COMPARISON OF A SIX MONTH NON-AGGRAVATING GYM REHABILITATION PROGRAMME VERSUS USUAL SURGICAL ADVICE POST LUMBAR DISCECTOMY. A PROSPECTIVE RANDOMISED CONTROLLED TRIAL WITH 3 YEAR FOLLOW UP

Barry Donaldson

A thesis submitted for the degree of

Doctor of Philosophy

at the University of Otago,

Christchurch School of Medicine and Health Sciences

New Zealand

2009

ABSTRACT

Introduction:

Lumbar discectomy is now the operation of choice for lumbo-sacral radicular syndrome. Few studies of high quality have been performed on the post-surgical management of these cases. All but two of the studies that have been reported compare one exercise regime to another.

The aim of this study was to compare long term outcomes of usual surgical advice, involving no formal post-surgical rehabilitation, with a non-aggravating six month gym rehabilitation programme following lumbar discectomy. This study is a prospective randomised controlled trial using a cohort followed for three years.

Method

The patients were computer-randomised into two groups. Group A, the control group, followed usual surgical advice, which was to resume normal activity as soon as pain allowed. Group B undertook the gym rehabilitation programme. Inclusion criteria were: age 17 to 65 years, good health and no major medical problems. The surgical level had to be L3, L4, or L5. Patients were excluded if they had central neurological disorders, communication difficulties, any condition making gym-based exercises unsafe or if the surgery was indicated for spinal infection, tumour or inflammatory disease. Participants were followed for a three year period using validated outcome measures (Roland-Morris Questionnaire and Oswestry Low Back Pain Index) and an annual Quality of Life (QoL) questionnaire. A sample of 40 participants per group provided the study with 80% power ($p < 0.05$) to detect a 3.5 point change in the RMQ and a 10% change in the ODI. The annual questionnaire reported information on number of doctor visits, other therapist visits, medication levels and time off work.

Results

Ninety-four participants were randomised, control n = 47 and trial n = 47. Eighty-seven participants completed the study. Randomisation achieved a balance of confounding factors, with the exception of work heaviness, where there were a greater number of participants in the ‘very heavy’ and ‘heavy’ categories in the trial group ($p < 0.01$). Thirty-nine of 47 participants completed the gym programme (83%). Functional outcome measures did not

show statistically significant differences between groups over the three year period. Key findings of cumulative 3 year data for the QoL questionnaire are on intent-to-treat analysis; fewer participants had doctor visits in the trial group ($p = 0.01$ (18% vs 5%)). In the per-protocol population; fewer participant doctor visits ($p = 0.03$ (range control 0 - 9 vs trial 0 - 3)), fewer total number of doctor visits ($p = 0.05$ (range control 0 - 22 vs trial 0 - 6)) and less medication use in the trial group in year two ($p = 0.04$ (49% control vs 28% trial)). In the per-protocol minus re-operation group; fewer doctor visits ($p < 0.01$ (range control 0 - 7 vs trial 0 - 2)) and fewer trial participants requiring medication in year three ($p = 0.05$ (37% control vs 17% trial)).

Discussion

The results from this study did not significantly differentiate the control and trial groups. There were some results that reached statistical significance from the QoL questionnaire. These results are similar to previous research comparing an exercise group with a non-exercise group after lumbar discectomy; the results from these previous studies were also equivocal. This indicates that further research is required to confirm what is best practice after lumbar discectomy surgery.

PREFACE

The physiotherapy profession is being challenged by medical colleagues and funding institutions to prove that physiotherapy interventions lead to functional improvement and are cost effective. These challenges were the driving force that initiated this study.

Lumbar discectomy accounts for almost 90% of all lumbar surgery. In recent years the procedure of lumbar discectomy has become increasingly popular and consequently absorbs significant resources. For these reasons, any form of management that demonstrates improved long term outcomes following lumbar discectomy, has the potential to save millions of health care dollars.

The principal investigator for this study has frequently managed post-surgical lumbar cases using a gymnasium-based physiotherapy rehabilitation programme. Anecdotally, the principal investigator has found that patients who were supervised through a rehabilitation programme tended to have fewer post-surgical complications than patients who had not had any formal post-surgical intervention.

To test whether these improved outcomes were attributable to the rehabilitation programme, a prospective randomised controlled trial was undertaken. The objective was to test whether a progressive non-aggravating gymnasium-based physiotherapy rehabilitation programme provided superior long term outcomes for post-surgical lumbar discectomy participants. This thesis is the final result of that trial.

PUBLICATIONS

There have been two publications from this study to date.

Referred Publications

Donaldson, B.L., Shipton, E.A., Inglis, G., Rivett, D., & Frampton, C. (2006). Comparison of usual surgical advice versus a non-aggravating six-month gym-based exercise rehabilitation programme post-lumbar discectomy: results at one year follow-up. Spine Journal: Official Journal of the North American Spine Society, 6 (4), 357-363.

Donaldson B.L., Rivett D., Shipton E.A., Inglis G., & Frampton C. (May 2008). Concepts of Exercise Prescription. Australasian Musculoskeletal Medicine, 13 (1), 28-39.

Conference Presentations

The results of this study have been presented by the author at six conferences thus far, including an invitation as a keynote speaker in October 2008 for the Australasian Musculoskeletal Medicine Conference. The topic of these presentations was, Comparison of usual surgical advice versus a non-aggravating six-month gym-based exercise rehabilitation programme post-lumbar discectomy. Donaldson, B.L., Shipton, E.A., Inglis, G., Rivett, D., & Frampton, C.

The one-year follow-up results were presented at The Australasian Spine Society conference, Auckland New Zealand, April 2005 and The New Zealand Orthopaedic Association biennial conference, Christchurch New Zealand, October 2005.

The three-year follow-up results were presented at The New Zealand Manipulative Therapists Association biennial conference, Rotorua New Zealand, August 2007; The New Zealand Orthopaedic Association biennial conference, Auckland New Zealand, October 2007; The Australasian Spine Society, Adelaide Australia, April 2008 and The Australasian Musculoskeletal Medicine conference, Melbourne Australia, October 2008 (Keynote presentation and workshop leader).

Prizes

This study won the prize for ‘best presentation’ at the New Zealand Manipulative Therapists Association conference in Rotorua New Zealand, 2007.

Study Grants

Study grants awarded for this research totalled $25,000. Five institutions supported this study. These institutions were; The New Zealand Accident Compensation Corporation, The New Zealand Society of Physiotherapists, The New Zealand Manipulative Therapists Association, The University of Otago Christchurch Campus and the Wishbone trust.

ACKNOWLEDGEMENTS

Whilst the undertaking of this thesis has been a huge learning experience, it has also been an extremely enjoyable journey. I am indebted to the team of people who have supported and worked with me through this process. It is the skill, knowledge and experience of these people that has made this journey a pleasurable one.

This team begins with Professor Ted Shipton. Some thesis candidates complain of poor contact with their supervisors; in the case of 'the Prof.', he was often the one questioning when the next piece of work would be ready for inspection. It was not uncommon for a significant passage of work to be sent to the Prof., only for it to have been thoroughly marked, re-worked and back in my inbox in 48 hours. The Prof.'s experience and knowledge has been invaluable.

Professor Darren Rivett was the physiotherapy supervisor for the team. Darren has been fantastic at keeping me on task as far as structure, method and content are concerned. I really appreciate Darren's experience and advice, always given so diplomatically when sometimes he must have been wondering what I was thinking.

Associate Professor Chris Frampton is simply a genius in the world of biostatistics. He is always a pleasure to work with and have a chat about rugby. I thoroughly appreciate how he expeditiously dealt with questions even when it must have been inconvenient for him. When preparing work for presentations or publications there have been many tight deadlines; Chris always came through.

Mr. Grahame Inglis is the Orthopaedic Surgeon from whose operating list the participants came. Grahame does not know the extent to which he has assisted in this study. Grahame was instrumental in the initial stages when the study method was designed, and has been a strong constant support all the way through. It was Grahame who encouraged me to publish the first year results and present these, and further results, at surgical conferences. Original research can be a lonely place at times and when the going got tough, a visit to Grahame was like a dose of adrenalin.

Finally, Dr Ruth Helms, Manager of Academic Programmes at University of Otago, Christchurch Campus, must be thanked. I call Ruth, Dr 'Can Do'. Being a physiotherapist in a medical programme there were issues with finding supervisors and departments. The study

also grew a bit big for a Masters and this caused further issues. All the way through, Ruth simply made things happen. I strongly recommend the University of Otago, Christchurch Campus as a place to complete research.

And finally to my wife Kirsty who stepped up to the mark when she was really needed. Thanks for holding the family together when my life went on hold to complete this thesis and thanks for the hours of patient proof reading.

To all the above named people, my sincerest thanks for your help and support. My thanks also to the institutions that supported this study financially.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ADLs	Activities of daily living
CBT	Cognitive behavioural therapy
CG	Control group
CSA	cross sectional area
DB	Dumbbell
EMG	Electromyography
FRP	Functional restoration programme
GP	General Practitioner
grp	group
ICC	Intraclass correlation
Kgs	kilograms
LBP	Low back pain
Lx	Lumbar
Max	maximum
MCIC	Minimally important clinical change
MCID	Minimum clinical important difference
Med	medication
N	Number
ODI	Oswestry disability index
OLBI	Oswestry low back index outcome measure
OMH	overall mental health
OPH	overall physical health
PAP	Post-activation potentiation
PI	Principal Investigator
QoL	Quality of life
RA1	Research assistant one
RA2	Research assistant two
R.O.M.	range of movement
RCT	Randomised controlled trial
Reps	Repetitions
RM	Roland-Morris outcome measure
RM	repetition maximum
RMQ	Roland-Morris Questionnaire
RTW	Return to work
SE	standard error
SF36	Short form 36 outcome measure
SLR	Straight leg raise
SPSS	Statistical Package for Social Sciences
TG	Trial group
Tot score	Total Score
VAS	Visual analogue scale
VO2 max	Volume of oxygen consumption
Vs	versus
wks	weeks
WL26	Work limitations 26 outcome measure
WL27	Work limitations 27 outcome measure
yrs	years

CHAPTER ONE
INTRODUCTION

1.1 Statement of Purpose

The current surgical management after lumbar discectomy in New Zealand is for patients to rest for a prescribed period of time post-surgery and then return to normal function as pain allows. No formal post-surgical physiotherapy rehabilitation programmes are encouraged; in fact, spinal surgeons discourage programmes of this nature.

Although spinal surgeons do not encourage post-surgical rehabilitation programmes, many insurance case managers do. Over a number of years prior to commencing this study, the author had the opportunity to work with lumbar discectomy patients from an early stage post-surgery. The author designed a periodized gymnasium rehabilitation programme for these patients. Over a period of several years the author anecdotally noted that patients undergoing the rehabilitation programme suffered fewer long term complications when compared to patients who had no post-surgical rehabilitation. The author acknowledges clinical bias both in this interpretation and also in the patient population that is being managed. Of all lumbar discectomies performed those that have follow-up complications are in the minority. Patients that have no post-surgical complications do not present for follow-up treatment. To establish whether or not rehabilitation results in fewer post-surgical complications is the primary motivating factor for initiating this research.

Low back pain (LBP) consumes significant resources on a world-wide basis each year (Frymoyer & Cats-Baril, 1991; Schroth, Schectman, Elinsky, & Panagides, 1992; van Tulder, Koes, & Bouter, 1995; Weinstein, Lurie, Tosteson et al., 2006). Any intervention that can be demonstrated to improve the management of this condition, even in a small way, has the potential to save millions of dollars annually.

In recent years, rates of lumbar surgery have dramatically increased (Frymoyer & Cats-Baril, 1991; Gray et al., 2006; Taylor, Deyo, Cherkin, & Kreuter, 1995; Weinstein, Lurie, Tosteson et al., 2006). Lumbar discectomy is the most common surgery performed on the lumbar spine (Gray et al.), consequently it consumes the most resources (van Tulder et al., 1995). The bulk of the overall costs associated with spinal surgery are the indirect ongoing entitlements that relate to, further medical supervision, investigation for post-surgical

complications, vocational and social assistance, and weekly compensation (Cats-Baril, 1991; van Tulder et al.).

To illustrate the New Zealand experience for the purpose of this thesis, all operating theatres performing spinal surgery in Canterbury, New Zealand (an area servicing a population of 400,000 people) were surveyed. This provided complete information for the cohort of the Canterbury population. Between 2000 and 2005, the number of lumbar discectomies performed increased by just over 250% from 156 (in 2000) to 396 (in 2005). Further data from the New Zealand Accident Compensation Corporation (ACC) (personal communication February 2008, Table 2.2) demonstrated that this rate of increase is consistent with other regions in New Zealand. It should be noted that Table 2.2 only includes lumbar discectomies covered by ACC and does not include privately funded surgery or surgery performed in public hospitals.

During the period 2002 to 2005, the ACC database put annual incidence rates of LBP at 300-350/100,000 for female and 190-200/100,000 for male (personal communication, ACC database, October 2006). These data concur with other studies that demonstrate higher rates of reporting of LBP for females than males (van Tulder et al., 1995). Conversely, the surgical data revealed higher rates for males than females. This was possibly due to the fact that more males were involved in the heavy work category, which is a known high risk factor for low back surgery (Carragee et al., 1999;Donceel & Du Bois, 1999; Hasenbring, Marienfeld, Kuhlendahl, & Soyka, 1994).

The literature demonstrates an increasing rate and cost of lumbar discectomy. It also reveals the majority of costs associated with this surgery to be indirect costs after surgery. In order to optimise surgical success and minimise costs and complications after surgery, health providers and funding bodies need to know what is ‘best practice’ for the management of lumbar discectomy following surgery. The question is, do patients need rehabilitation after lumbar discectomy or are they better to follow conservative surgical advice and return to pre-injury function via self management?

The aim of this study is to determine whether a non-aggravating periodized physiotherapy rehabilitation programme provides superior long-term outcomes after lumbar discectomy when compared with usual conservative surgical advice. In this instance ‘usual conservative surgical advice’ is to return to usual functional activities as pain allows. Patients are encouraged to return to normal function as quickly as possible. There is a lack of research

in the literature on the post-surgical management of lumbar discectomy. The purpose of this study is to add to the body of knowledge on this topic.

1.2 The Null Hypothesis

There is no difference in long-term outcome (three years) after lumbar discectomy between patients managed by usual conservative surgical advice and those patients utilizing a six month non-aggravating periodized physiotherapy rehabilitation programme.

1.3 Method

To test this hypothesis, a prospective randomised controlled trial with a three-year follow-up period was designed. Two groups of participants were compared within the study; these were the control and trial groups. The control group followed surgical advice that encouraged participants to return to normal function as soon as pain allowed; no formal rehabilitation was prescribed. The trial group was supervised through a non-aggravating six-month progressive gymnasium rehabilitation programme. Participants filled out baseline stratification information and functional questionnaires at six weeks post-surgery. These participants completed a number of validated outcome measures over the next three years. Outcome measures included the Roland-Morris 24 point Questionnaire, the Oswestry Low Back Index, the SF36, and an annual questionnaire related to a number of Quality of Life variables. A more detailed description of the method is provided in Chapter 3.

CHAPTER TWO
LITERATURE REVIEW

2.1 Introduction

This chapter reviews the anatomy and mechanisms of disc injury, and summarises the literature on the prevalence of low back pain in relation to surgery (specifically to lumbar discectomy). Surgery versus conservative therapy for lumbar discectomy, and the cost to society of low back pain (with the focus on lumbar discectomy), is then discussed. Complications of fusion and non-fusion lumbar surgery are also considered; the surgical procedures for lumbar discectomy are reviewed. The cognitive management of patients post lumbar discectomy is discussed. Looking at 'predictors of poor outcome' follows cognitive management, as they are closely related. Before reviewing the relevant studies on the post-surgical management of lumbar discectomy, a section outlining the key outcome measurements and the limitations associated with these 'tools' or 'instruments' is presented. Knowledge of outcome measurements assists in understanding the methods and results of the relevant papers. Finally, the principles of exercise rehabilitation and strength training are reviewed.

2.2 The anatomy of disc injury

What follows is a review of disc anatomy and the mechanisms involved with disc herniation. A lumbar disc consists of three primary structures; the nucleus, the annulus and the cartilaginous endplates (Bogduk & Twoomey, 1991; Marchand & Ahmed, 1990; McGill, 2007). The nucleus and annulus are separate but are not distinct entities. The outer layers of the nucleus blend with the inner layers of the annulus in an area known as the transitional zone (Bogduk & Twoomey; Marchand & Ahmed). Marchand and Ahmed report that it is difficult to ascertain whether the cartilaginous endplates (lying superior and inferior to the nucleus and annulus, and connecting the nucleus and annulus to adjacent vertebrae) are part of the disc or the vertebrae. Marchand and Ahmed consider the cartilaginous endplates to be part of the disc rather than the vertebrae; this is consistent with the views of Bogduk and Twoomey (1991) and McGill (2007).

The nucleus is a gel-like substance that consists of collagen fibrils suspended in a base of water and mucopolysaccharides. This semi-fluid mass provides both viscosity and elastic response when forces are applied to the disc (McGill, 2007). The chemical content of the nucleus allows it to resist and absorb pressures by deforming and by transmitting the applied pressures in all directions (Bogduk & Twoomey, 1991).

The annulus (radial to the transition zone) consists of between 15-25 lamellae made of collagen, elastin bundles and proteoglycan gel. The number and thickness of the lamellae for an individual disc depends on the circumferential location, the age and the spinal level (Bogduk & Twoomey, 1991; Marchand & Ahmed, 1990). The lamellae are not circumferential. The annular circumference consists of a series of lamellae that overlap each other and surround the nucleus (Schollum, Robertson, & Broom, 2008). The distinct layers of lamellae form "a cross-ply structure containing parallel arrays of collagen bundles aligned oblique to the spine axis, and alternating from left to right between successive lamellae" (Hirsch & Schajowicz, 1953; Pezowicz, Robertson, & Broom, 2005. p. 299). These layers of lamellae are bound together by a complex fibrous interconnecting network (Pezowicz, Robertson, & Broom, 2006; Schollum et al.), known as a translamellae bridging network (Schollum et al.).

Studies have found that when compressive loads were applied to the disc, the first structures that suffered damage were the vertebral endplates followed by the annulus (Brinckmann, Biggemann, & Hilweg, 1988; Yoganandan et al., 1988). Fracture of the vertebral endplates led to a reduction in hydrostatic pressure within the disc, that caused outward bulging of the outer annulus and inward bulging of the inner annulus (Adams & Dolan, 1995; Pezowicz, Schechtman, Robertson, & Broom, 2006). This 'bulging' led to microclefts and tears of the lamellae that resulted in delamination of the annular lamellae (Vernon-Roberts, Fazzalari, & Manthey, 1997). This delamination allowed nuclear material to seep through the annular lamellae (Pezowicz, Schechtman et al.; Veres, Robertson, & Broom, 2008). Pezowicz et al. hypothesised that the annulus was then weakened by the delamination to the point that subsequent flexion/extension loading of the annulus would be sufficient to "drive a single tear in one preferred direction" (p. 2902).

Damage to the annulus in the form of radial fissuring and herniation-related ruptures tends to concentrate in the para-central posterior annulus (Maezawa & Muro, 1992; Osti, Vernon-Roberts, Moore, & Fraser, 1992; Vernon-Roberts et al., 1997). The reason this

distinct pattern of failure occurs within the disc is unclear. Marchand and Ahmed (1990) report that contributing factors may be the low annular wall thickness and the relatively large number of incomplete lamellae in this area of the disc. A more recent study by Veres et al. (2008) suggested that the posterior region of the disc was susceptible to rupture and herniation due to its lack of ability to disperse hydrostatic pressure. Veres et al. hypothesised that the laminate architecture relating to the narrow annular wall and a large number of incomplete lamellae in this region, was the cause of this reduced ability.

The cartilaginous endplates, commonly referred to as vertebral endplates, encapsulate the nucleus and the inner aspect of the annulus. The outer edge of the annulus is not covered by the vertebral endplate, hence it connects directly to the outer edge of the vertebral body The vertebral endplate consists of hyaline and fibrocartilage. Hyaline cartilage occurs towards the vertebral body and is most evident in young discs. Fibrocartilage occurs towards the nucleus pulposis. As discs age, the amount of hyaline cartilage is reduced and replaced by fibrocartilage. The vertebral endplate is strongly attached to the annulus and nucleus via a network of collagenous fibres. The endplate is less strongly attached to the vertebral body; this is one reason that the vertebral endplate is considered to be part of the disc rather than part of the vertebral body (Bogduk & Twoomey, 1991).

All three (annulus, nucleus, and cartilaginous endplates), if damaged, are capable of causing back pain. The outer one third of the annulus is innervated (Bogduk, Tynan, & Wilson, 1981; Palmgren, Gronblad, Virri, Kaapa, & Karaharju, 1999). Stimulation of this part of the disc via mechanical probing elicits back pain (Kuslich, Ulstrom, & Michael, 1991). Studies using magnetic resonance imaging (MRI) and computer tomography (CT) discography, have demonstrated that internal annular tears that reach the outer one third of the annulus correlate with patients' low back pain (Aprill & Bogduk, 1992; Moneta et al., 1994).

Other studies have reported the presence of high intensity signal in the annulus on MRI scan to be a normal finding, as this signal is often found in asymptomatic individuals (Ricketson, Simmons, & Hauser, 1996; Stadnik et al., 1998; Weishaupt, Zanetti, Hodler, & Boos, 1998). Weishaupt et al. state that caution needs to be applied when correlating a zone of high intensity signal on MRI scan with the presence of pain. Two studies by Ricketson et al. and Stadnik et al. did not find a statistically significant correlation between the finding of a zone of high intensity signal on MRI scan and the reporting of pain by the patient. Therefore annular protrusion exhibited on MRI scan alone, is not a good predictor of back pain (Boden, Davis, Dina, Patronas, & Wiesel, 1990; Boos et al., 1995; Jensen et al., 1994; Weishaupt et

al., 1998). However, extrusion and sequestration of the lumbar disc with nerve root compression correlates well with low back pain and sciatica (Boos et al.; Maezawa & Muro, 1992; Stadnik et al.; Weishaupt et al.).

In 1988 Brinckmann, Biggemann and Hilweg published a paper that described seven different macroscopic classifications of endplate fracture. In the same year, Modic, Steinberg, Ross, Masaryk and Carter (1988) published a paper describing endplate fracture as a three-stage process; these stages are known as Modic changes type I, II, and III. The respective lesions on MRI scan are characterised as follows: Type I, low T2 and high T2 signals are associated with fissuring of the cartilaginous endplate and increased vascularity within the subchondral bone marrow; Type II, high T1 and T2 signals reflect fatty degeneration of the bone marrow; Type III, low T1 and T2 signals correlate with subchondral bone sclerosis (Kuisma et al., 2007). The Modic classification system is commonly reported in the literature.

Making a diagnosis of endplate fracture as the primary cause of low back pain is difficult due to endplate fractures often appearing in the presence of other spinal column pathology. Hence, it is difficult to determine the exact structure that is the primary source of pain. McGill (2007) states that “endplate fracture is a very common compressive injury and perhaps the most misdiagnosed” (p. 41).

Two studies have endeavoured to correlate endplate changes with low back pain and sciatica (Kuisma et al., 2007; Kjaer, Leboeuf-Yde, Korsholm, Sorenson & Bendix, 2005). Kuisma et al. investigated the associations of the frequency and intensity of low back pain and sciatica with Modic changes in two groups of middle-aged men. One group was associated with a high degree of occupational vibration (train drivers); the other group had sedentary occupations. The prevalence of Modic changes was similar in both groups even though the train drivers reported (on average) a higher level of sciatica. The study concluded that “Modic changes at the L5-S1 level, and Modic Type I lesions are more likely to be associated with pain symptoms than other types of Modic changes, or changes located at other levels” (p. 1121).

The study by Kjaer et al., (2005) used a sample population of 40-year-olds to look at the clinical relevance of abnormal MRI scan findings, their prevalence, and association with low back pain. They concluded that “most degenerative disc abnormalities were moderately associated with LBP” (p. 1173). However, the strongest associations were noted for Modic changes and anterolisthesis. Kuisma et al. (2007) reported a higher prevalence of Modic

changes in their sample population compared to those of Kjaer et al. The sample population used by Kuisma et al. was males only; the Kjaer study used both males and females. A recent study reported Modic changes to be more significantly related to the male population (Karchevsky et al., 2005). These studies demonstrate that Modic changes are commonly associated with symptomatic LBP, and illustrate the difficulty associated with clinically diagnosing the primary cause of LBP. This research supports the statement by McGill (2007), who reported that vertebral endplate injury may frequently be misdiagnosed.

Both genetic and mechanical factors play a role with annular herniation leading to extrusion and sequestration (Battie, Videman et al., 1995; Boos et al., 1995; Kelsey et al., 1984; Mundt et al., 1993). Battie et al. studied 115 identical male twin pairs, aged between 35 – 69 years, who all underwent MRI scanning. Participants were selected according to occupational material handling, sedentary work, exercise participation, vehicular vibration, and smoking. The disc degeneration scores revealed that 75% of disc degeneration in the upper lumbar spine and 50% of disc degeneration in the lower lumbar spine could be attributed to genetic and early environmental influences. Work type, exercise participation, vibration and smoking were found to have relatively little influence on levels of degeneration. The authors noted that a significant confounding factor for this study was the retrospective nature that required the participants to remember aspects of their past employment. However, the authors reported that despite this confounding factor, the results suggested that genetics do have a role in lifetime degenerative disc change.

Battie et al. (Battie, Haynor et al., 1995) performed another study in which a ‘blinded’ independent observer read a series of MRI scans from forty twin pairs. The observer was able to match the twin pairs by their scans, providing further evidence for the genetic role in lifetime disc degeneration. However, twin pairs are likely to have experienced similar environmental factors. Lifetime degeneration attributed to genetics or early environmental factors is a topic for future research.

To assist with the classification of mechanical risk factors that lead to annular tears or herniation of discs, in vitro studies have endeavoured to define specific loading combinations (Adams, Freeman, Morrison, Nelson, & Dolan, 2000; Adams & Hutton, 1982, 1983; Callaghan & McGill, 2001; Gordon et al., 1991; McNally, Adams, & Goodship, 1993; Simunic, Robertson, & Broom, 2004). The combined results of these studies demonstrated that compression, flexion/extension and axial rotation of the spine all played a part in causing these conditions.

Relating laboratory research to functional activities of everyday life indicates that tasks incorporating lumbar compression, flexion/extension and twisting motions could compromise the lumbar spine. The literature supports this concept. A case-control study by Kelsey et al. (1984) found that individuals in jobs that required lifting objects of more than 11.3 kilograms, on average 25 times or more per day, had a threefold risk of promoting lumbar disc prolapse. The risk was increased if the body was twisted during the lift or if the knees were not bent. Lifting objects of less than 11.3 kilograms or twisting in the horizontal plane without flexing, incurred no extra risk of lumbar disc prolapse. A similar study by Mundt et al (1993) that considered non-occupational risk factors (lifting children or other objects of 11.3 kilograms or more) found similar results to the Kelsey study. Mundt et al. also found positive correlations of increased risk of lumbar disc prolapse when the body was twisted during lifting with arms extended. This was consistent with the findings of McGill and Norman (1987), who reported that lifting with a neutral spine (compared to a bent spine) significantly reduced the shear forces through the L4-5 segment (McGill, 1997; McGill & Norman, 1987).

In addition to the effects that load and posture have on applying force through the spine, another factor to consider is the velocity of movement required to perform a task. Velocity of movement for the lumbar spine is measured in degrees per second. Fathallah, Marras and Parnianpour (1998) studied a population of workers involved in a range of occupations, grouped according to complexity and speed of movements. Injury records of corresponding jobs were retrospectively analysed. The authors found that the high and medium injury risk groups tended to be involved in jobs that required higher amounts of complex movements performed at speed (especially at end range flexion); the low risk group did not perform these tasks. The research by Fathallah et al. suggests that assessment of movement kinematics is important when formulating risk-management guidelines for specific tasks relating to low back disorders.

McGill (2007) summarised the functional risk factors for low back problems as follows: static work postures, seated work postures, frequent bending and twisting, lifting, pulling, and pushing, and vibration (especially seated). Almost every occupation from heavy to sedentary work includes at least one of these postures and/or movement patterns. Risk factors identified by McGill are supported by the literature. A summary of the literature reviewed on this topic suggests that too much of any one posture, movement or force increases the risk of low back disability. Moderation of movement, force and posture remains the best method of reducing risk of injury to the low back.

2.3 Prevalence of surgery for low back pain

Low back pain has a high incidence in the western world and absorbs significant resources. Lifetime prevalence rates of 80% and annual incidence rates of 40% are widely reported in the literature (Frymoyer, 1988; Santos-Eggimann, Wietlisbach, Rickenbach, Paccaud, & Gutzwiller, 2000; Taylor et al., 1995; Waxman, Tennant, & Helliwell, 2000). Some authors suggest that the incidence of LBP in society has always been high (Cats-Baril, 1991; Frymoyer & Cats-Baril, 1991). They explain the recent explosion in the reporting and treatment of LBP to be due to the perception that western medicine has become more successful in managing it.

The United States of America (USA) has the highest rate of lumbar surgery per capita in the world (Cherkin, Deyo, Loeser, Bush, & Waddell, 1994). For this reason many studies have focussed on lumbar surgery rates in the United States (Cherkin et al.; Frymoyer & Cats-Baril, 1991; Gray et al., 2006; Taylor et al., 1995; Weinstein, Lurie, Olson, Bronner, & Fisher, 2006). Cherkin et al. compared international back surgery rates with those of the United States. The international countries/regions in this study included Scotland, England, Manitoba, Sweden, New Zealand, Australia, Ontario, Norway, Finland, Denmark and the Netherlands. The results of this study are summarised in *Figure 2.1*.

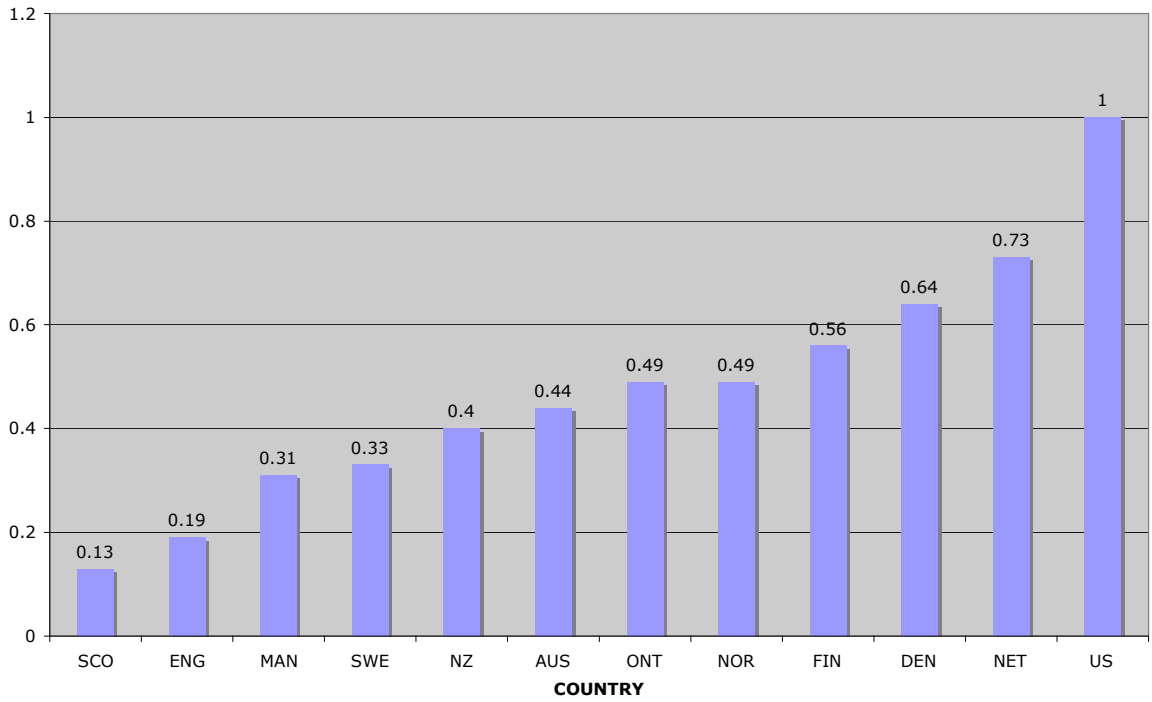


Figure 2.1. Ratios of back surgery rates in selected regions and countries to the back surgery rate in the United States (1988 - 1989): Cherkin et al., 1994. (Reproduced with permission from the author)

Frymoyer and Cats-Baril (1991) studied United States of America social security disability insurance awards from 1957 to 1976. LBP was compared with other chronic disabling diseases. From 1957 – 1976, hospitalisation rates increased as follows: heart disease 300%, schizophrenia 200%, lung cancer 500%, all causes category 250%, and LBP 2800%. The estimated cost of LBP in 1991 in the United States was 50 billion dollars.

