization compared

to prone lying. J Orthop Sports Phys Ther 39:4–11.

70. Terahata N, Ishihara H, Ohshima H, et al (1994) Effects of axial traction stress on solute transport and proteoglycan synthesis in the porcine intervertebral disc in vitro. Eur Spine J 3:325–330.
71. Ito K, Creemers L (2013) Mechanisms of intervertebral disk degeneration/injury and pain: a review. Glob spine J 3:145–152.
72. Urban JP, Roberts S (2003) Degeneration of the intervertebral disc. Arthritis Res Ther 5:120–130.
73. Battié MC, Videman T, Kaprio J, et al (2009) The Twin Spine Study: Contributions to a changing view of disc degeneration? Spine J 9:47–59.
74. Adams MA, McNally DS, Dolan P (1996) "Stress" distributions inside intervertebral discs. The effects of age and degeneration. J Bone Jt Surg - Br Vol 78:965–972.
75. Adams MA, Freeman BJ, Morrison HP, et al (2000) Mechanical initiation of intervertebral disc degeneration. Spine (Phila Pa 1976) 25:1625–36.
76. Carragee E, Alamin T, Cheng I, et al (2006) Are first-time episodes of serious LBP associated with new MRI findings?. Spine J Off J North Am Spine Soc 6:624–635.
77. Soukane DM, Shirazi-Adl A, Urban JP (2007) Computation of coupled diffusion of oxygen, glucose and lactic acid in an intervertebral disc. J Biomech 40:2645–2654.
78. Videman T, Levalahti E, Battie MC (2007) The effects of anthropometrics, lifting strength, and physical activities in disc degeneration. Spine (Phila Pa 1976) 32:1406–1413.
79. Wang Y, Videman T, Battie MC (2013) Morphometrics and lesions of vertebral end plates are associated with lumbar disc degeneration: evidence from cadaveric spines. J bone Jt surgeryAmerican Vol 95:e26.
80. Pappou IP, Cammisa Jr FP, Girardi FP (2007) Correlation of end plate shape on MRI and disc degeneration in surgically treated patients with degenerative disc disease and herniated nucleus pulposus. Spine J 7:32–38.
81. Kauppila LI, Tallroth K (1993) Postmortem angiographic findings for arteries supplying the lumbar spine: their relationship to low-back symptoms. J Spinal Disord 6:124–129.
82. Kauppila LI (1997) Prevalence of stenotic changes in arteries supplying the lumbar spine. A postmortem angiographic study on 140 subjects. Ann Rheum Dis 56:591–595.
83. Madsen R, Jensen TS, Pope M, et al (2008) The effect of body position and axial load on spinal canal morphology: an MRI study of central spinal stenosis. Spine (Phila Pa 1976) 33:61–7.
84. Kimura S, Steinbach GC, Watenpaugh DE, Hargens AR (2001) Lumbar spine disc height and curvature responses to an axial load generated by a compression device compatible with magnetic resonance imaging. Spine (Phila Pa 1976) 26:2596–600.
85. Cheung KM, Karppinen J, Chan D, et al (2009) Prevalence and pattern of lumbar magnetic resonance imaging changes in a population study of one thousand forty-three individuals. Spine (Phila Pa 1976) 34:934–940.
86. Izzo R, Popolizio T, D'Aprile P, Muto M (2015) Spinal pain. Eur J Radiol 84:746–756.
87. Maigne J-Y, Lapeyre E, Morvan G, Chatellier G (2003) Pain immediately upon sitting down

and relieved by standing up is often associated with radiologic lumbar instability or marked anterior loss of disc space. Spine (Phila Pa 1976) 28:1327–34.

88. Mauch F, Jung C, Huth J, Bauer G (2010) Changes in the lumbar spine of athletes from supine to the true-standing position in magnetic resonance imaging. Spine (Phila Pa 1976) 35:1002–7.
89. Hansson T, Suzuki N, Hebelka H, Gaulitz A (2009) The narrowing of the lumbar spinal canal during loaded MRI: the effects of the disc and ligamentum flavum. Eur Spine J 18:679–86.
90. Lee S-U, Hargens AR, Fredericson M, Lang PK (2003) Lumbar spine disc heights and curvature: upright posture vs. supine compression harness. Aviat Space Environ Med 74:512–6.
91. Harrison DE, Cailliet R, Harrison DD, et al (2002) Changes in sagittal lumbar configuration with a new method of extension traction: nonrandomized clinical controlled trial. Arch Phys Med Rehabil 83:1585–1591.
92. Wang Y-C, Jeng C-M, Wu C-Y, et al (2008) Dynamic effects of axial loading on the lumbar spine during magnetic resonance imaging in patients with suspected spinal stenosis. J Formos Med Assoc 107:334–9.
93. Wolfer LR, Derby R, Lee JE, Lee SH (2008) Systematic review of lumbar provocation discography in asymptomatic subjects with a meta-analysis of false-positive rates. Pain Physician 11:513–538.
94. Carragee EJ, Lincoln T, Parmar VS, Alamin T (2006) A Gold Standard Evaluation of the "Discogenic Pain" Diagnosis as Determined by Provocative Discography. Spine (Phila Pa 1976) 31:2115–2123.
95. Carragee EJ, Don AS, Hurwitz EL, et al (2009) 2009 ISSLS Prize Winner: Does discography cause accelerated progression of degeneration changes in the lumbar disc: a ten-year matched cohort study. Spine (Phila Pa 1976) 34:2338–45.
96. Rinkus KM, Knaub MA, Koes BW, et al (2008) Clinical and Diagnostic Evaluation of Low Back Pain. Semin Spine Surg 20:93–101.
97. Nitz WR, Reimer P (1999) Contrast mechanisms in MR imaging. Eur Radiol 9:1032–1046.
98. Klose U, Jiru F (2005) Principles of MR Spectroscopy and Chemical Shift Imaging. In: Landini L, Positano V, Santarelli M (eds) Adv. Image Process. Magn. Reson. Imaging. CRC Press, Taylor & Francis Group, Boca Raton, pp 369–411
99. Majumdar S, Link TM, Steinbach LS, et al (2011) Diagnostic tools and imaging methods in intervertebral disk degeneration. Orthop Clin North Am 42:501–11, viii.
100. Shah B, Anderson SW, Scalera J, et al (2011) Quantitative MR imaging: physical principles and sequence design in abdominal imaging. Radiographics 31:867–880.
101. Le Bihan D, Mangin JF, Poupon C, et al (2001) Diffusion tensor imaging: concepts and applications. J Magn Reson Imaging 13:534–546.
102. Beattie PF (2011) Diffusion-weighted magnetic resonance imaging of the musculoskeletal system: an emerging technology with potential to impact clinical decision making. J Orthop Sports Phys Ther 41:887–895.
103. Baliyan V, Das CJ, Sharma R, Gupta AK (2016) Diffusion weighted imaging: Technique

and applications. World J Radiol 8:785–798.

104. Li X (2011) MR T1ρ Relaxation Time Quantification in Cartilage. In: Link TM (ed) Cartil. Imaging Significance, Tech. New Dev. Springer-Verlag, Berlin, pp 159–169
105. Johannessen W, Auerbach JD, Wheaton AJ, et al (2006) Assessment of human disc degeneration and proteoglycan content using T1rho-weighted magnetic resonance imaging. Spine (Phila Pa 1976) 31:1253–1257.
106. Auerbach JD, Johannessen W, Borthakur A, et al (2006) In vivo quantification of human lumbar disc degeneration using T(1rho)-weighted magnetic resonance imaging. Eur Spine J 15 Suppl 3:S338-44.
107. Li X, Han ET, Newitt D, Majumdar S (2004) T1rho relaxation quantification using spiral imaging: a preliminary study. Conf Proc .Annual Int Conf IEEE Eng Med Biol Soc Eng Med Biol Soc Conf 2:1032–1035.
108. Nguyen-minh C, Riley L, Ho KC, et al (1997) Effect of degeneration of the intervertebral disk on the process of diffusion. AJNR Am J Neuroradiol 18:435–42.
109. Regatte RR, Akella S V, Lonner JH, et al (2006) T1rho relaxation mapping in human osteoarthritis (OA) cartilage: comparison of T1rho with T2. J Magn Reson Imaging 23:547–553.
110. Bolbos RI, Link TM, Ma CB, et al (2009) T1rho relaxation time of the meniscus and its relationship with T1rho of adjacent cartilage in knees with acute ACL injuries at 3 T. Osteoarthritis Cartilage 17:12–18.
111. Zuo J, Joseph GB, Li X, et al (2012) In vivo intervertebral disc characterization using magnetic resonance spectroscopy and T1ρ imaging: association with discography and Oswestry Disability Index and Short Form-36 Health Survey. Spine (Phila Pa 1976) 37:214–21.
112. Welsch GH, Hennig FF, Krinner S, Trattnig S (2014) T2 and T2* Mapping. Curr Radiol Rep 2:
113. Thompson JP, Pearce RH, Schechter MT, et al (1990) Preliminary evaluation of a scheme for grading the gross morphology of the human intervertebral disc. Spine (Phila Pa 1976) 15:411–415.
114. Jensen MC, Brant-Zawadzki MN, Obuchowski N, et al (1994) Magnetic Resonance Imaging of the Lumbar Spine in People without Back Pain. N Engl J Med 331:69–73.
115. Brinjikji W, Luetmer PH, Comstock B, et al (2015) Systematic literature review of imaging features of spinal degeneration in asymptomatic populations. AJNR Am J Neuroradiol 36:811–6.
116. Steffens D, Hancock MJ, Maher CG, et al (2014) Does magnetic resonance imaging predict future low back pain? A systematic review. Eur J Pain 18:755–765.
117. Brinjikji W, Diehn FE, Jarvik JG, et al (2015) MRI Findings of Disc Degeneration are More Prevalent in Adults with Low Back Pain than in Asymptomatic Controls: A Systematic Review and Meta-Analysis. Am J Neuroradiol 36:2394–2399.
118. Videman T, Battié MC, Parent E, et al (2008) Progression and determinants of quantitative magnetic resonance imaging measures of lumbar disc degeneration: a five-year follow-up of adult male monozygotic twins. Spine (Phila Pa 1976) 33:1484–90.

119. Videman T, Battié MC, Ripatti S, et al (2006) Determinants of the progression in lumbar degeneration: a 5-year follow-up study of adult male monozygotic twins. Spine (Phila Pa 1976) 31:671–8.

120. Elfering A, Semmer N, Birkhofer D, et al (2002) Risk factors for lumbar disc degeneration: a 5-year prospective MRI study in asymptomatic individuals. Spine (Phila Pa 1976) 27:125–34.

121. Li X, Majumdar S (2013) Quantitative MRI of articular cartilage and its clinical applications. J Magn Reson Imaging 38:991–1008.

122. Khalil JG, Nassr A, Maus TP (2012) Physiologic imaging of the spine. Radiol Clin North Am 50:599–611.

123. Pospiech J, Stolke D, Wilke HJ, Claes LE (1999) Intradiscal pressure recordings in the cervical spine. Neurosurgery 44:375–379.

124. Fujiwara A, An HS, Lim TH, Haughton VM (2001) Morphologic changes in the lumbar intervertebral foramen due to flexion-extension, lateral bending, and axial rotation: an in vitro anatomic and biomechanical study. Spine (Phila Pa 1976) 26:876–882.

125. Inufusa A, An HS, Lim TH, et al (1996) Anatomic changes of the spinal canal and intervertebral foramen associated with flexion-extension movement. Spine (Phila Pa 1976) 21:2412–2420.

126. Schmid MR, Stucki G, Duewell S, et al (1999) Changes in cross-sectional measurements of the spinal canal and intervertebral foramina as a function of body position: in vivo studies on an open-configuration MR system. AJRAmerican J Roentgenol 172:1095–1102.

127. Nilsson M, Lagerstrand K, Kasperska I, et al (2016) Axial loading during MRI influences T2-mapping values of lumbar discs: a feasibility study on patients with low back pain. Eur spine J 25:2856–2863.

128. Iwata T, Miyamoto K, Hioki A, et al (2013) In vivo measurement of lumbar foramen during axial loading using a compression device and computed tomography. J Spinal Disord Tech 26:E177-82.

129. Hioki A, Miyamoto K, Sakai H, Shimizu K (2010) Lumbar Axial Loading Device Alters Lumbar Sagittal Alignment Differently From Upright Standing Position. Spine (Phila Pa 1976) 35:995–1001.

130. Antonio Aguilera-Repiso J (2011) Pilot Study of the Dynamic Effects of Simulated Axial Loading Magnetic Resonance Imaging in Lumbar Disc Herniation. Open Spine J 3:9–11.

131. Buchbinder R, Blyth FM, March LM, et al (2013) Placing the global burden of low back pain in context. Best Pract Res Rheumatol 27:575–589.

132. Ehrlich GE (2003) Low back pain. Bull World Health Organ 81:671–676.

133. McKinnis LN (2014) Fundamentals of musculoskeletal imaging, 4th ed. F.A DAVIS, Philadelphia, PA

134. Karel YH, Verkerk K, Endenburg S, et al (2015) Effect of routine diagnostic imaging for patients with musculoskeletal disorders: A meta-analysis. Eur J Intern Med. doi: S0953-6205(15)00225-3 [pii]

135. McKenzie R, May S (2003) The lumbar spine mechanical diagnosis & therapy. Spinal Publications New Zealand

136. Tan Y, Aghdasi BG, Montgomery SR, et al (2012) Kinetic magnetic resonance imaging analysis of lumbar segmental mobility in patients without significant spondylosis. Eur Spine J 21:2673–9.

137. Zou J, Yang H, Miyazaki M, et al (2008) Missed lumbar disc herniations diagnosed with kinetic magnetic resonance imaging. Spine (Phila Pa 1976) 33:E140-4.

138. Jinkins JR, Dworkin JS, Damadian R V (2005) Upright, weight-bearing, dynamic-kinetic MRI of the spine: initial results. Eur Radiol 15:1815–1825.

139. Kanno H, Ozawa H, Koizumi Y, et al (2012) Dynamic change of dural sac cross-sectional area in axial loaded magnetic resonance imaging correlates with the severity of clinical symptoms in patients with lumbar spinal canal stenosis. Spine (Phila Pa 1976) 37:207–13.

140. Kong MH, Morishita Y, He W, et al (2009) Lumbar segmental mobility according to the grade of the disc, the facet joint, the muscle, and the ligament pathology by using kinetic magnetic resonance imaging. Spine (Phila Pa 1976) 34:2537–44.

141. Panjabi MM, Goel VK, Takata K (1982) Physiologic strains in the lumbar spinal ligaments. An in vitro biomechanical study 1981 Volvo Award in Biomechanics. Spine (Phila Pa 1976) 7:192–203.

142. Ombregt L (2013) A System of Orthopaedic Medicine. Churchill-Livingstone, Edinburgh

143. Leone A, Guglielmi G, Cassar-Pullicino VN, Bonomo L (2007) Lumbar intervertebral instability: a review. Radiology 245:62–77.

144. Modic MT, Ross JS (1991) Magnetic resonance imaging in the evaluation of low back pain. Orthop Clin North Am 22:283–301.

145. Higgins JPT, Green S, Collaboration C (2008) Cochrane handbook for systematic reviews of interventions. Wiley-Blackwell, Chichester, England ; Hoboken, NJ

146. Liberati A, Altman DG, Tetzlaff J, et al (2009) The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. J Clin Epidemiol 62:e1-34.

147. Mokkink LB, Terwee CB, Patrick DL, et al (2010) The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. J Clin Epidemiol 63:737–745.