Taylor, Deyo, Cherkin and Kreuter (1995) later measured hospitalisation rates for LBP in the United States from 1979 - 1990. The data was divided into non-surgical and surgical admissions. In 1979 the rate of surgical low back admissions was 102 per 100,000 capita. By 1990 the rate had climbed to 158 per 100,000, an increase of 55%. Of these surgical admissions, non-fusion rates climbed from 89 to 131 per 100,000, an increase of 47%. Fusion rates went from 13 to 26 per 100,000, an increase of 100%. The rate of fusion versus non-fusion was approximately 1:6, demonstrating the preference for non-fusion low back surgery. These statistics were further supported by Gray et al. (2006), who reported that lumbar discectomy accounts for 70-90% of all lumbar surgery.

Weinstein, Lurie, Olson et al., (2006) reported the lumbar spine surgical trends of Medicare beneficiaries over the age of 65 in the United States between 1992 and 2003. Rates

of lumbar discectomy/laminectomy and fusion all increased. The rate of surgical fusion rose significantly more than other options in this age group. Discectomy/laminectomy increased by 24% (from 1.7 per 1000 in 1992, to 2.1 per 1000 in 2003). Surgical fusion increased by nearly 400% in the same time period (from 0.3 per 1000 to 1.1 per 1000).

Weinstein et al. (2006) also found wide geographical variation in surgical rates. Rates of discectomy/laminectomy among Medicare patients varied almost eightfold between different states of America; fusion rates showed a 20-fold geographical variation. This significant variation could not be explained. However, the authors noted a marked increase in the fusion rate with the introduction of a particular device involved in surgical fusion. It was speculated that surgeons felt this device would provide superior outcomes. However, Weinstein et al. commented that randomised controlled trials should be performed prior to widespread implementation of new surgical techniques or devices to test whether improved post-surgical outcomes would result.

The reported rise in fusion surgery compared to discectomy/laminectomy could be related to the over-65 age group. It would be valuable to repeat this study in the under-65 age group to assess whether rates of surgical fusion have increased in this group or not. Weinstein et al. (2006) demonstrated no correlation between the numbers of orthopaedic and neurosurgeons in an area, and the number of lumbar operations performed. However, the data on ‘number of surgeons’ related to all orthopaedic and neurosurgeons performing all types of surgery. Therefore, from these data it was not possible to accurately correlate numbers of spinal operations with numbers of spinal surgeons.

The literature reports a marked increase in rates of spinal surgery. No studies have explained the reasons for this sudden popularity in spinal surgery during the past few decades. A possible explanation is that improved radiological imaging, patient selection and surgical techniques, have resulted in improved surgical outcomes. This may have led to funding providers, and the patient population, being more willing to undergo surgical options for low back pain. This being the case the question now is, when should the surgical option or the conservative therapy option be used?

2.4 Surgery compared with conservative therapy for lumbar disc prolapse

It is important to consider the evidence comparing surgery with conservative treatment for lumbar disc prolapse. A recent Cochrane review by Gibson and Waddell (2007) reported three randomised controlled trials (RCTs) comparing surgical treatment with some form of conservative treatment (natural resolution or placebo) (Buttermann, 2004; Weber, 1983; Weinstein, Tosteson et al., 2006). Another RCT by Osterman, Seitsalo, Karppinen & Malmivaara (Osterman, Seitsalo, Karppinen, & Malmivaara, 2006), compared microdiscectomy with conservative management.

Three non-prospective RCTs compared surgical discectomy with conservative management. Two were observational studies (Atlas, Keller, Chang, Deyo, & Singer, 2001; Weinstein, Lurie, Tosteson et al., 2006) and the third, a retrospective clinical audit (Saal & Saal, 1989).

Two key issues need consideration. Were the study populations comparable with similar signs and symptoms? If surgery was indicated, when was the optimum time for surgery? Some papers provided specific details on the signs and symptoms of their study populations (Buttermann, 2004; Osterman et al., 2006; Saal & Saal, 1989). Others were not as specific (Atlas et al., 2001; Weber, 1983; Weinstein, Lurie, Olson et al., 2006; Weinstein, Lurie, Tosteson et al., 2006). Gibson and Waddell (2007) reported that there was “a lack of scientific evidence on the optimal timing of surgery” (p. 1745). However, Postacchini (Postacchini, 1996) reviewed the literature on the results of surgery when compared to conservative management. Postacchini concluded “surgery should be undertaken if the patient does not respond after at least two months of conservative treatment” (p. 1383). The six papers listed by Gibson and Waddell plus the paper by Osterman et al. are discussed in the summary of this section, and summarised in Table 2.1.

Table 2.1
Summary of papers on the topic of surgery versus conservative therapy for lumbar disc prolapse.

Author & year of publication	Type of study	Sample Population	Duration of follow-up	Intervention	Outcome measures	Results
Weber 1983	RCT	126 patients with uncertain signs & symptoms for surgery	10 years	Lumbar discectomy vs. conservative management	A satisfaction questionnaire & a physical and psychological examination providing a rating of Good, Fair, Poor & Bad	An advantage for surgery up until 4 years post-surgery. This advantage then balanced between the groups from 4 to 10 years.
Weinstein et al. 2006	RCT	501 participants with confirmed radiological disc herniation & persistent signs and symptoms of radiculopathy for at least 6 weeks	2 years	Standard open discectomy vs. non-operative treatment	Primary: Medical outcome study, 36-item short form health survey bodily pain & physical function scales & modified ODI. Secondary: Sciatica Bothersomeness index, satisfaction, self-report improvement & employment status.	Advantage for surgery over non-operative treatment in the per-protocol group. Crossover in the intent-to-treat group blurred the results.
Butterman 1996	RCT	169 participants with a cross sectional MRI demonstrating a disc herniation greater than 25% of the spinal canal & no response to 6 weeks of conservative therapy	3 years	Lumbar discectomy vs. epidural injection	Self assessment VAS for low back pain & leg pain, ODI & neurological examination	A strong advantage for surgery over epidural injection
Osterman et al. 2006	RCT	56 participants with disc extrusion or sequester on CT scan, leg pain of 6 - 12 weeks duration & at least one neurological sign relating to weakness, tendon reflex or sensation	2 years	Microdiscectomy vs. conservative management	Leg & back pain VAS, a work disability VAS, ODI, a generic health-related quality of life measure 15D, Satisfaction VAS & clinical examination	No statistical difference between the groups
Weinstein et al. 2006	Observational trial	743 participants with lumbar disc herniation who self-selected their own treatment option	2 years	Standard open discectomy vs. non-operative treatment	Primary: Medical outcome study 36-item short form health survey bodily pain & physical function scales & modified ODI. Secondary: Sciatica Bothersomeness Index, satisfaction, self report improvement & employment status	A statistical advantage for the surgical option
Atlas et al. 2001	Observational trial	402 patients with lumbar disc herniation who self-selected their own treatment option	5 years	Standard open discectomy vs. non-operative treatment	Self report leg & LBP, functional status, satisfaction, employment & compensation	Results favoured the surgical cohort
Saal & Saal 1989	Retrospective clinical trial	64 patients with leg pain, SLR less than 60 degrees, CT scan demonstrating a significant disc bulge & EMG demonstrating radiculopathy	30 months after discharge	A physical rehabilitation programme	A questionnaire relating to levels of activity, pain work status & further medical care	No statistical difference between patients with neurological weakness or extruded discs when compared with the total study population.

Note: Abbreviations: ODI = Oswestry Disability Index; LBP = Low back pain; CT = Computer tomography; EMG = electromyography; SLR = straight leg raise; RCT = randomised controlled trial; VAS = visual analogue scale; vs. = versus; MRI = Magnetic Resonance Imaging

Weber (1983) conducted a prospective longitudinal randomised controlled trial on 280 patients with back pain and sciatica. This cohort was separated into three groups as follows: the first group of 67 patients required surgery due to the severity of their signs and symptoms; a second group of 87 patients were treated conservatively because of the lack of indication for surgical intervention; and a third group of 126 patients, whose signs and symptoms provided uncertain indications for surgery or conservative therapy.

The group of 126 participants with objective signs and symptoms of disc herniation (including positive radiography) were randomised to receive either surgical discectomy or conservative physiotherapy. This group was followed up for ten years. During the first year, participants filled out functional questionnaires at 3, 6, 9 and 12 months. The questionnaire was then repeated at 2, 3 and 4 years post surgery. The participants were re-examined at one-year post-surgery and then again at four years, and finally at ten years.

Weber (1983) reported an ‘advantage’ for the surgical group in the first year after surgery that gradually lessened over time until there was ‘no advantage’ at the ten year mark. At the four-year mark, the difference between the groups was not statistically significant. Weber did not report a p-value for this result but did report a trend in favour of the surgical group.

Weber’s study (1983) was credited at the time with profoundly changing the philosophical paradigm regarding the benefits of surgery (Bessette, Liang, Lew, & Weinstein, 1996). Bessette et al. re-analysed Weber’s study, stating that it was a landmark paper at the time. However, Bessette et al. also reported that if modern criteria relating to methodological design of RCTs were applied, a number of methodological flaws in the study were noted. They stated that caution is required when interpreting these results.

Some of the methodological issues that Bessette et al. (1996) described were, baseline participant demographics were not well documented and the two groups were not meaningfully compared. The assessor was not ‘blinded’. Co-interventions or other therapies that the subjects may have undergone were not mentioned. The key assessment measure was vague and subject to bias by both participant and assessor. There were a number of crossovers where participants in the conservative-care group had surgery. Seventeen (26%) of the conservative therapy group had surgery in the first year post-randomisation. The number of crossovers compromised the power of the statistical analysis.

Similarly, Gibson and Waddell (2007) critiqued Weber's (1983) study and concluded that Weber had made an inaccurate interpretation of the results, in deciding that conservative therapy was as effective in the long term as surgical discectomy. According to Weber, approximately 60% of the conservatively treated participants reported final results as 'Good' or 'Fair'. From this, Weber extrapolated that potentially 60% of the participants receiving surgery, did not need surgery. Due to the methodological issues identified by Bessette et al. (1996) and Gibson and Waddell, this should not be concluded. However, Weber did not actually state that conservative therapy produced equivalent or better results than surgery. He stated: "the results of surgical treatment were significantly better than the results in the conservatively treated group after one year of observation. During the following nine years of observation, this difference became less apparent" (Weber, 1983. p. 139).

Weinstein, Tosteson, Lurie, Tosteson et al. (2006) conducted a RCT comparing surgery with non-operative treatment for lumbar disc herniation. Patients with lumbo-sacral radicular syndrome (n = 1244) in 13 multidisciplinary spine clinics from eleven states of America were enrolled in a prospective study. Seven hundred and forty three patients opted to choose their own form of treatment but agreed to participate in an observational cohort. The other 501 patients agreed to enter a prospective RCT where they were assigned to either surgical discectomy or conservative management. Two hundred and forty five were assigned to surgery and 256 to conservative care. All participants were followed up for two years. The primary outcome measures used were the SF-36 bodily pain and physical functioning scales, and the modified Oswestry Disability Index. Secondary outcomes included the Sciatica Bothersomeness Index, satisfaction with symptoms, self-reported improvement, and employment status. There was a strong crossover trend in this study, with only 140 participants in the surgical group actually having surgery and 107 participants in the non-surgical group eventually having surgery. This led to a blurring of the intent-to-treat analysis.

Results of the intent-to-treat analysis were not statistically significant. However due to the significant crossover effect for both groups, the intent-to-treat analysis was somewhat annulled. Per-protocol analysis of the outcome measures demonstrated a statistically significant advantage for the surgical group up until two years post-surgery.

Strengths of this study included its multicentre design that covered a wide geographical area and involved a number of different surgeons. A weakness was the dilution of the intent-to-treat analysis due to the number of participants in the conservative treatment group who had surgery. The authors appropriately pointed out the weakness of the per-protocol analysis

versus the intent-to-treat analysis. However, the per-protocol analysis is clinically relevant. Clinical decisions are made on a case-by-case basis; many different variables contribute to the final decision. In the clinical setting, treatment options are not decided by a process of randomisation. Therefore the per-protocol results of this study are clinically relevant and should be considered when deciding between surgical and conservative treatment options.

Buttermann (1996) investigated the treatment of lumbar disc herniation by comparing epidural steroid injection with surgical discectomy. All 169 patients who entered the study underwent a period of six weeks conservative pain management and physical treatment. If at the end of the six-week conservative-management period the patient still had significant problems, they were randomised to either a surgical discectomy group or to an epidural steroid-injection group. A key factor in the inclusion criteria was that the disc herniation had to have a cross-sectional area greater than 25% of the spinal canal on the MRI scan. Therefore all of these patients had significant disc herniation.

Statistical power required the discectomy and the epidural injection groups to have 50 participants per group; recruitment continued until this had been achieved. This eventuated in a total of 169 participants. By the end of the six-week conservative-management period, 69 patients had improved to the point where intervention therapy was not required. The study followed three groups of patients (a surgical group, an epidural-injection group and a conservative-therapy group). All patients were followed up for three years. The follow-up measures were as follows: a self-assessment questionnaire (including a visual analogue scale for low back pain and leg pain), an Oswestry Disability Index, and a neurological examination. The results of this study demonstrated the benefit of surgery compared to epidural injection. During the follow-up periods, 92% - 98% of surgical patients reported rapid symptom loss and satisfaction compared to 42% - 56% of patients who had epidural injections. Twenty-seven patients in the epidural injection group eventually had surgery.

This study highlighted some interesting findings. Firstly, conservative management was successful in 69 patients with large disc herniation. These patients did not require further intervention. Clinical characteristics of patients who responded to epidural injection were that they had lower disability scores on the Oswestry Disability Index, tended to be older and were more likely to have a sequestered disc or an extruded disc herniation. The patients who failed to respond to the epidural injection usually had a poorly hydrated herniated disc on MRI scan. A shortcoming of this study was the lack of work status assessment. This is an important

factor in today's economic environment where decisions to operate or not are often decided on an economic basis.

Osterman et al. (2006) conducted a prospective RCT on 56 patients who had the following: disc extrusion or sequestration on CT scan, radiating leg pain between 6-12 weeks duration, and at least one neurological positive test (muscle weakness, tendon reflex, sensory change or limited straight leg raise). These participants were randomised to either conservative management or to lumbar discectomy. Patients with severe intolerable pain were not included on ethical grounds; this may have resulted in a biased cohort. However, this 'middle ground' group of patients is often the group in which the decision for surgery is most difficult.

The patients underwent a battery of tests, with leg pain intensity measured by a Visual Analogue Scale (VAS) being the primary outcome measure. Patients were assessed and outcomes recorded at the initial interview and at intervals over a two-year period. It was found that surgery provided a slight advantage in the early stages (up to one year). There was no statistical advantage at two years after surgery. The only statistically significant outcomes were leg pain at 6 weeks, and treatment satisfaction at 6 weeks, 6 months and 2 years; these all favoured the surgical group.

Closer analysis of the data shows that whilst statistical significance was not achieved in all other categories, the results favoured the surgical group. Eleven of the 28 control group patients swapped to the surgical group, which biased the intent-to-treat results. To counter this, a per protocol analysis was performed comparing all surgical cases versus conservative treatment cases; again no statistical benefit for the surgical group was demonstrated. However, for this per-protocol group the question remains: if the crossover cases had not been operated on, would their ongoing outcome measures have demonstrated a statistical difference between the groups? Furthermore, removing 11 of the initial conservative-treatment patients and adding these patients to the surgical group, caused a numerical imbalance between the groups (39 versus 17).

The power of the study required 56 patients for a clinically significant difference of 15 mm on the leg pain intensity VAS (28 patients in each group). At one and two years' follow-up, 41 and 50 patients respectively, returned their outcome measures. This reduced the number of completed outcomes and significantly diminished the power of the study. A weakness was that the study did not report on the six drop-outs. Six out of 56 is 12% of the

group, the loss of results from these drop-outs could have potentially affected the overall study results. The study used a VAS for participants to rate their own work ability; this form of outcome measure is open to bias. The study outcomes did not include work function, time off work, or return-to-work figures these data would have enhanced the study results. Osterman et al. (2006) concluded benefit from surgery to be rather modest during the first 2 years. They stated, “conservative therapy is a reasonable option for these patients” (Osterman et al. 2006. p. 2413). However, the authors themselves state that these results should be “interpreted with caution” (p. 2413).

The prospective observational study undertaken by Weinstein, Lurie, Tosteson, Skinner et al. (2006), reported the results of the 743 patients who chose to select their own therapeutic option. These patients consented to participate in a prospective observational cohort for the following two years. Five hundred and twenty eight patients opted for surgery and 191 for conservative therapy. On an intent-to-treat analysis, surgery showed advantages over conservative therapy in all outcome measures. According to the authors, the patients who entered the study did so because they were more likely to want surgery, and had greater disability scores in the outcome measures. Hence, these patients were generally more disabled than the patients who entered the prospective RCT conducted by Weinstein, Tosteson, Lurie, Tosteson et al. (2006). The authors pointed out the design deficiencies of an observational cohort. However, this study was a multicentre trial that involved a number of different surgeons over a wide geographical area. It could be argued that an observational trial such as this reflects the clinical scenario more closely than a RCT, as patients were able to select their own course of treatment. This concept is similar to per-protocol analysis where results are not distorted by crossover treatments.

The Maine Lumbar Spine study (Atlas et al., 2001), like the Weinstein, Lurie, Tosteson Skinner et al. (2006) observational cohort, was a prospective observational study. Final analysis included results from 402 patients, 220 of whom were treated surgically and 182 of whom opted for conservative care. These patients were followed up for five years. Atlas et al. reported that the surgically treated patients achieved better outcomes when compared to the conservatively treated patients. Improvement in their predominant symptoms (back and leg pain) at five years favoured the surgical group (70% vs. 56%). Similarly, a larger proportion of surgical patients compared to conservative-care patients reported satisfaction with their current status at the five-year follow up point (63% vs. 46%).

In reporting the weaknesses of this study, the authors highlighted that the initial patient diagnosis was based on 'physician assessment' and not on definitive imaging. This was a confounding factor. However, the authors explained that the goal of the study was to focus on outcomes in usual clinical care, and not to alter physician practice or decision making. The results of this study related to the 'real world scenario'.

Saal and Saal (1989) conducted a retrospective clinical audit on a group of patients who presented with a herniated lumbar disc and who had been treated at their clinic. Sixty-four patients met the inclusion criteria, and were sent questionnaires at approximately 30 months post-discharge; 58 people returned the questionnaires. There was no control or comparison group. Six patients required surgery. Of the remaining 52 patients, 50 reported 'Good' or 'Excellent' outcomes. Forty-eight patients returned to work (83% of the total group). Eighty-five percent (N=44) of the 52 non-surgical participants returned to their previous employment. The paper did not report the number of surgical candidates that returned to their pre-injury employment. Average sick-leave time was 3.8 (+/-1.0) months; 26 patients reported sick leave of one week.

Saal and Saal (1989. p. 434) made a number of conclusions, namely:

- *Lumbar herniated discs can be treated non-operatively with a high degree of success*.
- *Failure of passive non-operative treatment is not sufficient for the decision to operate.*
- *The presence of weakness does not adversely affect the outcome of non-operative treatment, and should not be used as overwhelming evidence that surgery is necessary.*
- *The presence of disc extrusion does not adversely affect the outcome of non-operative treatment and should not be used as overwhelming evidence that surgery is necessary.*
- *The premise that operative patients fare better in the first year is contrary to the results.*
- *Failure to improve with aggressive non-operative measures suggests the presence of stenosis, and should probably warrant greater decompression than purely disc excision or nuclectomy.*
- *High surgical volumes of "simple" disc excisions or nuclectomies probably represent over-treatment in a group that carries a favourable prognosis in the short and long term by non-operative treatment.*

- *The decision to operate should be based on the patient's level of function and whether that functional level can be improved by an aggressive active rehabilitation programme, rather than on imaging studies and/or physical findings*.

The general theme of these statements is in keeping with the literature; more recent research supports many of these ideas. However, it is beyond the scope of a retrospective audit to confirm these statements.

To answer the initial question, 'to operate or not to operate,' the literature outlines sound guidelines for when surgery should be considered. These guidelines are: cases of severe, constant, unremitting low back and/or leg pain with positive neurological findings; loss of bladder/sphincter control; signs and symptoms of cauda equina syndrome in the presence of disc herniation on MRI scanning. Reasons for not considering surgery in these cases are: extenuating psychosocial issues, obesity, or other co-morbidities that contraindicate surgery. Saal (1996) described in detail a comprehensive list of positive prognostic factors for successful outcomes with lumbar disc herniation when using non-operative care. These included both objective assessment findings and subjective psychosocial factors. The factors listed by Saal now have some support in the literature.

Saal and Saal (1989) wrote this paper 20 years ago. Surgeons' philosophies in regard to surgery, compared with conservative therapy, for lumbar disc prolapse have been modified by improved radiological investigations and increased awareness of psychosocial issues. Rates of surgery have continued to increase, the reasons for which are many and varied. New Zealand is no exception to this trend. Table 2.2 demonstrates the increase in numbers of ACC lumbar discectomies completed in New Zealand between years 2000 and 2007 (personal communication ACC database February 2008). Saal and Saal questioned whether over-treatment was a factor contributing to increased surgical rates; this deserves further investigation.

Table 2.2
The number of completed ACC Lumbar Discectomy procedures for each calendar year between years 2000 – 2007, and the region of New Zealand where these operations were performed.

Region Name	2000	2001	2002	2003	2004	2005	2006	2007
1. Auckland/Northland	292	338	385	437	441	434	541	588
2. Waikato/Bay Of Plenty	132	154	158	142	199	198	231	246
3. Bays Region	59	51	47	46	51	41	54	85
4. Central/Northwest	178	176	164	170	191	191	244	237
5. Canterbury/West Coast	168	186	177	208	220	264	352	355
6. Otago/Southland	86	112	113	147	147	139	165	195
7. Unknown	1	2	2	3		1		

Reviewing the literature provided no conclusive answer to the question of which option (surgery or conservative therapy) was best practice for dealing with patients with lumbo-sacral radicular syndrome. These papers fell into three groups. Group one supported conservative therapy more than surgery (Osterman et al., 2006; Saal & Saal, 1989). Group two reported that surgery provided superior outcomes in the first year after surgery but that the differences between groups were not statistically significant at four years post-surgery (Weber, 1983). Group three supported surgery over conservative therapy (Atlas et al., 2001; Weinstein, Lurie, Tosteson et al., 2006; Weinstein, Tosteson et al., 2006). Excluding the study by Saal and Saal, the literature reported that surgery provided an early advantage (in terms of pain relief) compared to conservative therapy, and as time passed, this advantage for the surgical group gradually reduced.

The literature demonstrates a marked rise in rates of lumbar discectomy in recent years. In the current economic environment, cost effectiveness of any medical intervention needs consideration. Hence, section 2.5 reviews literature on this topic.

2.5 Cost of low back pain with particular reference to surgery

A number of studies have estimated the gross costs of low back pain and lumbar surgery. Van Tulder, Koes and Bouter (1995) performed a ‘cost-of-illness’ study of back pain in the Netherlands for the year 1991. They reported the total cost of back pain to be 1.7% of the country’s Gross Domestic Product. Van Tulder et al. found that musculo-skeletal diseases formed the fifth most expensive disease category for hospital care and were the most expensive in relation to work absenteeism and disability.

This data is supported by Badley, Rasooly and Webster (1994) who analysed data from the 1990 Ontario Health Survey. This consisted of a stratified random sample of the household-dwelling population in Ontario (based on 45,650 individuals aged 16 years and over). Muscular-skeletal disease ranked first in prevalence in the following areas: the cause of chronic health problems, long-term disabilities, and consultations with health professionals. It ranked second for restricted activity days and for the use of drugs.

Van Tulder et al. (1995) found that in the Netherlands, one third of hospital care and one half of absenteeism and disablement costs were due to LBP. Direct medical costs for LBP were found to be US$367.6 million (200 million dollars for hospitalisation). Other costs included the following: US$4.6 billion for the indirect costs to the country's labour force of LBP, US$3.1 billion for absenteeism and US$1.5 billion for disablement. Indirect costs accounted for 93% and direct medical costs for 7% (in 1991) of the total cost of LBP in the Netherlands.

Illustrating the North American experience of LBP, Schroth, Schectman, Elinsky and Panagides (1992) estimated the cost of LBP in the United States to between 8 - 13 billion dollars per annum. Between 1980 and 1985, lumbar spine surgery rates in the United States increased by 53% (Deyo & Tsui-Wu, 1987; Hoffman, Wheeler, & Deyo, 1993). At this time, an estimated 200,000 lumbar discectomies were being performed annually. This number continued to grow. In the early 1990s, more than 250,000 elective lumbar spine operations were being performed annually, with lumbar discectomy being the most common (Gray et al., 2006; Taylor et al.,1994). Cats-Baril (1991) reported that inpatient care for LBP in the United States in 1990 accounted for 33% of the total medical costs. This is significantly less than van Tulder's (1995) findings of 54% for the same cost category. There are a number of variables relating to cultural and attitudinal beliefs and political systems that may have led to the differences found in this cost category by these two different groups of authors.

When comparing homogenous groups, differences in cultural, attitudinal and political beliefs may lead to variable outcomes and prognosis in relation to low back pain. A study by Burnett et al. (2009), and another by Beemsterboer, Stewart, Groothoff and Nijhuis (2008) reported differences in outcomes related to the perception of prognosis, and to the frequency of sick leave reporting. Furthermore, a study by Ihlebaek, Hansson, Laerum, Brage and Eriksen et al.(2006), and another by van Doorn (1995), demonstrated how different political systems and remuneration systems correlated with different outcomes in relation to time off

work (after episodes of low back pain). The results of these studies are discussed in the following paragraphs.

Using 382 female physiotherapy and nursing students, Burnett et al. (2009) assessed whether country of origin, (Australia, Taiwan, Singapore) and level of education, (year two, three and four) influenced the belief systems of these students in relation to low back pain. Statistically significant differences between the groups were found in relation to country of origin and year of study.

For example, the Australian students reported lower levels of 'fear of future consequences' and of 'fear avoidance' of physical movement when compared to the Taiwanese and Singaporean students. In these same categories, year three and four physiotherapy students in the Australian group had lower scores when compared to year two students. Physiotherapy students had lower levels of 'future consequence' and 'fear avoidance' than did nursing students. These differences were explained by the fact that year two Australian physiotherapy students received no education on spinal rehabilitation, whereas year three students had undergone 52 hours of education on this topic. In the Australian nursing curriculum the management of low back pain focused more on passive treatment measures such as rest and medication. The authors queried whether this type of educational emphasis had led to poorer results in terms of consequence of injury and fear avoidance of physical activity.

Beemsterboer et al. (2008) studied two socio-economically comparable but socio-culturally different groups in the Netherlands. The two groups were from the districts of Utrecht and South Limburg, and were homogenous in terms of age and profession. South Limburg had a reportedly less sober lifestyle with poorer health perception and higher rates of reported disability than that of Utrecht (Beemsterboer et al., 2008). The results demonstrated that South Limburg participants reported more frequent and longer episodes of sick leave. The authors concluded that to ensure optimal results when applying health policy on a national basis, local cultural variation needs to be considered.

Considering the effect of political systems on the reporting of low back pain, Ihlebaek et al. (2006) looked at two homogeneous socio-economic groups that lived under different political health care systems. These groups lived close together but were separated by the national borders of Sweden and Norway. The health care system in Sweden reimbursed 80-90% of the weekly wage when people were off work with low back pain; the Norwegian

health care system reimbursed 100% of the weekly wage. The study found a higher rate of medically certified low back pain people in Norway; it questioned whether this was due to the higher rate of weekly compensation in that country.

On a similar note in relation to weekly compensation, van Doorn (1995) studied homogeneous groups of professionals in the Netherlands (self-employed dentists, veterinarians, physicians and physical therapists) with different levels of disability insurance cover. The study focussed on low back pain. The results revealed a correlation between those professionals who had higher levels of cover (or whose weekly compensation commenced early after injury) when compared to professionals with lower levels of cover (or who had to wait for their insurance cover to commence). The study results found that individuals with more comprehensive insurance cover were more likely to report low back pain.

Weinstein, Lurie, Olson, Bronner, and Fisher (2006) studied United States lumbar spine surgery trends and regional variations in the over-65 years age group between 1992 – 2003. They found that although inpatient back surgery rates more than doubled during the period of the study, inflation adjusted spending for discectomy/laminectomy was found to decrease by 10% from $342 million to $306 million. This decrease may reflect reductions in hospital-stay time. Surgeons are now more proactive in having the patient mobilise early post-surgery to facilitate an earlier discharge.

In the paper by Weinstein et al. (2006), spending on lumbar fusion increased (by 500%) from $75 million in 1993, to $485 million in 2003. In 1992, lumbar fusion accounted for 14% of the total cost of lumbar spine surgery in the over-65 age group; by 2003, this figure had grown to 47% (an increase in both volume and cost). This increase may be attributed to ‘new techniques’ or instrumentation that provides improved outcomes; this promotes confidence in the population to progress to this type of surgery. Furthermore, fusion is more often an option in the older age group; this paper studied the over-65 years population. It would be useful to repeat this study in the under-65 years population to ascertain whether the same rise in fusion surgery occurred in this age group.

2.6 Cost effectiveness of lumbar discectomy

Clinicians and funding authorities currently support anecdotal evidence that lumbar discectomy surgery is cost effective. This is evident by the increasing prevalence of lumbar discectomy in the western world, and by the fact that this surgery continues to be supported by different governmental and private funding agencies. This section discusses issues related to the cost of low back pain, particularly in relation to lumbar discectomy, as well as some of the difficulties in measuring these costs.

The difficulty in measuring cost-effectiveness is, what to measure? There are many objective and subjective variables. The objective variables are quantifiable; they include the following: hospital stay time, surgical complications, physician visits, level of medication use, time off work, re-herniation rates and re-operation rates. These issues are discussed in detail by Amick, Lerner, Rogers, Rooney and Katz (2000). The authors reviewed health-related work outcome measures, their uses and recommended some measures for the purpose of measuring interventions utilised in musculo-skeletal medicine. These points were also discussed in the van Tulder, Koes and Bouter study (1995), designed to estimate the cost of low back in the Netherlands in 1991.

Subjective variables are qualitative; these include levels of pain, levels of function (at work and at home), and the patient's perception of improvement. The measure commonly used to quantify these variables is the 'Quality-Adjusted Life-Years' (QALY).

The purpose of the QALY is an attempt to measure not just years of life saved but also the quality of those life years (Cookson, Drummond, & Weatherly, 2009; Naidoo & Wills, 2000). 'Quality' of life includes freedom from pain, discomfort, and the ability to live independently (Naidoo & Wills, 2000). QALYs are important to consider when making decisions regarding the allocation of health care resources. The drawbacks of QALYs are that they often use data that has been collected for managerial or administrative reasons. Hence, this data is not specific to an individual's personal health. Some methods of designing QALYs focus specifically on one dimension of health whilst others use a broad spectrum approach (Cookson et al.; Naidoo & Wills). This makes the comparison of QALYs difficult. For these reasons Naidoo and Wills state that "What is important then is to be specific about why you wish to measure health, and to then go on to select the most approporiate means of doing so" (p. 68).