148. Jerosch-Herold C (2005) An Evidence-Based Approach to Choosing Outcome Measures: A Checklist for the Critical Appraisal of Validity, Reliability and Responsiveness Studies. Br J Occup Ther 68:347–353.

149. Bialocerkowski A, Klupp N, Bragge P (2010) How to read and critically appraise a reliability article. Int J Ther Rehabil 17:114–120.

150. Vandenbroucke JP, von Elm E, Altman DG, et al (2007) Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration. Ann Intern Med 147:W163-94.

151. Cornelius LR, van der Klink JJ, Groothoff JW, Brouwer S (2011) Prognostic factors of long term disability due to mental disorders: a systematic review. J Occup Rehabil 21:259–274.

152. Portney LG, Watkins MP (2009) Foundations of clinical research: applications to practice. Pearson/Prentice Hall, Upper Saddle River, N.J.

153. Eberhardt K, Ganslandt O, Stadlbauer A (2014) Functional and quantitative magnetic

resonance myelography of symptomatic stenoses of the lumbar spine. Neuroradiology 56:1069–78.

154. Ahn T-J, Lee S-H, Choi G, et al (2009) Effect of intervertebral disk degeneration on spinal stenosis during magnetic resonance imaging with axial loading. Neurol Med Chir (Tokyo) 49:242–7; discussion 247.
155. Keorochana G, Taghavi CE, Lee K-B, et al (2011) Effect of sagittal alignment on kinematic changes and degree of disc degeneration in the lumbar spine: an analysis using positional MRI. Spine (Phila Pa 1976) 36:893–8.
156. Karadimas EJ, Siddiqui M, Smith FW, Wardlaw D (2006) Positional MRI changes in supine versus sitting postures in patients with degenerative lumbar spine. J Spinal Disord Tech 19:495–500.
157. Okawa A, Shinomiya K, Komori H, et al (1998) Dynamic motion study of the whole lumbar spine by videofluoroscopy. Spine (Phila Pa 1976) 23:1743–9.
158. Hu JK, Morishita Y, Montgomery SR, et al (2011) Kinematic Evaluation of Association between Disc Bulge Migration, Lumbar Segmental Mobility, and Disc Degeneration in the Lumbar Spine Using Positional Magnetic Resonance Imaging. Glob spine J 1:43–8.
159. Kinder A, Filho FP, Ribeiro E, et al (2012) Magnetic resonance imaging of the lumbar spine with axial loading: a review of 120 cases. Eur J Radiol 81:e561-4.
160. Kim YK, Lee JW, Kim H-J, et al (2013) Diagnostic advancement of axial loaded lumbar spine MRI in patients with clinically suspected central spinal canal stenosis. Spine (Phila Pa 1976) 38:E1342-7.
161. Chung SS, Lee CS, Kim SH, et al (2000) Effect of low back posture on the morphology of the spinal canal. Skeletal Radiol 29:217–23.
162. Edmondston SJ, Song S, Bricknell R V, et al (2000) MRI evaluation of lumbar spine flexion and extension in asymptomatic individuals. Man Ther 5:158–64.
163. Fazey PJ, Takasaki H, Singer KP (2010) Nucleus pulposus deformation in response to lumbar spine lateral flexion: an in vivo MRI investigation. Eur Spine J 19:1115–20.
164. Fujiwara A, Tamai K, An HS, et al (2000) The relationship between disc degeneration, facet joint osteoarthritis, and stability of the degenerative lumbar spine. J Spinal Disord 13:444–50.
165. Hirasawa Y, Bashir WA, Smith FW, et al (2007) Postural changes of the dural sac in the lumbar spines of asymptomatic individuals using positional stand-up magnetic resonance imaging. Spine (Phila Pa 1976) 32:E136-40.
166. Kozanek M, Wang S, Passias PG, et al (2009) Range of motion and orientation of the lumbar facet joints in vivo. Spine (Phila Pa 1976) 34:E689-96.
167. Nazari J, Pope MH, Graveling RA (2012) Reality about migration of the nucleus pulposus within the intervertebral disc with changing postures. Clin Biomech (Bristol, Avon) 27:213–7.
168. Geisser ME, Haig AJ, Tong HC, et al (2007) Spinal canal size and clinical symptoms among persons diagnosed with lumbar spinal stenosis. Clin J Pain 23:780–5.
169. Amundsen T, Weber H, Lilleås F, et al (1995) Lumbar spinal stenosis. Clinical and radiologic features. Spine (Phila Pa 1976) 20:1178–86.

170. Zamani AA, Moriarty T, Hsu L, et al (1998) Functional MRI of the lumbar spine in erect position in a superconducting open-configuration MR system: preliminary results. J Magn Reson Imaging 8:1329–1333.

171. Weishaupt D, Schmid MR, Zanetti M, et al (2000) Positional MR imaging of the lumbar spine: does it demonstrate nerve root compromise not visible at conventional MR imaging? Radiology 215:247–253.

172. Kuo Y-WW, Hsu Y-CC, Chuang I-TT, et al (2014) Spinal traction promotes molecular transportation in a simulated degenerative intervertebral disc model. Spine (Phila Pa 1976) 39:E550-6.

173. Chung T, Yang H, Ahn SJ, Park JH (2015) Herniated Lumbar Disks : Real- time MR Imaging Evaluation during. 275:755–762.

174. Gilbert JW, Martin JC, Wheeler GR, et al (2010) Lumbar Disk Protrusion Rates of Symptomatic Patients Using Magnetic Resonance Imaging. J Manipulative Physiol Ther 33:626–629.

175. Vertinsky AT, Krasnokutsky M V, Augustin M, Bammer R (2007) Cutting-edge imaging of the spine. Neuroimaging Clin N Am 17:117–136.

176. Riesenburger RI, Safain MG, Ogbuji R, et al (2014) A novel classification system of lumbar disc degeneration. J Clin Neurosci. doi: 10.1016/j.jocn.2014.05.052; 10.1016/j.jocn.2014.05.052

177. Trattnig S, Stelzeneder D, Goed S, et al (2010) Lumbar intervertebral disc abnormalities: comparison of quantitative T2 mapping with conventional MR at 3.0 T. Eur Radiol 20:2715–2722.

178. Mwale F, Iatridis JC, Antoniou J (2008) Quantitative MRI as a diagnostic tool of intervertebral disc matrix composition and integrity. Eur Spine J 17 Suppl 4:432–40.

179. Parent E, Fritz J HB (2010) Development of a preliminary prediction rule to identify patients with low back pain responding to extension exercises. World Congr. Low Back Pain Pelvic Pain

180. Goh S, Tan C, Price RI, et al (2000) Influence of age and gender on thoracic vertebral body shape and disc degeneration: an MR investigation of 169 cases. J Anat 197 Pt 4:647–57.

181. Périé D, Curnier D (2010) Effect of pathology type and severity on the distribution of MRI signal intensities within the degenerated nucleus pulposus: application to idiopathic scoliosis and spondylolisthesis. BMC Musculoskelet Disord 11:189.

182. Oktay AB, Albayrak NB, Akgul YS (2014) Computer aided diagnosis of degenerative intervertebral disc diseases from lumbar MR images. Comput Med Imaging Graph 38:613–619.

183. Neubert A, Fripp J, Engstrom C, et al (2014) Validity and reliability of computerized measurement of lumbar intervertebral disc height and volume from magnetic resonance images. Spine J 14:2773–2781.

184. Schellhas KP, Pollei SR, Gundry CR, Heithoff KB (1996) Lumbar disc high-intensity zone. Correlation of magnetic resonance imaging and discography. Spine (Phila Pa 1976) 21:79–86.

185. Aprill C, Bogduk N (1992) High-intensity zone: a diagnostic sign of painful lumbar disc on

magnetic resonance imaging. Br J Radiol 65:361–369.

186. Lam KS, Carlin D, Mulholland RC (2000) Lumbar disc high-intensity zone: the value and significance of provocative discography in the determination of the discogenic pain source. Eur Spine J 9:36–41.
187. Law MW, Tay K, Leung A, et al (2013) Intervertebral disc segmentation in MR images using anisotropic oriented flux. Med Image Anal 17:43–61.
188. Lundon K, Bolton K (2001) Structure and function of the lumbar intervertebral disk in health, aging, and pathologic conditions. J Orthop Sports Phys Ther 31:291-303–6.
189. Fardon DF, Williams AL, Dohring EJ, et al (2014) Lumbar disc nomenclature: version 2.0: Recommendations of the combined task forces of the North American Spine Society, the American Society of Spine Radiology and the American Society of Neuroradiology. Spine J 14:2525–2545.
190. Boos N, Wallin A, Aebi M, Boesch C (1996) A new magnetic resonance imaging analysis method for the measurement of disc height variations. Spine (Phila Pa 1976) 21:563–570.
191. Dabbs VM, Dabbs LG (1990) Correlation between disc height narrowing and low-back pain. Spine (Phila Pa 1976) 15:1366–1369.
192. Kumar A, Beastall J, Hughes J, et al (2008) Disc changes in the bridged and adjacent segments after Dynesys dynamic stabilization system after two years. Spine (Phila Pa 1976) 33:2909–2914.
193. Tunset A, Kjaer P, Samir Chreiteh S, Secher Jensen T (2013) A method for quantitative measurement of lumbar intervertebral disc structures: an intra- and inter-rater agreement and reliability study. Chiropr Man Therap 21:26.
194. Pearson AM, Spratt KF, Genuario J, et al (2011) Precision of lumbar intervertebral measurements: does a computer-assisted technique improve reliability? Spine (Phila Pa 1976) 36:572–80.
195. Hurxthal LM (1968) Measurement of anterior vertebral compressions and biconcave vertebrae. Am J Roentgenol Radium Ther Nucl Med 103:635–644.
196. Miyakoshi N, Abe E, Shimada Y, et al (1999) Anterior decompression with single segmental spinal interbody fusion for lumbar burst fracture. Spine24 ()(pp 67-73), 1999Date Publ 01 Jan 1999 24:67–73.
197. Videman T, Battié MC, Parent E, et al (2008) Progression and determinants of quantitative magnetic resonance imaging measures of lumbar disc degeneration: A five-year follow-up of adult male monozygotic twins. Spine (Phila Pa 1976) 33:1484–1490.
198. Hancock MJ, Battie MC, Videman T, Gibbons L (2010) The role of back injury or trauma in lumbar disc degeneration: an exposure-discordant twin study. Spine (Phila Pa 1976) 35:1925–1929.
199. Pope MH, Hanley EN, Matteri RE, et al (1977) Measurement of Intervertebral Disc Space Height. Spine (Phila Pa 1976) 2:282–286.
200. Matteri RE, Pope MH, Frymoyer JW (1976) A biplane radiographic method of determining vertebral rotation in postmortem specimens. Clin Orthop Relat Res (116):95–98.
201. Terwee CB, Bot SDM, de Boer MR, et al (2007) Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol 60:34–42.

202. Hioki A, Miyamoto K, Shimizu K, Inoue N (2011) Test-retest repeatability of lumbar sagittal alignment and disc height measurements with or without axial loading: a computed tomography study. J Spinal Disord Tech 24:93–8.

203. Shao Z, Rompe G, Schiltenwolf M (2002) Radiographic changes in the lumbar intervertebral discs and lumbar vertebrae with age. Spine (Phila Pa 1976) 27:263–268.

204. Twomey L, Taylor J, Furniss B (1983) Age changes in the bone density and structure of the lumbar vertebral column. J Anat 136:15–25.

205. Frobin W, Brinckmann P, Biggemann M, et al (1997) Precision measurement of disc height, vertebral height and sagittal plane displacement from lateral radiographic views of the lumbar spine. Clin Biomech (Bristol, Avon) 12 Suppl 1:S1–S63.

206. Ferguson SJ, Ito K, Nolte LP (2004) Fluid flow and convective transport of solutes within the intervertebral disc. J Biomech 37:213–221.

207. Ferrara L, Triano JJ, Sohn MJ, et al (2005) A biomechanical assessment of disc pressures in the lumbosacral spine in response to external unloading forces. Spine J 5:548–553.

208. Horner HA, Urban JP (2001) 2001 Volvo Award Winner in Basic Science Studies: Effect of nutrient supply on the viability of cells from the nucleus pulposus of the intervertebral disc. Spine (Phila Pa 1976) 26:2543–2549.

209. McMillan DW, Garbutt G, Adams MA (1996) Effect of sustained loading on the water content of intervertebral discs: implications for disc metabolism. Ann Rheum Dis 55:880–7.

210. Rajasekaran S, Babu JN, Arun R, et al (2004) ISSLS prize winner: A study of diffusion in human lumbar discs: a serial magnetic resonance imaging study documenting the influence of the endplate on diffusion in normal and degenerate discs. Spine (Phila Pa 1976) 29:2654–67.

211. Race A, Broom ND, Robertson P (2000) Effect of loading rate and hydration on the mechanical properties of the disc. Spine (Phila Pa 1976) 25:662–669.

212. Beattie PF, Morgan PS, Peters D (2008) Diffusion-weighted magnetic resonance imaging of normal and degenerative lumbar intervertebral discs: a new method to potentially quantify the physiologic effect of physical therapy intervention. J Orthop Sports Phys Ther 38:42–9.

213. Videman T, Leskinen T, Stalhammar H, et al (1994) Effect of glycosaminoglycan polysulphate on the human intervertebral discs based on stature measurements. Clin Biomech (Bristol, Avon) 9:206–7.

214. Perie D, Curnier D (2010) Effect of pathology type and severity on the distribution of MRI signal intensities within the degenerated nucleus pulposus: application to idiopathic scoliosis and spondylolisthesis. BMC Musculoskelet Disord 11:189.

215. Cohen J (1992) A power primer. Psychol Bull 112:155–159.

216. Wong KWN, Leong JCY, Chan M, et al (2004) The flexion-extension profile of lumbar spine in 100 healthy volunteers. Spine (Phila Pa 1976) 29:1636–41.

217. Stelzeneder D, Kovács BK, Goed S, et al (2012) Effect of short-term unloading on T2 relaxation time in the lumbar intervertebral disc--in vivo magnetic resonance imaging study at 3.0 tesla. Spine J 12:257–64.

218. Welsch GH, Trattnig S, Paternostro-Sluga T, et al (2011) Parametric T2 and T2* mapping techniques to visualize intervertebral disc degeneration in patients with low back pain: initial results on the clinical use of 3.0 Tesla MRI. Skeletal Radiol 40:543–551.

219. Dimitriadis A, Smith F, Mavrogenis AF, et al (2012) Effect of two sitting postures on lumbar sagittal alignment and intervertebral discs in runners. Radiol Med 117:654–68.

220. Apfel CC, Cakmakkaya OS, Martin W, et al (2010) Restoration of disk height through non-surgical spinal decompression is associated with decreased discogenic low back pain: a retrospective cohort study. BMC Musculoskelet Disord 11:155.

221. Nowicki BH, Yu S, Reinartz J, et al (1990) Effect of axial loading on neural foramina and nerve roots in the lumbar spine. Radiology 176:433–437.

222. Abdollah, Vahid; Parent, EC; Adria, S; Battie M The Immediate Effects of Loading on Lumbar Imaging Findings: A Systematic Review. Unpubl. Work

223. Lagarias JC, Reeds JA, Wright MH, Wright PE (1998) Convergence Properties of the Nelder--Mead Simplex Method in Low Dimensions. SIAM J Optim 9:112–147.

224. Abdollah, Vahid; Parent, EC; Battie M Is the Location of the Signal Intensity Weighted Centroid a Reliable Measurement of Fluid Displacement within the Disc. J. Biomed. Eng.

225. Harrell FE (2001) Regression modeling strategies: with applications to linear models, logistic regression, and survival analysis. Springer, New York

226. Jones RW, Witte RJ (2000) Signal Intensity Artifacts in Clinical MR Imaging. RadioGraphics 20:893–901.

227. Ramos G, Martin W (1994) Effects of vertebral axial decompression on intradiscal pressure. J Neurosurg 81:350–3.

228. Magnusson ML, Aleksiev AR, Spratt KF, et al (1996) Hyperextension and spine height changes. Spine (Phila Pa 1976) 21:2670–2675.

229. Gerke DA, Brismee JM, Sizer PS, et al (2011) Change in spine height measurements following sustained mid-range and end-range flexion of the lumbar spine. Appl Ergon 42:331–336.

230. Owens SC, Brismee JM, Pennell PN, et al (2009) Changes in spinal height following sustained lumbar flexion and extension postures: a clinical measure of intervertebral disc hydration using stadiometry. J Manipulative Physiol Ther 32:358–363.

231. Chung TS, Yang HE, Ahn SJ, Park JH (2015) Herniated Lumbar Disks: Real-time MR Imaging Evaluation during Continuous Traction. Radiology 275:755–762.

232. Heuer F, Schmidt H, Wilke H-J (2008) Stepwise reduction of functional spinal structures increase disc bulge and surface strains. J Biomech 41:1953–1960.

233. Laible JP, Pflaster DS, Krag MH, et al (1993) A poroelastic-swelling finite element model with application to the intervertebral disc. Spine (Phila Pa 1976) 18:659–70.

234. Natarajan RN, Williams JR, Andersson GBJ (2004) Recent advances in analytical modeling of lumbar disc degeneration. Spine (Phila Pa 1976) 29:2733–41.

235. O'Connell GD, Johannessen W, Vresilovic EJ, Elliott DM (2007) Human Internal Disc Strains in Axial Compression Measured Noninvasively Using Magnetic Resonance Imaging. Spine (Phila Pa 1976) 32:2860–2868.

236. Johannessen W, Cloyd JM, O'Connell GD, et al (2006) Trans-Endplate Nucleotomy Increases Deformation and Creep Response in Axial Loading. Ann Biomed Eng 34:687–696.

237. Nachemson AL (1981) Disc pressure measurements. Spine (Phila Pa 1976) 6:93–7.

238. Wong AYL, Parent EC, Dhillon SS, et al (2015) Do Participants With Low Back Pain Who Respond to Spinal Manipulative Therapy Differ Biomechanically From Nonresponders, Untreated Controls or Asymptomatic Controls? Spine (Phila Pa 1976) 40:1329–1337.

239. Videman T, Battié MC, Gill K, et al (1995) Magnetic resonance imaging findings and their relationships in the thoracic and lumbar spine. Insights into the etiopathogenesis of spinal degeneration. Spine (Phila Pa 1976) 20:928–35.

240. Nachemson A (1960) Lumbar intradiscal pressure. Experimental studies on post-mortem material. Acta Orthop Scand 43:1–104.

241. Abdollah, Vahid; Chaggar, G; Parent E (2014) Validity and Reliability of Lumbar Disc Height Measurement: A Systematic Literature Review. Edmonton

242. Danielson BI, Willén J, Gaulitz A, et al (1998) Axial loading of the spine during CT and MR in patients with suspected lumbar spinal stenosis. Acta Radiol 39:604–11.

243. Willén J, Danielson B, Gaulitz A, et al (1997) Dynamic effects on the lumbar spinal canal: axially loaded CT-myelography and MRI in patients with sciatica and/or neurogenic claudication. Spine (Phila Pa 1976) 22:2968–76.

244. Willén J, Danielson B (2001) The diagnostic effect from axial loading of the lumbar spine during computed tomography and magnetic resonance imaging in patients with degenerative disorders. Spine (Phila Pa 1976) 26:2607–14.

245. Neubert A, Fripp J, Engstrom C, et al (2012) Automated detection, 3D segmentation and analysis of high resolution spine MR images using statistical shape models. Phys Med Biol 57:8357–8376.

246. Michopoulou SK, Costaridou L, Panagiotopoulos E, et al (2009) Atlas-based segmentation of degenerated lumbar intervertebral discs from MR images of the spine. IEEE Trans Biomed Eng 56:2225–2231.

247. Gervais J, Perie D, Aubin CE (2014) Sensitivity of MRI signal distribution within the intervertebral disc to image segmentation and data normalisation. Comput Methods Biomech Biomed Engin 17:1383–1390.

248. Claudia C, Farida C, Guy G, et al (2012) Quantitative evaluation of an automatic segmentation method for 3D reconstruction of intervertebral scoliotic disks from MR images. BMC Med Imaging 12:26.

249. Weiss KL, Storrs JM, Banto RB (2006) Automated spine survey iterative scan technique. Radiology 239:255–262.

250. Roberts N, Gratin C, Whitehouse GH (1997) MRI analysis of lumbar intervertebral disc height in young and older populations. J Magn Reson Imaging 7:880–886.

251. Adams MA (2013) The biomechanics of back pain, 3rd ed. Churchill Livingstone/Elsevier, Edinburgh ; New York

252. Gidwani R, Sinnott P, Avoundjian T, et al (2016) Inappropriate ordering of lumbar spine magnetic resonance imaging: are providers Choosing Wisely? Am J Manag Care 22:e68-

76.

253. Beattie PF, Butts R, Donley JW, Liuzzo DM (2014) The within-session change in low back pain intensity following spinal manipulative therapy is related to differences in diffusion of water in the intervertebral discs of the upper lumbar spine and L5-S1. J Orthop Sports Phys Ther 44:19–29.

254. Association AER, Association AP, Education NC on M in, Testing JC on S for E and P (1999) Standards for educational and psychological testing. American Educational Research Association, Washington, DC

255. Parent EC, Videman T, Battié MC (2006) The effect of lumbar flexion and extension on disc contour abnormality measured quantitatively on magnetic resonance imaging. Spine (Phila Pa 1976) 31:2836–2842.

256. Mallett R, Hagen-Zanker J, Slater R, Duvendack M (2012) The benefits and challenges of using systematic reviews in international development research. J Dev Eff 4:445–455.

257. Thompson KJ, Dagher AP, Eckel TS, et al (2009) Modic changes on MR images as studied with provocative diskography: clinical relevance--a retrospective study of 2457 disks. Radiology 250:849–55.

258. Espinoza Orías AA, Mammoser NM, Triano JJ, et al (2016) Effects of Axial Torsion on Disc Height Distribution: An In Vivo Study. J Manipulative Physiol Ther 39:294–303.

259. Kanno H, Ozawa H, Koizumi Y, et al (2016) Increased Facet Fluid Predicts Dynamic Changes in the Dural Sac Size on Axial-Loaded MRI in Patients with Lumbar Spinal Canal Stenosis. AJNR Am J Neuroradiol 37:730–5.

260. Kubosch D, Vicari M, Siller A, et al (2015) The Lumbar Spine as a Dynamic Structure Depicted in Upright MRI. Medicine (Baltimore) 94:e1299.

261. Liu N, Wood KB, Schwab JH, et al (2015) Utility of Flexion-Extension Radiographs in Lumbar Spondylolisthesis. Spine (Phila Pa 1976) 40:E929–E935.

262. Gallagher KM, Sehl M, Callaghan JP (2016) A radiographic assessment of lumbar spine posture in four different upright standing positions. Clin Biomech (Bristol, Avon) 37:131–6.

263. Kanno H, Ozawa H, Koizumi Y, et al (2016) Increased Facet Fluid Predicts Dynamic Changes in the Dural Sac Size on Axial-Loaded MRI in Patients with Lumbar Spinal Canal Stenosis. AJNR Am J Neuroradiol 37:730–5.

264. Takasaki H (2015) Comparable effect of simulated side bending and side gliding positions on the direction and magnitude of lumbar disc hydration shift: in vivo MRI mechanistic study. J Man Manip Ther 23:101–8.

265. Zhong W, Driscoll SJ, Tsai T-Y, et al (2015) In vivo dynamic changes of dimensions in the lumbar intervertebral foramen. Spine J 15:1653–9.

# APPENDIX I. IMAGE SEGMENTATION

The semi-automatic segmentation using the mid-sagittal image begins by clicking on the lumbar vertebrae starting with S1 to label the levels. A line is then drawn tangential to the superior endplate (Figure 17a). The left endpoint of the line is used for initializing the segmentation of the upper boundary of the disc. Twenty pixels below the point is then scanned (Figure 17a). The pixel with maximum signal intensity difference (MSID) is determined and is selected as the initial left endpoint of the disc-vertebra boundary (Figure 17b). Ten pixels above and below the next adjacent pixel in the horizontal direction are then scanned and the pixel with MSID is determined and is selected as the second point (Figure 17b). The same process continues for the rest of the points until the right end of the superior tangent line is reached (Figure 17b). For segmentation of the inferior boundary, similar to the superior border, a line is drawn tangential to the inferior endplate and the same processes are used for determining the disc vertebra boundary (Figure 17c).

To locate the anterior and posterior corners of the vertebrae, after the algorithm completes finding the points on the center portion of the disc-vertebra boundary, 20 pixels adjacent to the initial endpoints are scanned further anteriorly or posteriorly. The pixel with MSID in the superior-inferior direction is then determined and is selected as the vertebra endpoint on each side. A Gaussian operator using a 3×3 kernel for six pixels above and below the pixel with MSID is employed for spatial smoothing of the borders and to eliminate outlier points, which could be due to noise. Users are also able to modify the location of individual disc boundary pixels in the case of outlier errors (Figure 17).

After determining the slope of the best-fit lines passing through the points of the upper and lower boundaries, a line with the slope corresponding to the average of these two lines is then automatically generated in the middle, dividing the disc into superior and inferior halves (Figure 17c). Each pixel of the superior line is automatically connected to a corresponding pixel on the inferior line by following a line perpendicular to the third line bisecting the disc.

The midpoints of the line connecting the anterior endpoints of the disc vertebrae boundaries and of the corresponding line dorsally are connected using a disc bisector line (DBL). The bisector line is then extended 20 pixels at each end (Figure 17e). Multiple lines parallel to the extended line, with the same length, are then generated superiorly and inferiorly at one pixel interval, until the lines fully lie in the vertebral bones (Figure 17f). Adjacent pixels with MSID along those lines

parallel to the line bisecting the disc are selected as the ventral and dorsal boundaries of the disc (Figure 17g). A Gaussian operator is used in a manner similar to the smoothing strategy used for the disc-vertebra boundary. Users are also able to modify the location of individual boundary pixels in case there are outliers.

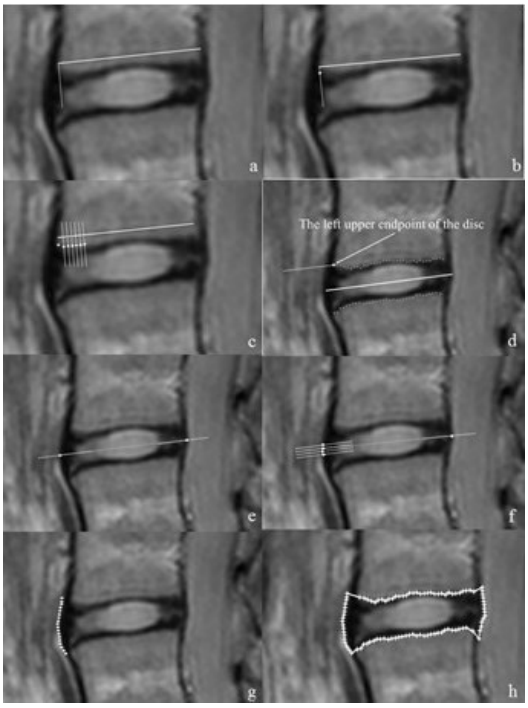


Figure 17 Disc segmentation process a:) drawing the first tangential line to the upper disc-vertebra boundary of the disc of interest and scanning 20 pixels below the left endpoint of the line, b:) determining the first point of the disc-bone boundary, c:) determining the upper boundary of disc pixel by pixel along the upper disc-vertebra boundary, d:) generating the disc bisector line (DBL) between the upper and lower boundary of the disc and determining the real vertebra endpoint, e:) determining anterior and posterior endpoint of disc along the DBL, f: determining the anterior boundary of disc pixel by pixel along the anterior contour, g:) anterior boundary of disc, h:) a fully segmented disc

**Intradiscal segmentation**

A similar algorithm identifying adjacent pixels with the greatest difference in signal intensity along lines parallel to the DBL is used to automatically segment the higher intensity nucleus region from the anterior and posterior annulus regions (Figure 6a&b). To segment the nucleus from adjacent structures above and below it, each pixel of the superior border of the disc

is connected to its equivalent pixel on the inferior border of the disc using a line perpendicular to the DBL, and pixels with MSID along those vertical lines are selected as the superior and inferior border of the nucleus (Figure 6c&d).

For segmentation of the anterior annulus, the initial first seven points of the superior and inferior boundary of the nucleus are scanned and the most anterior points to the straight regression lines passing through the superior and inferior boundary of the nucleus are determined. These points are considered the endpoints of the anterior annulus superiorly and inferiorly. The points are then connected to the corresponding points on the upper and lower boundary of the disc using lines perpendicular to the DBL. These two lines as well as points lying before endpoints are used for defining the posterior boundary of the anterior annulus (Figure 6e). The same strategies are used for determining the anterior boundary of the posterior annulus. Each disc is thus segmented into four ROIs: nucleus region, anterior and posterior annulus regions, and whole disc (Figure 6f).

**Disc height measurements**

The distance between the anterior-inferior corner of the vertebrae above the disc and the anterior-superior corner of the vertebra below, identified by the segmentation approach, represented the anterior disc height $(Ht_{Ant})$ (Figure 19a). Likewise, the distance between posterior inferior corner of the vertebra above and the posterior-superior corner of the vertebra below the disc represented the posterior disc height $(Ht_{post})$ (Figure 19a). The distance between the mid-point of the disc-vertebra boundaries above and below the disc (endplate bisectris) was quantified as the middle disc height $(Ht_{mid})$ (Figure 19a).

Disc height measurements were obtained by first dividing the disc area by the length of the DBL (anterior-posterior disc dimension) located between the anterior and posterior disc boundaries identified by the segmentation process (Figure 19b) [255]. The disc area was then computed using 80% (figure 3c) and 60% (Figure 19d) of the length of the DBL centered between the disc boundaries, and the disc heights were computed as the ratio of the area and A-P dimension of the disc corresponding to this proportion (Figure 19).