The quantifiable objective variables can be measured in dollar terms; funding agencies largely base their decisions on this. Current literature supports the concept that surgical lumbar discectomy achieves an earlier return to work when compared to conservative therapy (Osterman et al., 2006; Weber, 1983; Weinstein, Lurie, Tosteson et al., 2006; Weinstein, Tosteson et al., 2006). This provides a financial advantage to compensate for the cost of surgery.

Measuring return to work and ability to stay at work is a complex issue (Amick et al., 2000). Three options are available when patients return to work, namely: they return to their usual pre-injury tasks; they return to alternate or light duties (and never return to their pre-injury tasks); or they return to a new occupation that places less stress on their back. Whilst these latter options may be recorded as a full return to work, they do not constitute full return to pre-injury status. Clinically this situation is often witnessed but is difficult to quantify. Returning to alternate duties or to a different occupation often carries a cost for the patients or their employers. For the patients, alternate duties or a new job may not attract an equivalent rate of pay compared to their pre-injury jobs. There are costs for the employers, when injured staff re-enter the workforce but are unable to perform their pre-injury jobs. These include costs involved with retraining the injured person, as well as recruitment and replacement costs. In the literature, detailed data on the ability to return to work and stay at work after surgery is lacking.

Two periods of analysis for time off work after lumbar surgery can be considered. These are, initial time off in the immediate post-surgery period and any further time off work after the patient has returned to work due to low back problems. Most studies include detailed data on work absenteeism and follow up patients for one year (Alaranta et al., 1986; Burke, Harms-Constas, & Aden, 1994; Carragee et al., 1999; Donceel, Du Bois, & Lahaye, 1999). The study by Carragee et al. however, only looks at return to work immediately after surgery; it does not include data on further episodes off work once the patient has returned to work.

Malter, Larson, Urban and Deyo (1996) endeavoured to quantify the cost effectiveness of lumbar discectomy in terms of QALYs. They used a formula of:

$$\text{cost effectiveness} = \text{cost} / \text{effectiveness}$$

(where cost is the incremental cost of the intervention being evaluated relative to standard therapy, and effectiveness is the incremental benefit).

Malter et al. (1996) applied this formula to two different trials involving herniated discs (Javid et al., 1983; Weber, 1983). Lumbar discectomy over conservative therapy was favoured in Weber's study; chymopapain therapy over conservative therapy was favoured in the study by Javid et al.. Malter et al. concluded that lumbar discectomy was cost effective in comparison to conservative therapy and chymopapain therapy.

The fact that patients generally lose their leg pain immediately after surgery contributes to the philosophy that lumbar discectomy is a successful and worthwhile option. The patients' leg pain is usually the main site of pain and the main cause of their functional limitation. With the sudden abolition of leg pain, patients experience marked relief and quickly become more functional. It allows patients to quickly return to work, thereby enhancing the cost effectiveness of the procedure.

When measuring the success of two comparable interventions, cost effectiveness is a crucial outcome measure. Any ongoing costs due to post-intervention complications need to be considered in the cost analysis.

2.7 Complication rates of lumbar surgery, specifically lumbar discectomy

Post-surgical complication rates of 10% - 50% for lumbar discectomy are reported in the literature (Barrios, Ahmed, Arrotegui, & Bjornsson, 1990; Carragee et al., 1999; Danielsen, Johnsen, Kibsgaard, & Hellevik, 2000; Dolan, Greenfield, Nelson, & Nelson, 2000; Donceel & Du Bois, 1999; Donceel et al., 1999; Kjellby-Wendt & Styf, 1998; Manniche, Asmussen et al., 1993b; Yorimitsu, Chiba, Toyama, & Hirabayashi, 2001). These figures represent post-surgical complications from the following categories; medical, anatomical, quality of life factors and functional. Medical complications include pneumonia, myocardial infarction, pulmonary embolus, cerebrospinal fluid leakage and infection. Anatomical complications include recurrent disc herniation and arachnoiditis. Quality of life factors consist of time off work, change of work, and recreational limitations. Functional factors consist of the inability to sit, stand, or walk for prolonged periods, as well as low back and leg pain. This section discusses the literature surrounding these issues.

In 1988, Malter, McNeney, Loeser and Deyo (1998) performed a retrospective analysis of all patients undergoing lumbar surgery for degenerative conditions in the State of Washington, USA. In this study 6376 patient notes were reviewed 5 years after surgery. Post-

surgical complication rates of 18% were reported for fusion, and 7% were reported for laminectomy and non-fusion. The re-operation rates were 18.2% for fusion, 14.6% for non-fusion, and a combined re-operation rate of 15%. The resultant data forms a valuable contribution to the body of knowledge on this topic.

Graver, Haaland, Magnaese and Loeb (1999) performed a prospective study on 122 discectomy patients with a 7-year follow up period. They reported a 6% re-operation rate with 88% satisfaction rates. Five per cent of patients were partially satisfied; seven per cent were not satisfied. Seven years post-surgery, 67% had returned to full time work, 10% had returned to part time work, and 23% had not returned to work. Detailed information regarding return to work was gathered. Due to the study extending over a seven year period, some participants went from being students or homemakers into the work force and vice versa. Graver et al. outlined strict criteria for classifying whether an individual had returned to work or not. Seven years post-surgery, 33% of the participants had not returned to work, or had only returned to part-time work; this is noteworthy in economic terms. Extrapolating these figures across the population of a country would lead to millions (if not billions) of dollars in lost work-hours. Studies have been completed that support these findings (Badley et al., 1994; Cats-Baril, 1991; Schroth et al., 1992; van Tulder et al., 1995).

The key points arising from the Graver et al. (1999) study are:

- The ‘Clinical Overall Score’ (COS) was a reliable outcome measure.
- Pre-surgical psychological distress was a predictor of poor outcome at both one year and seven years post-surgery.
- The COS improved significantly from pre-surgery to one year post-surgery but deteriorated when measured at 7 years post-surgery.

It was postulated that the deterioration of the COS at seven years was due to normal lifetime joint degeneration.

Yorimitsu, Chiba, Toyama and Hirabayashi (2001) surveyed 72 patients who had undergone discectomies a minimum of 10 years previously. They reported a re-operation rate of 12.5%. At ten years post-surgery, 74.6% of the cohort still experienced low back pain and 12.7% had severe low back pain. The ‘Japanese Orthopaedic Association Score’ was the outcome measure used. Yorimitsu et al. found that of the 12.7% patient population that

experienced severe residual LBP, the majority were less than 35 years of age but had advanced disc degeneration prior to surgery. Patients with good pre-surgery disc height were less likely to have residual LBP post-surgery but were more likely to re-herniate their discs. The authors postulated the reason for this was that younger patients were physically more active from a recreational and employment perspective, and therefore, their post-surgical backs were exposed to greater forces than more sedentary older subjects.

In the study by Weber (1983), a statistically significant correlation occurred between psychosocial stress and lack of physical activity. Physically active patients obtained more favourable results when compared to less active patients. Patients who had greater than 3 months pre-surgical sick leave reported higher levels of LBP and sciatica at 4 years post-surgery compared with patients who had less than 3 months pre-surgical sick leave.

2.8 Surgical technique

A number of surgical procedures are used for the diagnosis and treatment of disc herniation. These range from less invasive techniques such as coblation therapy and laser discectomy, to the more invasive and common procedures of microdiscectomy and open discectomy. Coblation therapy and laser discectomy are not popular due to generally poorer results and limited applications. In a Cochrane review by Gibson and Waddell (2007), it was commented that coblation therapy and laser discectomy techniques should be regarded as research techniques. The most common procedures are the more invasive techniques of microdiscectomy, percutaneous discectomy and open discectomy.

In this latter group of procedures, there is further differentiation between the less invasive endoscopic procedure and more invasive open discectomy. The open discectomy is regarded as the gold standard technique (Kambin & Brager, 1987; Kambin & Nass, 2003). It is a widely applicable technique at all lumbar disc levels and for all types of disc herniation. It has been in vogue for much longer than endoscopic techniques and is the most researched. A growing body of literature compares the results of endoscopy to open discectomy (Gibson & Waddell, 2007; Hermantin, Peters, Quartararo, & Kambin, 1999; Huang, Hsu, Li, & Cheng, 2005; Kambin & Nass).

2.8.1 Microdiscectomy versus Open Discectomy

Some studies have demonstrated an advantage for microdiscectomy over open discectomy. However, all these studies applied the technique to selected patient groups (Hermantin et al., 1999; Huang et al., 2005; Kambin & Nass, 2003). In 1999, Hermantin et al. published the results of a prospective RCT comparing microdiscectomy with open discectomy. There were strict inclusion criteria for the trial; patients had to have single level intracanalicular herniation between L2/L3, L3/L4, L4/L5 and L5/S1, not exceeding half the antero-posterior diameter of the spinal canal, and, in addition, have no central or lateral stenosis. The advantages of microdiscectomy as noted by the authors, were that patients only required local anaesthetic and were managed as day surgery patients. The open discectomy patients all spent one night in hospital due to receiving a general anaesthetic. The microdiscectomy patients used fewer narcotics post-surgery. The mean time to return to work post-surgery was 49 days for the open discectomy patients versus 27 days for the microdiscectomy patients.

An early retrospective clinical trial involving 150 cases, performed between 1982 and 1986, that compared microsurgery with open discectomy was published by Barrios, Ahmed, Arrotegui, Bjornsson and Gillstrom (1990). In this study, the advantages in favour of microsurgery were fewer incorrect level operations and shorter hospital stay time. The paper reported that microsurgery patients returned to ‘usual activities’ in half the time when compared to the open discectomy group. Statistical analysis supporting this statement was not reported. Two large tables itemised several variables for every individual patient in raw data format. The ‘return to work’ data was not mentioned in the paper.

Disadvantages for the microsurgery group included six dural tears versus four dural tears in the open discectomy group. Three microsurgery patients had re-operations within one month of initial surgery due to incomplete nucleus removal at the initial surgery. Other than these complications, almost identical success rates for both groups were reported; no other significant differences occurred in any other variables (Barrios et al., 1990).

Key points noted in this study are as follows, the surgery was completed between 1982 and 1986; 150 retrospective clinical cases (75 per group) were selected to enter the trial; exact inclusion and exclusion criteria were not described; and the open discectomy group had more patients with loss of reflexes when compared to the microsurgical group (47 in the open discectomy group, 33 in the microsurgery group). This indicated that the open discectomy group was generally more disabled. Furthermore, since the completion of this study 23 years

ago, advances in radiological scanning have virtually abolished the problem of operating on the wrong level, and hospital stay times have markedly reduced.

In a more recent study by Huang, Hsu, Li, and Cheng (2005), 22 patients were randomised (between April 2002 and June 2003) to receive either a microdiscectomy or an open discectomy. Inclusion criteria meant that patients did not have significant motor deficit or sphincter disturbance. This was a selected sample population and therefore was not necessarily representative of the general lumbar discectomy population. The accepted definition of signs and symptoms for lumbo-sacral radicular syndrome due to lumbar disc prolapse is “irradiating pain over an area of the buttocks or legs served by one or more spinal nerve roots of the lumbar vertebrae or sacrum, combined with phenomena associated with nerve root tension or neurological deficit” (Ostelo, de Vet, Waddell et al., 2004. p. 3). It is usually confirmed by radiological evidence; these symptoms usually result in motor and sphincter deficit. The sample population in the study by Huang et al. did not fit these criteria. Ostelo et al. noted that differences in inclusion criteria might partly account for the wide variation in reported success rates over different studies.

Huang et al. (2005) primarily focussed on systemic cytokines as a measure of tissue damage. As expected, the microsurgery option resulted in less tissue damage and less intra-operative blood loss than open discectomy; however, the operation time was greater in microsurgery. At the 18 months follow up, levels of satisfaction were essentially the same. Hospital stay time was longer for the open discectomy patients than for the microdiscectomy patients. No follow-up outcome measures related to work ability were reported; this is a weakness of this study.

The studies that compared microsurgery with open discectomy tended to use selected patient populations. Some authors reported microsurgery to be advantageous in certain cases or types of disc herniation. This has not been conclusively proven, particularly in relation to long-term functional outcomes. In the Cochrane review by Gibson and Waddell (2007), it is stated, “The choice of microdiscectomy or standard discectomy at present probably depends more on the training and expertise of the surgeon and the resources available, than on scientific evidence or efficacy” (p. 1745).

2.9 Predictors of poor outcome following spinal surgery

According to Hasenbring, Marienfeld, Kuhlendahl and Soyka (1994), the primary predictors of poor outcome following spinal surgery fell into three categories, biological, psychological and social. In this section, the factors involved in each category are discussed with support from the literature.

2.9.1 Biological predictors of poor outcome

Hasenbring et al. (1994) measured the predictive value of psychological, somatic and social variables leading to chronic pain in 111 discectomy patients. A combined analysis using all variables provided the most accurate likelihood ratios. Patients' pain responses at 6 weeks and 6 months post-surgery were measured; the variables were tested to predict early retirement from work at 6 months. The results demonstrated a sensitivity of 90% and a specificity of 83%, with a total of 86% correct predictions for ongoing pain at 6 weeks post-surgery. At 6 months post-surgery, the sensitivity was 87% and the specificity was 87% (with a combined ratio of 87%). The only relevant somatic predictor was disc displacement. A smaller disc displacement led to a greater chance of patients complaining of ongoing pain. A secondary finding was that patients with a variable scoliosis prior to surgery had a greater probability of suffering pain at 6 months post-surgery.

Previous injury is a well-known predictor of ongoing problems in a variety of musculo-skeletal conditions including low back pain and surgery. Three prospective studies researched the prevalence of low back injury in the community and looked at the value of psychosocial factors in predicting the reporting of ongoing pain. Previous history of LBP was found to be a strong predictor of future problems (Bigos et al., 1991; Burton, Tillotson, Main, & Hollis, 1995; Waxman et al., 2000).

Two studies looking specifically at predictive factors in relation to discectomy patients appear to contradict each other. Donceel, Du Bois and Lahay (1999) studied the effect of a rehabilitation-orientated approach in insurance medicine, using 710 post discectomy patients. Previous back operations were found to be a high risk factor for further problems. However, Graver et al. (1999) followed 122 patients for 7 years post-surgery; they found that a history of previous surgery had no influence on prognosis. This finding is unusual. Previous surgery in the lumbar spine implies a previous injury of the lumbar spine leading to that surgery;

previous injury is a high risk factor for any type of musculo-skeletal injury. Graver et al. reported that 6% of the total follow-up of 114 patients had undergone previous surgery. This is a total of seven patients; this low number of repeat surgical cases might explain their findings.

Other biological factors include a good disc height. Yorimitsu et al. (2001) found that young patients under the age of 35 years with good disc height had less residual LBP post-surgery but were more likely to re-herniate their discs. The classification of 'good disc height' (used by Yorimitsu et al.) has been defined by Fujimura, Wakano, and Hijikata (1974). The method used to define the disc height ratio (used by Yorimitsu et al.) was described by Mochida, Nishimura and Nomura (1994). The two papers by Fujimura et al. and Mochida et al. were both published in Japanese and therefore are not discussed in this thesis. Yorimitsu et al. postulated that younger patients placed greater strains on their backs from both the work and the recreational perspectives. Yorimitsu et al. are the only authors to mention this finding in the literature. This might be clinically relevant and deserves further investigation.

Studies have demonstrated that weak back muscles lead to increased stress on the osseous-ligamentous structures of the spine, resulting in early degenerative changes or instability (Cooper, St Clair-Forbes, & Jayson, 1992; Dolan & Adams, 1998; Hides, Stokes, Saide, Jull, & Cooper, 1994). Hides et al. demonstrated a segmental loss of multifidus cross sectional area (CSA) in patients with acute LBP. Multifidus is a segmental stabiliser of the lumbar spine (Cholewicki & McGill, 1996; Panjabi, 1989). If in acute back pain patients this muscle was found to be wasted, it could be assumed that lumbar discectomy patients (who have had symptoms for some time whilst waiting for surgery) would be more affected. This hypothesis is supported by Cooper et al. (1992): they demonstrated a significant difference in paraspinal and psoas CSA between acute and chronic back pain subjects. The chronic pain subjects revealed a greater loss of CSA in these muscles compared with the acute pain subjects. Further work by Hides, Richardson and Jull (1996) demonstrated that multifidus did not automatically regain its normal cross sectional area after the back pain ceased.

Lumbar fusion is fraught with a much higher rate of post-surgical complications than non-fusion surgery (Deyo, Ciol, Cherkin, Loeser, & Bigos, 1993; Donceel & Du Bois, 1998; Frymoyer, Hanley, Howe, Kuhlmann, & Matteri, 1978; Malter et al., 1998; White, von Rogov, Zucherman, & Heiden, 1987). The literature describes early and late post-surgical complications. The former occurs intra-operatively or in the immediate postoperative period. It includes dural tears, spinal level errors, nerve root damage, myocardial infarction, wound

infection, and vascular and respiratory problems (Barrios et al., 1990; Malter et al.; Pappas, Harrington, & Sonntag, 1992). The latter is the 'failed back' syndrome or failed functional outcome. This leads to more surgery, to extended time off work, and to ongoing pain and dysfunction.

In 1988, Malter et al. (1998) conducted a retrospective study on a large population in the State of Washington. Spinal surgery subjects were followed up for 5 years after surgery. Intra-operative complication rates of 18% for fusion and 7% for non-fusion surgery were found. Re-operation rates were found to be 18.2% for fusion and 14.6% for non-fusion surgery.

Barrios, Ahmed, Arrotegui and Bjornsson (1990) found that symptom relief provided by auto traction pre-surgery was an important predictor of good outcome post-surgery. However, this does not suggest that all potential lumbar discectomy patients be given traction therapy prior to surgery simply to test their prognosis after surgery.

Age and gender comprise the final biological variables described in the literature. Older subjects are prone to higher rates of complications (Donceel & Du Bois, 1998). Females are reported to have poorer surgical outcomes compared with males (Donceel & Du Bois; Graver et al., 1999).

2.9.2 Psychological predictors of poor outcome

Psychosocial status is widely reported in the literature to be the best predictor of ongoing pain (Burton et al., 1995; Graver et al., 1999; Hasenbring et al., 1994; Junge et al., 1996; Weber, 1983). Depression is the strongest psychological indicator for ongoing pain in both surgical and non-surgical cases (Hasenbring et al., 1994). Other variables include: catastrophising, fear avoidance, heightened awareness of pain, self-belief of a poor outcome in relation to ongoing pain and decreased functional ability (den Boer, Oostendorp, Beems, Munneke, & Evers, 2006).

Burton et al. (1995) conducted a prospective survey of 252 LBP patients who presented to primary care providers. The purpose of the study was to determine the relative value of clinical and psychosocial variables for early identification of patients with a poor prognosis. These patients completed a battery of psychosocial instruments and underwent clinical examination. The patients were reassessed one year after presentation. Eighteen percent

showed significant psychological distress at presentation. Discriminant models successfully allocated 76% of cases to recovered/not-recovered groups, largely on the basis of psychosocial factors evident at presentation. Psychosocial factors and cognitive beliefs were found to be seven times more accurate in predicting ongoing LBP than the usual clinical testing procedures (range of movement, straight leg raise, strength, and palpation).

Hasenbring et al. (1994) highlighted a number of psychological variables associated with the likelihood of ongoing pain. These included depression, pain coping strategies and non-verbal expressions of pain. Pain coping strategies at both ends of the spectrum were identified. These ranged from 'fear avoidance' to 'grit your teeth and work through it'. 'Going with the flow', that is, a sensible graduated return to normal function, correlated best in terms of pain management. Patients who expressed pain non-verbally, by showing pain via facial expression or bodily gestures, tended towards chronic pain. Patients who demonstrated such behaviours did so because they had learnt that this type of behaviour gained the most support from their spouse or family. Hasenbring et al. found the best instrument for predicting ongoing pain to be a combination of biological, psychological and social variables; psychosocial factors were more important than biological factors. In relation to early retirement from work, psychosocial data provided the greatest accuracy for prediction of early retirement. The addition of biological data did not improve its accuracy.

2.9.3 Social predictors of poor outcome

There are several social factors that are reported to influence post-surgical pain and function. These include, time off work prior to surgery, work place relationships, work heaviness categories, post-surgical management, long hospital stay times and differing medical and insurance systems (Dvorak, Valach, Fuhrimann, & Heim, 1988; Hasenbring et al., 1994; Ihlebaek et al., 2006; Junge et al., 1996; Schade, Semmer, Main, Hora, & Boos, 1999; van Doorn, 1995).

Work variables form the most important social variables that correlate with return to normal function, and return to work (Carragee et al., 1999; Donceel & Du Bois, 1999; Junge et al., 1996; Weber, 1983). Time off work prior to surgery is proportional to time off after surgery. The more time patients have off prior to surgery, the more time off work they are likely to require after surgery. This point is supported by data collected from New Zealand's Accident Compensation Corporation (ACC Statistics Department data between 2001 – 2006,

October 2006, personal communication). Work place harmony or friction is another key variable affecting return to work post-surgery (Bigos et al., 1991; Dvorak, Valach et al., 1988). The individual who is happy at work will have less time off post-surgery, compared to those who are unhappy at work. Work heaviness is the final variable relating to post-surgical complications. Donceel and Du Bois (1998) reported higher complication rates for patients in heavy work categories. This research is supported by other studies. These have found that patients in light work categories (particularly involving sitting or sedentary work) return to work sooner post-surgery than patients involved in heavy work (Carragee et al.; Hasenbring et al., 1994; Weber, 1983).

Pro-active post-surgical management is widely supported in the literature as being essential for a good prognosis after surgery (Carragee et al., 1999; Danielsen et al., 2000; Dolan et al., 2000; Donceel et al., 1999; Kjellby-Wendt & Styf, 1998; Manniche, Asmussen et al., 1993b; Manniche et al., 1994; Manniche, Skall et al., 1993a). Carragee et al. found a correlation between activity restriction and pain behaviour. Individuals who had limited activity levels post-surgery reported increased levels of pain. Weber (1983) found that patients who maintained a high level of physical activity post-surgery achieved superior outcomes, when compared with patients with low levels of physical activity. Weber also found that patients with psychosocial problems were less likely to be physically active. Longer hospital stay times correlated with a poor prognosis and longer time off work (Donceel & Du Bois, 1998). Another study by Donceel et al. (1999) found a correlation between poor prognosis and cigarette smoking.

The final outcome of any medical condition can be influenced by the medical system in which it is managed (Burnett et al., 2009; Ihlebaek et al., 2006). This point is frequently made in the literature. However, no controlled trials comparing different umbrella medical systems specifically related to lumbar discectomy were found. Cats-Baril (1991) reported that inpatient care costs of lumbar surgery were 33% of the total direct medical cost in the United States. Van Tulder (1995) reported inpatient costs to be 54% of the total direct medical cost in the Netherlands. Insurance and welfare systems that compensate people for being off work result in longer post-surgical rehabilitation time frames (Caspar, Campbell, Barbier, Kretschmmer, & Gotfried, 1991; Ihlebaek et al., 2006; van Doorn, 1995). A number of European countries have liberal social security systems. The post-surgical return-to-work time frames for those countries are generally greater than in countries that are not so generous (Caspar et al.; Ihlebaek et al.).

New Zealand is a case in point. During the 1970s to late 1990s, people injured by accident were covered under the national insurance agency, ACC; they were entitled to 'off work weekly compensation' for as long as they reported pain directly associated with the accident. During this time, New Zealand built a multimillion dollar annual deficit from people on weekly compensation. In the 1996-97 financial year, the cost of new back claims for ACC was $NZ35,000,000 (Preventing low back injuries: A literature review, ACC publication 1998). In recent years New Zealand's legislation has changed. Pain on its own was no longer an acceptable reason to stay on weekly compensation. Since this change in legislation, the length of post-injury time off work has reduced with a reduction in cost of claims. In the 2000-01 financial year the cost of new back claims for ACC was $NZ15,542,175 (ACC claim statistics, April 2006, personal communication). Referring to all injuries, not specifically back injuries, a recent newsletter from ACC (ACC Newsletter, April 2008, issue 110) reported that in 1998 the average time off work for an injury was more than six months, whereas in 2008 the average time off work after injury was less than three months. This is an example of the influence of non-medical factors on post-injury rehabilitation. Whilst it may be argued that improved medical management during this ten year time period may have generated this improvement, it would be surprising if a reduction of this amount (for time off work) was solely due to improvement in medical management.

To determine rates of post-surgical complications, both success and failure after surgery need to be measured. To achieve this, appropriate outcome measures are required. An understanding of the different factors involved with outcome measures is essential if accurate and reliable information is to be achieved.

2.10 Outcome Measures

Selection of outcome measures for clinical research requires consideration of the following factors (Deyo et al., 1998; Bombardier., 2000):

- The internal statistical structure of outcome measure instruments.
- The domains the instruments actually measure, and discussion of the instruments used.
- The advantages and disadvantages of hard and soft data.
- The impact of psychosocial issues on outcome measure results.

- The four primary categories measured in spinal outcome research (back specific function, pain, work disability and patient satisfaction).

This section discusses these factors and the outcome measures most commonly used in low back pain research, particularly post-surgery.

2.10.1 Internal statistical structure

According to Deyo et al. (1998), five separate statistical domains are required for internal statistical structure. These are, sensibility, feasibility, reliability/validity, reproducibility and responsiveness.

Sensibility ensures that the outcome measure or 'instrument' actually measures what it is meant to measure. It includes relevant items and excludes irrelevant items. Sensibility ensures that the instrument is suited for the study purpose, such as, individuals versus groups, qualitative versus quantitative, appropriateness for the research setting, primary care versus outpatient care, urban care versus rural, telephonic communication versus written, and suitability for the patient population, age, gender, pathology, language, and culture. Feasibility means that the instrument is easy to use, easy to understand, cost effective, and that the results are interpretable and presentable.

Reliability and validity are two terms frequently mentioned in the literature and often have different meanings or definitions. Deyo et al. (1998) describes reliability as having three different categories, internal consistency, reproducibility, and responsiveness.

Internal consistency is the ability to correlate one instrument against another. Intraclass correlation (ICC) or Chronback's alpha score are used to measure internal consistency. These scores range between 0 and 1, with 1 being the perfect correlation. Ostelo and de Vet (2005) describe this domain as the criterion validity. Examples of poor criterion validity/ internal consistency are provided by Korres, Loupassis and Stamos (1992). The 'use of analgesics' outcome measure is rated by Finneson and Cooper (1979) and by Waddel et al. (1988) as 'good'. Stauffer and Coventry (1972) and Naylor (1974) in turn label this outcome measure as 'fair.' Korres et al. describe another example of where 'modification of employment' is rated as 'good' by Barr et al. (1967), but 'fair' by both MacNab (1971) and De Palma and Prabhakar (1966). Several authors cite this as a major problem when different research projects performed on the same topic are compared (Bombardier, 2000; Bombardier, Hayden,

& Beaton, 2001; Korres et al., 1992; Ostelo & de Vet, 2005; Patrick et al., 1995; Pengel, Refshauge, & Maher, 2004).

In the literature, several different titles occur for very similar concepts. Such titles include the following, reproducibility, intra-relater reliability, reliability, agreement, inter-relater reliability, and content validity. The key point is that the author defines exactly what they mean by each concept, and are consistent in their use.

Reproducibility is sometimes reported in the literature as 'intra-relater reliability'. This domain measures the ability of the instrument to reproduce the same outcome or score at different time intervals. Ostelo and de Vet (2005) state that reproducibility includes two concepts, namely, reliability and agreement. Reliability parameters (correlation coefficients) are important for measurements between individuals. Agreement parameters (standard error of measurement) are important when measuring changes in health status. Agreement represents the lack of measurement error. Reliability represents the extent to which individuals can be distinguished from each other, despite measurement error (Ostelo & de Vet).

Deyo et al. (1998) describe inter-relater reliability as the ability of the instrument to provide equal scores in the same population when administered by different researchers. Ostelo and de Vet (2005) describe inter-relater reliability as content validity. They define this as the ability of the instrument to actually measure what it is meant to measure. For example, psychometric measures use psychometric questions, whilst functional measures test functional ability.

The final domain listed by Deyo et al. (1998) is 'responsiveness.' This is universally accepted in the literature as the ability of the instrument to detect a subtle but clinically relevant change. Responsiveness is measured by the minimally clinically important difference (MCID) or the minimally clinically important change (MCIC). Ostelo and de Vet (2005) reported that the MCID indicated the difference between patients, and that the MCIC measured any change in the individual's health status. Many studies have been performed to ascertain accurate estimates of MCICs and MCIDs. Nevertheless, Bombardier, Hayden and Beaton (2001) reported that the current literature is unable to define the specific limits of MCICs and MCIDs for any instruments; this is due to the lack of studies providing comparable validation.

Selecting an instrument that accurately measures the outcome required by an individual research topic is important. Psychometric instruments are obviously used to measure psychometric outcomes. Many instruments measure multiple categories. The SF36 includes different aspects of a single category (Ware, 2000). The Oswestry Disability Index is considered to measure function but incorporates a degree of psychometric pain testing as well. The Roland-Morris Questionnaire remains a purely functional instrument (Bombardier, 2000).

Deyo et al. (1994) have described the different types of instruments that are applicable in low back research. These include the following:

- **Physiological measures:** spinal range of movement; muscle EMG activity; spinal fluid endorphin levels; and muscle strength and endurance.
- **Anatomical measuring instruments:** measuring solid fusion mass; disc height; and vertebral displacement.
- **Complications:** report of drug side effects; neurological deficits; infections; dural tears; and cardiopulmonary complications.
- **Physical examination:** neurological deficits; and straight leg raise.
- **Health related quality of life:** symptoms of pain (duration, severity, and frequency); and neurological deficits.
- **Functional status:** activities of daily living; psychological function; recreational activities; social function; health perceptions; and well-being.
- **Role Function:** employment status; disability compensation; days of work absenteeism; and days of limited activity.
- **Costs:** direct medical costs involving hospital care; costs of surgery and repeat surgery if required; and indirect costs relating to compensation, imaging, other investigations, assistive devices, physical or occupational therapy.
- **Satisfaction with treatment:** results of treatment; treatment providers; and the fulfilment of expectations.
- **Mortality:** death due to surgery.

Instruments that specialise in measuring one specific category tend to have greater criterion validity. An instrument that produces a Chronback's alpha score of between 0.7 and 0.9 is regarded as reliable (Roland & Fairbank, 2000).

When selecting an instrument for research purposes, the researcher must weigh a variety of different factors. Some are purely logistical, such as how quickly and easily it is for the participants to fill out and return the instrument. The easier it is for the participant, the greater the compliance, and the more valid and successful the research results. Considering these issues, researchers often opt for generic instruments with satisfactory correlation scores but with superior compliance rates and information over a broad number of categories (Deyo et al., 1994).

The final question the researcher needs to consider is the population being measured. The method and instruments used must be appropriate for that population. Aspects to consider include the geographical area or spread, socio-economic factors, language, attitudes and the particular categories measured (emotional, functional, work status) (Davidson & Tolich, 2003; De Vaus, 2002).

The literature states that definitive statements about the superiority of one outcome measure over another cannot be made because of the lack of comparison of different measures across similar populations (Bombardier, 2000; Korres et al., 1992; Ostelo & de Vet, 2005; Pengel et al., 2004). Because no one particular outcome measure has been proven to provide superior results, researchers usually opt for the most commonly used instruments.

2.10.2 Outcome measure instruments

In 2000, a panel of five experts proposed a set of appropriate categories and outcome measures for use in low back pain. The results of this work are outlined in Bombardier's paper published in Spine (Bombardier, 2000). By providing such a set of guidelines for the use of specific instruments in low back pain research, greater comparison of international research becomes possible. The categories and outcome measures suggested by the authors were, back specific function (using SF36 version 2); pain (using the Bodily Pain Scale of the SF36, or the Chronic Pain Grade); work disability (using the Work Status Questionnaire); and back specific satisfaction (using the Patient Satisfaction Scale and the Global Questionnaire). These widely used outcome measures are discussed in the following section.