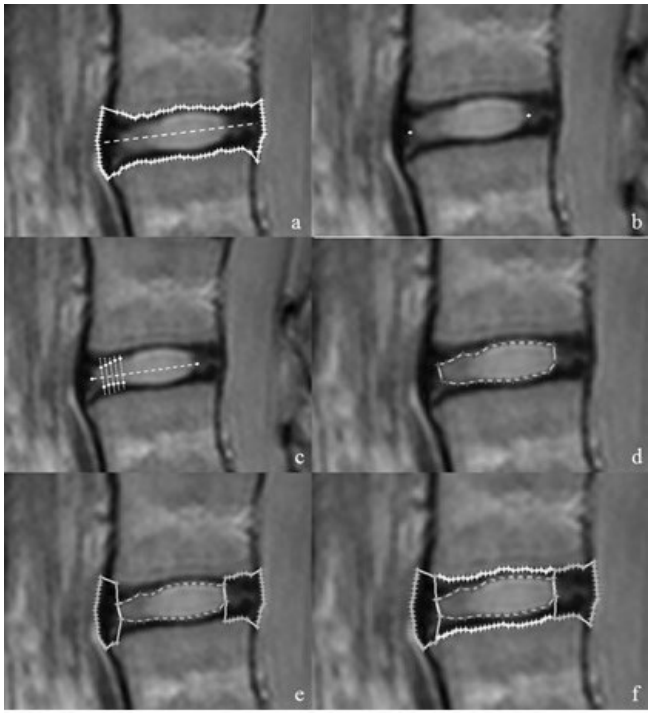


Figure 18 Segmentation of the regions of interest within the disc a:) generating disc bisector line, b:) determining the anterior and posterior endpoint of the nucleus, c:) determining the upper and lower boundary of the nucleus, d:) segmented nucleus, e:) segmenting anterior and posterior annulus f: a fully segmented disc

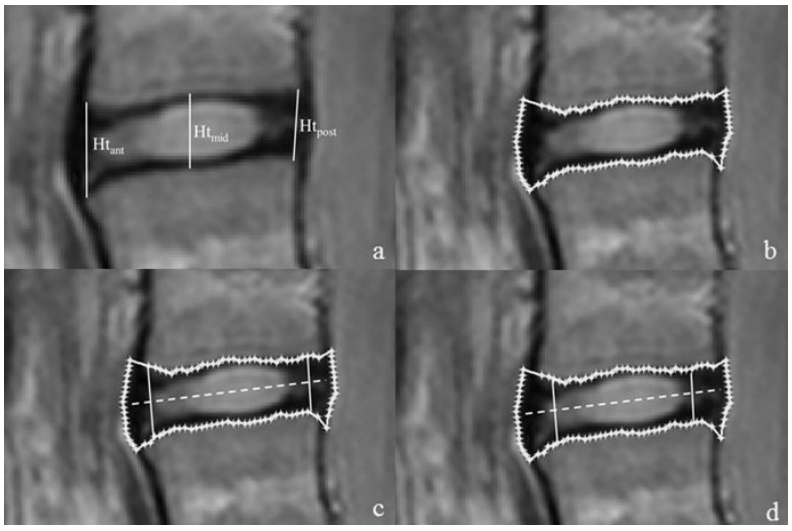


Figure 19 a:) Measuring disc height using Dabbs' (average of anterior ($Ht_{ant}$) and posterior ($Ht_{post}$) disc heights, Hurxtal's ($Ht_{mid}$) and combined method (average of all three), 3b:) area-based method using 100% disc width, 3c:) 80% disc width used, 3d:) 60% disc width used

The disc height is then quantified based on Dabbs' (mean of the anterior and posterior height of the disc) [191], Hurxthal's 2nd, height measured at the midpoint of the disc ($Ht_{mid}$) [195], and a combination of both methods (the mean of anterior, posterior and height measured at the midpoint of the disc) [196].

## APPENDIX II. SEARCH STRATEGIES

**MEDLINE**

1. Intervertebral Disc/me [Metabolism]

2. exp Intervertebral Disc Degeneration/me [Metabolism]

3. exp Intervertebral Disc Displacement/me [Metabolism]

4. 1 or 2 or 3

5. exp Intervertebral Disc/

6. ((disc or discs or disk or disks) adj2 (intervertebral or lumbar or low back)).ti,ab.

7. lumbar vertebrae/

8. ((lumbar or low back) and (vertebra or vertebrae or vertebral or spine or spinal)).mp. [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier]

9. 5 or 6 or 7 or 8

10. metabolism/ or absorption/ or biological transport/ or biological transport, active/ or facilitated diffusion/ or energy metabolism/ or glycolysis/ or osmoregulation/ or water-electrolyte balance/ or water loss, insensible/ or oxygen consumption/

11. (metaboli* or absorption or transport or diffusion or glycolysis or solute* or osmoregulation or water or oxygen or nutrient* or nutrition* or glucose).ti,ab.

12. magnetic resonance imaging/ or exp diffusion magnetic resonance imaging/

13. exp Magnetic Resonance Spectroscopy/

14. (signal intensity or mri or magnetic resonance or (fluid* adj3 flow*)).ti,ab.

15. finite element analysis/ or (strain or stress or finite element).mp.

16. 10 or 11 or 12 or 13 or 14 or 15

17. 9 and 16

18. 4 or 17

19. Traction/

20. (traction or compression or decompression or load* or unload* or weight bearing or bending or flexion or extension or rotation).ti,ab.

21. exp Weight-Bearing/

22. 19 or 20 or 21

23. 18 and 22

24. limit 23 to (editorial or interview or letter or news)

25. 23 not 24

26. (cervical not (lumbar or thoracic or thorax or low back)).mp.

27. fracture*.mp.

28. tuberculosis.mp.

29. 26 or 27 or 28

30. 25 not 29

31. limit 30 to english

32. limit 31 to humans

33. limit 31 to animals

34. 32 or 33

35. 31 not 34

36. limit 35 to ("in data review" or in process or "pubmed not medline")

37. 34 or 36

38. limit 37 to case reports

39. case series.mp.

40. 38 and 39

41. 37 not 38

42. 40 or 41

43. exp Materials Testing/

44. (hematoma or angiography or viscoelastic or surg* or implant* or leg*1 or lower extremit* or osteoporosis or cadaver* or scoliosis or kyphosis or postoperative or post surg* or postsurg* or failure test* or decompress* or spinal cord or trauma*).ti.

45. exp Rodentia/ or rat.mp. or rats.mp. or sheep.mp. or ovine.mp. or mouse.mp. or mice.mp. or gerbil*.mp. or guinea pig*.mp. or pig*1.mp. or swine.mp. or porcine.mp. or bovine*.mp. or cow*1.mp. or cattle.mp. or dog*1.mp. or canine*.mp. [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier]

46. 43 or 44 or 45

47. 42 not 46

48. (proceeding* or Symposi*).m_titl.

49. 47 not 48

50. Cadaver*.mp. or exp Cadaver/

51. In vivo.mp.

52. 50 and 51

53. 50 not 52

54. 49 not 53

55. 54 not Proceeding.mp. [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier]

56. 55 and Titanium.mp. [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier]

57. 55 not 56

58. 57 and cage.mp. [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier]

59. 57 not 58

60. 59 and screw.mp. [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier]

61. 59 not 60

62. Microsurgery.mp. and 61 [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier]

63. 61 not 62

64. 63 and Laminectomy.mp. [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier]

65. 63 not 64

66. 65 and cat.mp. [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier]

67. 65 not 66

68. 67 and cancer.mp. [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier]

69. 67 not 68

70. 69 and Malignant.mp. [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier]

71. 69 not 70

**EMBASE**

1. exp *Intervertebral Disk/

2. exp *intervertebral disk hernia/

3. exp *intervertebral disk degeneration/

4. ((disc or discs or disk or disks) adj2 (intervertebral or lumbar or low back)).ti,ab.

5. lumbar vertebrae/

6. ((lumbar or low back) and (vertebra or vertebrae or vertebral or spine or spinal)).mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword]

7. or/1-6

8. *disk diffusion/ or *diffusion coefficient/ or *gas diffusion/ or *gas exchange/ or *oxygen diffusion/ or exp *metabolism/ or *absorption/ or *water absorption/ or *energy absorption/ or exp *transport at the cellular level/ or *energy metabolism/ or *glycolysis/ or *osmoregulation/ or exp *electrolyte balance/ or *oxygen consumption/

9. (metaboli* or absorption or transport or diffusion or glycolysis or solute* or osmoregulation or water or oxygen or nutrient* or nutrition* or glucose).ti,ab.

10. *nuclear magnetic resonance imaging/ or *diffusion tensor imaging/ or *diffusion weighted imaging/ or *echo planar imaging/ or *functional magnetic resonance imaging/ or *susceptibility weighted imaging/

11. *nuclear magnetic resonance spectroscopy/

12. (signal intensity or mri or magnetic resonance or (fluid* adj3 flow*)).ti,ab.

13. *finite element analysis/ or (strain or stress or finite element).ti,ab.

14. 8 or 9 or 10 or 11 or 12 or 13

15. 7 and 14

16. *Traction therapy/

17. (traction or compression or decompression or load* or unload* or weight bearing or bending or flexion or extension or rotation).ti,ab.

18. exp *Weight-Bearing/

19. 16 or 17 or 18

20. 15 and 19

21. (cervical not (lumbar or thoracic or thorax or low back)).mp.

22. fracture*.mp.

23. tuberculosis.mp.

24. case series.mp.

25. exp Materials Testing/

26. (hematoma or angiography or viscoelastic or surg* or implant* or leg*1 or lower extremit* or osteoporosis or cadaver* or scoliosis or kyphosis or postoperative or post surg* or postsurg* or failure test* or decompress* or spinal cord or trauma*).ti.

27. or/21-27

28. limit 29 to english

29. exp *Intervertebral Disk/

30. exp *intervertebral disk hernia/

31. exp *intervertebral disk degeneration/

32. ((disc or discs or disk or disks) adj2 (intervertebral or lumbar or low back)).ti,ab.

33. lumbar vertebrae/

34. ((lumbar or low back) and (vertebra or vertebrae or vertebral or spine or spinal)).mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword]

35. or/29-34

36. *disk diffusion/ or *diffusion coefficient/ or *gas diffusion/ or *gas exchange/ or *oxygen diffusion/ or exp *metabolism/ or *absorption/ or *water absorption/ or *energy absorption/ or exp *transport at the cellular level/ or *energy metabolism/ or *glycolysis/ or *osmoregulation/ or exp *electrolyte balance/ or *oxygen consumption/

37. (metaboli* or absorption or transport or diffusion or glycolysis or solute* or osmoregulation or water or oxygen or nutrient* or nutrition* or glucose).ti,ab.

38. *nuclear magnetic resonance imaging/ or *diffusion tensor imaging/ or *diffusion weighted imaging/ or *echo planar imaging/ or *functional magnetic resonance imaging/ or *susceptibility weighted imaging/

39. *nuclear magnetic resonance spectroscopy/

40. (signal intensity or mri or magnetic resonance or (fluid* adj3 flow*)).ti,ab.

41. *finite element analysis/ or (strain or stress or finite element).ti,ab.

42. 36 or 37 or 38 or 39 or 40 or 41

43. 35 and 42

44. *Traction therapy/

45. (traction or compression or decompression or load* or unload* or weight bearing or bending or flexion or extension or rotation).ti,ab.

46. exp *Weight-Bearing/

47. 44 or 45 or 46

48. 43 and 47

49. (cervical not (lumbar or thoracic or thorax or low back)).mp.

50. fracture*.mp.

51. tuberculosis.mp.

52. case series.mp.

53. exp Materials Testing/

54. (hematoma or angiography or viscoelastic or surg* or implant* or leg*1 or lower extremit* or osteoporosis or cadaver* or scoliosis or kyphosis or postoperative or post surg* or postsurg* or failure test* or decompress* or spinal cord or trauma*).ti.

55. exp Rodentia/ or rat.mp. or rats.mp. or sheep.mp. or ovine.mp. or mouse.mp. or mice.mp. or gerbil*.mp. or guinea pig*.mp. or pig*1.mp. or swine.mp. or porcine.mp. or bovine*.mp. or cow*1.mp. or cattle.mp. or dog*1.mp. or canine*.mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword]

56. or/49-55

57. 48 not 56

58. limit 57 to english

59. 58 not cat.mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword]

60. 59 not Proceeding.mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword]

61. 60 not Symposium.mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword]

62. 61 and titanium.mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword]

63. 61 not 62

64. 63 and stent.mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword]

65. 63 not 64

66. 65 and screw.mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword]

67. 65 not 66

68. 67 and cage.mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword]

69. 67 not 68

70. 69 not Microsurgery.mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword]

71. 70 not Laminectomy.mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword]

72. 71 and Cancer.mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword]

73. 71 not 72

74. 73 and Malignant.mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword]

75. 73 not 74

76. cadaver*.mp. or exp cadaver/

77. 76 and 75

78. 77 not invivo.mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword]

79. 77 not in vivo.mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword]

80. 75 not 79

**Scopus**

TiTLE-ABS-KEY("intervertebral disc*" or "intervertebral disk*"or "lumbar disc*" or "lumbar disk*" or "lumbar vertebra*" or "lumbar spine" or ("low back" w/2 disc*) or ("low back*" w/2 disk*) or ("low back" w/2 vertebra*))

AND TITLE-ABS-KEY(metaboli* or absorption or transport or diffusion or glycolysis or solute* or osmoregulation or water or oxygen or nutrient* or nutrition* or glucose or "signal intensity" or mri or "magnetic resonance" or (fluid* pre/3 flow*) or strain or stress or "finite element")

AND TITLE-ABS-KEY(traction or compression or decompression or load* or unload* or "weight bearing" or bending or flexion or extension or rotation)

AND NOT TITLE-ABS-KEY(cervical or fracture* or tuberculosis or rodent* or rat or rats or mouse or mice or gerbil* or "guinea pig*" or bovine or pig or pigs or porcine or cattle or cow or cows or hematoma or angiography or viscoelastic or surg* or implant* or leg or legs or "lower extremit*" or osteoporosis or cadaver* or scoliosis or kyphosis)

**CINAHL**

( (MH "Intervertebral Disk/ME" or MH "Intervertebral Disk Displacement/ME" or MH "Lumbar Vertebrae/ME") OR (( MH "Intervertrebral Disk" or MH "Lumbar Vertebrae") OR (disc or discs or disk or disks) w2 (intervertebral or lumbar or low back) OR (lumbar or low back) w2 (vertebra or vertebrae or vertebral or spine or spinal)

AND ( MH "Metabolism+" OR MH "Absorption" OR MH "Biological Transport+" OR (metaboli* or absorption or transport or diffusion or glycolysis or solute* or osmoregulation or water or oxygen or nutrient* or nutrition* or glucose OR MH "Magnetic Resonance Imaging+" OR MH "Magnetic Resonance Spectroscopy" OR MH "Finite Element Analysis" OR "signal intensity" or mri or "magnetic resonance" or fluid* w3 flow* OR strain or stress or "finite element" )))

AND ( MH "Traction" OR MH "Weight-Bearing" OR traction or compression or decompression or load* or unload* or "weight bearing" or bending or flexion or extension or rotation ) )

NOT ( fracture* or tuberculosis OR rodent* orodent* or rat or rats or mouse or mice or gerbil* or "guinea pig*" or bovine or pig or pigs or porcine or cattle or cow or cows OR hematoma or angiography or viscoelastic or surg* or implant* or leg or legs or "lower extremit*" or osteoporosis or cadaver* or scoliosis or kyphosis )

**Web of Science Core Collection**

TS=("intervertebral disc*" or "intervertebral disk*"or "lumbar disc*" or "lumbar disk*" or "lumbar vertebra*" or "lumbar spine" or ("low back" near2 disc*) or ("low back*" near2 disk*) or ("low back" near2 vertebra*))

AND TS=(metaboli* or absorption or transport or diffusion or glycolysis or solute* or osmoregulation or water or oxygen or nutrient* or nutrition* or glucose or "signal intensity" or mri or "magnetic resonance" or (fluid* near/3 flow*) or strain or stress or "finite element")

AND TS=(raction or compression or decompression or load* or unload* or "weight bearing" or bending or flexion or extension or rotation)

NOT TS=(cervical or fracture* or tuberculosis or rodent* or rat or rats or mouse or mice or gerbil* or "guinea pig*" or bovine or pig or pigs or porcine or cattle or cow or cows or hematoma or angiography or viscoelastic or surg* or implant* or leg or legs or "lower extremit*" or osteoporosis or cadaver* or scoliosis or kyphosis)