2.10.2.1 Back specific function and pain.

Back specific function and pain will be discussed together, as the outcome measures used to measure them often combine both categories of function and pain. Korres et al. (1992) tested 15 different evaluation methods post-discectomy. They found that different evaluations produced different results in the same population. Post-discectomy, patients reported high (subjective) levels of satisfaction but low levels of satisfaction in relation to objective functional factors.

Korres et al. (1992) explained this discrepancy between subjective and objective results. Because patients compared their post-operative pain state with their pre-operative pain levels, they reported the level of improvement rather than their actual current subjective pain status. The patients still have pain post-surgery, but at much lower levels than pre-surgery. Their objective status, however, is reported accurately, as this measures what they are able to do or not do, at that particular point in time. It was found by Korres et al. that patients often managed their problems using lifestyle adaptation. They did not return to sport, or were choosing less physical sporting options, and modifying (or completely changing) their employment status.

The Roland-Morris Questionnaire (RMQ), Oswestry Disability Index (ODI) and SF36 are all strongly supported in the literature as being reliable, repeatable and valid (Beurskens, de Vet, & Koke, 1996; Bombardier et al., 2001; Brouwer et al., 2004; Fairbank & Pynsent, 2000; Hakkinen et al., 2003; Hutchinson, Laing, Waran, Hutchinson, & Hollingworth, 2000; Korres et al., 1992; Ostelo & de Vet, 2005; Ostelo, de Vet, Knol, & van den Brandt, 2004; Patrick et al., 1995; Pengel et al., 2004; Porchet et al., 2002; Roland & Fairbank, 2000; Sigl et al., 2006; Wittink, Turk, Carr, Sukiennik, & Rogers, 2004) An important factor when using these instruments is the MCIC.

The MCIC for the Roland-Morris Questionnaire ranges from 1 - 8 points (Beurskens et al., 1996; Ostelo, de Vet, Knol et al., 2004; Roland & Fairbank, 2000; Stratford, Binkley, Riddle, & Guyatt, 1998). A study that assessed 226 patients who complained of low back pain for less than 6 weeks, found that a MCIC of 1 - 2 points was clinically significant in patients with low initial scores. For patients with high initial scores, 7 - 8 points became significant (Stratford et al., 1998). Five points was the optimal cut off. In another study by Stratford et al. (1996), the authors noted that improvement in patients with initial scores of 4 or lower, and deterioration in patients with scores of 20 or higher, could not be detected with a high degree

of confidence. It is difficult to measure improvement in patients with minimal disability, or deterioration in patients who are already significantly disabled.

Beurskens et al. (1996) performed a study similar to that of Stratford et al. (1998). The Beurskens study followed 81 patients who had complained of low back pain for at least 6 weeks. They reported a range between 2.5 - 5 points for the MCIC of the Roland-Morris Questionnaire. Patrick et al. (1995) in turn reported a MCIC between 2 - 3 points using the 23-item RMQ on patients with sciatica.

Davidson and Keating (2002) compared five low back disability instruments for responsiveness and MCIC. The instruments tested included, the RMQ, Modified ODI, SF36 physical functioning scale, Quebec Back Pain Scale, and the Waddell Disability Index. To achieve a 90% confidence interval, Davidson and Keating reported a MCIC between 10.5 - 15% for the modified ODI, 8.6 - 9.5 for the RMQ, 16 - 22 for the SF-36 physical functioning scale, and 62 - 66 for the SF36 role limitation scale. The RMQ produced an ICC of 0.53 over a period of 6 weeks (lower than other studies reported). The authors commented that the RMQ did not have sufficient scale width to detect improvement or deterioration. For the most useful functional measures, the Modified ODI, the SF36 physical functioning scale, and the Quebec Back Pain Disability Scale, were recommended.

However, the results of Davidson and Keating (2002) have been questioned by Riddle and Stratford (2002). They argued that Davidson and Keating's results were based on small numbers, particularly the examination of subgroups within their study. Riddle and Stratford explained that when working with small numbers, one or two outliers could skew results. Davidson and Keating had compared intraclass coefficients rather than confidence intervals. Riddle and Stratford argued that intraclass coefficients are a point measure. Therefore, Davidson and Keating's results could have been due to "random variation associated with making a point estimate" (p. 513). Davidson and Keating's results are at odds with the extensive evidence in the literature that supports the reliability of the Roland-Morris scale.

Specifically relating to post-discectomy patients, Ostelo et al. (2004) tested six functional status questionnaires on 105 post-discectomy patients. The outcome measures tested consisted of the following: the Roland-Morris Questionnaire, the Modified RMQ, the short form RMQ physical functioning scale, the role limitations scale of the SF36, and the Main Complaint measure. The authors reported the three Roland-Morris questionnaires to be superior to the other measures. They recommended the 24-item RMQ measure to be the most

suitable for this specific post-surgical population. This study found the MCIC with 95% probability beyond measurement error to be 5.4 points, with 3.5 points the optimal cut-off that produced a sensitivity of 94.6% and a specificity of 88.2%.

In relation to the MCIC of the Oswestry Disability Index, Ostelo and de Vet (2005) reviewed two studies published by Beursken et al. (1996) and by Hagg, Fritzell and Nordwall (2003). The Beursken study assessed the optimal cut-off for patients who had improved, and patients who had remained stable, at 4 - 6 points on the ODI scale. Ostelo and de Vet pointed out that by using this method to produce this cut-off mark, potential existed for some stable patients to be falsely classified as improved.

The second study, by Hagg et al. (2003), followed 289 low back pain patients who were treated both surgically and non-surgically. A difference of 10% was required to detect clinically relevant change with a 95% probability beyond measurement error. Bombardier et al. (2001) reported that an objective MCIC could not be determined, because of the lack of cross-referenced studies on similar populations. Current literature supports this statement (Boos, 2003; Fritz & Irrgang, 2001; Hagg et al., 2003; Hakkinen, Ylinen, Kautiainen et al., 2003; Ostelo & de Vet, 2005; Pengel et al., 2004; Sigl et al., 2006).

2.10.2.1 (a) Time intervals in relation to outcome measurement instruments

Reviewing the literature highlights some secondary information regarding outcome measures. Both Brouwer et al. (2004) and Ostelo and de Vet (2005) described the consequence of time intervals on the reproducibility of outcome measures. Brouwer et al. found that the RMQ produced an Intra Class Correlation of 0.91, with a time interval of 2 weeks between tests. However, Brouwer et al. also reported studies (Davidson & Keating, 2002; Patrick et al., 1995) that demonstrated ICCs of 0.42 and 0.66, with time intervals greater than 6 weeks. Brouwer et al. concluded that greater time intervals led to lower ICC scores and were less reliable. Von Korff, Jensen and Karoly (2000) in contrast stated that “key parameters of chronic recurrent pain have acceptable levels of validity for at least a three month recall period” (p. 3149).

There appears to be some contention in the literature as to exactly how long, the period of accurate recall actually is. There seems to be consensus that self-report measures would provide accurate information for up to six weeks, and potentially for up to three months; there

is less evidence supporting the three-month time frame. After three months, there is consensus that the reliability of self-report measures diminishes.

2.10.2.1 (b) Time intervals in relation to memory.

Time intervals are directly related to patients' memory of pain. This topic was studied by Linton and Melin (1982) using 12 chronic pain patients. These patients filled out baseline pain measures prior to entering a treatment period; the measures were repeated on discharge 3 - 11 weeks later. On discharge, the patients were asked how much pain they had at baseline. The results showed that the patients recalled markedly more pain than was rated on baseline measurement. Linton and Melin concluded by stating that, "Caution is warranted when using post-hoc pain measures with chronic pain patients" (p. 284).

Two other studies measured the ability of people to recall time off work due to different causes, the first for the previous six months and the other for the previous year. One of the causes studied was low back pain (Burdorf, Post, & Bruggeling, 1996; Severens, Mulder, Laheij, & Verbeek, 2000). Both studies found memory recall up to 2 months to be reasonably accurate. After 2 months, memory accuracy diminished significantly. The Severens study reported that 95% of the study group could remember their sick leave with complete accuracy at 4 weeks and 87% at 2 months. At 6 months however, only 57% were accurate. At one year only 51% remained accurate. The mean proportional differences of 'remembered time off' work at 2 weeks, 4 weeks, 2 months, 6 months and one year were 32.9%, 35.2%, 45.3%, 34.9%, and 113.6% respectively. These figures clearly demonstrate a wide discrepancy of memory recall. These studies cast doubt on the accuracy of questionnaires used with a follow-up time frame of more than two months.

2.10.2.1 (c) The association between pain and disability outcome measures

Hakkinen et al. (2003) surveyed 145 post-discectomy patients at two months and at 14 months post-surgery. The purpose of the study was to assess the prognostic value of some preoperative and early post-operative indicators in the prediction of disability 14 months after lumbar disc surgery. The outcome measurements used were, the VAS, Oswestry and Million disability indices, Beck depression score and work status. Hakkinen et al. found that overall the patients progressed well; yet severe pain was reported by 5% of patients at 2 months post-surgery and by 8% at 14 months. When this pain outcome data was related to the disability outcome measures, 7% of patients reported severe disability at two months and 8% reported

severe disability at 14 months. This demonstrated good correlation over the one-year period. This study suggested early post-surgical outcome to be a reliable indicator of outcome one year on. Clinically this is important, as it offers a method of predicting which patients may require early intervention, or may require more specific post-surgical management.

In patients with sciatica, Porchet et al. (2002) tested the association between pain and disability scores, and the association between pain and the radiological assessment of lumbar disc disease. They found a statistically positive linear correlation between the leg pain scales, the Roland-Morris, the Prolo disability scores, and the SF36 scores regarding physical functioning, physical role, and bodily pain. All other categories of the SF36 demonstrated no significant correlation to the severity of disc disease. Porchet et al. stated, "The positive correlation between disability status and imaging findings validates both assessment methods" (p. 1253). They concluded that disability outcome measures are useful, and add an extra dimension to the assessment of patients with sciatica.

In support of Porchet et al. (2002), Hutchinson et al. (2000) studied a cohort of first and second time lumbar discectomy patients. They assessed whether the outcome differed (was poorer) after the second surgery, than after the first surgery. Hutchinson et al. used the Roland-Morris, and two versions of the SF36 and analogue pain scales that were specifically modified for low back pain and leg pain. The study found a positive correlation between patients' pain scale reports and disability scores. Poorer scores were obtained in the group undergoing two operations. This is expected, and is confirmed by the literature (Biondi & Greenberg, 1990; Connolly, 1992; Herron, 1994; North, Ewend, Lawton, Kidd, & Piantadosi, 1991; Ozgen, Naderi, Ozek, & Pamir, 1999; Stewart & Sachs, 1996). Hutchinson reported that the SF36 revealed significantly worse general health scores in the patients undergoing two operations compared with the patients undergoing one operation. All patients had poorer scores when compared with the normal population. The Porchet study found no correlation between disc disease and the generic SF36 scores. However, Hutchinson et al. used the SF36 scale to measure general health not specific disc disease. The Hutchinson study provides support for using patient completed outcome measures in assessing outcomes in lumbar disc surgery.

In contrast to Hutchinson et al. (2000), Yen, Eaton and Maxwell (1993) found only a weak correlation between pain and impairment as did Korres et al. (1992). This weak correlation between pain and a patient's ability to function is often seen clinically. For example, a patient with an enjoyable work environment and who is well motivated, will return

to work and function efficiently, even in the face of severe pain. On the other hand, in a patient with a number of psychosocial problems, small measures of pain are often catastrophised, resulting in significant impairment and disability.

2.10.2.1 (d) Advantages and disadvantages of hard and soft outcome measures

Hard outcome measures relate to objective findings that are quantifiable. Examples of these include blood tests, radiological findings, range of movement, strength, time off work, levels of medication, numbers of operations, medical costs, etc. Soft outcome measures refer to subjective questionnaires based on quality of life factors, such as psychological function, social function, levels of pain, activities of daily living, etc (Deyo et al., 1994).

Some authors have stated that self-report questionnaires provide adequate and accurate data on a patient's physical ability (Bombardier, 2000; Deyo et al., 1994; Hakkinen et al., 2003; Lee, Simmonds, Novy, & Jones, 2001; Porchet et al., 2002). Others authors have queried these findings (Schiphorst-Preuper et al., 2008; Smeets, van Geel, Kester, & Knottnerus, 2007). Both Schiphorst-Preuper et al. and Smeets et al. reported low correlations between self-report measures and physical capacity in LBP populations.

Self-report questionnaires are open to personal bias (Deyo et al., 1994). In relation to this, Johns (1994) conducted a study where employees were asked to report their own periods of absenteeism, and whether their levels of absenteeism were more or less than the company average. In the same study, managers were asked to report whether the group of employees under their management reported more or less absenteeism compared with the company average. In both cases the employees and managers self-reported lower levels of absenteeism than actually occurred. Johns explained this inaccurate reporting as not so much a case of deceit on behalf of the employee or manager, but a perception of wanting to be better than others in their peer group.

The use of hard data only has its own limitations. The work of Yen et al. (1993) and Korres et al. (1992) demonstrated that for a number of reasons, the correlation between 'hard' data results and patient function is not always accurate. The psychological state of mind can often override measures of strength and range of motion. For example, 40% of the normal population have bulging discs on MRI scans, yet remain completely asymptomatic (Bernard, 1990; Boden et al., 1990; Jensen et al., 1994; Simmons, Emery, McMillin, Landa, & Kimmich, 1991). This is an example of data providing a false positive.

'Return to work' outcomes can provide hard data. However, this data can be skewed by confounding factors, such as environmental or psychosocial issues that affect how soon a patient will return to work, or in what capacity they will return.

Factors required for hard data include reliability and reproducibility. Deyo et al. (1998) reported on the inter-observer variability of lumbar spine interpretation by musculoskeletal radiologists compared with two self-administered questionnaires (sickness impact profile and medical history). The Kappa scores for the radiologists' diagnoses are as follows: any abnormality = 0.51; facet joint sclerosis = 0.33; and any narrowed discs = 0.49. The Kappa scores for the two self-administered questionnaires were as follows: sickness impact profile = 0.87 and medical history = 0.79. In this case, the so-called soft data was more reliable and more reproducible than the hard data.

2.10.2.2 Work Disability

In this section it needs to be noted that 'time off work' is reported using different wording. Some authors refer to time off work as 'sick leave', whilst others refer to weekly compensation or simply 'off work' time. In this thesis, any phrases used relate specifically to time off work due to low back pain.

Ostelo and de Vet (2005) reviewed methods of collecting work disability/work status data. They highlighted a number of pertinent factors. For example, should a person's usual working week consist of four days per week, the number of days absent should be four days not five if this person had one week off work. Episodes off work can be interpreted differently. If a person was off work and returned to work for a day before taking more time off work, was this classified as one or two episodes off work? Some authors required a defined period back at work before a new episode of sick leave could be reported. Steenstra, Anema, Bongers, de Vet and van Mechelen (2003) advocated a four-week resumption of work without any sick leave to signify that the last 'off work' episode had ended.

A study published by Amick et al. (2000. p. 3152) detailed the reasons why collection of 'off work' data was important, namely:

- *To assess productivity loss in clinical trials*
- *To evaluate the effectiveness of health services*
- *To target injury and re-injury prevention programmes*

- *To evaluate the effectiveness of work re-organisation projects such as ergonomic changes*
- *To improve provider-worker and provider-safety engineer interaction*.

Amick et al. (2000) outlined the type of baseline work data that needed to be collected, namely; was the person employed at their usual job, part time or full time? Were they performing their usual job on a full time basis, or on a graduated return-to-work basis? Did a graduated return-to-work involve part-time full duties and/or part-time light duties? Was the person on paid leave or unpaid leave? Had the person returned to a different job due to the injury? Was the person off work, or had they re-entered the work place in a different capacity due to this injury or other health problems? Was the person capable of working at full production, or were they working below par or on different quotas? A person not in paid employment could be retired, a student, a homemaker, or on a disability allowance for another health problem. Outcome measures would have to assess when (and how) that person regained full pre-injury function.

A list of the WL26 items is in appendix A in the article by Amick et al. (2000). Professor Amick stated that the WL26 has since been upgraded to the WL27 and is now his preferred option (B. C. Amick, November 2007, personal communication).

The study by Amick et al. (2000) highlighted the political aspect of work outcome measures. They argued that research on work outcomes could not be separated from politics. Work outcome data had the potential to be used in different ways, depending on the political context. The studies by Ihlebaek et al. (2006) and van Doorn et al. (1995) discussed in section 2.5 provide examples of how different political systems influence work outcome data. Cultural mores and governmental policy can often be confounding factors when trying to compare research and interventions across borders (Beemsterboer et al., 2008; Burnett et al., 2009).

Ostelo and de Vet (2005) stated that ‘time off work’ recorded with company or occupational health registries, provided the most reliable and valid data, as this data was recorded prospectively. Self-report questionnaires and interviews could provide useful data as well. A study by Fredriksson, Toomingas, Torgen, Thorbjornssen and Kilbom (1998) compared self-report sick leave due to musculo-skeletal reasons with registered sick leave. They found that the validity of retrospectively collected self-report sick-leave data could be used as a measure of musculo-skeletal morbidity in analysing work related conditions.

In contrast to Fredriksson et al.(1998), a study by Dasinger, Krause, Deegan, Brand and Rudolf (1999) compared administrative data with self report data in 433 low back injury patients who had been injured 1 to 3.7 years previously. Gross differences were found between the self-report measures, and the administrative insurance data. Administrative data consistently under-reported the amount of sick leave reported by the workers. They conceded that in America, where workers compensation can be a contentious and emotional issue, there could be a tendency for workers to overestimate their sick leave. They felt that their study would have been enhanced by collecting 'third party' data, for example, from doctors or from an independent source involved in the process. By doing this, the self-report data or the administrative data could have been verified. This would have enhanced the results. Their study therefore questioned the accuracy of administrative data for 'off-work time' in LBP patients. Prior to this study, administrative data was thought to be highly accurate. Dasinger et al. concluded that further research was required to validate administrative sick leave data.

The overestimation of reported sick leave by employees in the study by Dasinger et al. (1999) contrasted with the underestimation of sick leave by employees in a study by Johns (1994). The Dasinger study took place in America; the study by Johns was based in Canada. Differences between employee groups reporting sick leave may be either due to different cultures or political systems in their respective countries, or due to different research method.

In countries that provide 'sick leave' compensation for workers, the person signing the sick leave form and his/her attitude towards the worker being off work has been found to determine the amount of time off that workers receive (Arrelov et al., 2007; Englund, Tibblin, & Svardsudd, 2000; Swartling, Peterson, & Wahlstrom, 2007; Swartling, Alexanderson, & Wahlstrom, 2008). In clinical research, RCTs with adequate population sizes can balance these variations in attitude. Here, the investigator attempts to spread equal balance of 'doctor attitude' across any research cohort, and to minimise bias.

Ostelo and de Vet (2005) reported that 'minimally clinically important change' has not been defined for sick leave. This category is primarily socio-economic. The social costs of people not working were difficult to quantify, yet important in terms of other illnesses, and loss of confidence and social contact. Work absence and earlier 'return to work' were important.

Indirect medical costs for musculoskeletal injury far outweigh direct medical costs (van Tulder et al., 1995). The bulk of these indirect costs are weekly compensation. Therefore, for

any clinical trial testing a medical intervention, time off work becomes an important outcome measure.

2.10.2.3 Patient satisfaction

The final domain mentioned by Bombardier (2000), is patient satisfaction. This is a contentious area of research, with clinicians sometimes feeling threatened by patient feedback, particularly when they feel that they are being judged in areas over which they have no control (Hudak & Wright, 2000). For example, a satisfaction questionnaire may ask questions about patient parking, accessibility, availability, cost etc; all of these influence the patient's overall experience of the service, but are out of the control of the supervising clinician.

Clinicians can feel that their clinical judgement is being questioned should they receive poor results from satisfaction questionnaires (Fitzpatrick, 1991; Sitzia & Wood, 1997). In some situations, this may well be the case, particularly when the outcome has been less than desirable. The primary problem with satisfaction questionnaires is the validity of the questionnaires themselves. Ambiguity about the meaning of 'satisfaction' is the dominant weakness (Hudak & Wright, 2000). For example, some patients report a positive or negative experience with the health care system but then report opposing views in terms of satisfaction with that same health care system (Dougall, Russell, Rubin, & Ling, 2000).

Patient satisfaction, and measuring that satisfaction, are important in terms of gaining the desired outcome. Studies have demonstrated that satisfied patients participate in the treatment process, co-operate with clinicians (and the system in general), and continue to access appropriate medical services (Aharony & Strasser, 1993; Carr-Hill, 1992; Ferris et al., 1992). Dissatisfied patients, however, make treatment less effective by neglecting to seek care when needed, or by refusing to comply with prescribed treatment. An example of the latter situation is where the initial response after the procedure was excellent but the final outcome depended on further interventions. If the provider/patient relationship is frictional, the patient may not wish to seek further contact with that provider and may miss out on what could be important interventions or steps in the process of achieving a successful final outcome.

Hudak and Wright (2000) reported on the advantages and disadvantages of patient satisfaction measures. They described two major categories involved with satisfaction measures, content and method. Content is the focus or substance; method refers to how the measure is administered. Each category has four sub-categories, these are discussed in the

following sections with reference to the paper by Hudak and Wright (2000) and other supporting literature.

The four sub-categories of Content are:

- The use of global or multi-dimensional questionnaires.
- Whether the focus of the questionnaire is on the care process or the outcome.
- Whether the measure is generic or designed to be used on a specific population or disease.
- Whether the measure is direct or indirect.

The four sub-categories of Method are:

- Is the measure factual or affective?
- Are the questions open or closed?
- Is the measure self-administered or interviewer administered?
- What is the type of response format?

2.10.2.3 (a) Content: global versus multi-dimensional questionnaires

A global questionnaire measures the general or overall satisfaction with a service. Global questionnaires usually consist of one or two questions, for example, 'how satisfied were you with your outcome after surgery?' The visual analogue scale when used for purposes of satisfaction is an illustration of the use of a global measure. Multi-dimensional questionnaires consist of multiple items or dimensions relating to different aspects of care. The Patient Satisfaction Questionnaire (PSQ) contains 43 items that reflect seven dimensions of service quality. They include hospital care, access to care, availability of services, technical quality of care, interpersonal care, communication, and financial implications.

The strengths of global measures are that they are simple, direct, easy to construct, easy to administer and usually cost effective. However, their weakness is that they do not provide variety or detailed information. A global measure does not nominate a particular subject on which the patient should focus. The researcher does not know what aspect of the treatment process the patient was thinking about when answering the question. Different people will think about different aspects of their care. As a result when comparing different individuals, the data provided is not reproducible.

Multi-dimensional measures focus the patient on specific aspects or dimensions of the care process; they provide more reliable and valid data. However, they are more complicated to construct and more time consuming to fill out. Standardised versions are available but these may not cover the particular dimensions required by the researcher, therefore, their usefulness is limited. How the measures are constructed and administered significantly affects the information that is produced. Standardised measures have usually undergone the rigours of psychometric testing for validity and reliability.

The content of existing satisfaction measures was reviewed by Ware, Snyder, Wright and Davies (1983). These authors found the most common dimensions included were: interpersonal manner, technical quality, accessibility, convenience, finances, efficacy and outcome, continuity, physical environment and availability. Important points of satisfaction are not well represented in current satisfaction measures. These include: being treated with respect; being involved in treatment decisions; the care-giver showing interest and concern; and the care-giver taking time to listen and explain (Cleary & Edgman-Levitan, 1997). The paper by Ware et al. was published many years ago (in 1983). Hudak and Wright (2000) made no mention of these dimensions being used in modern satisfaction measures. A recent Medline search revealed no recent reviews of current satisfaction measures. Whether current satisfaction measures adequately cover these important dimensions is a topic for future research.

Failure to communicate information about medical conditions and treatment options causes high levels of dissatisfaction (Carr-Hill, 1992; Locker & Dunt, 1978). Global measures mask dissatisfaction as these tend to skew scores towards higher levels of satisfaction (Blais, 1990; Ferris et al., 1992; Locker & Dunt, 1978).

When the outcome of primary care was compared with patient satisfaction scores, it was found that patients were disproportionately influenced by dimensions that caused dissatisfaction (Otani, Harris, & Tierney, 2003). Efforts aimed at reducing levels of dissatisfaction will provide greater benefits. Otani et al. suggested that policy makers should focus on and improve those areas that cause dissatisfaction; they state that this would be the best use of resources to improve patient satisfaction and outcomes.

2.10.2.3 (b) Content: care versus outcome

It is generally agreed that ‘care’ and ‘outcome’ should be assessed separately. The primary reason for this is that it is impossible for patients themselves to separate these issues. The easiest method of separating care and outcome for patients is to focus the questions on each separate area. Separating these two issues enhance research and measurements of specific clinical interventions. For example, the post-surgical care that a patient receives on a hospital ward can impact on that patient’s outcome, yet may not be related to the surgical result. By questioning the patient on aspects of care and outcome, these dimensions of the whole process can be measured independently.

Correlation between outcome and quality of care remains weak (Cleary & Edgman-Levitan, 1997; Ware, 1981). In the spine literature thus far, satisfaction with outcome has been primarily measured using global measures (Hazard, Haugh, Green, & Jones, 1994; Ljunggren, Weber, Kogstad, Thom, & Kirkesola, 1997; Torstensen et al., 1998). Future research needs to incorporate the use of multi-dimensional measures.

2.10.2.3 (c) Content: generic and disease/population specific measures

Generic measures are advantageous, as they are often readily available and have been tested psychometrically; the data they produce is valid and reliable. However, the data they collect is usually broad in nature and non-specific. This can limit their effectiveness. Disease-specific or population-specific measures are more likely to provide relevant data for a particular population or disease. When selecting a satisfaction measure, it is important to check that the measure includes dimensions that are relevant to the purpose of collecting the data.

2.10.2.3 (d) Content: direct versus indirect measures

Direct measures ask questions about the patient’s own direct experience with a service or intervention. Indirect measures ask about the patient’s attitudes towards the system in general. Direct measures are classified as specific (or micro measures), whereas indirect measures are classified as general (or macro) measures. The Patient Satisfaction Questionnaire is an indirect measure, whereas the Client Satisfaction Questionnaire is a direct measure.

Direct and indirect measures tend to be used for assessing different dimensions of satisfaction (Pascoe, Attkisson, & Roberts, 1983; Stewart & Wanklin, 1978). Direct measures are often used to assess services that are received or not received (Guyatt et al., 1995; Nork,

Hu, Workman, Glazer, & Bradford, 1999); these tend to produce greater levels of satisfaction compared to indirect measures (Hall & Dornan, 1988; Lebow, 1974). Because of their specificity, direct measures are favoured by clinicians for measuring the results of specific interventions (Ferris et al., 1992; Pascoe et al., 1983; Ware & Hays, 1988).

2.10.2.3 (i) Method: factual and affective.

Factual measures focus on the facts of what occurred, whilst affective measures emphasise the patient's perception of what occurred (Cleary & Edgman-Levitan, 1997; Ware, 1981). These concepts are sometimes couched in terms of reports and ratings. A report is an objective description of what actually took place, whereas a rating is a subjective score of the patient's perception of the process or intervention. An example is shown in the following questions:

'How long did you spend with the physician today?' (Factual)

'Did your physician spend enough time with you today' (Affective)

'Outcomes' are better measured by factual methods, as outcomes are more often black and white; the intervention either achieved what it set out to do, or it did not. Ratings or affective measures are used more often when measuring 'care'. These measures are influenced by the care the patient received, and by the patient's preferences and expectations. The type of measure used should reflect the type of data required by the researcher.

2.10.2.3 (ii) Method: open and closed questions

Open questions allow patients to express themselves in areas they feel are important to them. Patients will sometimes not consider all the different dimensions that need to be considered, which can lead to important data being missed. Closed questions allow questionnaires to focus the patient's thoughts, and can direct the patient's answers towards particular dimensions. Closed questions may highlight patient dissatisfaction in areas that the patient does not feel are important or worth mentioning. Patient satisfaction measures have problems with acquiescence, as some people always wish to please (Ware, 1978). Closed questions are a way of limiting this problem. The danger of closed questions is that they need to cover a broad variety of dimensions, otherwise, they too are susceptible to missing important data and biasing the patient's response. Hudak and Wright (2000) suggested that a combination of both types of questions provided the best option.

2.10.2.3 (iii) Method: self-administered versus interviewer-administered

Interviewer-administered measures tend to produce higher levels of satisfaction, probably because of acquiescence bias, but are more costly to administer (Ferris et al., 1992; LeVois, Nguyen, & Attkisson, 1981). Interviewer-administered measures are useful for sourcing information from non-respondents. With careful interviewer technique, they can also seek reasons or areas of dissatisfaction that may not be reported as effectively when using self-administered measures. When using interviewer-administered measures these factors need to be considered (Hudak & Wright, 2000).

2.10.2.3 (iv) Method: response formats

Response formats are the available answers for the patients to select. These formats may be numerical data, selected words, or open answers (in which the patients have to write their thoughts). The literature remains unclear as to the best method.

Ware and Hays (1988) demonstrated that a five point rating scale (excellent, very good, good, fair, poor) yielded lower mean scores, greater response variability and higher correlations with measures of behavioural intention, than a six point scale (extremely satisfied, very satisfied, somewhat satisfied, neither satisfied nor dissatisfied, somewhat dissatisfied, very dissatisfied). Greater numbers of available answering options tend to minimise acquiescence bias and highlight areas of dissatisfaction.

Satisfaction measures have methodological problems with selection bias, high-undifferentiated levels of satisfaction, and acquiescence bias. Satisfaction is usually measured at the completion of a programme or intervention. The problem with this is that dissatisfied patients may have already exited the programme and are thus not available to survey (Blais, 1990; Larsen, Attkisson, Hargreaves, & Nguyen, 1979; Pascoe et al., 1983). In this case, it is imperative that at least a sample of all non-respondents is followed up to enquire about their reasons for withdrawal. Blais reported that non-respondents have lower levels of satisfaction, but that satisfaction alone was not necessarily the primary reason for their withdrawal. Measuring satisfaction at the end of a longitudinal study that has followed participants for a number of years, may result in problems of memory recall (Burdorf et al., 1996; Linton & Melin, 1982; Severens et al., 2000).

High levels of undifferentiated data provide no significant information for the researcher. The reasons for obtaining this data are either due to selection bias, the use of

global measures or that the intervention is truly successful (produces high levels of satisfaction). To counter this problem, multi-dimensional measures and response formats with a wide variety of answers are of benefit.

Acquiescence bias is where the patient tends to agree with the statement, regardless of its content (Ware, 1978). The wording of statements is crucial; if the statement is positively worded, the measure tends to produce high levels of satisfaction, whereas if the wording is negative, the converse occurs.

To control for this bias, the literature suggests using neutrally worded statements, Ross, Steward, & Sinacore (1995) suggested using the adjectives 'poor, fair, good, very good, and excellent'. This format does not use the word 'satisfaction'. There is debate in the literature as to whether the word 'satisfaction' needs to be used in the response format, if satisfaction is what is being measured. Hudak and Wright (2000) commented on the fact that many patient satisfaction measures did not use the word 'satisfaction'. They stated that "a lack of clarity concerning the meaning of satisfaction and its relation to other measures, has been highlighted as a major weakness in this field of enquiry" (p. 3167). Unless the questions have been specifically designed to assess acquiescence bias, the degree to which it will affect a measure is not known (Ross et al.; Ware, 1978; Winkler, Kanouse, & Ware, 1982).

Generally it is accepted that there is a degree of acquiescence bias in all scores. Certain populations are known to produce higher levels of acquiescence; they are older, less well educated, less affluent, and poorer in health. However, Hudak and Wright (2000) stated that where there was a difference of one standard deviation or more in the scores, a true difference between the groups was represented. When satisfaction scores are assessed, these factors need to be considered.

In general, the literature is unable to recommend any particular satisfaction measure for the reasons previously discussed. Some authors caution against standardised measures because they are not context-specific (Carr-Hill, 1992). However standardised measures are easy to use, allow for comparisons across settings and studies, and have been exposed to the rigours of psychometric testing. Hudak and Wright (2000) suggested that the best option is a standardised quantitative measure with sections of open-ended questions, or even a qualitative follow-up.

Hudak and Wright summarised by making the following points:

- Where possible use a multi-dimensional measure, supplemented by a global measure.
- Consider the items of the measure to ensure that it will provide relevant information.
- Assess satisfaction of care separate to satisfaction of outcome.
- Direct measures are preferable.
- Use both open and closed questions.
- Always test a sample of non-respondents.

In conclusion, modern paradigms relating to outcome measures of patients for specific interventions, accept that the equation of patient function and well-being comprises more than just pure physiological and anatomical measures. Psychosocial issues play an important part as well. When assessing patient function and well-being, a number of outcome measures need to be employed, providing hard and soft data that cover a variety of different dimensions; this is in order to yield an accurate and reliable result (Deyo et al., 1994; Lee et al., 2001). An awareness of these issues is advantageous when critically appraising comparable literature on the topic of post-surgical management of lumbar discectomy.

2.11 The Post-surgical management of lumbar discectomy.

2.11.1 Introduction

In 2004, a Cochrane review on the topic of post-surgical management of lumbar discectomy was performed by Ostelo et al.(2004). In this review, the authors based their conclusions on 13 studies, six of which they defined as being of high quality. To be deemed high quality, at least five validity criteria had to be fulfilled.

These criteria were as follows:

- Method of randomisation
- Concealment of treatment allocation
- Dropout rate during the intervention period

- Withdrawals during follow-up
- Co-interventions avoided or equal
- Blinding of patients
- Blinding of outcome assessment
- Blinding of care providers
- Intention-to-treat analysis
- Compliance
- Similarity of baseline characteristics
- Adequate length of follow-up (more than six months)

Further Medline searches since this review have found three other studies published by Erdogmus et al. (2007), Filiz, Cakmak and Ozcan (2005), and Yilmaz et al. (2003). These, plus all 13 studies from the Cochrane review, are discussed in this section. A paper included in the Cochrane review by Rothhaupt, Laser, Ziegler and Liebig (1997) is only available in German, and is not discussed. The 15 papers that are reviewed are separated into two groups according to their quality rating by the Cochrane review. The key elements of these papers have been tabulated. Table 2.3 summarises the seven ‘high quality’ papers plus the paper by Erdogmus et al., as this paper also fits the Cochrane criteria for ‘high’ quality.

Table 2.3
Details of the seven high quality papers according to the Cochrane Review criteria 2004

Author & year of publication	Study type	Population sample & number	Intervention	Intervention begins post-surgery	Duration of intervention	Key Outcome Measures	Duration of follow-up	Results
Manniche C et al. 1993a	RCT	N = 96; Age 18 - 70 yrs; first time lumbar disc protrusion	Low intensity exercise versus high intensity exercise	4 - 5 weeks	6 weeks	Disability scores and days off work	One year	Favoured high intensity group at six weeks; no statistical difference at one year
Donceel P, Du Bois M & Lahaye D 1999	RCT	N = 60; insurance medical assessors and 710 participants; Participants aged between 15 - 64 yrs; must be working and have not been off work for more than 1 year prior to surgery	Early return to work using CBT vs. usual management	6 weeks	Until return to work is achieved or until one year	Return to work rates at one year post-surgery	One year	Strongly in favour of CBT group (p = 0.002); at one year 10.1% in intervention group and 18.1% in control group still off work
Dolan P, Greenfield K, Nelson RJ & Nelson IW 2000	RCT	N = 20; Age 18 - 60 years; radiological evidence of disc prolapse with sciatica of less than one-year duration.	Exercise vs. no exercise	6 weeks	4 weeks	Pain and disability using VAS and low back outcome score; Multidimensional Health Locus of control; Modified Somatic Perception Questionnaire; Zung Depression Scale; Posture and Mobility; Paraspinal muscle activity EMG; Biering Sorenson back muscle fatigue	One year	In favour of the intervention group
Kjellby-Wendt G, Styf J 1998	RCT	N = 60; Drop-outs = 12; results based on 48; Age 16 - 70 years; symptoms of disc herniation who did not respond to conservative therapy and who had not had previous back surgery	Comparison of two exercise groups	Day one	12 weeks	Hamstring length; lumbar ROM; patient satisfaction; sick leave	Patient satisfaction at two years; other outcomes at one year	At one year no significant difference in sick leave; other outcome measures slightly favoured the Early Active Training (EAT) group; at two years the EAT group had higher satisfaction levels
Manniche C et al. 1993b	RCT	N = 62; Age 18 - 74 years; first time lumbar surgery; Patients global assessment was good, fair or unchanged	Resisted hyperextension vs. resisted extension to neutral	14 - 60 months	12 weeks	Patients global assessment; Low back pain rating scale	One year	No statistical difference
Danielson JR, Kibsgaard SK, Hellevi E 2000	RCT	N = 63; Age 20 - 60 years and no co-morbidities that would impede physical training	Gymnasium based vs. usual conservative home based exercise programme	4 weeks	8 weeks	Roland-Morris; VAS pain diagram; Wonca's functional status; sick leave	One year	Favoured the gymnasium exercise group
Erdogmus CB et al. 2007.	RCT	N = 120; First time Lx discectomy; pre-op symptoms of less than 6 months duration; no other co-morbidities.	Physiotherapy exercise programme vs. sham therapy neck massage vs. no treatment	One week	12 weeks	Low back pain rating scale; satisfaction; compliance with exercise programme; socioeconomic and psychological parameters; return to work rates.	18 months	No statistical difference between groups

Note: Abbreviations: RCT = randomised controlled trial; N = number; CBT = Cognitive Behavioural Therapy; vs = versus; VAS = Visual Analogue Scale; Lx = Lumbar; EMG = Electromyography; ROM = Range of Movement; yrs = years

Table 2.4 summarises the nine low quality papers. The studies by Filiz et al. (2005) and Yilmaz et al. (2003) are included with the low quality Cochrane review papers because their follow-up periods are three months.

2.11.2 Analysis of papers meeting the high quality criteria of the Cochrane review 2004

The authors of the seven high quality papers (Table 2.3) are Danielson et al. (2000), Dolan et al. (2000), Donceel et al. (1999), Erdogmus et al. (2007), Kjellby-Wendt et al. (1998), Manniche et al. (1993a) and Manniche et al. (1993b). The inclusion/exclusion criteria for these studies are all very similar. The inclusion criteria have been described in Table 2.3. The exclusion criteria for all of these studies were previous lumbar surgery, infection, other medical co-morbidities and any psychosocial or mental health problems that may impede post-surgical rehabilitation.

Six of the studies compared an exercise programme with either another exercise programme or with a control group. One paper by Donceel et al. (1999) compared a cognitive behavioural therapy programme with 'usual management'. Of the six papers that used post-surgical exercise programmes, Dolan et al. (2000) and Erdogmus et al. (2007), compared an exercise regime with a control group that undertook no formal post-surgical rehabilitation. The study by Dolan et al. was a pilot study. The other four studies by Danielson et al. (2000), Kjelly-Wendt et al. (1998), and Manniche et al. (1993a & b), all compared one exercise programme against another. Because the control groups in these studies also performed a post-surgical exercise regime, it is not possible to report whether the rehabilitation programmes being tested in these studies were any better than 'usual post-surgical conservative care' or a 'no exercise' group. Therefore, the results of these studies offer limited evidence to support post-lumbar discectomy exercise programmes.

Of the two studies that compared a post-surgical exercise trial group with a no-exercise control group, Dolan et al. (2000), at one year post-surgery, reported results that favoured the exercise group. The categories that achieved statistical significance between the groups in the Dolan study were 'pain', and some muscle endurance recordings. Dolan et al. also reported greater improvements in lumbar and hip R.O.M. in the exercise group but the authors did not provide statistical analysis for these data. The authors also found strong correlations between pain, disability, psychological function and spinal function. They noted that functional measures were more highly correlated with pain diaries than visual analogue scales in the

exercise group. In the control group, the converse was true, functional measures correlated better with the visual analogue scale than participant pain diaries. The authors make no comment as to why this finding may have occurred.

The study by Erdogmus et al. (2007) involved a large sample size of 120 participants, the exercise programme was of three months duration and the authors followed participants for 18 months. This is the longest follow-up time period for any of the studies included in the 'high quality' category of the review by Ostelo et al. (2004). The results of the Erdogmus study showed no statistical difference between the physiotherapy exercise, the sham neck massage and control/no treatment groups. These results suggest that the physiotherapy programme in question did not provide superior long term outcomes when compared to sham neck massage or to a no-treatment group.

Issues that lead to confounding in evidence-based medicine affected all the studies. For example, a large number of drop-outs affecting the sample size and statistical power (Dolan et al. 2000 and Danielson et al. 2000); high drop-out numbers and validity of outcome measures (Kjellby-Wendt et al. 1998); and insufficient intervention periods. Manniche et al. (1993a) reported that a longer intervention period was required to obtain significant long-term gains.

However, these studies have raised many points of interest. For example, Manniche et al. (1993a) used aggressive long lever exercises focussed on the lumbo-pelvic region at an early stage post-surgery. These patients were no worse off for doing this exercise. During the 1980s and early 1990s, the emphasis after surgery was more on rest than activity. This philosophy would have been considered proactive at this time. Long lever exercises resulted in high internal tissue loads (McGill, 2007). Wise, Uhl, Mattaqcola, Nitz and Kibler (2004) studied gleno-humeral muscular activity on a series of supported versus unsupported, and short lever versus long lever exercises. The unsupported and long lever exercises incorporated higher levels of muscular activity. Wise et al. (2004) stated that when designing a progressive exercise programme, short lever exercises should precede long lever exercises. Long lever exercises were usually reserved for advanced stage rehabilitation, as a means of work hardening prior to the patient returning to full function.

The study by Manniche et al. (1993a) challenges this research. Current post-surgical paradigms encourage early activity (Carragee et al., 1999; Saal, 1990; Saal & Saal, 1989), which is the complete opposite of the 1980s philosophy. These studies, including the study by

Manniche et al., have pioneered the thinking towards the modern concept of early mobilisation post-surgery.

A key finding in the study by Donceel et al. (1999), was the ability to rank positive and negative variables relating to return to work (RTW). The positive variables (in descending order) were, younger age, pro-active post-surgery management, male gender, lack of radiating leg pain prior to surgery and not smoking. The negative variables relating to RTW were firstly, previous back surgery, and secondly, a longer period of work incapacity prior to surgery. This knowledge is valuable in terms of management and avoidance of potential complications post-discectomy.

The positive variable, namely, 'lack of radiating leg pain prior to surgery' is noteworthy. Different surgical criteria seemed to apply in Belgium. According to Ostelo, de Vet, Waddell et al. (2004) radiating leg pain is a symptom of lumbo-sacral radicular syndrome. These authors reported that lumbar-discectomy was usually reserved for patients exhibiting severe, unremitting low back and leg pain, associated with increasing motor blockade, and bladder and rectal sphincter problems.

The study by Erdogmus et al. (2007) reported no statistical difference between the physiotherapy and sham treatment groups at end point of analysis. However, raw data demonstrated a definite advantage for the physiotherapy group compared to both the sham-treatment and no-treatment groups, in terms of percentage of participants that had returned and had not returned to work. No statistical analysis was reported for this data. Return to work remains an important economic indicator of success or failure for any intervention.

Table 2.4
Details of the nine low quality papers according to the Cochrane Review criteria 2004

Author & year of publication	Study type	Population sample & number	Intervention	Intervention begins post-surgery	Duration of intervention	Key Outcome Measures	Duration of follow-up	Results
Amaranth H, et al. 1986	RCT	N = 212; First time Lx discectomy; < 55yrs; not retired	Intensive back school vs. usual care	4 weeks	2 weeks	Occupational handicap; sick leave; subjective pain report	One year	No statistical difference
Brennan GP, et al. 1994	Quasi RCT	N = 37; First time lumbar discectomy; age 19 - 51 yrs; no co-morbidities that would inhibit aerobic exercise; postero-lateral herniations only	Aerobic exercise vs. floor and stretching exercises	4 weeks	12 weeks	Weight; % fat; VO2 max; pain intensity; pain bothersome test; Beck depression score; Activity Pattern Indicator ratio	3 months	Participants were able to increase their aerobic fitness without increasing pain; Low statistical power voided other results
Johannsen F, et al. 1994	Prospective Clinical trial	N = 40; First time Lx discectomy; age 18 - 65 yrs; no other co-morbidities	Supervised vs. unsupervised exercise programme	6 weeks	12 weeks	Trunk flexion/extension strength; spinal mobility; back pain disability score; drug therapy; sick leave	6 months	No statistical difference
Burke SA, Harms-Constas CK, Aden PS. 1994	Prospective Clinical trial	N = 397; LBP either surgical or non-surgical; unable to work; off work not less than 6 weeks; no significant psycho-pathology; no interfering co-morbidities	FRP vs. usual care	Not reported; assumed to be within 6 weeks	Until a number of set criteria were fulfilled	A variety of tests that, combined, formulated a comprehensive work capacity assessment.	One year	P < 0.0001 in favour of the FRP
Kitteringham C. 1996	Clinical trial; Pilot study	N = 12; First time discectomy; age 20 - 65 yrs; unilateral referred pain, positive limited SLR.	Optimal SLR regime vs. usual SLR regime	Day one	7 days	Angle of SLR; 101 numerical rating scale; Oswestry Low Back Pain Index; Lumbar Flexion / Extension R.O.M	6 weeks	No statistical difference; a trend in favour of the usual care group
Filiz M, Cakmak A, Ozcan E. 2005	RCT	N = 60; First time discectomy; age 20 – 50 yrs; with good post-surgical result and no co-morbidities that would complicate rehabilitation	Intensive supervised exercise programme vs. home programme vs control non-exercise group	4 weeks	8 weeks	Low back and abdominal muscle endurance; Schober's test; VAS; Modified Oswestry Disability Index; Low Back Pain Rating Scale; Beck Depression Inventory	3 months	Favoured the intensive exercise group
Ostelo RW, et al. 2003	RCT	N = 105; First time discectomy; age 18 - 65 yrs; no leg / back pain restricting ADLs or work	Behavioural Graded Activity programme vs. usual care	6 weeks	3 months	Global Perceived Effect; Roland-Morris Score; Pain Catastrophizing Scale; Tampa Scale; back & leg pain severity; ROM; General Health; Social Functioning	3 months	No statistical difference
Timm KE. 1994	RCT	N = 250; Had a L5 laminectomy within the last one year; ongoing LBP or leg pain not below the knee; CT scan demonstrated disc abnormality; no abnormal psychosocial factors	Comparison of electro-therapy, joint mobilisation, low tech exercises, high tech exercises and control group with no intervention	An L5 laminectomy within the last year and pain for at least the last 6 months	8 weeks	Lumbar ROM test; Modified-modified Schubert test; the Cyber lift test; Sweaty Low Back Index; Achieved period of pain relief;	One year	Results strongly favoured active treatment approaches over passive treatment
Wilma F et al. 2003	RCT	N = 42; First time discectomy; age 20 - 60 yrs; no other co-morbidities.	Supervised dynamic lumbar stabilisation exercises vs. home programme & control group no exercise	Within one month	8 weeks	VAS pain; Modified Oswestry Index; Beck Depression Scale; spinal mobility; weight lifting capacity and body strength	12 weeks	Results favour the supervised exercise group, however no between group p-values were reported

Note: Abbreviations: RCT = randomised controlled trial; N = number; CBT = Cognitive Behavioural Therapy; vs. = versus; VAS = Visual Analogue Scale; Lx = Lumbar; FRP = Functional Restoration Programme; R.O.M. = range of movement; VO2 max = volume of oxygen consumption; ADLs = Activities of Daily Living; yrs = years; SLR = Straight Leg Raise; CT = computer tomography

2.11.3 Analysis of papers meeting the low quality criteria of the Cochrane review 2004

The results of these nine 'low quality' papers (Table 2.4) have to be considered with caution due to weaknesses in the study method and design. Some of the authors recognised their study's weaknesses and reported them (Kitteringham, 1996, Filiz et al., 2005, Brennan et al., 1994).

Four papers reported no statistical difference between comparison groups. The paper by Kitteringham demonstrated that usual care provided superior results compared with the intensive mobilising regime (Alaranta et al., 1986; Johannsen et al., 1994; Kitteringham, 1996; Ostelo et al., 2003).

The study by Burke et al. (1994), that tested a functional restoration programme against a comparison group, demonstrated very strong results in favour of the intervention group. However, this was a clinical trial with selected participants, any interventions applied to the comparison group were not described. The percentage of completed follow up questionnaires at one year ranged from 93% in the treatment group, 80% in the comparison group, and 60% in the drop out group. These confounding factors must be considered when examining the results of this study.

Ostelo et al. (2003) conducted a RCT that compared a behavioural graded activity programme (a similar approach to the functional restoration programme used by Burke et al. (1994)), with 'usual physiotherapy' and found no difference between the groups. It would be interesting to test the management regime used by Burke et al. in a RCT.

Filiz et al. (2005) and Yilmaz et al. (2003) both reported significant results in favour of the intervention groups at three months post-surgery. However, it has been shown in other studies (Erdogmus et al., 2007; Kjellby-Wendt & Styf, 1998; Manniche, Asmussen et al., 1993b; Manniche, Skall et al., 1993a) that these early results are sometimes negated with time. By one year post-surgery the early benefits in favour of the exercise groups were nullified. A short follow-up period significantly compromised any results that may have been achieved.

In spite of the shortcomings of these studies some interesting points arise. Alaranta et al. (1986) and Johannsen et al. (1994) both reported no statistical difference between their comparison groups. Close analysis of these studies reveals that in both cases, their

comparison groups received similar interventions; the primary intervention group received a more intensive version relative to the control group. In the Alaranta et al. study, both groups received similar information in the form of a 'Back School'. The difference between the groups was that the intervention group had this information given to them in a more intensive format. Johannsen et al. followed a similar structure to that of Alaranta et al. the comparison groups were given similar exercises; the intervention group was supervised, whilst the control group performed their exercises as a home programme. The non-significant findings of these studies support other research that has demonstrated less intensive or more general programmes are as effective, if not more effective, than intensive programmes (Ferreira et al., 2007; Kitteringham, 1996; Manniche, Asmussen et al., 1993b).

In reference to the study by Ostelo et al. (2003), two points are worthy of mention. The study did not allow the physiotherapists in the 'usual care' group to use specific cognitive behavioural tools or management and the authors reported that many physiotherapists in the Netherlands used a biomechanical model of disease.

Usual physiotherapy management consists of cognitive therapy (in terms of educating patients about their pain), reducing fear avoidance, reducing catastrophisation, and advising how best to manage the pain. By removing these aspects of normal physiotherapy treatment, the study was not using 'usual care physiotherapy'. The purpose of this study was to show that a behavioural graded cognitive therapy programme might provide superior results when compared to the biomechanical approach. Even if this had eventuated, the results would only be applicable to physiotherapists who adhere solely to the biomechanical paradigm.

During the process of this study, Ostello et al. (2003) taped treatment conversations between patients and therapists. This was to ascertain whether or not cross-over treatment was occurring. The authors found this did occur but not to a significant extent. This means that some physiotherapists were using cognitive therapy, even when they had been instructed not to do so. This implies that physiotherapists may use more cognitive behavioural therapy than is realised.

Ostelo et al. (2003), reported that the physiotherapy 'usual care' group achieved a statistically significant advantage in the Global Perceived Effect outcome measure, when compared with the behavioural graded programme. However, after 'adjusted' statistical analysis, this statistical advantage for the 'usual care' group was lost. The paper did not

clearly describe the meaning of ‘adjusted’ statistical analysis. The result of the Global Perceived Effect would have been interesting if physiotherapists in the ‘usual care’ group had been allowed to use their usual method of cognitive behavioural management.

2.12 Concepts of Exercise Prescription

2.12.1 Introduction

Patients with pain, particularly persistent (or chronic) pain, often exhibit low levels of activity and fitness. The literature reports that, should these patients increase their activity and fitness levels, a corresponding improvement in their own sense of personal well-being would be experienced (McAuley et al., 2000; McAuley et al., 2005; North, McCullagh, & Tran, 1990; Ruuskanen & Ruoppila, 1995; Tsutsumi et al., 1998; Watanabe, Takeshima, Okada, & Inomata, 2001). A recent review by Warburton, Nicol, and Bredin (2006) documented the health benefits of increased physical activity across a wide variety of chronic diseases, such as diabetes, cancer, obesity, heart disease, depression and bone and joint disease.

Post-lumbar-discectomy patients have generally experienced pain and decreased physical function for extended periods of time. As a consequence, some patients experienced increased levels of psychosocial stress and decreased physical ability (den Boer et al., 2006; Dvorak, Valach et al., 1988; Kahanovitz, Viola, & Gallagher, 1989). By increasing physical fitness and strength, these negative consequences can be minimised (Geiger, Todd, Clark, Miller, & Kori, 1992). The essence of any post-surgical exercise programme is to achieve these goals without increasing the pain.

There are several concepts of exercise prescription that are central to the design of the exercise programme used in this thesis. These concepts will be discussed in relation to post surgical and pain patients with support from the literature. These concept questions and points are as follows:

- How does exercise improve function?
- The cognitive effects of exercise.
- Equipment options and how they are used.
- Methods of strength training.

- Principles of periodization.
- Principles of exercise prescription for the injured person.

This section is summarised by discussing how the above concepts have been integrated into the exercise programme used in this thesis.

2.12.2 How does exercise improve function?

Research has demonstrated that the application of load on a muscle leads to an increased ability for that muscle to withstand load, and generate greater force resistance. As a result, it gets stronger (Ahtiainen, Pakarinen, Alen, Kraemer, & Hakkinen, 2003; Campos et al., 2002; Edstrom & Grimby, 1986; Gruber & Gollhofer, 2004; Hakkinen, Alen, Kraemer et al., 2003; Hakkinen, Kallinen, Komi, & Kauhanen, 1991; Hakkinen et al., 1996; Kraemer, Fleck, & Evans, 1996; Kraemer et al., 2003; Kraemer & Ratamess, 2004; Taaffe & Marcus, 1997). Although this concept is an accepted fact, the question remains. Does an increase in strength lead to an improvement in function? And if so how does this occur?

A person's functional ability is a sum of biomechanical and psychosocial factors. It is well known that improved strength or range of movement alone does not necessarily lead to improved function (Korres et al., 1992; Yen et al., 1993). However, if there are no negative psychosocial factors present, several studies have demonstrated a direct correlation between strength and function (Albright et al., 2001; Brill et al., 1999; Canning, Ada, Adams, & O'Dwyer, 2004; Hurley & Scott, 1998; O'Reilly, Jones, Muir, & Doherty, 1998; Rainville, Sobel, Hartigan, Monlux, & Bean, 1997; Vuori, 2001).

Research has demonstrated that resistance training increases internal ligament and tendon strength, osseoligamentous and osseotendonous intersection strength, and bone mineral density (Staff, 1982). The literature reports that damaged tendons and ligaments regain normal maximal strength levels faster as a result of increased physical activity (Tipton, Matthes, Maynard, & Carey, 1975). These factors, along with pre-season strength and conditioning training, have been shown to decrease 'in season' rates of injury (Abbott & Kress, 1979; Cahill & Griffith, 1978; Hejna & Rosenberg, 1982) and have also been shown to decrease overuse injuries, such as swimmer's shoulder and tennis elbow injuries (Fleck & Falkel, 1986). Decreased rate of injury and quicker recovery time leads to improved functional performance.

Lumbar discectomy patients often present in a physically weakened state after having suffered severe low back pain for extended time periods (Mayer, Smith, Keeley, & Mooney, 1985). They often undergo a period of recovery that does not include any formal rehabilitation immediately post-surgery (Mayer et al., 1989). Mayer et al. (1985) tested saggital trunk strength on a group of chronic low back pain patients, and compared this data with normal subjects. These authors found that "strength deficits are a major factor in the deconditioning syndrome associated with chronic low back pain" (p. 727). On a group of spinal surgery patients three months post-surgery, Mayer et al. (1989) conducted a study using computer tomography that measured muscle area/density and isokinetic trunk strength. This study found the post-surgical mean trunk strength of these patients to be 50% below those of normative control subjects. On qualitative analysis, there was a correlation between weaker trunk strength and reduced area and density of trunk muscles particularly in the erector spinae, in the rectus abdominus and to a lesser degree in the psoas.

Ylinen et al. (2003) conducted a similar study, and compared cervical discectomy patients with normal controls; they tested strength and range of movement. The authors found this patient group to be significantly weaker than the matched control group. They reported a direct correlation between pain and disability scores and decreased muscle strength.

The findings of Mayer et al. (1985, 1989) and Ylinen et al. (2003) suggested that patients who suffered chronic pain also suffered a consequent loss of strength; this led to what Mayer et al. described as the 'deconditioning syndrome'. These studies support the theory that patients who have suffered chronic pain do physically decondition, and do lose strength. These studies provided evidence to support the theory that decreased strength is related to decreased function. Therefore, by increasing strength, it is possible to increase functional performance. Specifically pertaining to the post-lumbar-discectomy population, studies have demonstrated that restoring the patients' whole body strength improved their functional activity levels (Alaranta et al., 1986; Danielsen et al., 2000; Manniche, Skall et al., 1993; Ostelo, de Vet, Waddell et al., 2004).

2.12.3 The cognitive effects of exercise

Although the relationship between strength and function has been confirmed, many other studies have demonstrated a stronger relationship between psychological dysfunction, (involving fear avoidance, catastrophisation, depression) and disability than between pain and

weakness, and disability (Foa & Kozak, 1986; Philips, 1987; Picavet, Vlaeyen, & Schouten, 2002; Simmonds, Kumar, & Lechelt, 1996). If exercise is going to be effective in reducing disability, it must positively influence these psychological factors. The question is, does exercise do this and how?

Research comparing levels of energy expenditure and psychological well-being in older adults has found a direct correlation. High activity levels were associated with high levels of personal well-being, mood improvement and decreased levels of depression. Low activity levels were associated with low levels of well-being (North et al., 1990; Tsutsumi et al., 1998; Watanabe et al., 2001). This research was further supported by Ruuskanen and Ruoppila (1995), who found that low activity levels were associated with high depression scores.

The psychological effect of long-term exercise training in a group of spinal injury patients was assessed by Hicks et al. (2003). These authors found that exercise for this group resulted in significant improvements in both physical and psychological well-being. They postulated the mechanism for these positive changes was threefold:

- Exercise induced changes in pain mediate positive changes in quality-of-life factors.
- Exercise improves sense of control and mastery of physical functioning.
- Exercise usually involves social interaction.

These three points were also highlighted in other studies (Martin, 2001; McCain, Bell, Mai, & Halliday, 1988; Rejeski, Ettinger, Martin, & Morgan, 1998). The study by McCain et al. compared an exercise and control group with similar amounts of social interaction. The exercise group demonstrated greater benefits in psychological well-being and quality-of-life factors when compared to social interaction alone.

A study by Sculco, Paup, Fernhall and Sculco (2001) further demonstrated the lack of correlation between pain, psychological dysfunction and disability. In their study, the authors followed a group of chronic low back pain patients for 30 months. A ten-week aerobic exercise programme was compared to a non-exercise control group. Both groups underwent normal conservative treatment for their LBP. Significant differences in favour of the exercise group were found across all measures of medication levels, therapy visits and time off work. The psychological outcome measures improved in the exercise group, but interestingly, the

perception of pain in this group was no different from the beginning of the intervention compared to the end. This study highlighted the fact that pain was only one component of general well-being and functional status. These patients improved because they felt generally better, even though their pain levels had not changed. During the study, the exercise and control groups were offered a chance to cross over into the other group. None of the exercise group crossed over to the non-exercise group; their preference was to continue exercising.

Fear avoidance is a key indicator of chronic pain, catastrophisation and depression (George, Fritz, Bialosky, & Donald, 2003; Vlaeyen, Kole-Snijders, Boeren, & van Eek, 1995). Many studies have demonstrated that a reduction in fear avoidance facilitates functional improvement (Frost, Klaber-Moffett, Moser, & Fairbank, 1995; Lively, 2002; Mannion, Muntener, Taimela, & Dvorak, 1999; Pfingsten, 2001; Rainville et al., 1997).

However, this may not always be the case. A study by George et al. (2003) highlighted that not all patients respond positively to fear avoidance education. Patients with high fear avoidance scores responded well to fear avoidance therapy or education, but patients with low fear avoidance scores were made worse with this type of therapy. The authors postulated that people with low fear avoidance scores are ‘confronters’ and did not need this type of intervention. In fact, this extra intervention may have detracted ‘confronters’ from actually confronting the situation and ‘getting on with it’. Another explanation could have been that confronters did not think about the factors involved with fear avoidance but when these points were raised during therapy, they then developed some of these factors.

Exercise is an excellent method of reducing fear avoidance (Kernan & Rainville, 2007). Kernan and Rainville conducted a study on chronic low back pain patients. The mean duration of symptoms was 28 months. These patients were prescribed a non-pain contingent quota of exercise. Follow–up at one year post-intervention revealed a statistically significant reduction in fear avoidance beliefs. It is hypothesised that by slowly introducing patients to an exercise regime, where they feel in control of the situation, a therapist is able to gradually increase levels of exercise and physical activity. Patients then begin to experience the sensation of improved well-being. This was demonstrated in the study by Sculco et al. (2001). This regime improves patients psychological well-being and allows them to feel confident to try more advanced exercises or functional activities. This is how patients literally climb out of the downward spiral of poor psychological well-being and decreasing physical function. This concept was demonstrated in a study by Lindstrom et al. (1992). These authors compared

‘usual care plus a graded physical activity programme’ with just ‘usual care’ in a population of subacute low back pain patients. The addition of a graded exercise programme provided superior results in terms of psychological and physical function.

A key issue in overcoming fear avoidance and encouraging patients to undertake exercise, is a full explanation of how this form of therapy is going to help them (Pfingsten, Hildebrandt, Leibing, Franz, & Saur, 1997; Poulter, 1999). Information and advice assist by placing a positive framework around sensitive issues (Burton, Waddell, Tillotson, & Summerton, 1999). Providing programmes that are well written and illustrated has been shown to significantly increase patient compliance (Schneiders, Zusman, & Singer, 1998). Exercise programmes aim to use the benefits of exercise to promote ‘wellness’ and inhibit illness behaviour (Cohen & Rainville, 2002; Rainville et al., 1997; Sollner & Doering, 1997). In following such programmes, patients feel empowered to take control over their pain and the problems that have previously been controlling them.

2.12.4 Principles of rehabilitation programmes for low back pain

A number of issues need to be considered when discussing rehabilitation programmes for low back pain. These include, the anatomical structure involved; the underlying pathology; the irritability of the pathology; the appropriate exercise prescription for that pathology (bearing in mind the force loading on the spine); the patient’s psychological state of mind; and the patient’s social environment McGill (2007). This section discusses these issues within the context of this thesis. It is in no way a complete summary of this complex and significant area of research. The discussion of rehabilitation programmes for low back disability will be limited to the physical aspects of this topic.

In 1992, Panjabi (1992a) reported that spinal stability was dependent on three subsystems, the spinal column, the spinal muscles, and the neural control unit. Panjabi further defined these structures as passive and active; the passive system was the spinal column consisting of the vertebrae, discs, ligaments, facet joint capsules and the passive mechanical properties of muscles. The active system consisted of spinal musculotendonous units and the neural feedback system. The neural feedback system consisted of motion transducers in the ligaments, tendons and muscles, and in the neural control centres.