**Biosis Previews**

TS=("intervertebral disc*" or "intervertebral disk*"or "lumbar disc*" or "lumbar disk*" or "lumbar vertebra*" or "lumbar spine" or ("low back" near2 disc*) or ("low back*" near2 disk*) or ("low back" near2 vertebra*))

AND TS=(metaboli* or absorption or transport or diffusion or glycolysis or solute* or osmoregulation or water or oxygen or nutrient* or nutrition* or glucose or "signal intensity" or mri or "magnetic resonance" or (fluid* near/3 flow*) or strain or stress or "finite element")

AND TS=(traction or compression or decompression or load* or unload* or "weight bearing" or bending or flexion or extension or rotation)

AND TA=(Hominidae)

NOT TS=(cervical or fracture* or tuberculosis or hematoma or angiography or viscoelastic or surg* or implant* or leg or legs or "lower extremit*" or osteoporosis or cadaver* or scoliosis or kyphosis)

## APPENDIX III: SYSTEMATIC LITERATURE REVIEW DATA EXTRACTION FORM

**Click here to choose a reviewer**

**Click here to enter a date.**

**Study description**

| | |
|---|---|
| ID No. | |

**STUDY SELECTION CRITERIA**

| **Inclusion** | **Exclusion** |
|---|---|
| ☐ English | ☐ Other languages |
| ☐ Focused on effects of loading | ☐ NOT focused on effects of loading |
| ☐ Original study | ☐ Narrative review |
| ☐ Disc or Vertebra outcomes measured for specific segment | ☐ Animal Studies |
| ☐ Imaging, mechanical loading or EMG\ | ☐ Cadaveric Studies |
| | ☐ Un-validated FEM model |
| | ☐ Surgical technique |
| | ☐ Whole spine |
| | ☐ Occupational loading |

**DECISION:** ☐ Included ☐ Excluded ☐ Unclear

**Study type information extraction**

| | |
|---|---|
| Goal of the study | |
| Type of the study | Click here to choose the study type |
| Timeline of study | Click to select the timeline of the study |

**Subject selection criteria.**

| | |
|---|---|
| Inclusion Criteria? | ☐ Not reported ☐ Unclear ☐ Yes, Specify: |
| Exclusion Criteria? | ☐ Not reported ☐ Unclear ☐ Yes, Specify: |
| Subject recruitment | ☐ Random ☐ Consecutive ☐ Volunteers ☐ Purposeful. |

**Groups (definition) or define reported group and subgroups**

| Group name: | | | |
|---|---|---|---|
| No of participants | | | |
| Age | | | |
| Diagnosis | | | |
| Pain duration | | | |
| Pain intensity | | | |
| Pain location | | | |
| Height | | | |
| Weight | | | |
| Gender | # of male:<br># of female: | # of male:<br># of female: | # of male:<br># of female: |

**Loading information (copy as needed, if multiple methods were employed)**

| | |
|---|---|
| Type of loading | Click here to choose type of loading |
| Describe mode of application | |
| Duration of loading | Click here to choose duration of loading |
| Magnitude of loading (if applicable ) | Click here to choose the magnitude of loading |
| Time of loading | Click here to choose time of loading |
| Precondition of loading (type, duration and magnitude of preconditioning) | |

**Method**

| | |
|---|---|
| Which method was employed? | Click here to choose which method was used |
| Which level was measured (specify)? | ☐ Disc ☐ Vertebra ☐ Paraspinal muscles |
| When the measurements were acquired relative to loading, specify | |
| Measurement outcomes | |

**List of measurements outcomes (for every subgroup/time-point/measure)**

Mean:

SD:

SE:

Range:

Min:

Max:

Other, specify:

**Results of statistical test of differences within (effect of loading) and between groups (effect of loading):**

Name of the test used:

Variables compared:

Groups or times compared:

P-value (for each comparison of interest):

Report of the mean difference and variability for each pairwise comparison:

**Results of statistical test of correlations between effects of loading and other variables of interest:**

Name of the association test used:

Variables tested:

Group tested:

Association estimate (correlation coefficient, regression equations)

P-values:

**Results of statistical test for diagnostic accuracy**

Name of the diagnostic statistical test used:

Name of the gold standard test:

Name of the loading variable of interest:

Diagnostic test estimate:

P-value or confidence interval:

| Criterion | Yes | No | NA | Comments |
|---|---|---|---|---|
| **Subjects recruitment** | | | | |
| Are the characteristics of the participants included in the study clearly described?<br>☛ *In cohort studies and trials, inclusion and/or exclusion criteria should be given. In case-control studies, a case-definition and the source for controls should be given. Single case studies must state source of patient* | ☐ | ☐ | ☐ | |
| Were the demographic characteristics of the sample reported for each group analyzed? | ☐ | ☐ | ☐ | |
| Were the subjects asked to participate in the study representative of the entire population from which they were recruited?<br>☛ *The study must identify the source population for participants and describe how the participants were selected.* | ☐ | ☐ | ☐ | |
| Were those subjects who were prepared to participate representative of the entire population from which they were recruited?<br>☛ *The proportion of those asked who agreed should be stated.* | ☐ | ☐ | ☐ | |
| Was an attempt made to blind study subjects to the intervention (loading condition) they received?<br>☛ *For studies where the participants would have no way of knowing which intervention they received, this should be answered yes. Retrospective, single group = NO; not described (UTD) if > 1 group and blinding not explicitly stated.* | ☐ | ☐ | ☐ | |
| Were study subjects randomized to intervention groups?<br>☛ *Studies which state that subjects were randomized should be answered yes except where method of randomization would not ensure random allocation.* | ☐ | ☐ | ☐ | |
| Were participants' characteristics stable during research (on the effect of loading)? | ☐ | ☐ | ☐ | |
| **Examiners** | | | | |
| Were the training and qualifications of the rater(s) reported? | ☐ | ☐ | ☐ | |
| Was/were the rater(s) blinded to the results of the comparator test when comparing different test measurements? | ☐ | ☐ | ☐ | |
| Was/were the rater(s) blinded to the results of previous measurements performed by the same or different rater(s) (e.g. blinded to pre-loading condition)? | ☐ | ☐ | ☐ | |

| Criterion | Yes | No | NA | Comments |
|---|---|---|---|---|
| Was the randomized intervention assignment concealed from health care staff until recruitment was complete and irrevocable?<br>☛ *All non-randomized studies should be answered no. If assignment was concealed from participants but not from staff, it should be answered no.* | ☐ | ☐ | ☐ | |
| **Methodology** | | | | |
| Are the exposures/interventions of interest clearly described?<br>☛ *Treatments (Loading conditions) and placebo (where relevant) that are to be compared should be clearly described.* | ☐ | ☐ | ☐ | |
| Was the sample size included in the analysis adequate? | ☐ | ☐ | ☐ | |
| Is replication of the assessment procedure possible? (description sufficiently detailed) | ☐ | ☐ | ☐ | |
| Are the distributions of principal confounders in each group of subjects to be compared clearly described?<br>☛ *A list of principal confounders is provided. YES = age, severity* | ☐ | ☐ | ☐ | |
| Was there adequate adjustment for confounding in the analyses from which the main findings were drawn?<br>☛ *In nonrandomized studies if the effect of the main confounders was not investigated or no adjustment was made in the final analyses the question should be answered as no. If no significant difference between groups shown then YES* | ☐ | ☐ | ☐ | |
| **Outcomes** | | | | |
| Is the hypothesis/aim/objective of the study clearly described? Must be explicit (Only focus on objective related to the study of the effect of loading). | ☐ | ☐ | ☐ | |
| Validity reported for main outcome measure | ☐ | ☐ | ☐ | |
| **Handling Missing Data (Concurrent and Criterion Validity)** | | | | |
| Compliance acceptable in all groups (80% acceptable) | ☐ | ☐ | ☐ | |
| Was the percentage of missing items given (Only for the analysis of the effect of loading)? | ☐ | ☐ | ☐ | |
| Withdrawal/dropouts rate described and acceptable<br>☛ *Maximum 20% drop out rate (Only for the analysis of the effect of loading)* | ☐ | ☐ | ☐ | |

| Criterion | Yes | No | NA | Comments |
|---|---|---|---|---|
| Have the characteristics of participants lost to follow-up been described?<br>• *If not explicit = NO. RETROSPECTIVE –if not described = UTD; if not explicit re: numbers agreeing to participate = NO. Needs to be >85% (only for analysis of the effect of loading).* | ☐ | ☐ | ☐ | |
| Was compliance with the intervention/s reliable?<br>• *Where there was non-compliance with the allocated treatment or where there was contamination of one group, the question should be answered no.* | ☐ | ☐ | ☐ | |
| Was there a description of how missing items were handled? | ☐ | ☐ | ☐ | |
| Were hypotheses regarding correlations or mean differences formulated a priori (i.e. before data collection)? | ☐ | ☐ | ☐ | |
| Was the expected direction of correlations or mean differences included in the hypotheses? | ☐ | ☐ | ☐ | |
| **Statistical Analysis** | | | | |
| Have actual probability values been reported (e.g. 0.035 rather than <0.05) for the main outcomes except where the probability value is less than 0.001? (On effect of loading. | ☐ | ☐ | ☐ | |
| Sample size described for each group | ☐ | ☐ | ☐ | |
| Were design and statistical methods adequate for the hypotheses to be tested? | ☐ | ☐ | ☐ | |
| Has confidence interval for PRE_ Post loading or change in outcomes from before to after loading been reported? | ☐ | ☐ | ☐ | |
| Have effect sizes for outcomes been reported or can be computed by reviewer? | ☐ | ☐ | ☐ | |
| **Results** | | | | |
| Are the main findings of the study clearly described?<br>• *Simple outcome data (including denominators and numerators) should be reported for all major findings so that the reader can check the major analyses and conclusions. (Only related to the effect of loading.)* | ☐ | ☐ | ☐ | |
| Have all important adverse events that may be a consequence of the intervention been reported? | ☐ | ☐ | ☐ | |

| Criterion | Yes | No | NA | Comments |
|---|---|---|---|---|
| ☛ *This should be answered yes if the study demonstrates that there was a comprehensive attempt to measure adverse events (complications but not an increase in pain). (Only related to the effect of loading).* | | | | |

## APPENDIX IV: LEVELS OF EVIDENCE SUMMARY STATEMENTS GROUPED BY QUANTITATIVE MEASUREMENTS OF THE EFFECT OF LOADING

| Level of evidence | N (Healthy; LBP) | Changes | Loading conditions |
|---|---|---|---|
| **Anterior disc height** | | | |
| Conflicting | 3 (2;1) [90, 156, 167] | Any change | Unloaded vs. loading |
| No evidence | | Any change | Unloaded vs postural |
| No evidence | | Any change | Unloaded vs compression device |
| No evidence | | Any change | Unloaded vs end range movements |
| Limited | 1 (1;0)[162] | ↓ | Among end range movements |
| Limited | 2 (2;0)[90, 167] | No difference | Among postural or device loading |
| **Cumulative disc height** | | | |
| Limited | 1 (0;1)[83] | ↓ | Unloaded vs. loading |
| No evidence | | Any change | Unloaded vs postural |
| Conflicting | 1 (0;1) [83] | Any change | Unloaded vs compression device |
| No evidence | | Any change | Unloaded vs end range movements |
| Limited | 1 (0;1)[83] | No difference | Among end range movements |
| No evidence | | Any change | Among postural or device loading |
| **Posterior disc height** | | | |
| Conflicting | 3 (2,1) | Any change | Unloaded vs loading |
| Conflicting | 3 (2,1) | Any change | Unloaded vs postural |
| Limited | 1 (1,0) | ↓ | Unloaded vs compression device |
| No evidence | | Any change | Unloaded vs end range movements |
| Limited | 1(1,0)[162] | No difference | Among end range movements |

| | | | |
|---|---|---|---|
| Limited | 2(2,0)[90, 167] | No difference | Among postural or device loading |
| **Anterior distance of the nucleus position** | | | |
| Limited | 1(1,0)[167] | ↑ | Unloaded vs loading (sitting or standing) |
| Limited | 1(1,0)[167] | ↑ | Unloaded vs postural (Sitting or standing) |
| No evidence | | Any change | Unloaded vs compression device |
| No evidence | | Any change | Unloaded vs end range movements |
| Limited | 1(1,0)[162] | ↓ | Among end range movements (extension vs flexion) |
| No evidence | | Any change | Among postural or device loading |
| **Anterior-posterior diameter of the nucleus** | | | |
| Conflicting | 1(1,0)[167] | Any change | Unloaded vs loading |
| Conflicting | 1(1,0) [167] | Any change | Unloaded vs postural |
| No evidence | | Any change | Unloaded vs compression device |
| No evidence | | Any change | Unloaded vs end range movements |
| No evidence | | Any change | Among end range movements (extension vs flexion) |
| No evidence | | Any change | Among postural or device loading |
| **Change in position of location of peak signal intensity mid coronal NP profile vs supine unloaded** | | | |
| No evidence | | Any change | Unloaded vs loading |
| No evidence | | Any change | Unloaded vs postural |
| No evidence | | Any change | Unloaded vs compression device |
| Limited | 1(1,0)[163] | ↑<br>Towards contralateral side to SB | Unloaded vs end range movements |
| No evidence | | Any change | Among end range movements (extension vs flexion) |
| No evidence | | Any change | Among postural or device loading |

| | | | |
|---|---|---|---|
| **Post distance of the nucleus** | | | |
| Conflicting | 1(1,0)[167] | Any change | Unloaded vs loading |
| Conflicting | 1(1,0)[167] | Any change | Unloaded vs postural |
| No evidence | | Any change | Unloaded vs compression device |
| No evidence | | Any change | Unloaded vs end range movements |
| No evidence | | Any change | Among end range movements (extension vs flexion) |
| No evidence | | Any change | Among postural or device loading |
| **AP diameter of the disc** | | | |
| Limited | 1(1,0)[167] | ↑ | Unloaded vs loading |
| Limited | 1(1,0)[167] | ↑ | Unloaded vs postural (sitting or standing) |
| No evidence | | Any change | Unloaded vs compression device |
| No evidence | | Any change | Unloaded vs end range movements |
| Limited | 1(1,00)[161] | No difference | Among end range movements (Neutral, flexion, extension, L and R Rotation ) |
| No evidence | | Any change | Among postural or device loading |
| **Left parasagittal, right parasagittal, mid coronal disc diameters** | | | |
| No evidence | | Any change | Unloaded vs loading |
| No evidence | | Any change | Unloaded vs postural (sitting or standing) |
| No evidence | | Any change | Unloaded vs compression device |
| No evidence | | Any change | Unloaded vs end range movements |
| Limited | 1 (1;0)[161] | No difference | Among end range movements (Neutral, flexion, extension, L and R Rotation ) |
| No evidence | | Any change | Among postural or device loading |
| **Change in Posterior bulging vs unloaded supine** | | | |
| Limited | 1(1,0)[90] | No difference | Unloaded vs loading |
| Limited | 1(1,0)[90] | No difference | Unloaded vs postural |

| | | | |
|---|---|---|---|
| Limited | 1(1,0)[90] | No difference | Unloaded vs compression device |
| No evidence | | Any change | Unloaded vs end range movements |
| No evidence | | Any change | Among end range movements (extension vs flexion) |
| Limited | 1(1,0)[90] | No difference | Among postural or device loading |
| **Dural sac cross-sectional area** | | | |
| Conflicting | 8(3,5) [83, 88, 130, 139, 154, 159, 160, 165] | Any change | Unloaded vs loading |
| Conflicting | 3(2,1)[88, 139, 165] | Any change | Unloaded vs postural |
| Conflicting | 5(1,4)[83, 130, 154, 159, 160] | Any change | Unloaded vs compression device |
| No evidence | | Any change | Unloaded vs end range movements |
| Conflicting | 1(1,0)[161] | Any change | Among end range movements (extension vs flexion) |
| No evidence | | Any change | Among postural or device loading |
| **Range of motion** | | | |
| Limited | 2(0,2)[153, 164] | No difference | Any Loading (combining end range, postural and device loading |
| Limited | 1(0,1)[153] | No difference | Unloaded vs loading |
| No evidence | | Any change | Unloaded vs postural |
| Limited | 1(0,1) [153] | No difference | Unloaded vs compression device |
| No evidence | | Any change | Unloaded vs end range movements |
| Limited | 1(0,1)[164] | No difference | Among end range movements (extension vs flexion) |
| No evidence | | Any change | Among postural or device loading |
| **Abnormal tilting movement/ Translatory instability/ Rotatory instability** | | | |

| | | | |
|---|---|---|---|
| No evidence | | Any change | Unloaded vs loading |
| No evidence | | Any change | Unloaded vs postural |
| No evidence | | Any change | Unloaded vs compression device |
| Limited | 1(0,1)[164] | No difference | Unloaded vs end range movements |
| No evidence | | Any change | Among end range movements (extension vs flexion) |
| No evidence | | Any change | Among postural or device loading |
| **Lumbar lordosis** | | | |
| Limited | 3(2,1) | No difference | Unloaded vs loading |
| Limited | 3(2,1) | No difference | Unloaded vs postural |
| Limited | 1(1,0)[83] | No difference | Unloaded vs compression device |
| Limited | 2(1,1)[83, 165] | No difference | Unloaded vs end range movements |
| Limited | 1(1,0)[165] | No difference | Among end range movements (extension vs flexion) |
| Limited | 1(0,1)[83] | No difference | Among postural or device loading |