In another study, Panjabi (1992b) redefined spinal instability in terms of the neutral zone. Panjabi described the neutral zone as, “that part of the range of physiological

intervertebral motion, measured from the neutral position, within which the spinal motion is produced with a minimal internal resistance. It is the zone of high flexibility or laxity" (p. 391).

Further research has demonstrated that injury and intervertebral disc degeneration increase the size of the neutral zone (Kaigle, Holm, & Hansson, 1995; Mimura et al., 1994; Panjabi, Abumi, Duranceau, & Oxland, 1989), and that simulated muscle forces are able to decrease the size of the neutral zone (Kaigle et al.; Panjabi et al.; Wilke, Wolf, Claes, Arand, & Wiesend, 1995). These studies implied that an increased neutral zone was associated with injury however, by altering the active system controlling the neutral zone, the effect of the injury could be reversed. The authors reported the limitations of their results, due to these studies all being performed on animal and human cadaver specimens.

This research suggests that a primary aim when designing rehabilitation programmes for low back problems, is to reduce the size of the neutral zone; that in clinical terms means stabilising an unstable back. A detailed description of the clinical presentation for various directional lumbar segmental instability patterns and other methods of diagnosis were presented by O'Sullivan (2000) and McGill (2007). It is not in the scope of this thesis to discuss the method and science leading to the diagnosis of low back instability.

Cholewicki and McGill (1996) stated that spinal stability was derived from spinal stiffness. Spinal stiffness is dependent on the proper functioning of the active and passive subsystems described by Panjabi (1992). Injury or dysfunction in either system may lead to reduced stiffness and reduced stability (Cholewicki & McGill).

For example, tissue damage, such as vertebral endplate or disc damage may lead to passive system dysfunction. Active system dysfunction may be caused by abnormal motor patterns or abnormal motion control (McGill, Grenier, Kavcic, & Cholewicki, 2003). These authors stated that 'motor patterns' were related to the pattern of muscular activation when performing a task. For example, 'abdominal in-drawing' might be dominated by the deep abdominal muscles or by the rectus abdominus muscle; 'motion patterns' referred to "the kinematic description of body segments" (p. 354). An example of a different motion pattern to perform the same task, is using either knee extension or hip extension to move from sitting to standing. The key point to note from this research is that injury may be both a cause, and a

result of altered motor, and/or motion patterns (McGill et al., 2003; O'Sullivan et al., 1997; O'Sullivan et al., 2003).

Examples of pain provokative motor patterns in low back pain populations are evident in research conducted by Dankaerts, O'Sullivan, Burnett and Straker (2006). These authors compared a group of LBP patients with a group of pain-free controls. The objective was to compare trunk muscle activation during unsupported sitting. The study used a classification system described by O'Sullivan (2000) to sub-classify the low back pain group. One of the sub-classifications of this system is the active extension pattern. Patients who fit this category tended to experience pain when they performed any activities that applied extension forces to the spine (O'Sullivan, 2000). Dankaerts et al. found that these same patients, when sitting, had hyperactive back extensor muscle and transverse fibres of internal oblique muscles. These patients were over-using their extensor muscles, causing increased extension and stress on their already painful lumbar spines; they were inadvertently increasing their pain by using this pain provocative motor pattern.

McGill and Norman (1987) gave an example of a pain provocative motion pattern. In this study, one person was asked to lift a weight with a flexed lumbar spine, and again with a neutral lumbar spine, while the shear forces through L4-5 were measured. The shear forces applied to the spine in flexion were nearly ten times that of the neutral spine.

Several authors have advocated specific exercise regimes to correct pain provocative motor patterns prior to resistance exercises being applied (Hides et al., 1996; Hodges & Richardson, 1996; McGill & Karpowicz, 2009; O'Sullivan, 2000; Richardson & Jull, 1995). Increasing patients' strength might not necessarily have improved their lifting technique or sitting posture. If the patients with poor lifting technique increased their strength, they might have been encouraged to lift heavier weights and exposed themselves to further harm. If the patients with pain provocative sitting posture increased the strength of their erector spinae muscles, greater forces would be applied to an already stressed spine; this is likely to aggravate those patients' pain.

The biomechanical model suggested that improving the coordination/co-contraction of the muscles acting on the spine led to increased spinal stiffness and therefore stability (Allison et al., 1998; Cholewicki & VanVliet, 2002; Hodges, 1999; Kaigle et al., 1995; Kavcic,. Grenier, & McGill, 2004a; McGill, 1991; Panjabi et al., 1989; Richardson & Jull, 1995;

Wilke et al., 1995). An area of debate in the literature is the use of local and global muscles, and their respective effect on spinal column stiffness and stability.

Local muscles attach directly to the spine. They tend to be small and are positioned deep relative to the superficial muscles; their primary function is to provide intersegmental stability (Kavcic et al., 2004a; Panjabi et al., 1989). Global muscles do not have direct attachment to the spine, but because of their directional moments of force, they apply significant torque to the spine. Global muscles tend to be larger, more superficial muscles, and act as general trunk stabilisers; they are not capable of direct segmental influence (Kavcic et al., 2004a; O'Sullivan, 2000).

Acccording to Panjabi (1989), a simplistic biomechanical analysis of the local muscles suggested that they were not able to exert sufficient force on the lumbar spine to provide functional stiffness and stability. However, Panjabi then outlined a number of arguments that countered this viewpoint. According to Panjabi the short muscles close to the centre of rotation had an important role to play in providing fast feedback to the neuromuscular system, facilitating an efficient and smooth spinal stabilising control system. Panjabi argued that because of the local muscle system's ability to provide intersegmental control, they were likely to be the primary stabilisers of the spine.

This argument by Panjabi (1989) was supported in a study conducted by Cholewicki and McGill (1996). These authors found that an increase of between 1-3% in maximum voluntary contraction for the mulitifidus and for the lumbar erector spinae muscles was enough to provide stability to the neutral spine. Although the local muscles are not able to apply significant torque to the spine, it appears that large forces are not required when the spine is in neutral, and not significantly loaded. It is hypothesised that the local muscle system provides a stable platform from which the global muscles can act (Allison et al., 1998; Hodges & Richardson, 1996; Wilke et al., 1995).

The importance of multifidus was further demonstrated in a study conducted by Hides, Stokes, Saide, Jull and Cooper (1994). In this study a group of acute low back pain patients was compared with a group of painfree controls. The focus of this study was to assess whether acute low back pain had an affect on multifidus. The authors found the low back pain patients all had marked asymmetry of the cross sectional area (CSA) of the multifidus muscle, with the CSA being reduced on the side of pain. This asymmetry was not seen in the control

group. In 24 of the 26 acute low back pain patients the level of multifidus wasting coincided with the clinically determined symptomatic level. Another finding of this study was that the muscle wasting was specific to one level only. A correlation between specific, one level, unilateral multifidus muscle wasting and acute low back pain was demonstrated. However, the amount of muscle wasting did not correlate with the patients' reported severity of symptoms. This is not unexpected as the literature reports numerous factors that affect the reporting of pain, including financial incentives, biomechanical, psychosocial, and cultural factors (Donceel et al., 1999; Bigos et al., 1991; Burton et al., 1995; Waxman et al., 2000;Graver et al.,1999; Hasenbring et al., 1994; Junge et al., 1996; Weber, 1983; Dvorak et al., 1988; Ihlebaek et al., 2006; Schade et al., 1999 & van Doorn, 1995).

After the first onset of acute low back pain, episodic reoccurrence is often reported in the literature (Hesterbaek, Leboeuf-Yde, Manniche, 2002; Hesterbaek et al., 2003). In a review of low back pain reoccurrence rates Hesterbaek et al. (2003) reported an average reoccurrence rate of 60% (range 44%-78%) during the one year period immediately after the first episode of low back pain. To test whether multifidus wasting had an influence on back pain or specifically the prevention of back pain, Hides, Richardson and Jull (1996) conducted a RCT that compared two groups of acute low back pain patients. The control group underwent medical management only while the trial group was supervised through a specific multifidus muscle rehabilitation programme as well as medical management. The reoccurrence rate of low back pain for the control group in the year following their first episode of acute low back pain was 80%, in keeping with previous results reported in the literature. The reoccurrence rate for the trial group was 30%, significantly lower than expected reoccurrence rates. The authors concluded that the higher level of reoccurrence in the control group might be due to the injured segment being exposed to "decreased muscle support and loss of fine motion controls" (Proceedings of the Australian National Physiotherapy Congress 1996; pp43).

Transverse abdominus is another local muscle that acts on the lumbar spine and has been extensively researched. A review by Hodges (1999) outlined the results of several studies that investigated the effect of transverse abdominus in relation to lumbar stability. The initial investigations asked whether transverse abdominus was activated when the spine was moved (Hodges and Richardson 1997a & 1997b). These studies proved that transverse abdominus was an active stabiliser of the spine and that it acted on a feed forward neural loop. This meant that transverse abdominus activated prior to the body moving. From these data the

authors concluded that the central nervous system caused transverse abdominus to activate and stabilise the spine, thereby providing a stable base from which the limbs were able to act.

Once Hodges and colleagues had demonstrated that transverse abdominus was an active spinal stabiliser the next question was whether the muscle reacted to load variation. Hodges and Richardson (1997c) utilised limb speed to increase and decrease load on the spine. They found that at high and intermediate speeds transverse abdominus was active whereas at low speed the muscle did not increase levels of activation.

Hodges and colleagues hypothesised that transverse abdominus functions as a spinal stabiliser in conjunction with the diaphragm and the pelvic floor by increasing intra-abdominal pressure (Hodges & Richardson 1997a; Hodges, Sapsford & Pengal 2007). Hodges and Richardson conducted a study comparing the muscle activation timing of transverse abdominus and the diaphragm. Subjects performed rapid arm movements while connected to an electromyograph. The authors reported that these muscles activated simultaneously when the arm was moved. In another study, Hodges, Sapsford and Pengal found that pelvic floor muscle activity occurred in advance of the prime mover of the arm (deltoid muscle) when the arm was moved. The authors described this muscle activity as a "component of the pre-programmed anticipatory postural activity"(pp362).

In a further study to investigate the function of transverse abdominus Hodges and Richardson (1999b) compared the reactions of the prime mover of the shoulder (deltoid muscle), superficial trunk muscles and transverse abdominus during arm movements. In this study subjects were given specific instructions as to what to expect. However, these instructions were deliberately disorganised; the signals used provided correct information, no information or incorrect information. In the conditions where the instructions were not given or were incorrect the reaction times of the deltoid and the superficial trunk muscles were delayed. In contrast, the activation of transverse abdominus was unchanged. In Hodges' review (1999), he stated that these findings "suggested that transverse abdominus was controlled independently of the other trunk muscles and provided further evidence that this muscle contributed to stiffness of the spine in a general manner"(pp82).

These studies demonstrated that transverse abdominus does play a role in spinal stability. However, all of these studies were conducted on normal healthy individuals. The question then asked was, did low back pain affect the function of transverse abdominus?

Hodges and Richardson conducted two studies (1998; 1999a), in which they had subjects perform similar arm movement patterns to the studies that had been previously conducted and from which normative information was obtained. However, in this latter series of studies all subjects were suffering from low back pain. The authors found that the muscle firing patterns were significantly altered compared with normative data. They reported that the onset of activation for transverse abdominus was delayed, the level of activation was reduced, the activation was more phasic, the threshold for activation was increased and transverse abdominus no longer activated independently of the superficial trunk muscles.

The conclusions that can be taken from these studies (Hodges & Richardson 1998, Hodges & Richardson 1999a; Hodges et al. 2007) are that transverse abdominus does play a role in spinal stabilisation and that it can be affected by low back pain. A study by O'Sullivan, Twoomey and Allison (1997) utilised transverse abdominus training in a group of spondylolisthetic/spondylytic patients. This was a randomised controlled trial in which the control group underwent usual medical care while the trial group underwent a specific deep abdominal training regime. The trial group demonstrated reduced pain and increased function that was still present 30 months after intervention. Actual transverse abdominus function was not specifically measured in this study, however, the results of this study provided indirect support for the relationship of transverse abdominus and low back pain.

To further demonstrate the complexity of this topic, Crisco and Panjabi (1991) studied the lateral stability of the spine. In this study it was found that the multisegmental global muscles provided the primary spinal stability relative to the local intersegmental muscles, for this motion pattern. Other research has demonstrated that global muscles provide the bulk of spinal stability; O'Sullivan commented on this in a paper published in 2000, studies by Kavcic et al. (2004a) and Panjabi et al. (1992a, 1989) have provided evidence of this.

O'Sullivan (2000) stated that whilst the global muscle system provided the spine with the bulk of its stiffness and stability, the local muscle system was necessary to maintain segmental stability. The consensus within the literature supported Panjabi's (1992a) original concept of spinal stability being dependent upon the osseoligamentous spine, supporting muscles, and neural control system If any part of these three subsystems was affected, the whole of the system may be affected.

Recent research on the topic of low back rehabilitation programmes has provided some guidelines to assist the clinician in formulating appropriate treatment programmes. McGill and O'Sullivan both advised that pain provocative motor patterns needed to be corrected before any resisted load was applied (McGill & Karpowicz, 2009; O'Sullivan, 2000). Exercises that generated high muscle activation levels but that exerted low loads through the spine, were preferable (Cholewicki & McGill, 1996). Spinal stability was dependent upon numerous actions that took place within the passive and active subsystems at any single instant in time (Kavcic et al., 2004a). Therefore focussing on just one area, or one muscle group, in the hope that this will resolve a problem was probably not going to succeed (Kavcic et al., 2004a).

Kavcic, Grenier and McGill (2004b) investigated a series of exercises to quantify their muscle activation levels and lumbar spine compression loads, in relation to their stabilising effect. The authors separated these exercises into four categories, 1) high compression/high stability, 2) lower compression/lower stability, 3) higher compression/ moderate stability, and 4) lower compression/lower stability. They stated that clinical knowledge of this information assisted clinicians in treating patients. Depending on the irritability of the patient's injury, this information allowed the clinician to judge whether low load, slower progressions should be made, or whether an athlete required high stress more demanding exercises in preparation for return to explosive sport.

Suni et al. (2006) conducted a RCT with a one-year follow-up using exercises that aimed to reduce the size of the neutral zone. The 106 participants were all railroad workers in Finland. The trial group exercised twice a week (one session was supervised), while the control had no intervention and did their 'own usual physical activity' (with or without exercise). The trial group reported a reduction in pain intensity and improved expectations about their future work abilities. However, there were no improvements in trunk muscle endurance, or flexibility. The authors suggested this was due to insufficient training stimulus. Even though participants were meant to complete the exercise programme twice a week, the authors reported that compliance reduced over the long-term. Because there was no difference in the trunk muscle endurance and flexibility categories, the authors deduced that the improved outcome results were due to a reduction in the size of the neutral zone. This conclusion remains theoretical, as the neutral zone was not quantified in these participants. This research applied the theories based on biomechanical modelling directly to the real-life situation.

The spine is a mobile structure with a large range of motion and the relationship between spine motion and spine stability is complex (McGill et al., 2003). The fact that direct measurement of in-vivo human specimens is not available to researchers means that much of this research is based on biomechanical models. All authors in this area acknowledge the limitations related to the use of biomechanical models (Allison et al., 1998; Cholewicki & VanVliet, 2002; Kavcic et al., 2004b; McGill, 1991; Panjabi et al., 1989; Wilke et al., 1995).

There is consensus among researchers that dysfunction in the active or passive subsystem described by Panjabi (1992), might lead to lumbar spine instability; that in turn may lead to pain. To resolve this problem, the spine needs to be stabilised; there are exercise-based regimes capable of doing this, but there is no panacea. There has been debate in the literature as to which muscles are the most important. More recent research has proposed that no single muscle group is more important than the other. The most important outcome is achieving spine stability using programmes individualised for the patient. It is debatable as to whether specific exercises provide superior long-term benefit compared with general exercise programmes (Ferreira et al., 2007; Kitteringham, 1996; Manniche, Asmussen et al., 1993; McGill & Karpowicz, 2009; O'Sullivan, 2000). In relation to general versus specific exercise programmes, a key issue is the underlying diagnosis or classification. The critical aspect for the clinician is deciding what to use and when; further research in this area should help provide the answers.

In this section the principles of rehabilitation programmes for low back problems have been discussed. Most of these principles do not significantly utilise gymnasium equipment. This thesis was based on a programme using gymnasium equipment; therefore, it is appropriate to discuss the different variables that apply to resistance training using gymnasium equipment.

2.12.5 Methods of strength training

There are a wide variety of methods and equipment available for use in resistance training. Some of these are, free form body weight exercises, elastic bands, Swiss balls, machine pin stack weights, hydraulic resistance machines, free weight bars, and dumbbells (Baechle & Earle, 2000; Burke, Culligan, Holt, & MacKinnon, 2000). Of the many different forms of

applying resistance, although each method requires a specific technique, the key principles of strength training remain the same for all.

There are two broad categories that apply to equipment-based resistance training, namely, fixed form and free form exercises. Fixed form exercises use equipment that offer a more stable base, such as machine resistance equipment. Free form exercises use a more unstable base, such as free weights, or the use of unstable platforms (Behm & Anderson, 2006; Spennewyn, 2008).

There are advantages and disadvantages for both forms of exercise. Advantages for machine resistance equipment are, 1) safety, as less skill is required to maintain control of a machine weight compared with a free weight, 2) ease of use, 3) higher levels of confidence using machine weights, 4) the selection of resistance levels is made easy with a pin stack, and 5) the resistance levels are easily quantifiable (Baechle & Earle, 2000; Behm & Anderson, 2006). The disadvantages are, 1) the equipment is often expensive to use or purchase, 2) the lack of muscle co-contraction, 3) a lower requirement of proprioception and balance input, 4) a tendency to isolate muscle groups, 5) linear movements that do not reflect functional activities, and 6) the possibility of poor technique (Behm & Anderson; Spennewyn, 2008).

The advantages of free form exercises are, 1) use of whole body parts requiring coordination, 2) co-contraction and synchronization of muscle groups, 3) increased gains in strength and balance, 4) easy to access (often can be performed in the home environment), and 5) the ability to more closely mimic functional requirements (Behm & Anderson, 1995, 2006; Milner-Brown, Stein, & Lee, 1975; Rutherford & Jones, 1986; Spennewyn, 2008; Stone et al., 1998). Disadvantages of free form exercise include the requirement of higher skill and the risk of injury, particularly if using weights on unstable surfaces or performing complex exercises, for example, an olympic clean-and-jerk lift (Baechle & Earle, 2000; Behm & Anderson, 2006). According to Behm and Anderson (2006) an examination of the literature concluded that a combination of fixed and free form exercises provided the best results in terms of musculoskeletal health and rehabilitation.

The literature has reported that resistance training offers a number of cardio-respiratory, musculoskeletal and rehabilitative benefits (Kraemer, Ratamess, & French, 2002; Pollock et al., 1998). It has also been reported that general resistance exercise programmes could be

beneficial for low back pain patients (Ferreira et al., 2007; Kell & Asmundson, 2009). This thesis study made use of gymnasium resistance training equipment and free weights, therefore, this discussion will focus on training with this specific equipment.

The key determinants of weight training programmes are volume and intensity. Volume relates to sets and repetitions used; intensity relates to the percentage of repetition maximum used. A repetition is one movement of the exercise. A set is the number of repetitions performed between rest periods. For example, 'three sets of ten repetitions' means that the trainer performs ten repetitions, then has a rest for a prescribed amount of time; the trainer completes a second set of repetitions, rests, and completes the final third set of repetitions. Repetition maximum is a commonly used term; it refers to the quantity of load that can be moved in one maximal movement. Repetition maximum is abbreviated to 1 RM or RM.

The three primary functions of weight training are, to improve muscle endurance, and increase strength and power (Dorgo, King, & Rice, 2009; Kraemer & Ratamess, 2000). Many purist strength trainers do not adhere to the philosophy of using weight training to improve endurance, as it is argued that endurance is better obtained using sport or function-specific tasks (Naclerio, Colado, Rhea, Bunker, & Riplett, 2009, Rhea et al., 2008). Strength is the ability of the muscle to resist force, or be able to move bigger loads (Cronin, 1997). Power equals force times velocity (McGill, 2007). Power training is used to develop explosive speed; it is primarily sport-specific but could also be required for some occupations (Naclerio et al.).

There is consensus in the literature that sets of 15 or more repetitions achieve muscle endurance more than strength and power (Campos et al., 2002; Kraemer, Adams et al., 2002; Kraemer & Ratamess, 2004). There is much debate in the literature regarding the optimal number of sets and repetitions. There are two main sides to this debate; they are multiple sets versus a single set.

2.12.5.1 The single set paradigm

According to Smith and Bruce-Low (2004) the single set paradigm is largely based on the work of Arthur Jones who authored many early articles on resistance training protocols. Some of Jones' articles are listed in the paper by Smith and Bruce-Low. The articles listed were not published in peer reviewed scientific journals which made them difficult to source for review. Jones subscribed to the philosophy that one set of 8 - 12 repetitions per exercise was all that

was required to optimise strength and power gains. This theory is supported by many authors on this topic (Carpinelli & Otto, 1998; Winett, 2004; Winett & Carpinelli, 1999).

A meta-analysis on this topic reported 48 papers that supported the theory that single set training was advantageous when compared with multiple set training (Winett, 2004). Smith and Bruce-Low (2004) published a paper that critiqued many studies supporting the theory of 'multiple' sets. Step by step they dismantled the 'science' behind these papers and concluded that there was no evidence for the efficacy of high volume, multiple set training regimes.

A paper published by Munn, Herbert, Hancock and Gandevia (2005) found that three sets produced superior results compared with one set. This paper was critiqued by Winett (2006). Winett used a variety of different statistical analyses on Munn's data, all of which demonstrated an advantage for the multiple set training groups. Winett concluded that there was no evidence to support the benefit of multiple set training over single set; this is quite paradoxical. Authors that support single set training write in quite a passionate manner, inviting suspicion of bias.

In 1998, the American College of Sports Medicine wrote a position stand on "The Recommended Quantity and Quality of Exercise for Developing and Maintaining Cardio-respiratory and Muscular Fitness and Flexibility in Adults" (Pollock et al., 1998. p. 975). Winett and Carpinelli (1999) stated that this position stand was considered a landmark document, "because of its emphasis on different training modalities, the specification of frequency, intensity, and duration of exercise, and the recognition of a broad continuum of exercise doses that produce health benefits" (p. 916). Winett and Carpinelli questioned the research behind some of the statements made in the position stand, particularly in reference to multiple sets versus single set resistance training.

2.12.5.2 The multiple set paradigm

The 'position stand' from the American College of Sports Medicine (Pollock et al., 1998), endorsed multiple set training. Multiple set training is where more than one set of repetitions of the same exercise is performed during one training session. For example, athletes will complete eight different exercises; for each exercise they will complete 3 sets of 10 repetitions per exercise, all in the same training session. This 'multiple set' philosophy is currently the most common training paradigm; it has recently been endorsed by the American College of Sports Medicine (Ratamess et al., 2009).

Within the multiple set philosophy it is generally accepted that sets of repetitions ranging between 8 – 12, hypertrophy muscle, and sets between 1 - 10 primarily strengthen muscle (Atha, 1981; Campos et al., 2002; Kraemer & Ratamess, 2004; McDonagh & Davies, 1984). However, there is a multitude of different set and repetition compositions that are used within the multiple set philosophies. The 'best' combination of sets and repetitions for achieving maximal muscle strength and bulk is a common topic of research.

Examples of different combinations or systems of training are summarised in an article by Henry Drought (1992). A more recent protocol, not mentioned in the article by Drought, is the 'Daily Undulating Programme' described by Newton et al. (2002). This protocol used variations in volume and intensity over a 7 - 10 day period; it attempted to train various components of the neuromuscular system (strength, power and endurance) within that 7 - 10 day cycle.

2.12.5.3 Periodization

The concept of 'Periodization' was first described by Selye (1976). Selye used the term 'General Adaptation Syndrome' to describe the human body's process of adaptation to any form of stress; the stress might be a pathogen, or an external mechanical force such as, resisted exercise performed on a regular basis. Selye's 'General Adaptation Syndrome' consisted of three separate phases, namely, alarm, resistance and exhaustion. In reference to resistance training, Kraemer and Ratamass (2004) described these three stages as shock, adaptation and staleness.

Shock is the body's initial reaction to a stimulus; this is not always performance enhancing. Often this is associated with increased lactic acid production and microtrauma that cause muscle stiffness and inhibit functional performance. Adaptation occurs when the body adapts to the stimulus. In this phase, the body devises ways of managing the stimulus. Specifically the body's metabolic, hormonal, and muscle systems increase their activity levels, leading to improved muscle function. This is a performance-enhancing phase. Staleness occurs when the body has adapted; the stimulus no longer represents a threat, and the body remains at status quo.

Periodization is universally accepted in strength training as an essential component for 'progressive goal attainment.' In strength training terminology, 'progressive goal attainment' is shortened to 'progression'. Many studies have proven the advantage of periodized

programmes compared to non-periodized programmes for achieving progression (Kraemer, 1997; Kraemer et al., 2003; Kraemer et al., 1997; Marx, Ratamess, & Nindl, 2001; O'Bryant, Byrd, & Stone, 1988; Stone, O'Bryant, & Garhammer, 1981; Stone et al., 1997; Stowers et al., 1983; Willoughby, 1993). Progression is achieved using three key factors, progressive overload, variation, and specificity (Kraemer & Ratamess, 2004).

Progressive overload can be achieved by:

- Increasing load
- Increasing repetitions
- Varying repetition speed
- Varying rest periods
- Increasing volume
- Any combination of the above

Variation means the use of a variety of different training methods. These can be any combination of different sets and repetitions, such as pyramid sets, super sets, circuit sets, and different training phases (endurance, hypertrophy, strength).

Specificity refers to the detail in the programme. A beginner weightlifter should start by using a general, holistic, non-specific programme whereas elite athletes require specific programmes. These programmes are very specific in terms of sets, repetitions, loading levels, and rest periods; they often require high levels of skill. There is a continuum from beginner to elite whereby over a period of years, the beginner develops into the elite. Along this continuum, the athlete’s programme gradually becomes more specific (Kraemer & Ratamess, 2004).

Whilst progression is the primary goal, periodization is the means of achieving this goal. In reference to strength training, periodization consists of three separate cycles. These are the macro-cycle, meso-cycle and micro-cycle (Smith, 2003; Stone, 1990).

Macro-cycles are measured in months (or up to one year); meso-cycles relate to weeks; micro-cycles relate to days. Macro-cycles involve different stages of strength training, such as endurance, strength and power. Meso-cycles focus on a specific time period in the macro-cycle. A meso-cycle may relate to how the endurance phase of training is designed. The

microcycle is the specific daily design of the training programme within that mesocycle (Smith, 2003; Stone, 1990).

Macro-cycle changes incorporate different aspects of training, such as aerobic training, sport-specific training, or function-specific training. The timing of each different phase used in a macro-cycle programme is important. This can relate to how long each phase is performed, or it can relate to when (during the year) each phase is brought into the programme (Smith, 2003; Stone, 1990).

Examples of meso-cycle variation are, split programmes compared to daily undulating programmes, changes in sets and repetitions, or changes in exercise selection. Micro-cycle variation may include changes in sets and repetitions, speed of movement, or type of action (meaning concentric, eccentric or isometric). There are a variety of different methods available. The key is variation; variation leads to progression (Fleck, 1999; Stone et al., 1981).

Kraemer and Ratamess (2000) stated that there are numerous methods to periodize training programmes. They reported that one particular model of periodization, used by eastern European weightlifters, has been the focus of several investigations and is known as the 'classic model of periodization'. This model consists of five primary phases; hypertrophy, strength, power, peaking, and active rest. The hypertrophy phase builds muscle mass and tolerance; this is in preparation for subsequent mesocycles that require higher intensity and lower volume, and focus on increasing strength and power. The peaking phase is used to maximise strength and power prior to competition. The active rest phase is a recovery phase after a period of heavy training or competition; it is designed to prevent overtraining.

Table 2.5
Training volume and rest intervals for all phases of training for the 'classic model of Periodization'

	No. of Training Sets	Repetitions	Load % RM	Rest Intervals
Conditioning	3 - 4	10 - 20	65%	30 - 40 seconds
Hypertrophy	3 - 5	8 - 12	70 - 80%	1 - 2 minutes
Strength	3 - 5	6 - 10	75 - 85%	2 - 3 minutes
Absolute Strength	3 - 5	1 - 3	85 - 95%	5 minutes

Power	3 - 5	3 - 5	30 - 80%	2 - 3 minutes

The description of each phase outlined in Table 2.5 is an example of using guidelines from the literature (Cronin & Crewther, 2004; Kraemer & Ratamess, 2004; Stone et al., 1998). Experienced weight trainers whose musculo-skeletal systems have been conditioned to sustain very heavy loads, are the only population group that should perform at absolute strength (Kraemer & Ratamess; Ratamess et al., 2009; Stone et al., 1981). The power phase also requires a high level of skill; it should not be performed by inexperienced trainers (Baechle & Earle, 2000; Ratamess et al.). The wide range of percentage RM used in the power phase is because the percentage RM depends on the exercise being performed (Baker, Nance, & Moore, 2001a, b; Mcbride, Triplett-Mcbride, Davie, & Newton, 2002; Wilson, Newton, Murphy, & Humphries, 1993).

In summary, the conditioning phase uses light loads and high repetitions. The hypertrophy phase requires moderate to heavy loads, and a medium number of repetitions. The strength phase requires heavy loads and low repetitions. The power phase uses loads between 30 - 80%, with 3 - 5 repetitions, depending on the exercise; absolute strength uses very heavy loads, and low repetitions between 1 and 6.

2.12.5.5 Training time frames

Depending on the training goal, time frames for each individual phase can vary considerably. The literature demonstrates that using the same programme for more than eight weeks, leads to staleness, and a plateau of progress (Kraemer & Ratamess, 2004; Stone et al., 1998). Acceptable time frames for individual programmes are between four to eight weeks (Stone et al.). This allows for the shock and adaptation phase of periodization, without entering into the staleness phase. Once muscle stiffness has been overcome in the ‘shock phase,’ initial gains in the first four weeks of any programme are due to neural adaptation (Hakkinen et al., 2003; Hakkinen & Hakkinen, 1995; Hakkinen et al., 1991; Kraemer et al., 1995; Phillips, 2000; Ploutz, Tesch, Biro, & Dudley, 1994; Staron et al., 1994). After four to eight weeks, further improvements in strength are considered to relate to muscle hypertrophy (although improved neural activity is still a factor) (Kraemer et al., 1995; Phillips, 2000; Staron et al.).

Untrained individuals responded better to initial training than trained individuals (Hakkinen, Komi, Alen, & Kauhanen, 1987; Hakkinen, Pakarinen, Alen, Kauhanen, & Komi,

1987, 1988a, b; Ratamess et al., 2002; Rhea, Alvar, Burkett, & Ball, 2003). Kraemer and Ratamass (2004) reviewed the literature and stated that muscular strength increases approximately 40% in the 'untrained', 20% in the 'moderately trained,' 16% in the 'trained,' 10% in the 'advanced,' and 2% in the 'elite,' over periods ranging from four weeks to two years. Due to the significant gains in the initial stages of training for 'untrained' individuals, less programme variation is required (Stone et al., 1998). This has the added advantage of not over-complicating early programmes; it allows the beginner trainer to develop the confidence to exercise in a new environment. Periodized programmes should still be used to optimise progress (Stone et al., 1981).

2.12.5.6 Rest Periods

The final aspect of periodization is rest periods. These are crucial for maintaining a fresh attitude, optimising progression, and avoiding injury (Stone, 1990). Rest periods fit into the macro-cycles, meso-cycles, and micro-cycles of any periodized programme.

In the macro-cycle, rest periods can range from weeks to months; these can be active or passive. An active rest period is where the trainer uses a different training method to train one aspect of the body, whilst resting another. For example, the trainer stops an upper body strength programme for a month, and focuses on aerobic fitness, using a running programme. A passive rest period is where the trainer has a complete rest from all exercise for a specified time period.