## APPENDIX V: PRISMA 2009 CHECKLIST

| Section/topic | # | Checklist item | Reported on page # |
|---|---|---|---|
| **TITLE** | | | |
| Title | 1 | Identify the report as a systematic review, meta-analysis, or both. | 1 |
| **ABSTRACT** | | | |
| Structured summary | 2 | Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number. | 2 |
| **INTRODUCTION** | | | |
| Rationale | 3 | Describe the rationale for the review in the context of what is already known. | 3 |
| Objectives | 4 | Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS). | 5 |
| **METHODS** | | | |
| Protocol and registration | 5 | Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number. | |
| Eligibility criteria | 6 | Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale. | 5 |
| Information sources | 7 | Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched. | 5 |
| Search | 8 | Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated. | 6 |
| Study selection | 9 | State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis). | 6 |

| Section/topic | # | Checklist item | Reported on page # |
|---|---|---|---|
| Data collection process | 10 | Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators. | 7 |
| Data items | 11 | List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made. | 8 |
| Risk of bias in individual studies | 12 | Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis. | 7 |
| Summary measures | 13 | State the principal summary measures (e.g., risk ratio, difference in means). | 8 |
| Synthesis of results | 14 | Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., $I^2$) for each meta-analysis. | 8 |

| Section/topic | # | Checklist item | Reported on page # |
|---|---|---|---|
| Risk of bias across studies | 15 | Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies). | 7 |
| Additional analyses | 16 | Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified. | NA |
| **RESULTS** | | | |
| Study selection | 17 | Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram. | 9 |
| Study characteristics | 18 | For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations. | 9 |
| Risk of bias within studies | 19 | Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12). | 27 |
| Results of individual studies | 20 | For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot. | 30 |
| Synthesis of results | 21 | Present results of each meta-analysis done, including confidence intervals and measures of consistency. | 42 |
| Risk of bias across studies | 22 | Present results of any assessment of risk of bias across studies (see Item 15). | |
| Additional analysis | 23 | Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]). | |
| **DISCUSSION** | | | |
| Summary of evidence | 24 | Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers). | 42 |
| Limitations | 25 | Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias). | 47 |
| Conclusions | 26 | Provide a general interpretation of the results in the context of other evidence, and | 48 |

| Section/topic | # | Checklist item | Reported on page # |
|---|---|---|---|
| | | implications for future research. | |
| **FUNDING** | | | |
| Funding | 27 | Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review. | NA |

# APPENDIX VI: ADDENDUM: THE IMMEDIATE EFFECTS OF LOADING ON LUMBAR IMAGING FINDINGS: UPDATE OF A SYSTEMATIC REVIEW

Vahid Abdollah; Eric C Parent; Steffen Adria; Michele C Battié

## Objective

The objective of this addendum was provide an update of the systematic review presented in chapter 3. This was deemed necessary since a number of relevant papers had been published since the review had been conducted to inform the planning of the experimental studies of this Ph.D.

## Methods

***Literature search and study selection.*** The search was updated by adding to the original search results from a search from August 10 2014 until December 22, 2016, from Medline (Ovid), EMBASE (Ovid), Scopus, Web of Science Core Collection, BIOSIS Previews and CINAHL (EBSCO). The rest of the methodology for this update was the same as what has been presented in Chapter 3.

## Results

To illustrate the impact of the newly published evidence, changes resulting from the updating of the literature review in the result section are presented in italic and underlined.

***Studies included.*** After the update, overall the search identified *4768 references* after removing duplicates. After titles and abstracts screening, *4589* were excluded and *179* papers were included for full-text screening (Figure 20). After screening, a total of *30 papers* met selection criteria (*9 more than presented in Chapter 3*). The major reasons for exclusion were not focusing on loading (n=71), employing less than 10 subjects (n=*15*) narrative reviews (n=14), and *descriptive studies without statistical analysis (15)* (Figure 20).

***Demographic information.*** Prospective case series studies were the most frequent study type. *Two* were prospective case-control study [153, 258], and two were retrospective case-series studies [154, 155] (table 23). Most reported age varying between 20 and 77 years. *Fifteen* studies included participants with LBP [83, 87, 124, 127, 130, 140, 154, 156–160, 173, 259–261], and *three studies* included both participants with and without pain [153, 157, 258]. Only *six* reported on the chronicity [154, 156, 157], or duration of pain [87, 139, 258] (table 23). Only one study

reported the intensity of pain [139], and five reported the location of the pain [139, 154–156, 160] for subjects with LBP.

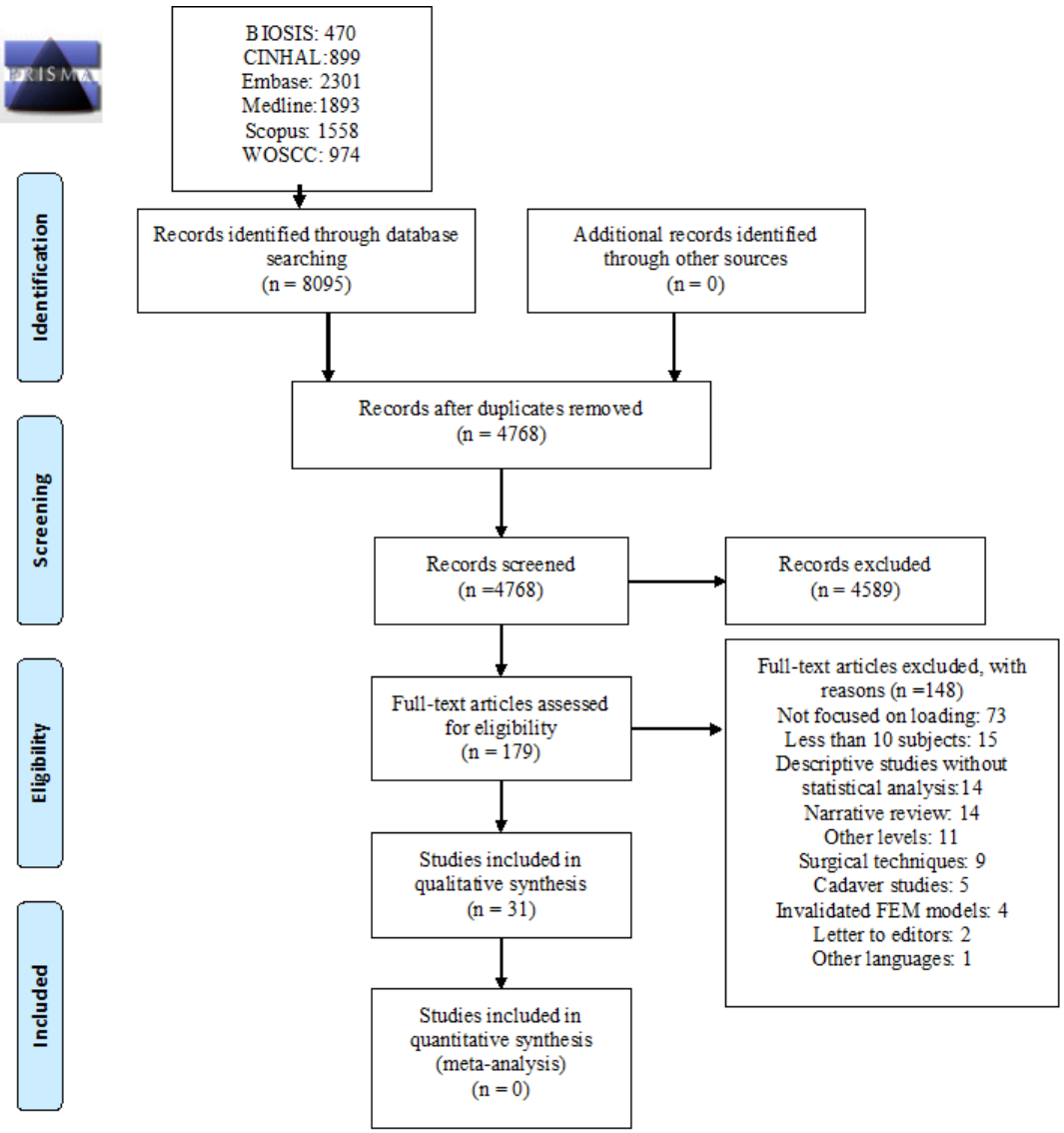


Figure 20 PRISMA flow diagram

**Table 23 Description of study type and study participants in the newly included studies.**

| Authors | Study Type | Recruitment Strategy and Selection Criteria | Number of Subjects and Groups | Descriptive | | | Diagnosis | Pain (Duration, Intensity, Location) |
|---|---|---|---|---|---|---|---|---|
| | | | | Age (years) | Height (cm) | Weight (kg) | | |
| Chung et al. [173] | Prospective Case-series | Consecutive, diagnosis of lumbar disk herniation after CT or MR without contraindications for MRI, surgery, or hypertension. | 48 (13♂, 35♀) | 39.5 (22-64) | | 62.7 (44-90) | LBP | |
| Gallagher et al. [262] | Prospective Case-series | Volunteer, exclusion criteria were any previous history of LBP that required medical intervention or time off work >3 days, previous lumbar or hip surgery, employment requiring prolonged static standing during the past | 17 (9♂, 8♀) | Non-pain developer Male (4): 22.0 (2.2) Female (4): 21.5 (0.6) Pain developer | Non-pain developer Male (4): 1.82 (0.10) Female (4): 1.68 (0.07) Pain developer | Non-pain developer Male (4): 88.9 (12.7) Female (4): 62.0 (12.4) | Healthy | |

| Authors | Study Type | Recruitment Strategy and Selection Criteria | Number of Subjects and Groups | Descriptive | | | Diagnosis | Pain (Duration, Intensity, Location) |
|---|---|---|---|---|---|---|---|---|
| | | | | Age (years) | Height (cm) | Weight (kg) | | |
| | | 12 months, ,and inability to stand for >2 hours, had a radiographic investigation within the past year or were exposed to radiation in their occupation, pregnancy | | Male (5): 22.8 (1.9) Female (4): 22.3 (1.5) | Male (5): 1.82 (0.09) Female (4): 1.65 (0.05) | Pain developer Male (5): 77.6 (6.9) Female (4): 56.3 (7.7) | | |
| Espinoza Orías et al. [258] | Prospective Case-control | Volunteer, recurrent LBP with ≥2 episodes lasting ≥6 weeks without prior surgery for back pain, age >60 years, claustrophobia or other contraindication to MRI and CT, severe osteoporosis, severe | 81 | 38.3±9.2 | | | LBP Healthy | Recurrent LBP (at least 2 episodes lasting more than 6 weeks. |

| Authors | Study Type | Recruitment Strategy and Selection Criteria | Number of Subjects and Groups | Descriptive | | | Diagnosis | Pain (Duration, Intensity, Location) |
|---|---|---|---|---|---|---|---|---|
| | | | | Age (years) | Height (cm) | Weight (kg) | | |
| | | disc collapse at multiple levels, severe central or spinal stenosis, destructive process involving the spine, litigation, or compensation proceedings, extreme obesity, congenital spine defects, and previous spinal injury Asymptomatic: no LBP or previous spinal surgery, younger than 60 years, without claustrophobia or other contraindication to MRI and CT | | | | | | |

| Authors | Study Type | Recruitment Strategy and Selection Criteria | Number of Subjects and Groups | Descriptive | | | Diagnosis | Pain (Duration, Intensity, Location) |
|---|---|---|---|---|---|---|---|---|
| | | | | Age (years) | Height (cm) | Weight (kg) | | |
| Kanno et al. [263] | Prospective Case-series | Consecutive, inclusion: neurogenic intermittent claudication and leg pain or numbness with neurologic signs, radiographically confirmed lumbar spinal canal narrowing on cross-sectional imaging Exclusion: previous lumbar spine surgery, spondylolysis, disc herniation, severe osteoporosis, scoliosis, polyneuropathy, arterial insufficiency, and inflammatory/crystalline arthropathies, congenital | 93 (60♂, 33♀) | 68±10 | 160±9 | 64±11 | LSCS | |

| Authors | Study Type | Recruitment Strategy and Selection Criteria | Number of Subjects and Groups | Descriptive | | | Diagnosis | Pain (Duration, Intensity, Location) |
|---|---|---|---|---|---|---|---|---|
| | | | | Age (years) | Height (cm) | Weight (kg) | | |
| | | spinal anomalies and spinal deformities due to spinal trauma, infection, or tumor | | | | | | |
| Kubosch et al. [260] | Prospective Case-series | Inclusion: scheduled for transforaminal lumbar interbody fusion L5/S1 diagnosed with degenerative spondylolisthesis L5/S1 via conventional X-ray, segment infiltration, and probatory corset Exclusion: previous surgical interventions and malignancies | 15 | | | | LBP | |
| Liu et al [261] | Prospective Case-series | Consecutive, Exclusion: previous thoracolumbar | 68 (24♂, 44♀) | 61 (20-85) | | | Lumbar spondyloli | |

| Authors | Study Type | Recruitment Strategy and Selection Criteria | Number of Subjects and Groups | Descriptive | | | Diagnosis | Pain (Duration, Intensity, Location) |
|---|---|---|---|---|---|---|---|---|
| | | | | Age (years) | Height (cm) | Weight (kg) | | |
| | | surgery, spinal trauma, tumor, ankylosing spondylitis, multilevel spondylolisthesis and retrolisthesis | | | | | sthesis instability | |
| Nilsson et al. [127] | Prospective Case-series | Inclusion: spine surgery candidates referred from unit<br>Exclusion: any radicular nerve symptoms, previous disc surgery or contraindications for MRI | 11<br>(7♂, 4♀) | ♂41<br>(25-51)<br>♀44<br>(25-69) | | | LBP | |
| Takasaki [264] | Prospective Case-series | Volunteers, inclusion: no LBP history and no contraindication for MRI | 20<br>(10♂, 10♀) | 24.8 (4.0) | | | Healthy | |

| Authors | Study Type | Recruitment Strategy and Selection Criteria | Number of Subjects and Groups | Descriptive | | | Diagnosis | Pain (Duration, Intensity, Location) |
|---|---|---|---|---|---|---|---|---|
| | | | | Age (years) | Height (cm) | Weight (kg) | | |
| | | Exclusion: movement loss of the back or lateral deformity of the spine, lumbar disc degeneration (Grades III–V) | | | | | | |
| Zhong et al. [265] | Prospective Case-series | Volunteers, exclusion: current or prior LBP, anatomic abnormalities or any spinal disorders | 10 (5♂, 5♀) | 40-60 | | | Healthy | |