Rest periods in the meso-cycle are often (but not always) associated with the frequency of training days per week. The authors who support single set training, propose training one muscle group to failure 1-2 times per week (Smith & Bruce-Low, 2004). Supporters of multiple set training promote 2 - 3 sessions per week for the novice, 3-4 sessions per week for the intermediate and 4-5 sessions per week for the advanced trainer (Ratamess et al., 2009). Heavy training loads are usually associated with twice-weekly training sessions; lower loads are usually associated with three times per week sessions (or training every second day). Light endurance sessions can be performed up to 5 - 6 days per week. The days between sessions can be used as passive rest/recovery days, or as active rest days, by incorporating some other form of exercise (running, cycling, Pilates, Yoga). Split programmes are also a form of rest, training one body part whilst resting another.

In the micro-cycle, rest periods are the intervals between sets. Kraemer and Ratamass (2004) provided a guideline as follows: 30 - 40 seconds rest between sets in the endurance phase; 1 - 2 minutes rest between sets in the hypertrophy phase; 2 - 3 minutes rest between sets in the strength and power phases; and up to 5 minutes rest between sets in the absolute strength phase (Kraemer, Adams et al., 2002; Pincivero, Lephart, & Karunakara, 1997; Robinson et al., 1995). Kraemer (1997) conducted a study with college football players who performed leg press and bench press exercises at 10RM; initially these were performed with three-minute rest periods between three sets, then with a one-minute rest between sets. Using the three-minute rest period, all participants completed three sets of ten repetitions. When the rest period was decreased to one minute, the mean number of repetitions per set decreased from the initial ten to eight, then seven for the last set. In strength phase, studies have demonstrated superior strength gains using longer rest periods between sets; three minutes was proved to be optimal (Pincivero et al.; Robinson et al.).

2.12.6 Principles of exercise prescription for the injured person

Exercises must not aggravate or increase pain; the ‘no pain no gain’ attitude is not appropriate for the injured person (McGill, 2007). Studies have demonstrated that pain reduces compliance in relation to exercise programmes (DeAnna et al., 2006; Der Ananian, Wilcox, Saunders, Watkins, & Evans, 2006). DeAnna et al. investigated factors that predicted compliance in Gulf War veterans affected by fatigue, musculoskeletal pain and cognitive problems. The strongest predictors of compliance in the follow-up period were less pain and greater age. Der Ananian et al. studied the compliance to exercise programmes among people with arthritis of the knee. They found that pain was the most commonly mentioned barrier to exercise. These studies demonstrate that when prescribing exercise programmes for patients with injuries, it is important that the exercise programme itself does not cause or increase pain, as this may lead to reduced patient compliance.

While it is important that exercise programmes do not cause pain, a degree of muscle stiffness due to microtrauma is acceptable. Microtrauma, due to increased loading on muscles, causes protein break down within the muscle filaments which leads to an increase in actin and myosin filaments. This process along with sarcomere addition, leads to muscle hypertrophy (Baechle & Earle, 2000). Microtrauma is a normal physiological side effect of exercise. It occurs when muscles adapt to external loading that leads to increased strength (Zatsiorsky, 1995).

Explanation of microtrauma avoids unnecessary fear or surprise when patients experience the resultant sensation of muscle stiffness (Friedrich, Gittler, Arendasy, & Friedrich, 2005). Friedrich et al. conducted a study with chronic low back pain patients comparing a usual exercise programme with the same exercise programme plus a cognitive motivational educational programme. The study demonstrated a long-term benefit in favour of the exercise plus motivation group; this was statistically significant in terms of work ability at five years post-intervention. The work by Friedrich et al. highlights the importance of patient education when prescribing exercise programmes.

Selected exercises must be simple. The exercise movements should be away from and towards the body (Wise et al., 2004). Long lever exercises and complex exercises that require high levels of skill should be avoided (Kraemer & Ratamess, 2004). Theoretically, large multi-joint muscle groups should be exercised first, followed by small single joint muscles. This order of exercise enhances strength gains (Kraemer & Ratamess; Ratamess et al., 2009; Stone et al., 1981).

It is important to avoid targeting the injured area in the early or acute phase. The injured area could be any region of the body, depending on the injury site. In the case of the study investigated in this thesis, it was the low back. In this situation by training the upper and lower body, the lower back and core of the body strengthen by providing a stable base, while the upper and lower body exercises are executed. This concept is evident in the research conducted by Hodges & Richardson (1997) and Hodges, Cresswell, & Thorstensson (1998). In this research Hodges et al. studied the effect on core abdominal muscles when an upper or lower limb is moved. These authors provided evidence indicating that core muscles are exercised indirectly, when the upper or lower limbs are moved or exercised.

Ploutz, Tesch, Biro and Dudley (1994) completed a study where one quadriceps muscle of the participants was trained twice per week for six weeks; the other quadriceps muscle was used as the control muscle. They found that after six weeks, the quadriceps muscle that had been trained had increased its 1 RM strength by 14%; the quadriceps muscle that had no training increased its 1 RM by 7%. The authors explained this objective gain in the untrained muscle due to increased neural activity in the cortex of the brain. Neural brain activity in the motor cortex is an important component of muscle strengthening (Enoka, 1988; Hakkinen & Hakkinen, 1995; Schantz, 1983). This research demonstrated that it is possible to indirectly work one area of the body, whilst focussing specifically on other body areas. In exercise

programmes, while the exercises used may focus on the upper or lower body, they may still have an indirect effect on the core or low back areas of the body.

Warm-ups are important for preparing muscles prior to exercise; warm-downs prevent excessive lactate irritation (Beedle & Mann, 2007; Holt & Lambourne, 2008; Stewart & Sleivert, 1998; Tessitore et al., 2008). By increasing blood flow, the body experiences a flushing effect that eliminates excess lactic acid from the blood (Tessitore et al.). In the hypertrophy and strength phases, the pre-warm-up sets are important for preparing the muscles to lift or push heavy weights (Davis, Wood, Andrews, Elkind, & Davis, 2008). Using a pre-warm-up set, the muscle begins the process of motor unit recruitment. This is known as post-activation potentiation (PAP) (Robins, 2005). PAP is most commonly used in complex training, where heavy resisted exercise is used in combination with an explosive or pliometric functional exercise such as, a vertical jump (Weber, Brown, Coburn, & Zindor, 2008). By performing a pre-warm-up set of higher repetition and lower weight, PAP better prepares the muscle to withstand greater force.

Correct technique is essential when performing exercises (Zatsiorsky, 1995). Sometimes when a patient is working hard to lift a weight, there is a tendency to use ‘cheat’s tricks’. ‘Cheat’s tricks’ are where the patient will put his or her body weight behind the movement by twisting or leveraging the body. ‘Cheat’s tricks’ cause injuries and should not be permitted. If a patient has to use ‘cheat’s tricks’ to lift the weight, the weight is too heavy, and should be decreased (Zatsiorsky).

Injured patients must ‘train not strain’. Whilst performing each exercise, patients should ‘brace their abdominals’ as described by Grenier and McGill (2007). They demonstrated that this abdominal brace technique provided the lumbar spine with maximal spinal stability, and greater stability when compared to the abdominal hollowing technique as described by other authors (Hodges, 1999; Urquhart, Hodges, Allen, & Story, 2005). The Grenier and McGill study is also supported by Brown, Vera-Garcia and McGill (2006). By bracing the abdominals during each exercise, the patient is subconsciously training these muscles to function in this way in everyday life. It is hypothesised that when they are performing activities of daily living, these muscles are more likely to automatically contract, and confer support to the spine. However, no literature could be found concerning the hypothesis that functional training of the abdominals, leads to an improvement in functional performance of the lumbar spine.

Finally, it is important to train with symmetry. This means that all exercises should be performed bilaterally rather than unilaterally. Bilateral exercises tend to balance the forces through the spine. Unilateral exercises lead to imbalanced forces being applied to the spine. This occurs if patients allow their bodies to twist in order to leverage their bodyweight behind the lift. As stated, this is a 'cheat's trick'; it can lead to stress and strain on the body. Bilateral exercises limit this tendency. There is sometimes a slight advantage in performing an exercise either unilaterally or bilaterally in terms of strength gain. However, research has demonstrated that this advantage goes both ways (depending on the exercise being performed), and overall, there is no significant difference (Hakkinen, Kallinen, Linnamo et al., 1996; Hakkinen, Kraemer, Kallinen et al., 1996).

2.13 Summary

A review of the literature indicated increasing rates of lumbar surgery. Lumbar discectomies comprise 70-90% of all lumbar surgery. This correlates with an increasing cost to funding agencies and government budgets. It is important to insure this increasing trend in surgical rates delivers value for money.

Only one study on the topic of lumbar discectomy has attempted to report on its cost effectiveness (Malter et al., 1996). This study study was based on historical data. No prospective studies in the literature report cost effectiveness for lumbar discectomies; such a study would be a valuable addition to the knowledge base. The literature reports that approximately 90% of the total cost of lumbar discectomy is related to indirect, rather than to direct medical costs. Hence, when analysing the total cost effectiveness for lumbar discectomy, these indirect costs should be included. Further research is needed to determine the cost effectiveness of lumbar discectomies.

Despite the gap in the knowledge base regarding the cost effectiveness of lumbar discectomy, this surgical procedure is common, and according to the literature, is becoming more popular. Some authors ask the question, which is the most appropriate course of action for patients with lumbar disc injuries; lumbar discectomy or conservative physiotherapy rehabilitation programmes? Analysis of the literature indicates an early advantage in the first year post-surgery for lumbar discectomy that gradually balances out over subsequent years.

This early advantage post-surgery allows patients to return to normal function faster than conservative therapy. However the literature also reports a 10-40% complication rate post-surgery. Therefore, interventions that reduce post-surgical complication rates will be of benefit.

Post-surgical complications have been categorised under three headings, namely, biological, psychological and social. A biological complication refers to an event or consequence from the operation (such as a dural tear or post-surgical infection). Psychological complications relate to the patient's state of mind, for example, depression, catastrophisation, fear avoidance, and attitudinal factors (such as the burst and bust mentality). Social complications involve the patient's own social interactions; these mainly involve the influence of family, friends and work environments.

Where a complication becomes a problem is less well defined, and depends on the person assessing the issue. For example, researchers require a defined outcome measure to achieve statistical significance, clinicians may focus on functional outcomes and insurance assessors consider issues of cost. A variety of outcome measures have therefore been devised that measure different aspects of post-surgical complications.

Closely associated with outcome measures are 'predictors of poor outcomes post-lumbar discectomy'. The ability to predict how well a patient may recover after surgery, helps the surgeon decide on performing surgery in the first instance. If surgery is performed on a candidate with known risk factors, a more comprehensive post-surgical management regime may be indicated. The literature reports psychosocial issues, rather than biological issues, to be the best predictors of post-surgical complications.

Several studies have compared one post-surgical exercise rehabilitation programme with another. All reported that post-surgical rehabilitation exercise programmes improved surgical outcomes. However, only two studies that satisfied the Cochrane review high quality criteria, and that compared an exercise group with a group that did not perform any formal post-surgical rehabilitation programme have been published. This being the case the next question is, what type, or structure of programme will provide the best outcome? To answer this question a knowledge of exercise physiology and exercise prescription is required.

In New Zealand, the current philosophy post-lumbar discectomy is for the patient to mobilise as pain allows, and return to normal function as soon as possible. Formal

physiotherapeutic rehabilitation regimes are not encouraged. Based on a review of the literature, it was decided to test a progressive non-aggravating gymnasium-based exercise programme against usual surgical advice (return to normal activities as soon as pain allows). The exercise regime was a fully periodized programme of six months duration. The participants were followed with validated outcome measures, and an annual questionnaire for three years post-surgery. The goal of this study was to determine whether this exercise programme would provide superior long-term results when compared with usual surgical advice. The specific method used in this study is presented in the following chapter.

CHAPTER THREE
METHODS

3.1 Introduction

Chapter Three describes the methods undertaken to conduct a randomised controlled trial that investigated the post-surgical management of lumbar discectomy. The process to obtain ethical consent is outlined; this is followed by a description of the sample population. This describes the inclusion and exclusion criteria, the recruitment process, statistical analysis, and the collection of outcome data. Section 3.6 outlines the instructions provided to control group participants. Section 3.7 describes the surgical technique used for all participants. Section 3.8 details the instructions used and implementation of the rehabilitation programme for the trial group participants. Sections 3.9, 3.10 and 3.11 provide a detailed description of each component of the rehabilitation programme.

3.2 Ethical consent

Ethical consent to conduct a RCT that investigates the post-surgical management of lumbar discectomy for an 18-month period was made in November 2001; this was approved by the Upper South A Regional Ethics Committee in November 2002 (Appendix 1). Previous studies on the topic of post-surgical management of lumbar discectomy patients had followed participants for a maximum of one year. Participants in this study diligently completed and returned their outcome-measure documentation. At 18 months, it was decided to extend data measurements for a further 18 months. This potentially could provide worthwhile information and enhance the body of knowledge on this topic. An application to follow participants in this study for a further 18 months was submitted to the ethics committee and accepted in September 2003 (Appendix 1).

3.3 The sample population

The participant population was sequentially selected from the surgical list of one orthopaedic surgeon over a 20-month period that extended from February 2002 to October 2003. One hundred and thirty-nine patients were screened for the study. Ninety-four of these 139

satisfied the inclusion criteria. Fifty-five male and 39 female patients were selected; they ranged in age from 17 years to 64 years.

3.3.1 Inclusion criteria

- Patients undergoing lumbar discectomy (level of operation that involves discs at L3/4, L4/5, or L5/S1 levels).
- Patients should not have received weekly compensation for a period of more than two weeks, for a previous injury that is unrelated to the current situation. Participants may however, have been on weekly compensation for some time due to the injury that has led to this operation.
- Patients aged between 17 and 65 years.
- Patients of good health with no other major medical problems that would potentially prevent the individual from participating in a gymnasium rehabilitation programme.

3.3.2 Exclusion criteria

- Patients younger than 16 years, or older than 65 years.
- Patients with communication difficulties (e.g. neurological dysfunction making language difficult).
- Patients that were physically disabled, and were unable to perform gym-based exercise programmes.
- Patients with central neurological disorders (e.g. cerebral palsy, spina bifida).
- Patients with other major medical problems (e.g. osteoporosis, severe respiratory conditions, arthritic conditions, diabetes, and any medical condition where it was unsafe to prescribe gym-based exercises).
- Patients that were referred for postoperative rehabilitation by their General Practitioner, Surgeon or Case manager.

- Patients having undergone surgery for reason of infection, tumour or inflammatory disease (arthritis, such as ankylosing spondylitis).

3.3.3 Recruitment

At the follow-up appointment two weeks post-surgery, the patient was asked by the orthopaedic surgeon to consider participating in this study. If the patient agreed to consider this option, the next step was an appointment with the principal investigator (PI). A consent form with their contact details (Appendix 2) was signed and sent to the PI.

The PI then telephoned the patient and provided a brief outline of the study. If the patient was still interested in participating in the study, an appointment was made by the PI to meet with the patient. The purpose and requirements of the study were then further explained. The patient was given an information sheet (Appendix 3) to read at their leisure. If he or she consented to participate in the study, the patient signed the ‘Consent Form For Participants’ (Appendix 4). The patient was then considered a participant in the study.

The participants mailed their signed consent forms (Appendix 4) and 'Personal Information' forms (Appendix 4) along with their baseline measures, to research assistant 1 (RA1) at six weeks post-surgery. Patients not wishing to participate in the study advised the PI and returned the incomplete outcome measures and the unsigned consent form.

A computer-generated randomisation table randomised the study population using permuted blocks. This table was produced by the study biostatistician and sent directly to RA1. The table consisted of a series of codes beginning with the letter A or B followed by a three digit number. The letters designated the group allocation and the numbers related to the individual participant. On receipt of the signed consent form and baseline measures, RA1 randomised the participant by assigning that participant to the next sequential code on the table.

The PI was then informed of the participants’ coding. The PI contacted the participants to notify them of their group allocation. If the participants were randomised to the control group (CG), the PI informed them that they would have no further contact with the PI until the study was completed. They would receive a summary of the results at the completion of the study. They were advised that they would have regular contact with the RA1, regarding the sending and receiving of the outcome measures.

Research assistant 1 coordinated the sending and receiving of outcome measures. Once the completed outcome measures had been received by RA1, they were then passed on to the PI. The PI coordinated the scoring of the outcome measures which were completed by another trained research assistant (RA2). Because the outcome measures were coded, RA2 was blinded to the group and participant scores. To ensure accurate scoring by RA2, the PI took random samples of the completed scores and matched them to their codes. Once the outcome measures had been scored, they were entered into an excel data sheet in preparation for statistical analysis using SPSS version 13.0 (Hilbe, 2005).

The outcome measures were recorded at 6, 14, 23, 32, 58, 84, 104, 130 and 156 weeks. The outcome measures used were, the Roland-Morris 24 point Questionnaire (RM), the Oswestry Low Back Index (OLBI), the Short Form 36 (SF36) (Appendix 5) and, a questionnaire (designed by the PI) that focussed on Quality of Life variables (Appendix 6). The latter questionnaire was completed at the end of year three post-surgery; it included questions on patients' satisfaction with both their surgical result and with their post-surgical management. Questions relating to satisfaction were not included in the Quality of Life questionnaire for years one and two. Only the final Quality of Life questionnaire (three years after surgery) included questions on satisfaction.

The questionnaire designed by the PI was completed at 58 weeks, 104 weeks and 156 weeks. All the other outcome measures were completed at 6, 32, 58, 84, 104, 136, and 156 weeks. In addition, the RM and the OLBI were completed at 14 and 23 weeks. The SF36 was not completed by participants during the 14 and 23 week time intervals, due to the PI being conscious of patient compliance and not wishing to overload patients early in the study. The Roland-Morris and Oswestry were quicker to complete than the SF36. It was the PI's understanding that these subjective questionnaires would still provide sufficient functional data at these time points. From 32 weeks after surgery, the outcome measures were completed once every six months until three years post-surgery.

3.4 Statistical analysis

T-tests, Chi-square tests and Mann-Whitney U tests were used to compare baseline demographic and clinical variables between randomised groups, in order to confirm that randomisation had led to two well-matched groups. Repeated measures analysis of variance was used to compare the changes in key outcome assessments between treatment groups at

different assessment times. Additional ordinal and categorical outcome data were compared between groups using Mann-Whitney U tests and Chi-square tests. Associations between functional measures (OLBI, RM & SF36) and Quality of Life scales and subscales were assessed using Pearson's correlation coefficients. A p-value < 0.05 was taken to indicate statistical significance. The error measures on all figures in the results chapter (Chapter Four) represented standard error measures (SEMs)

3.5 A description of the outcome measures used and the between group comparisons performed.

Two types of outcome measures have been used in this study. They were validated subjective functional questionnaires and a non-validated Quality of Life (QoL) questionnaire. The functional questionnaires were the RM, OLBI and the SF36. The QoL was an annual questionnaire that collected data on:

- Doctor visits related to low back pain
- Other therapist visits
- Medication use
- Time off work
- Satisfaction with treatment
- Level of exercise/activity since surgery

Each of these annual questionnaire categories also collected subset information. For example, if a patient visited their doctor for a low back problem, the annual questionnaire asked how many visits that patient had made in the previous three months for this problem. The same is true in the ‘other therapist’ category. In this category the participant was able to name the type of therapist visited (physiotherapist, chiropractor, or osteopath) and the number of visits made. The medication use category asked patients to record the number of episodes during the previous three months in which they had required medication for their low back pain (for a period of two days or more). Asking the participants whether they had used medication for two days or more, was an attempt to obtain more accurate information. This was to discount occasional medication use and to highlight those episodes when the participant’s low back pain had flared up. This remains an assumption and is a limitation of

this questionnaire. The time off work question related to any further time off work taken since returning to work after surgery. The number of days off work was recorded. All time off work data was checked with the participants' medical records. Prior specific consent was obtained for the PI to check participants' medical records (Appendix 8). Participants were also asked whether they had undertaken any form of regular exercise and if so, to provide a brief description of that exercise.

'Satisfaction' was only questioned in the final three-year post-surgery annual questionnaire. The participants were asked how satisfied they were with the result of their surgery and how satisfied they were with their post-surgical management.

The results of each outcome measure will be discussed in Chapter Four. These include the subjective functional questionnaire and annual questionnaire results of the three primary study populations; intent-to-treat (section 4.3), per-protocol (section 4.4) and per-protocol minus operation (section 4.5). Results of the satisfaction questionnaire are included in these summaries.

Two further populations of the total cohort are also discussed:

- Participants within the control group that performed no exercise at all throughout the whole three year trial period (N = 9), compared with the gym rehabilitation group
- Trainer A versus trainer B.

'Trainer A versus trainer B' relates to the two groups in the trial group who were supervised by different physiotherapists. Trainer A is the PI of this study; trainer B is another physiotherapist who is familiar with this specific rehabilitation programme. This comparison is important as, in a limited manner, it tests the generalisability of the rehabilitation programme. However, the author acknowledges that having used only one other physiotherapist does limit generalisability. Trainer A supervised those participants who preferred to use commercial facilities local to where they worked or lived. Trainer B supervised any participants who were able to complete the programme in the rehabilitation gymnasium of the clinic where both physiotherapists worked. These participants lived locally and it was their preference to either attend the physiotherapy rehabilitation gymnasium or a commercial gymnasium in the area. All commercial gym fees were paid by a study grant so that the participants incurred no cost. Participants who completed the programme at the physiotherapy clinic owned by the PI did not pay a fee and no fee was charged to study

funding or any funding institution for these participants. The PI accepted this as a cost of completing the study.

3.6 Instructions given to control group participants.

All communication between the principal investigator and control group participants/patients occurred prior to randomisation, with the exception of a follow-up telephone call by the PI immediately after randomisation, thanking the participant for entering the study. During this conversation, the PI re-affirmed that no further communication would be received from him until the trial was completed. This method was designed to avoid unnecessary chance of bias or contamination of the control group by the PI.

During the recruitment process, all participants were informed that if they were randomised to the control group, they were to follow any medical advice. No restrictions were applied to control group participants regarding their post-surgical management, in terms of whether or not to perform their own exercise regime. Participants randomised to the control group were informed that they were allowed to perform their own exercise programmes but should discuss this issue with their surgeon or their general practitioner first. Some control group participants did exercise after their surgery. This data was collected in the annual QoL survey; these results are reported in Chapter Four.

The control group participants were informed that they would have regular contact with RA1 through the sending and receiving of the outcome measures. Should they experience any problems relating to their low back after discharge from the surgeon, their surgeon or general practitioner must be contacted, and the medical advice followed.

3.7 The surgical technique

The surgical technique used for all operations in this study was a standard open discectomy. Surgery was performed via a small 3 cm incision using the Spengler technique (Balderston et al., 1991). To improve visualisation, an endoscopic light was placed at the end of the main retractors. The endoscopic light pipe and its effect are pictured in *Figures 3.1* and *3.2*.

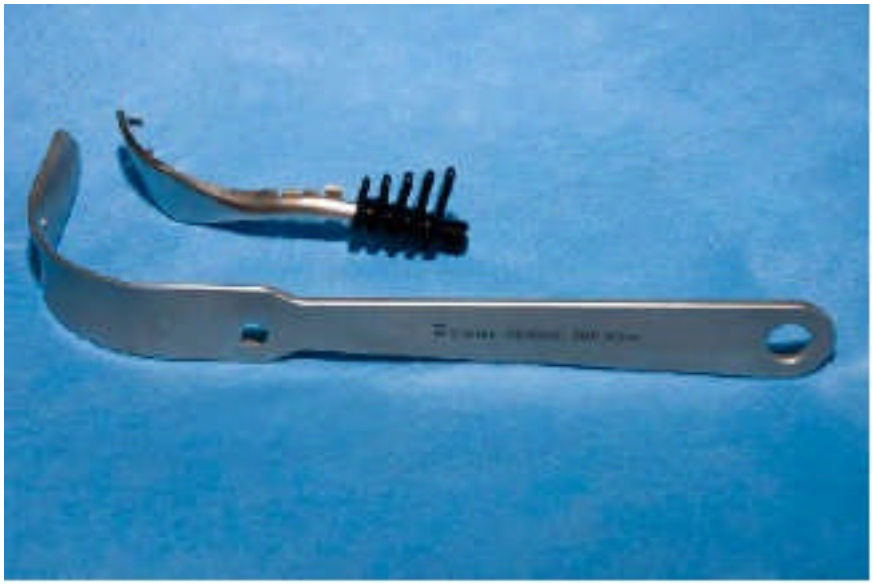


Figure 3.1. The light pipe

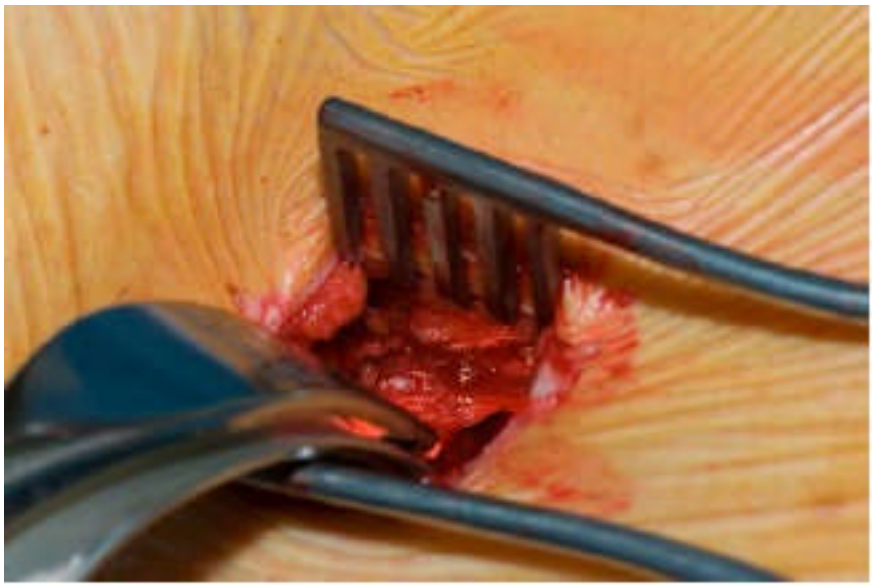


Figure 3.2. The effect of the light pipe in situ

3.8 The method of instruction utilized for the trial group.

Once participants had been randomised to the trial group, the PI arranged an appointment at a gymnasium of their preference, and a membership was purchased on their behalf if the participant attended a commercial gymnasium. The physiotherapy trainers demonstrated the first conditioning programme (Table 3.1) to the participant (Appendix 7) at six weeks post-surgery.

Table 3.1
An outline of all the exercises used in the 26 week long programme, including sets and repetitions for each phase. The Swiss ball exercises used in the programme are not included in this table.

Conditioning 3x 10,12,15	Hypertrophy 12-10-8			Strength 10-8-6		
(8 weeks)	A (3 weeks)	B (3 weeks)	C (3 weeks)	A (3weeks)	B (3 weeks)	C (3 weeks)
Warm up	Warm up	Warm up	Warm up	Warm up	Warm up	Warm up
Seated bench press	Seated bench press	Flat DB flys / Pec deck	Inclined DB press	Seated bench press	Flat DB flys / Pec deck	Inclined DB press
Prone pull	Prone pull	Seated row	One armed pulls	Prone pull	Seated row	One armed pulls
Prone fly	Prone fly	Inclined prone fly	Reverse pec deck	Prone fly	Inclined prone fly	Reverse pec deck
Front pull down	Front pull down	Close grip pull down	Upright Row	Front pull down	Close grip pull down	Upright Row
Leg press	Leg press	Lunges	Backward lunge	Leg press	Lunges	Backward lunge
Pully hamstring curls (standing)	Pully hamstring curls (standing)	Butt kick	Hamstring curl (lying)	Pully hamstring curls (standing)	Butt kick	Hamstring curl (lying)
Dumbbell bicep curl	Dumbbell bicep curl	Bicep curl with a bar	Inclined DB bicep curl	Dumbbell bicep curl	Bicep curl with a bar	Inclined DB bicep curl
Tricep push down	Tricep push down	Elbow extension	DB press	Tricep push down	Elbow extension	DB press
Warm down	Warm down	Warm down	Warm down	Warm down	Warm down	Warm down

Note: DB = Dumbbell

Only when the trainer was convinced that the participant was confident and comfortable to execute all exercises with the correct technique, were they allowed to perform the programme themselves; in all cases this took one session. The participant was introduced to a floor instructor or gym manager, with whom any queries could be clarified. The participants were advised that they had to complete the programme three times per week. They were also taught how to increase the level of resistance themselves. The method of resistance progression taught to the participants is described in Section 3.9.2.1. The participants were advised that should they experience any pain or problems when performing the exercises, the trainer was to be contacted immediately. The importance of the exercises not irritating or aggravating their low back pain was emphasised.

Three weeks after the initial gym appointment, a follow-up appointment was arranged with the trainer. The purpose of this appointment was to check that the participant was confident they were performing the exercises correctly and satisfied with their current

progress. If a participant needed to have an exercise technique checked, the trainer corrected the exercise if required.

This follow-up appointment served to begin the participant's abdominal regime as well. Abdominal training was a component of this rehabilitation programme. If the participant had a reduced lordosis, they were prescribed a deep abdominal training regime, as described by Richardson and Jull (1995). This regime aimed to illicit co-contraction of the transverse abdominus and multifidus muscles. The regime commenced with participants' performing an indrawing of the deep abdominal muscles in a four point kneeling position. Once this action was learned and performed correctly the exercise was progressed to a prone lying position using a biofeedback pressure cuff under the low abdominal region. The pressure cuff measured a decrease in pressure when the the abdominal wall was drawn in. The aim of the exercise was to achieve a decrease of up to a maximum of 10 mmHg on the pressure cuff, to hold this reduction in pressure for up ten seconds and to repeat this manoeuvre ten times.

If the participant had a normal or increased lordosis, a different deep abdominal regime was prescribed, as proposed by Sahrmann (Diagnosis and Treatment of Movement Impairment Syndrome. Shirley A Sahrmann, pp 373-377). This exercise required the participant to lie supine with knees bent and their feet flat on the floor. In this position the participant performed the same abdominal indrawing manoeuvre as described in the Richardson and Jull (1995) exercise. After practising this manoeuvre for approximately one week, or, when the participant could correctly manage this indrawing technique, the exercise was progressed. The progressed exercise required the participant to maintain the pelvic neutral position, with the abdominals drawn in, while extending one leg, lowering it to the floor and then bringing the leg back into the start position. This was then repeated with the other leg. The exercise required the participant to repeat this manoeuvre ten times for each leg whilst holding the pelvic neutral position at all times. These abdominal training regimes were progressed throughout the programme, each time the trainer met with the participant. The decision of the trainer to use the 'Richardson and Jull' regime for reduced lordosis and the 'Sahrmann' regime for normal or increased lordosis, was based purely on the clinical experience of each trainer; this is a limitation of this study and is reported as such in section 5.7 of this thesis.

Abdominal training was in addition to the swissball core stability exercises prescribed (Fig 3.3 & 3.4) and the abdominal brace. The use of the abdominal brace was taught to

participants at the first demonstration of the initial conditioning programme. Participants were taught to brace their abdominals when performing all exercises in the rehabilitation programme. The abdominal brace used in the programme is that described by Grenier and McGill (2007) .

The gymnasium rehabilitation programme was split into three different phases:

- Phase one - Conditioning phase (of 8 weeks duration)
- Phase two - Hypertrophy phase (9 weeks)
- Phase three - Strength phase (9 weeks)

The total duration of the rehabilitation programme was 26 weeks. Therefore, eight weeks after the initial gym programme was demonstrated, the participant attended a session with the trainer, in which the 'hypertrophy' phase was explained.

At this session, the overall structure of the programme was discussed. This allowed the participants to understand their current status and achievements, and what was planned for the rest of the programme.