Abbreviations and symbols: LBP=Low back pain; MRI=Magnetic Resonance Imaging; ♂=males; ♀=females; LSCS= lumbar spinal canal stenosis; wks=weeks; mths= months; yrs= years

**Table 24 Description of the imaging methods and loading strategies tested in the newly included studies**

| Studies | Imaging Modalities | Type of Loading | Mode of Application | Duration of Loading | Magnitude of Loading | Time of Day of Loading | Precondition of Loading |
|---|---|---|---|---|---|---|---|
| Chung et al. [173] | MRI | Traction | Custom made device | 30 min during MRI | 30 kg | | |
| Gallagher et al. [262] | X-ray | Standing | Postural | ? | | | |
| Espinoza Orías et al. [258] | MRI/CT | Rotation to right in supine | Positional | During imaging | | | |
| Kanno et al. [263] | MRI | Compression | DynaWell | During MRI | 50% BW | | 5 min. before loading |
| Kubosch et al. [260] | MRI | Standing/lying | Postural | During MRI | | | |
| Liu et al [261] | MRI/X-ray | Standing/flexion-extension in standing/lying supine | Positional | During imaging | | | |
| Nilsson et al. [127] | MRI | Compression | DynaWell | During MRI | 50% BW | 11-13 | 20 min. lying before imaging |

| Studies | Imaging Modalities | Type of Loading | Mode of Application | Duration of Loading | Magnitude of Loading | Time of Day of Loading | Precondition of Loading |
|---|---|---|---|---|---|---|---|
| Takasaki [264] | MRI | Supine/supine side bending | Positional | During MRI | | | 30 min. lying between images |
| Zhong et al. [265] | X-ray video fluoroscopy | Active flexion/extension in standing | Positional | During fluoroscopy | | | |

Abbreviations: MRI= Magnetic resonance imaging, Min.= minutes, hrs=hours

***Imaging and loading modalities used.*** MRI was employed alone in *23* studies. X-ray fluoroscopy was used in *three* studies [157, 166, 265] and *three* studies employed plain x-ray imaging [87, 164, 262], *two* studies employed MRI with CT-scans [258] or X-ray [261] (table 24). *Thirteen* studies employed positional loading in the form of end range positions of the lumbar spine during imaging including: flexion/extension [83, 87, 124, 140, 157, 161, 165, 166, 261, 265], side-bending [163, 264] or rotation [161, 166, 262]. The DynaWell® compression device was employed to load the lumbar spine with 40 to 50% compression body weight (BW) in *eight* studies [87, 90, 127, 130, 139, 154, 160, 263], while a mechanical bracing device to load with 75% BW compression load [153], one nonmagnetic compression device (50% BW) [159], *and one nonmagnetic traction device* [173] were used in three other studies. Postural loading was used in *eight* studies (standing/sitting/lying [88, 155, 156, 167, 260, 262] or kneeling [90]). In all *but 6 of 31* included studies, loading was applied only during the imaging. These *five* studies applied loading before and during imaging with pre-imaging loading durations of >5min [154, 263], and ≤20min [127, 130, 153]. Only *three* studies described standardised preconditioning before imaging [88, 90, 167] with one more specifying time of examination to control the loading history [165].

***Methodological quality.*** None of the included studies met the 70% high-quality threshold (table 25). *Six studies scored between 60% and 65%* [127, 231, 261–264], *three scored between 50% and 60%* [87, 164, 258], and the rest scored less than 50% [87, 88, 130, 139, 140, 155, 156, 165, 260, 265]. Lack of a well-defined subject's recruitment (selection bias) and examiners' training strategy were the most common sources of bias. Only *13/31* studies reported the qualification of the examiners [87, 88, 127, 130, 140, 154, 155, 159, 160, 163, 231, 263, 264]. Eleven studies employed a consecutive recruitment strategy to recruit subjects with back pain [83, 130, 139, 140, 154, 155, 164, 173, 261, 263] and without back pain [166]. The rest employed volunteers with pain [90, 127, 156, 160, 260], with sciatic or neurogenic claudication [159, 263], or without pain [88, 153, 157, 161–163, 165, 167, 262, 264, 265] or both with and without pain [153, 157, 258]. Study hypothesis or objective and validity of the outcome measured were reported in all studies, except two with unclear outcomes reporting strategy [161, 164]. Only *two studies* met all requirements for handling missing data [160, 262]. A possible common source of bias was the majority of studies failing to formulate correlation and mean difference-testing hypotheses a priori. Only one study had a well-defined statistical analysis strategy [165], and *five* met 80% of

the statistical analysis requirements [83, 88, 156, 165, 264]. No study met all the requirements for reporting results because none reported on important adverse events related to loading.

**Table 25 Quality appraisal of the studies included**

| Authors | Subjects Recruitment /7* | Examiners /4* | Methodology /5 | Outcomes /2 | Missing Data /8 | Statistical Analysis /5 | Results /2 | Overall Score /33* | Overall Score (%)* |
|---|---|---|---|---|---|---|---|---|---|
| Chung et al [231] | 4 | 1 | 3 | 2 | 6 | 3 | 2 | 21 | 64 |
| Gallagher et al. [262] | 3 | 2 | 3 | 2 | 8 | 3 | 1 | 20 | 61 |
| Espinoza Orías et al. [258] | 2 | 0 | 4 | 2 | 5 | 3 | 2 | 18 | 54 |
| Kanno et al. [263] | 4 | 3 | 3 | 2 | 6 | 2 | 1 | 21 | 63 |
| Kubosch et al. [260] | 2 | 0 | 1 | 2 | 0 | 3 | 1 | 9 | 27 |
| Liu et al [261] | 4 | 0 | 5 | 2 | 6 | 3 | 1 | 21 | 63 |
| Nilsson et al. [127] | 4 | 3 | 2 | 2 | 6 | 3 | 1 | 21 | 64 |
| Takasaki [264] | 4 | 1 | 3 | 2 | 6 | 4 | 1 | 21 | 64 |
| Zhong et al. [265] | 3 | 0 | 2 | 2 | 6 | 1 | 1 | 15 | 45 |

Overall score: sum of all scores

*One question from the subject domain and two from the examiner domain could be scored “not relevant” and excluded from score calculations. The affected scores were calculated as the ratio of the number of relevant criteria met to the total number of relevant criteria and reported out of the original number of questions for the score.

***Measurement outcomes.*** *Twenty-seven* quantitative measurements of the effect of loading on the lumbar spine were identified and level evidence summary statements were formulated for each measure and loading conditions (table 26). As no study reached the high-quality threshold, the highest possible level of evidence was limited (table 26).

***Unloaded vs. any type of postural or axial loading.*** When formulating summary statement combining the evidence of both postural and axial loading, there was limited evidence of a statistically significant increase of the anterior distance of the nucleus positions from the anterior limit of the disc [167], anterior-posterior diameter of the disc [167], *and $T_2$ time* [127] in response to compression loading, and *of an increased spine length in response to traction* [173] (table 26). *There was also limited evidence of decrease of the diameter of the intervertebral foramen in response to compression loading* (table 26) [260]. There was also limited evidence of no change in posterior disc bulge [90], range of motion [153], and lumbar lordosis angle in response to loading (table 26). The evidence was conflicting on the effects of loading on the cumulative [83], anterior and posterior disc heights [90, 156, 167], anterior-posterior diameter of the nucleus [167], posterior distance of the nucleus position [167], DCSA [83, 88, 130, 139, 154, 159, 160, 165], *area of the intervertebral foramen* [260]*, and volume of the spinal canal* [260]*, and spine length (between 20 and 30 minutes of traction)* [173] (table 26).

***Unloaded vs. axially loaded spine using any devices.*** When focusing only on axial loading using compression device compared to the supine unloaded position, there was limited evidence of a statistically significant *increase in the $T_2$ time* [127], and a decrease in the posterior disc height in response to compression loading (table 26) [90]. *There was also limited evidence of increased spine length in response to traction* [173]. There was limited evidence of no change in response to compression loading for the cumulative disc height [83], posterior disc bulging [4], range of motion [153], and lumbar lordosis angle [83] (table 26). The evidence was conflicting on the effects of loading on the DSCA [83, 130, 154, 159, 160].

**Table 26 Levels of evidence for summary statements for each quantitative measurement of the effect of lumbar loading.**

| Level of evidence | From n studies (n healthy; n LBP) | Changes | Measurement Construct | Loading conditions compared |
|---|---|---|---|---|
| **Unloaded vs. any type of loading (axial loading/postural loading)** | | | | |
| Limited | 1(1,0)[167] | ↑ | Anterior distance of the nucleus positions from the anterior limit of the disc | Standing; sitting; lying supine |
| Limited | 1(1,0)[167] | ↑ | Anterior-posterior diameter of the disc | Standing; sitting; lying supine |
| *Limited* | *1(0,1)*[127] | ↕ | *$T_2$ time* | *Supine with 50% BW compression* |
| *Limited* | *1(0,1)*[173] | ↕ | *Spine length* | *Supine with 30 kg traction* |
| *Limited* | *1(0,1)*[260] | ↕ | *Diameter of the intervertebral foramen* | *Supine; standing* |
| Limited | 1(1,0)[90] | No difference | Change in the posterior bulging | Kneeling; 50% BW axial loading |
| Limited | 1(0,1)[153] | No difference | Range of motion | Supine with 75% BW compression |
| *Limited* | *6(4,2)*[83, 88, 127, 165, 260, 262] | *No difference* | *Lumbar lordosis angle* | *Standing; supine with 40% and 50% BW compression; sitting* |
| *Conflicting* | *1(0,1)*[173] | *Any changeAny change* | *Spine length* | *Difference between 0, 10, 20 and 30 mins. of supine traction with 30 kg traction* |
| *Conflicting* | *1(0,1)*[260] | *Any changeAny change* | *Area of the intervertebral foramen* | *Supine; standing* |
| *Conflicting* | *1(0,1)*[260] | *Any changeAny change* | *Volume of the spinal canal* | *Supine; standing* |

| | | | | |
|---|---|---|---|---|
| Conflicting | 1 (0;1)[83] | Any change | Cumulative disc height | Supine with 40% and 50% BW compression |
| Conflicting | 3(2;1) [90, 156, 167] | Any change | Anterior disc height | Sitting; kneeling; supine with 50% BW compression; standing/sitting |
| Conflicting | 3(2,1)[90, 156, 167] | Any change | Posterior disc height | Sitting; kneeling; supine with 50% BW compression; standing/sitting |
| Conflicting | 1(1,0)[167] | Any change | Anterior-posterior diameter of the nucleus | Standing; sitting; lying supine |
| Conflicting | 1(1,0)[167] | Any change | Posterior distance of the nucleus positions from the posterior limit of the disc | Standing; sitting; lying supine |
| Conflicting | 8(3,5)[83, 88, 130, 139, 154, 159, 160, 165, 263] | Any change | Dural sac cross-sectional area | 50% BW compression standing; supine with 40% and 50% BW compression standing/sitting |
| **Unloaded vs Axial Loading with Devices** | | | | |
| *Limited* | *1(0,1)*[127] | ↕ | *$T_2$ time* | *Supine with 50% BW compression* |
| *Limited* | *1(0,1)*[173] | ↕ | *Spine length* | *Supine with 30 kg traction* |
| Limited | 1 (1,0)[90] | ↓ | Posterior disc height | Supine with 50% BW compression |
| Limited | 1(0,0)[83] | No difference | Cumulative disc height | Supine with 40% and 50% BW compression |
| Limited | 1(1,0)[90] | No difference | Change in the posterior bulging | Supine with 50% BW compression |

| | | | | |
|---|---|---|---|---|
| Limited | 1(0,1) [153] | No difference | Range of motion | Supine with 75% BW compression |
| Limited | 1(1,0)[83] | No difference | Lumbar lordosis angle | Supine with 40% and 50% BW compression |
| Limited | 3(2,1) [83, 88, 165] | No difference | Lumbar lordosis angle | Supine with 40% and 50% BW compression |
| Conflicting | 5(1,4)[83, 130, 154, 159, 160] | Any change | Dural sac cross-sectional area | Supine with 40% and 50% BW compression |
| **Unloaded vs. Postural Loading** | | | | |
| Limited | 1(1,0)[167] | ↑ | Anterior distance of the nucleus positions from the anterior limit of the disc | Standing; sitting |
| Limited | 1(1,0)[167] | ↑ | Anterior-posterior diameter of the disc | Standing; sitting |
| Limited | 1 (0;1)[83] | ↓ | Cumulative disc height | Standing |
| *Limited* | *1(0,1)*[260] | ↓ | *Diameter of the intervertebral foramen* | *Supine; standing* |
| Limited | 2(2;0)[90, 167] | No difference | Anterior disc height | Kneeling standing; sitting |
| Limited | 1(1,0)[90] | No difference | Change in the posterior bulging | Kneeling |
| Limited | 3(2,2)[83, 88, 165, 262] | No difference | Lumbar lordosis angle | Kneeling standing; sitting |
| *Conflicting* | *1(0,1)*[260] | *Any change* | *Area of the intervertebral foramen* | *Supine; standing* |
| *Conflicting* | *1(0,1)*[260] | *Any change* | *Volume of the spinal canal* | *Supine; standing* |
| Conflicting | 3(2,1)[90, 156, 167] | Any change | Posterior disc height | Standing; sitting; kneeling |
| Conflicting | 1(1,0) [167] | Any change | Anterior-posterior diameter of the nucleus | Standing; sitting |
| Conflicting | 1(1,0)[167] | Any change | Posterior distance of the nucleus positions from the posterior limit of the disc | Standing; sitting |
| Conflicting | 3(2,1)[88, 139, 165] | Any change | Dural sac cross-sectional area | Standing; sitting |
| **Among postural or axial loading** | | | | |
| Limited | 1(0;1)[83] | No difference | Cumulative disc height | Supine with 40% and 50% BW compression |
| Limited | 2(2,0)[90, 167] | No difference | Posterior disc height | Kneeling; 50% BW axial loading |

| | | | | |
|---|---|---|---|---|
| | | | | Standing; sitting |
| Limited | 1(1,0)[90] | No difference | Change in the posterior bulging | Kneeling; 50% BW compression |
| Limited | 1(0,1)[83] | No difference | Lumbar lordosis angle | Supine with 40% and 50% BW compression |
| Conflicting | 1(0,1)[173] | Any change | Spine length | difference between 10, 20 and 30 mins. supine with 30 kg traction |
| **Unloaded vs end range movements** | | | | |
| Limited | 1(1;0)[162] | ↓ | Anterior disc height | Passive flexion/extension, rotation in supine |
| Limited | 1(1,0)[163] | ↑ Towards contralateral side to SB | Change in the location of peak signal intensity of the nucleus in the mid-coronal plane | Active side-bending in supine |
| *Limited* | *1(1,1)* [261] | ↕ | *Motion segment instability* | *Supine; active flexion/extension in standing* |
| *Limited* | *1(1,1)* [258] | ↕ | *Disc height* | *Supine; passive supine rotation* |
| Limited | 1(0,1)[164] | No difference | Abnormal tilting movement/ translatory instability/ rotatory instability | Active flexion/extension in lateral decubitus |
| *Limited* | *2(1,1)*[83, 165] | *No difference* | *Lumbar lordosis angle* | *Active flexion/extension in sitting*<br>*passive extension in supine with 50% BW compression* |
| *Conflicting* | *1(1,1)* [258] | *Any change* | *Height of different zones of the disc* | *Passive supine rotation* |

| | | | | |
|---|---|---|---|---|
| *Conflicting* | *1(1,0)*[264] | *Any change* | *Lateral flexion angle* | *Supine side bending* |
| *Conflicting* | *1(1,0)*[264] | *Any change* | *Segmental rotation angle* | *Supine side bending* |
| *Conflicting* | *1(1,0)*[262] | *Any change* | *Intervertebral angle* | *Maximum active lumbar extension positions* |
| **Among end range movements** | | | | |
| Limited | 1(1;0)[162] | ↓ | Anterior disc height | Passive flexion/extension in supine |
| Limited | 1(1,0)[162] | ↓ | Anterior distance of the nucleus positions from the anterior limit of the disc | Passive flexion/extension in supine |
| *Limited* | *1(1,0)*[265] | ↓ | *Area of the intervertebral foramen* | *Active flexion/extension in standing* |
| *Limited* | *1(1,0)*[265] | ↓ | *Width of the intervertebral foramen* | *Active flexion/extension in standing* |
| Limited | 1(1,0)[162] | No difference | Posterior disc height | Passive flexion/extension in supine |
| Limited | 1(1,00)[161] | No difference | Anterior-posterior diameter of the disc | Passive flexion/extension in supine |
| Limited | 1 (1;0)[161] | No difference | Left and right parasagittal, mid coronal disc diameters | Passive flexion/extension in supine |
| Limited | 1(0,1)[164] | No difference | Range of motion | Active flexion/extension in lateral decubitus |
| Limited | 1(1,0)[165] | No difference | Lumbar lordosis angle | Active flexion/extension in sitting |

| | | | | |
|---|---|---|---|---|
| Conflicting | 1(1,0)[161] | Any change | Dural sac cross-sectional area | Passive flexion/extension in supine |
| *Conflicting* | *1(1,0)*[265] | *Any change* | *Height of the intervertebral foramen* | *Active flexion/extension in standing* |

***Unloaded vs. postural loading (sitting/standing).*** When focusing only on axial compression using postural loading, there was limited evidence of a statistically significant decrease in the cumulative disc height [83] and *diameter of the intervertebral foramen* [260] in response to loading (table 26). There was also limited evidence of a statistically significant increase in the anterior distance of the nucleus positions from the anterior limit of the disc and anterior-posterior diameter of the disc in response to loading (table 26) [167]. There was limited evidence of no difference in anterior disc height [90, 167], changes in the posterior disc bulging [90] and lumbar lordosis angle [83, 88, 165]( table 26). The evidence was conflicting about the posterior disc height [90, 156, 167], the anterior-posterior diameter of the nucleus [167], the posterior distance of the nucleus positions from the posterior limit of the disc [167], DSCA [88, 139, 165], *area of the intervertebral foramen and volume of the spinal canal* [260].

***Comparison among different postural and axial loading conditions.*** Few studies compared multiple loaded conditions [90]. There was limited evidence of no difference of the cumulative (Passive flexion/extension in supine; 40% BW, 50% BW compression) [83] and posterior disc heights (Kneeling; 50% BW axial loading; standing; sitting; lying supine) [90, 167], change in the posterior bulging (kneeling; 50% BW axial loading) [90] and lumbar lordosis angle (passive flexion/extension in supine; 40% and 50% BW on compression) [83].

***Unloaded vs. end range movements.*** Only *eight* studies compared unloaded spine images with end-range movement conditions. There was limited evidence of a statistically significant decrease of the anterior disc height in response to going from lumbar flexion to extension in the supine position (table 26) [162]. *There was also limited evidence of a statistically significant increase motion segment instability in response to flexion-extension in standing* [261] *and disc height in response to supine rotational loading* [258]. There was also limited evidence of a statistically significant shift of the location of the nucleus peak signal intensity toward the contralateral side in the mid-coronal plane during side bending [163]. There was also limited evidence of no difference in abnormal tilting movement, translatory and rotatory instability measurements [164] as well as in lumbar lordosis angle [83, 165]. *The evidence was conflicting about the height of different zones of the disc with supine rotation* [258], *as well as for the lateral flexion angle* [264], *segmental rotation angle* [264], *and the intervertebral angle* [262] *in response to side-bending movements in table 26.*

***Among end range movements.*** Only *six* studies compared measurements among different end-range movements. There was limited evidence of a statistically significant increase in the anterior disc height and the anterior distance of the nucleus positions from the anterior limit of the disc going from a flexed position to an extended position [162]. *There was also limited evidence of a statistically significant decrease in the area and width of the intervertebral foramen going from a flexed position to an extended position* [265]. There was limited evidence of no change for the posterior disc height (supine extension vs. flexion) [162], anterior-posterior, left and right parasagittal, mid-coronal diameter of the disc (among supine neutral, flexion, extension, rotation to the right and left) [161], range of segmental motion (standing flexion vs. extension) [164], and lumbar lordosis angle (among supine unloaded, normal sitting, sitting flexed, sitting extended and standing) [165]. The evidence was conflicting on the DSCA changes on neutral, extension and rotation to the right and left compared to the flexion position [161], *and height of the intervertebral foramen comparing active flexion and extension in standing* [265].

*Among participants with and without pain. There was limited evidence of no difference in the disc height changes in response to rotation in supine between participants with and without LBP* [258]. *There was limited evidence of no difference in the lumbosacral lordosis angle, lumbar lordosis angle, L1-2 and L5-S1 intervertebral joint angle in response to different standing positions (standing on level ground, sloped surface, one leg elevated, and maximum lumbar spine extension) between participants with and without LBP* [262].

**Discussion**

The present systematic review identified important research gaps. Interestingly, for numerous quantitative measurements of the effect of loading on the lumbar spine, no evidence could be found suggesting further research is needed. Indeed, depending on the loading conditions compared, *only 1 to 12 of the 27 measurements* were used in at least one study. This review found no moderate or strong evidence about the effect of loading or of end-range movement positions on lumbar imaging quantitative findings due to the insufficient quality of the included studies. Further, only *four* studies included both healthy and participants with LBP [153, 157, 258, 261, 262], *but only two studies reported the statistical significance of the comparisons of the effects of loading between groups* [258, 262]. Therefore, the clinical value of some of the proposed measures

still remains unclear and more high-quality studies among matched patients, and asymptomatic participants are needed to determine the clinical value of these measurement constructs. We can therefore, *for most measurements,* only rely on comparing studies conducted with only one type of participants to understand the difference in response to loading of participants with and without pain.

A number of conflicting (*1 to 8 measurements*) and limited (*2 to 8 measurements*) evidence summary statements concerning the effects of loading and movements on the lumbar spine were observed depending on the loading conditions compared. This conflicting evidence could be due to heterogeneity in the studies in terms of the population studied (e.g. with or without pain, different ages) or because of other heterogeneous methodological considerations. The limited evidence summary statements often showed no effect of the loading conditions compared (*4 to 6* measurements depending on the loadings compared). Overall, the observation of conflicting results and the small number of measurement constructs that detected limited evidence of consistent effects of loading may suggest that loading has a relatively small effect on the quantitative measurements examined on the lumbar spine to date.

*Eight* measurement constructs were compared among studies where both studies focused only on either healthy or LBP participants were available (table 23). For those comparisons, the results of *two* measurement constructs were consistent showing no effect of loading in both participants with and without pain from different studies *and two showing an increase in both groups*. For the remaining *four* other instances, results were conflicting when combining results from participants with and without pain. When trying to formulate levels of evidence summary statements separately for participants with and without pain, a statement on two outcomes remained conflicting in both types of patients. For the DSCA, there was limited evidence of a decrease in postural loading in those with LBP and the results remained conflicting in healthy participants. For DSCA, there was limited evidence of a decrease in healthy with combined loading studies and with device compression while evidence remained conflicting in those with LBP. For posterior disc height comparing unloaded to any loading there was limited evidence of no difference with loading in LBP and the evidence remained conflicting in healthy participants. Other measurement constructs were evaluated only on participants with back pain or asymptomatic participants, and not both (table 23).

The studies compared were heterogeneous, which may have contributed to the frequent conflicting evidence summary statements and limiting our ability to observe consistent effects of loading. *Twenty-seven* different measurement constructs were observed. The sample size and age varied widely. *Twelve* studies only included healthy participants, *15* only LBP participants, and *three* both, with and without pain (table 23). More than *17* different loading conditions (table 24) and three imaging modalities were inventoried. Only three studies employed preloading of the spine before imaging. A few studies conducted manually digitised measurements using OSIRIX [83, 156, 165, 167, 263], while the rest employed other software or did not report software details. Three studies employed automatic measurements [140, 155, 158], one study employed semi-automatic measurements [153], and the rest employed manual measurements.

Difficulty in formulating stronger levels of evidence summary statements is partially due to the lack of quality research on this topic. Most of the studies reviewed did not report the degree to which assessors were trained before conducting measurements or whether the measurements extracted presented adequate reliability. Further, *12 of the 30* included studies included samples with fewer than 30 participants suggesting that a limited power may have limited the authors' ability to observe statistically significant results.

It is unclear if compression devices are limited regarding the maximum load that can be applied to the spine safely. In the present review, only one study applied a compression load equal to 75% BW [153], others employed compression loads less than or equal to 50% BW, with only *five studies* controlling for how long before the imaging the loads were applied. The relatively small number of consistent effects of loading observed with supine compression devices (only decreased posterior disc height and *increased $T_2$ time*) versus when using postural loads (increased anterior distance of nucleus, A-P diameter of disc and decreased cumulative disc height *and diameter of the intervertebral foramen*) may be due to the employing less than 80% BW for supine compression. Overall, for most measurements, the results of this review, were not in favour of the hypothesis that axial loading using postural or device compression could induce significant morphologic changes in the lumbar spine. Further research using a wide array of quantitative measurements and loading conditions is needed to determine if loading during imaging may increase the specificity for clinically relevant lumbar pathology.

The evidence was conflicting about the effects of loading on the imaging presentations of DSCA. Clinically, imaging is not always correlated with the severity of LSCS symptoms [168, 169]. Subjects with significant narrowing of the spinal canal, may not demonstrate any symptoms, while subjects with severe symptoms may not demonstrate any significant narrowing of the LSCS [139]. Identifying imaging findings correlating with clinical symptoms in LSCS would thus be beneficial to achieve a more accurate diagnosis and help plan a more appropriate treatment. Our review yielded conflicting results on the effects of loading on the LSCS, and only one study was identified evaluating the effects of loading specifically on LSCS symptoms [139]. The results indicated a statistically significant negative correlations between the DCSA in axially loaded MRI with both walking distance and the Japanese Orthopedic Association score [139]. Our review did not yield any evidence on the possible effects of extension motion compared to an unloaded supine position on the LSCS. Although some researchers have suggested that the extension could trigger LSCS symptoms by a reduction of the dural sac size [170, 171]. Further high-quality studies are therefore needed to figure out the possible correlation between the change in the DCSA caused by the loading and the severity of the symptoms in patients with LSCS.

Although imaging the lumbar spine while it is loaded using sensitive quantitative measurements may help capture abnormal vertebral motion, and thus increase the sensitivity of imaging for diagnosis of LSI, there were no studies included in this review examining the effect of compression on LSI measurements. Further, only *two studies* of on participants with LBP examined the effect of end-range movements on quantitative LSI measurements [164, 261]. *Therefore, there is still insufficient evidence to formulate strong level of evidence summary statements related to the correlation between LSI imaging findings in response to loading and patients' symptoms (table 23). Only one study evaluated the difference between those with and without pain* [261]. *The results indicated that the percentage of patients with back pain was twice as high in the group with higher mobility in flexion-extension compared to asymptomatic participants* [261]. *The lack of pain likely facilitates higher mobility in flexion-extension* [261].

Further, despite encountering *19* loading conditions, *our review yielded only one study that evaluated the effects of traction* [173]. *The results indicated that spine length increased after 10 and 30 minute of traction compared to unloaded stage, but the evidence was conflicting when difference among different time-points examined* [173]. Traction has been employed traditionally

to reduce pain and discomfort among participants with LBP and sciatica [47], affect the fluid content, promote molecular transport in the IVDs [172], and reduce the size of herniated material [173]. Therefore, traction theoretically could be a loading condition useful to identify biomarkers of relevant low back pathology. Recently, Chung et al. study found a significant elongation of the lumbar spine and a decrease of the size of disc herniation after 30 minutes of traction loading [173]. As no studies about the in-vivo effects of traction on other measurements or comparing participants with and without pain have been conducted, the gap of knowledge still exists.

Although few studies have employed MR imaging to quantify effects of loading on the disc fluid content [22, 127], *our review only yielded one study that evaluated the effects of compression loading on the disc mean $T_2$ time in participants with LBP using $T_2$-mapping* [127]. *The results found a significant increase in $T_2$ values of the whole disc and subsections* [127]. *The IVDs are expected to lose or not change fluid content when undergoing compression* [32]. *The intended uniaxial compression load may not be equal to overcome osmotic pressure or have resulted in a complex loading condition throughout the spine. $T_2$-value changes were correlated with degeneration grade, changes in disc angle and lumbar level* [127]. As there are no studies about the in-vivo effects of all other forms of loading on the disc fluid content, the gap of knowledge related to disc fluid distribution in response to loading still exists.

***Study limitations.*** This review solely included studies published in English, and no search was conducted of the grey literature. These two factors might have caused a potential bias in selecting relevant studies. As discussed earlier, the results were very heterogeneous which prevented meta-analysis. Unfortunately, the literature is not sufficiently rich to limit the review to studies involving head-to-head comparisons of patients with and without low back pain. This review focused only on quantitative measurements of the effect of loading on the lumbar spine. Studies using discrete subjective ratings have been published to describe the effect of loading on the lumbar spine with some studies suggesting that loading helps identify relevant clinical findings [174]. However, discrete subjective ratings have routinely been criticised for their lack of reliability justifying excluding such measurements from this review.

## Conclusion

The heterogeneous results highlighted inconsistent evidence regarding the effects of loading on the lumbar spine. The review did not yield any moderate or strong evidence because of the insufficient quality scores of the included studies. For many measurement constructs, no evidence was identified to draw a solid conclusion regarding their possible effects of loading on the lumbar spine. More than half of the limited evidence observed on cumulative, anterior and posterior disc height, the anterior distance of the nucleus positions from the anterior limit of the disc, change in the posterior bulging, abnormal tilting movement, translatory and rotatory instability, range of motion, and lumbar lordosis was of no effect of loading. The results highlighted that the gap of knowledge regarding the effects of loading on the imaging presentations of the lumbar spine still exists. Particularly, there is a lack of research on whether the response to loading could help increase the specificity of imaging for LBP. Therefore, further high-quality studies, comparing responses in subjects with symptoms of LBP and LSCS to matched healthy subjects, are needed to establish a strong correlation between imaging findings and patients' symptoms. Our review also identified a lack of research on many quantitative measurements such as the effects of loading on the disc fluid content. Further, high-quality studies are needed to quantify the effects of loading on additional quantitative imaging measurements and to establish a strong correlation between these measurement constructs and patients' symptoms.