The participant then progressed through to the 'hypertrophy A'programme (Table 3.1 & Appendix 7). 'Hypertrophy A' is a replica of the conditioning programme in relation to the exercises performed, but the sets and repetitions are changed to 'pyramid style' training. This consists of one set of 12 repetitions, one set of 10 repetitions, and one set of 8 repetitions. As the repetitions are decreased, the weight is increased. Because exercises in the 'hypertrophy A' programme are the same as those used in the conditioning programme; participants were allowed to progress to a higher level of resistance on familiar exercises. They were then changed to a completely new programme of exercises in the 'hypertrophy B' programme.

The 'hypertrophy A' programme was performed three times per week for three weeks. The trainer then demonstrated the 'hypertrophy B' programme (Table 3.1 & Appendix 7), and set appropriate resistance levels. At this session (providing the participant's low back was not causing problems), two Swiss ball exercises were demonstrated (*figures 3.3* & *3.4*) (Corning Creager, 1994). These exercises were then performed in conjunction with the gymnasium rehabilitation programme each time the participant completed the programme. If the participant still had low back pain or irritation, these Swiss ball exercises were not used until

later in the programme, when the participant's back pain had settled. Three weeks later, the trainer returned to the gym to demonstrate and set the resistance levels for the 'hypertrophy C' programme (Table 3.1 & Appendix 7). This programme was performed three times per week for three weeks.

Compliance (performing the programme three times per week) was important. To ensure that sufficient work had been performed prior to progressing the participants to the next step in the programme, the participants were asked to date an attendance card each time they completed an exercise session. The trainer checked these cards at successive appointments. The participants had to complete at least 22 sessions within an eight week time period to complete the conditioning phase. Each programme in the hypertrophy and strength phases was performed nine times in a three to four week time period. The results of participant compliance are discussed in Chapter Five.

Figure 3.3. Shows the Swiss ball 'bridge' exercise where the participant raises and lowers each leg. This exercise consists of 3 sets of 10 leg movements per leg. The three photos show how the exercise is progressed by drawing in the arms, narrowing the exercise base.

Figure 3.4. This illustrates the Swiss ball hip extension exercise. To progress this exercise the arms are brought together in midline to narrow the exercise base. The exercise consists of 5, 5

second holds with the leg in the air and may be progressed to 5, 10 second holds as the participants improved their ability.

After nine weeks 'hypertrophy programmes A, B, and C' were completed; the final strength phase of the rehabilitation programme was then commenced (Table 3.1 & Appendix 7). The exercises in each programme of the strength phase were identical to the exercises of the programmes used in hypertrophy phases. These programmes were also completed three times per week, and three weeks per programme.

The change in this phase related to sets and repetitions. Instead of sets and repetitions of 12-10-8 used in the hypertrophy phase, sets of 10-8-6 are used in the strength phase. By decreasing the repetitions by two repetitions per set, the participants were able to increase the amount of weight lifted. Because the participants were familiar with all of the exercises (after practising them for three weeks in each hypertrophy phase), no further demonstration was required.

During this phase, the participants received one session in the clinic with the physiotherapy trainer that focussed on lifting technique. Participants were taught how to use their own body weight to best advantage when lifting and bending.

After completing 'strength C' programme, the participants received one final session with the physiotherapist, at which maintenance of their strength and fitness was discussed. The participants were encouraged to maintain the strength and fitness achieved during the previous six months. They were encouraged to do this using their own preferred methods, such as yoga, pilates, gym classes, circuit classes, swimming, tennis, or by continuing with their previous gym workouts. They were thanked for their ongoing participation in the study.

3.9 Explanation and discussion of the gym programme

The following is a detailed description of the three components of the physiotherapy rehabilitation programme, with particular emphasis on the gym programme.

There are three main components to the programme:

- A 26 week long gym based progressive non-aggravating exercise programme
- An abdominal regime

- One session relating to lifting and bending techniques

3.9.1 The 26 week physiotherapy rehabilitation programme

The points listed below outline the key issues relating to this programme to help orientate the reader. The exercise science relating to these points has been reviewed in Section 2.12. The structure of the gym rehabilitation programme (exercises, sets and repetitions, time frames) for each programme is outlined in Table 3.1.

- The programme was non-aggravating; at no time should the exercises have irritated or aggravated low back symptoms.
- The programme was progressive; the exercise resistance levels were gradually increased throughout the programme.
- Initially the exercises very specifically avoided placing any direct stress or strain on the low back. The first time the low back performed any stabilising work was in week 15 (exercises; Dumbbell Press & Upright Row; Appendix 7, hypertrophy C); even here the low back was kept in a neutral position. Throughout the whole programme, the participants were taught to keep their low back in a neutral position whilst performing the exercises.
- The initial focus of the programme was to strengthen the upper and lower body. Lumbar discectomy patients have usually had their problem for some time, leading to a generalised physical weakening. When the patient tries to perform any activity of a heavy nature (e.g. lifting and carrying a full wood basket), the low back tends to take the load, leading to increased stress and strain on it. By regaining usual upper body and lower body strength, these parts of the body are able to perform in a more normal way; this leads to less strain on the back.
- By working the upper and lower body, the trunk or core area of the body is strengthened; the body is stabilised to perform other exercises. This happens in an indirect way.
- All the exercises used short lever motions that were either moving directly away from or towards the body. No long lever lateral type movements were used. Therefore, the exercises were simple movements; they were easy to learn and less likely to cause injury. In the ‘hypertrophy phase B’ programme and in the ‘strength phase B’ programme, the ‘pec deck’ exercise could be interchanged with a ‘flat

fly'. The flat fly is a long lever exercise; this can be used as a method of work hardening, or as a method of gradually applying increased stress to the body (in a controlled manner). The flat fly was a simple movement; the participants were in a well-supported position and lying on their back. If there was any concern about performing this exercise, the 'pec deck' was used instead.

- Initially the resistance set was well within the capability of the patient.

3.9.2 Phase one: Conditioning

The Conditioning phase consisted of a warm up, a warm down, and eight exercises. The usual warm up and warm down consisted of rowing on a rowing ergometer and exercycling.

3.9.2.1 Method of progressing resistance levels in Conditioning phase

- Initially three sets of ten repetitions were performed for each exercise, with a 20 - 30 second rest between sets. To ascertain the initial starting weight, the participant was asked to perform an exercise at a low setting on each piece of equipment. This level of resistance was perceived by the physiotherapy trainer to be well within the participant's ability; it initially allowed the participant to feel the movement without significant resistance. The resistance level was then gradually increased to the point whereby the participant was aware of having to work to complete the movement, but was not stressed by doing so.
- When 3x10 repetitions could be performed comfortably, the repetitions were increased to 3x12.
- When 3x12 was comfortable, the repetitions were increased to 3x15.
- When 3x15 was comfortable the participant increased the resistance level, and decreased the repetitions to 3x10. The level of increase managed by the participant was either one weight level (this meant moving the pin on the stack to one weight level heavier or a half weight level if this was too much) or if using free weights, a suitable weight was added.
- On this increased resistance level the participant worked up through the repetitions to 3x15, before repeating the cycle.

This method of resistance level progression is a crucial aspect of the programme.

3.9.3 Phase two: Hypertrophy

At this point the overall structure of the programme (concepts of technique, exercise selection, periodisation and neuromuscular adaptation) were explained in detail to the participants. This was not done earlier, to avoid information overload. Learning the concept of a whole gymnasium-training programme is a considerable burden for many people, especially if they had never used weight resistance equipment previously.

Technique is crucial, as the best-planned programmes will fail if the exercise technique is not performed correctly. The conditioning phase allowed the participants to perfect their technique whilst training at a sub-maximal resistance level. It gave the participants time to become confident and comfortable with the exercises and the equipment. Exercises that focussed on the whole body were selected due to their simplicity of movement, and their ability to be performed with the spine in a neutral or supported position.

A periodized programme consists of five separate phases, namely: conditioning, hypertrophy, strength, pure strength and power. Depending on the authors these phases are sometimes summarised as strength endurance and strength. However, the recent 'position stand' from the American College of Sports Medicine has used conditioning, hypertrophy, strength, pure strength/peaking and power as their preferred terminology (Ratamess et al., 2009).

The concept of periodization was explained to the participants. After completing the conditioning phase, there were the hypertrophy and strength phases to complete, each consisting of nine weeks of training to complete the programme. Pure strength and power phases were briefly described for completeness only. These phases are designed for serious or competitive athletes. If participants wished to pursue this type of training they would need the assistance of a personal trainer.

The hypertrophy and strength phases were of nine weeks duration each; they were split into three blocks of three weeks. Each three week block was referred to as hypertrophy A, hypertrophy B, and hypertrophy C (or strength A, strength B, and strength C). In the hypertrophy programmes, each three-week block consisted of a completely new set of exercises. In the strength phase, the same exercises were used in each block but with different repetitions (Table 3.1). This concept of regularly changing the exercise programmes maximises the effects of neuromuscular adaptation.

In the hypertrophy phase, the resistance level was set so that participants could just manage to lift the weight 12, 10, and 8 times with correct technique and with no cheating. Performing the exercises without using 'cheat's tricks' is crucial, as this prevents over-exertion. When participants were able to lift or push the weight 13, 11 and 9 times, whilst maintaining good technique, they were then allowed to increase the level of resistance.

Standard resistance levels based on percentage maximums for a particular phase (as outlined in Table 2.5) were not used for setting resistance levels in the rehabilitation programme. These percentage maximums are based on 1RM resistance levels; they are not appropriate for an injured population due to the risk of further injury. Because 'hypertrophy A' programme exercises were identical to those used in the initial conditioning programme, the participants were familiar with these exercises and their resistance levels, as well as the sets and repetitions. This facilitated the PI in setting appropriate resistance levels for the 'hypertrophy A programme'.

As the participants progressed on to 'hypertrophy B and C' programmes with different exercises, the PI developed a clinical appreciation of each participant's strength level. When a new exercise was demonstrated, the same routine (used in the conditioning phase) was used to set the resistance level. A relatively light resistance was initially used so that the participant was able to 'feel' the new movement. The participants were asked if this resistance level felt light or medium. The resistance level for each set of 12, 10 and 8 repetitions was set accordingly.

Each exercise was increased on an individual basis. As soon as they were able to increase any individual exercise, they did so. This applied to all exercises throughout the duration of the rehabilitation programme.

In hypertrophy and strength phases, a warm-up set of 15 repetitions was performed prior to the first training set of 12 repetitions. This set of 15 repetitions was a warm up for each exercise, and used higher repetitions and lower resistance levels. The resistance level for this 'warm-up set' was approximately half to two thirds the level of resistance at which the first set of 12 repetitions was performed.

A simplistic summary of the meaning of neuromuscular adaptation and its application in the rehabilitation programme was given. Research on cognitive education has shown that if participants are aware of a reason for an activity and the expected goal, they are more likely to

complete the task (Burton et al., 1999; Ostelo et al., 2003; Pfingsten et al., 1997; Poulter, 1999).

3.9.4 Phase three: Strength

At the end of the hypertrophy phase, the final phase of the rehabilitation programme, strength, was explained. The strength phase consisted of three separate programmes, of three weeks duration each. Exercises in the strength phase replicated exercises performed in the hypertrophy phase. The only difference was that in the strength phase, sets and repetitions of 10-8-6 were used.

Using sets and repetitions of 10-8-6, a guideline was given to set the resistance level the participants lifted. The weight that participants lifted (or pushed) in the hypertrophy phase using 8 repetitions, was the starting resistance level in the strength phase with 10 repetitions. For example, in the Prone Pull exercise for the hypertrophy phase; if a participant lifted 15kgs, 17½kgs, and 20kgs for their sets of 12, 10 and 8 repetitions, respectively; in the strength phase the participant might lift 20kgs, 22½kgs, and 25kgs for 10, 8, and 6 repetitions, respectively. By this stage, the participants already knew all of the exercises, and were familiar with formulating levels of resistance.

At the end of the strength phase, the participants were given advice on how to maintain their heightened level of fitness and strength achieved. They were encouraged to maintain fitness and health in whatever way they found to be the most enjoyable and convenient (e.g. running, racket sports, swimming, yoga, pilates).

3.10 The Abdominal Regime

The abdominal regime included in this rehabilitation programme was described in Section 3.3. This regime was incorporated into the gym programme. Furthermore two Swiss ball exercises, previously described in Section 3.3, were incorporated into the programme at the time that the ‘hypertrophy B programme’ was taught.

3.11 Lifting Technique

During the strength phase, one session was undertaken to teach correct lifting and bending techniques. Fears were allayed about ‘bending their back’. The majority of participants had injured their backs using a bending and twisting motion. Since then, they had been extremely cautious when bending their back or excluded it all together.

The physiotherapist explained that bending one’s back was a normal movement. This concept was reinforced with the use of an articulated spine. The importance of using the human body in the manner for which it is designed, was explained. This included bending and flexing the spine. Many participants felt that bending their back would cause injury.

The participants were taught how to bend and flex with their body weight pivoting around a central axis of movement. This was to achieve body balance through movement. Unfurling from the bent position without straining their back was demonstrated.

During this session, the participants practised performing back-bending movements in the reassuring presence of the physiotherapist. During the course of the programme, trust between the participant and the therapist developed, particularly with a successful outcome. The participants felt confident to perform movements that they previously thought of as being dangerous. In a short space of time (40 minutes), the participants’ fear of bending was annulled. Participants learnt that bending was not going to cause them to re-injure their back. The final phase of rehabilitation would then be completed with restoration of full, normal, pain-free movement and function.

CHAPTER FOUR
RESULTS

4.1 Introduction

Chapter Four presents the results of a prospective randomised controlled trial comparing usual conservative surgical advice with a six-month non-aggravating physiotherapy rehabilitation programme.

4.2 The Sample Population

One hundred and thirty nine participants, all derived from the operation list of one orthopaedic surgeon, were screened for entry into the trial. Forty-five participants did not meet the inclusion criteria, which left 94 participants for randomisation to the study. Forty-seven participants were randomised to the control group. This group followed usual conservative surgical advice, which was to return to normal activities as soon as pain allowed. The other 47 participants entered the trial group and underwent a fully periodized six-month gymnasium-based physiotherapy rehabilitation programme.

Of the 94 participants randomised, one control group participant was immediately lost to follow-up due to moving to another part of the country before completing the baseline outcome information. An additional six participants were lost during the following three years, three from each group. The reasons for participants in the control group withdrawing from the study were, one withdrew for personal reasons (at 2 months), one developed depression and was admitted to hospital (at 6 months), and another went overseas (at 18 months). Reasons for withdrawal in the trial group were, one for personal reasons (at 2 months) and two not wanting to complete further outcome measures (both at 18 months).

Eight participants had a second discectomy for re-prolapse and a further participant had surgery for multiple pelvic fractures following a motorbike accident. Of the eight who had a second discectomy, five were in the trial group and three were in the control group. None of the causes for re-prolapse were associated with the rehabilitation programme. Two of the participants in the trial group who re-prolapsed had accidents in the community. One slipped and fell down stairs, this participant had completed six weeks of the rehabilitation

programme. The other injured himself after sitting on his ride-on mower for several hours and then bent over and lifted the mower up with one hand whilst endeavouring to release barbed wire that had entangled around the blade; this participant had completed 18 weeks of the programme. In both instances the participants were immediately aware of significant back pain, similar to their pre-surgery situation. The other three re-prolapses that occurred in the trial group happened in the second year of the study and were associated with house work, lifting at work, and lifting a barbecue. The three control group participants who re-prolapsed did so in the second (N=2) and third year (N=1) of the study. The causes of these re-prolapses were: lifting whilst doing house work, training (rope lunging) a horse, and lifting at work. All nine of the participants who had further surgery continued in the study.

Table 4.1 .

Table comparing demographic and clinical features of the control and trial groups prior to intervention for the intent-to-treat population. P-values are derived from t-tests and Chi-square tests as appropriate.

	Control Group N = 46	Trial Group N = 47	p-value
Age (years) • Mean • Range	 41 17 - 63	 42 25 - 63	0.41
BMI • Mean • Range	 27.5 18.6 - 41.2	 28.1 20.7 - 37.0	0.58
Hospital stay time (nights) • Median (range)	 1 (1 - 2)	 1 (0.5 - 2)	0.41
Time to surgery (weeks) • Median (range)	 16 (1 - 104)	 12.5 (0 - 156)	0.13
Time off prior to surgery (weeks) • Median (range)	 2 (0 - 16)	 3 (0 - 45)	0.48
Return to work after surgery (days) • Median (range)	 37 (0 - 168)	 35 (0 - 168)	0.65
	Number (%)	Number (%)	
Gender • Male • Female	 28 (61%) 18 (39%)	 26 (57%) 21 (43%)	0.47
Ethnicity • New Zealand European • Non New Zealand European	 43 (93.5%) 3 (6.5%)	 43 (91.3%) 4 (8.7%)	0.22
Operation Funding • Private • ACC / Workers Comp	 10 (21.7%) 36 (78.3%)	 10 (21.3%) 37 (78.7%)	0.96
Previous History • No previous hx • Significant hx • Previous surgery	 28 (60.9%) 14 (30.4%) 4 (8.7%)	 31 (66%) 12 (25.5%) 4 (8.5%)	0.86
Work Heaviness • Very heavy • Heavy • Medium • Light • Sedentary	 1 (2.2%) 3 (6.5%) 13 (28.3%) 22 (47.8%) 7 (15.2%)	 6 (12.8%) 12 (25.5%) 12 (25.5%) 10 (21.3%) 7 (14.9%)	0.01

Randomisation led to an equal balance between the groups with the exception of the work heaviness category (Table 4.1). The trial group had higher numbers of participants in the ‘very heavy’ and ‘heavy’ categories compared with the control group (18 to 4) and fewer numbers in the ‘light’ category (10 to 22). This imbalance was statistically significant.

The per-protocol population consisted of all participants from the control group minus withdrawals (N=3), and all participants from the trial group who completed the programme; this resulted in a per-protocol group of 85 participants. Five of the 47 participants randomised to the trial group did not complete the rehabilitation programme. One withdrew for personal reasons. Two participants had unplanned occupation changes immediately after surgery which prevented them from participating in the rehabilitation programme. One participant fell down stairs and underwent further surgery, and another participant had a baby and was unable to regularly participate in the programme due to childcare arrangements failing to occur as planned.

A third population in this study was the ‘per-protocol minus operation group’. This cohort included all control group members minus those participants who were lost to follow-up (N = 3) and also minus those who had further surgery (N = 4); three discectomies and one motor vehicle accident participant. This left 39 participants in the control group for this cohort; the trial group consisted of the 42 participants who completed the rehabilitation programme, minus the five participants who had further surgery, leaving 37 participants.

The per-protocol minus operation population was also an important cohort to analyse, as by removing the participants who had follow-up surgery, a number of confounding factors were removed. These participants were suffering considerable symptoms prior to their second surgery which adversely affected the outcome measure results. This resulted in increased doctor and other treatment-provider visits, increased medication use, time off work, and poor scores in the functional outcome measures. If the participants did not have to wait long for their second discectomy, the effect on the outcome measures was limited. However, if the participants waited for an extended period, this had a significant bearing on the outcome measure results. The ‘per-protocol minus operation’ cohort provided the ‘cleanest’ database, with minimal confounding, to compare the pure effect of the rehabilitation programme versus surgical advice.

4.3 Intent-to-treat analysis

4.3.1. Roland-Morris, Oswestry low back index and SF36

No statistically significant results were found in the intent-to-treat population for the RM and OLBI, (Table 4.2). This data is presented graphically in *Figures 4.1* (Roland-Morris) and *4.2* (Oswestry Low Back Index).

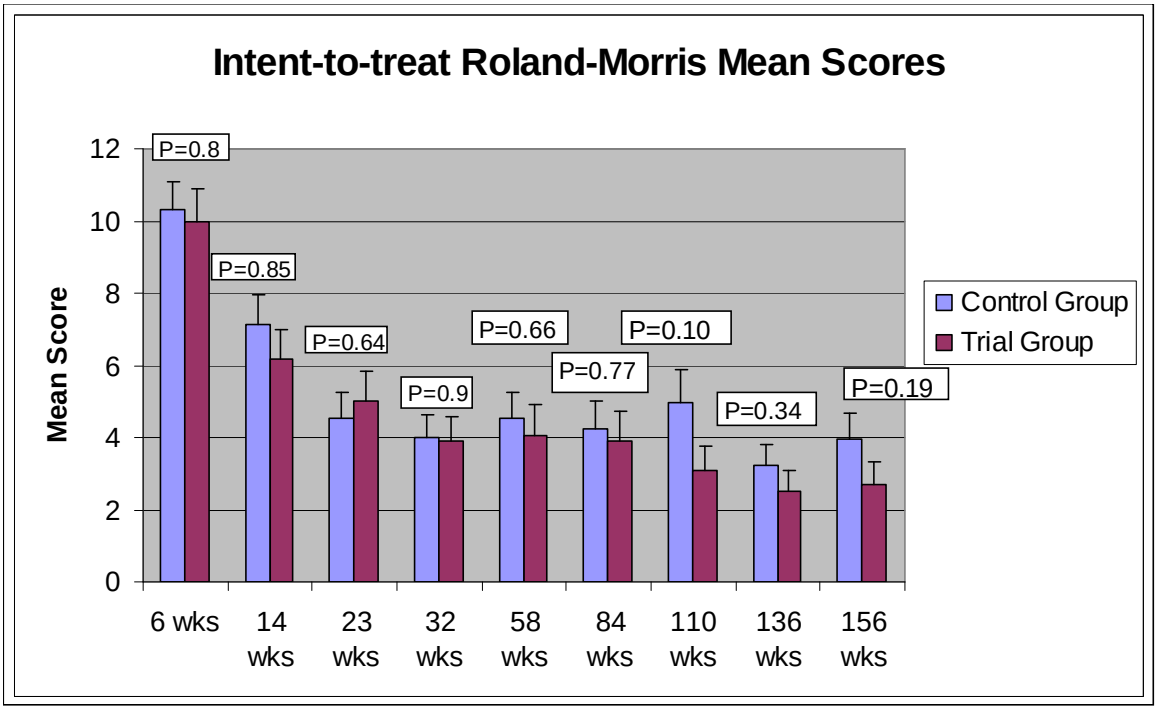


Figure 4.1. The mean Roland-Morris scores for the control and trial groups during the three-year follow-up period.

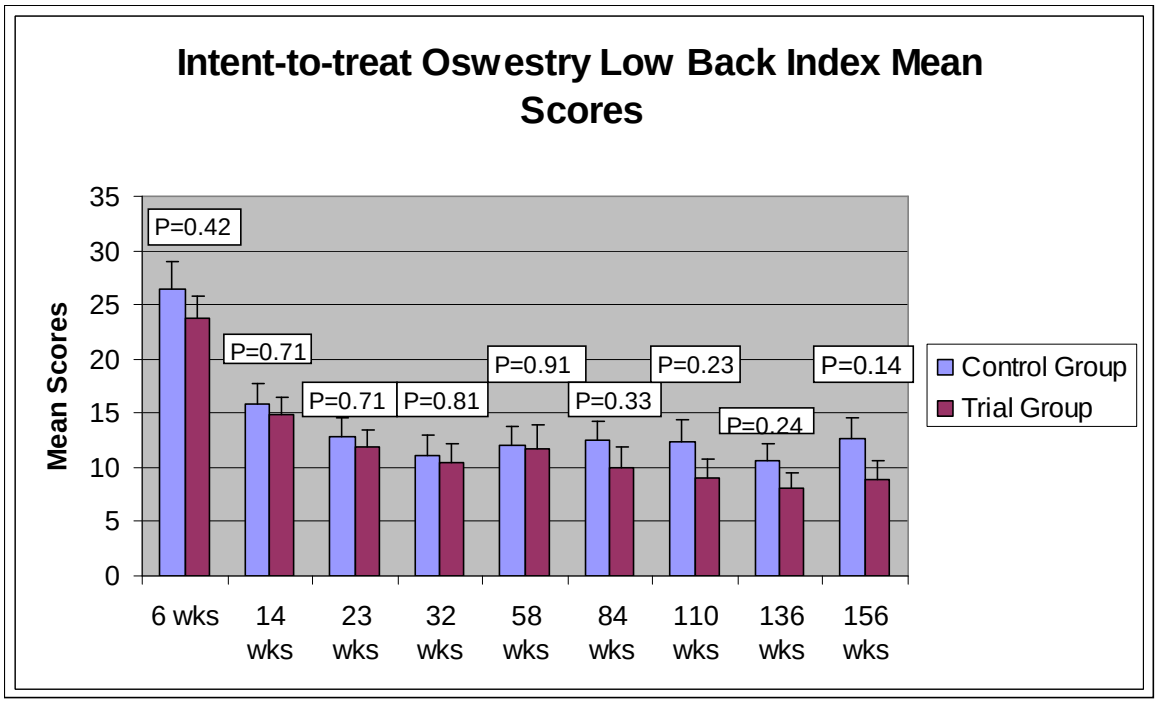


Figure 4.2. The mean scores for the Oswestry low back index for the control and trial groups during the three-year follow-up period.

Table 4.2

Intent-to-treat; means, standard errors, and number of completed outcome measures for the Roland-Morris and Oswestry Low Back Index at all time points.

Group RM		6 wks	14 wks	23 wks	32 wks	58 wks	84 wks	110 wks	136 wks	156 wks
Control	Mean (SE)	10.29 (.80)	7.13 (.81)	4.53 (.71)	4.02 (.64)	4.53 (.74)	4.23 (.76)	4.98 (.92)	3.23 (.62)	3.97 (.71)
Trial	Mean (SE)	9.98 (.94)	6.19 (.84)	5.02 (.75)	3.90 (.67)	4.03 (.91)	3.91 (.79)	3.07 (.67)	2.50 (.55)	2.71 (.63)
Number		88	88	87	85	85	87	84	81	81
P-value		0.80	0.85	0.64	0.90	0.66	0.77	0.10	0.38	0.19
Group OLBI		6 wks	14 wks	23 wks	32 wks	58 wks	84 wks	110 wks	136 wks	156 wks
Control	Mean (SE)	26.48 (2.54)	15.76 (1.90)	12.77 (1.76)	11.11 (1.80)	12.00 (1.84)	12.50 (1.77)	12.28 (2.10)	10.67 (1.62)	12.62 (1.93)
Trial	Mean (SE)	23.79 (2.13)	14.81 (1.71)	11.91 (1.54)	10.50 (1.76)	11.66 (2.25)	10.05 (1.76)	9.02 (1.68)	8.05 (1.51)	8.86 (1.65)
Number		86	87	87	85	86	87	84	81	81
P-value		0.42	0.71	0.71	0.81	0.91	0.33	0.23	0.24	0.14

Note: RM = Roland-Morris; OLBI = Oswestry low back index, wks = weeks

In the SF36 data (Table 4.3), six scores reached the threshold of statistical significance, these scores are highlighted in Table 4.3. Analysis of the intent-to-treat results for the RM, OLBI and SF36 questionnaires found no significant difference between the groups (Tables 4.2 & 4.3).

Table 4.3
Intent-to-treat: means, standard errors, number of completed outcome measures and p-values for the SF36 outcome measure at all time points.

Physical Function		6wks	32wks	58wks	84wks	110wks	136wks	156wks
Control	Mean(SE)	55.4(2.2)	78.2(1.9)	78.6(2.1)	79.1(2.1)	77.0(2.7)	82.6(2)	81.5(2.2)
Trial	Mean(SE)	54.1(2.5)	77(2.30)	78.7(2.4)	80.6(2.2)	80.0(2.4)	85.6(1.6)	86.1(1.7)
Number		92	86	88	86	80	80	79
P-value		0.78	0.79	0.97	0.72	0.56	0.40	0.24
V40 Role Physical								
Control	Mean(SE)	19.7(3.8)	70.7(4.1)	68.3(4.4)	63.4(4.5)	79.3(4.1)	75.6(3.8)	71.8(4.1)
Trial	Mean(SE)	10.0(3.0)	64.4(4.6)	70.9(3.9)	68.6(4.4)	64.7(4.6)	77.4(3.9)	78.1(4.0)
Number		92	86	88	86	80	80	79
P-value		0.16	0.48	0.78	0.56	0.10	0.81	0.44
Body Pain								
Control	Mean(SE)	45.8(2.3)	62.9(2.0)	64.5(2.4)	63.6(2.2)	68.2(2.9)	69.9(2.3)	65.5(2.7)
Trial	Mean(SE)	44.8(1.9)	67.1(2.2)	66.9(2.3)	69.4(2.0)	67.1(2.6)	74.4(2.2)	72.1(2.4)
Number		92	86	88	86	80	80	79
P-value		0.82	0.31	0.61	0.17	0.84	0.32	0.19
General health								
Control	Mean(SE)	76.2(1.6)	74.4(1.9)	73.5(2.1)	71.2(1.9)	71.9(2.3)	73.9(2.5)	71.6(2.4)
Trial	Mean(SE)	77.3(1.7)	79.1(1.9)	76.5(1.9)	74.1(2.0)	75.4(2.5)	75.8(2.3)	77.4(2.1)
Number		92	86	88	86	80	80	79
P-value		0.75	0.22	0.47	0.46	0.46	0.68	0.21
Vitality								
Control	Mean(SE)	43.2(1.8)	55.4(1.7)	57.2(2.1)	55.2(2.1)	54.2(2.3)	57.6(2.1)	56(2.7)
Trial	Mean(SE)	51.4(2.0)	61.3(1.9)	65.8(1.9)	62.1(2.1)	63.2(2.2)	66.0(1.7)	65.8(2.1)
Number		92	86	88	86	80	80	79
P-value		0.03*	0.11	0.04*	0.11	0.05*	0.03*	0.05*
Social function								
Control	Mean(SE)	58.0(2.7)	79.9(2.0)	78.3(2.3)	80.2(2.4)	83.8(2.7)	83.3(2.4)	83.3(2.3)
Trial	Mean(SE)	58.1(2.9)	83.8(2.0)	85.2(2.3)	86.3(1.9)	84.6(2.6)	82.9(2.3)	87.5(2.0)
Number		92	86	88	86	80	80	79
P-value		0.99	0.35	0.14	0.16	0.88	0.93	0.34
Role Emotional								
Control	Mean(SE)	55.3(5.0)	79.0(4.2)	74.1(3.9)	72.9(4.5)	83.7(4.4)	83.8(3.5)	82.9(3.2)
Trial	Mean(SE)	57.0(4.4)	80.8(3.3)	84.3(2.8)	78.3(3.9)	82.9(3.7)	85.4(3.4)	88.3(3.1)
Number		92	86	88	86	80	80	79
P-value		0.86	0.81	0.10	0.52	0.92	0.82	0.39
Mental Health								
Control	Mean(SE)	68.1(1.8)	74.0(1.8)	74.8(1.8)	74.5(1.8)	75.3(1.9)	78.2(1.8)	76.8(1.9)
Trial	Mean(SE)	71.6(2.0)	79.9(1.6)	81.4(1.4)	80.9(1.6)	82.0(1.4)	80.5(1.6)	80.2(1.9)
Number		92	86	88	86	80	80	79
P-value		0.36	0.09	0.04	0.07	0.05*	0.49	0.37
Overall Physical Health								
Control	Mean(SE)	48.1(1.8)	68.3(1.9)	68.4(2.3)	66.5(2.0)	70.1(2.5)	71.9(2.0)	69.3(2.2)
Trial	Mean(SE)	47.5(1.6)	69.8(2.0)	71.8(2.2)	71.0(2.1)	70.1(2.3)	75.8(1.9)	75.9(2.0)
Number		92	86	88	86	80	80	79
P-value		0.87	0.70	0.46	0.28	0.10	0.32	0.12
Overall Mental Health								
Control	Mean(SE)	60.2(2.0)	72.5(1.8)	71.6(2.0)	70.8(2.1)	73.8(2.2)	75.3(1.9)	74.1(2.1)
Trial	Mean(SE)	63.1(2.0)	77.0(1.7)	78.8(1.7)	76.4(2.0)	77.6(2.0)	78.1(1.8)	79.8(1.8)
Number		92	86	88	86	80	80	79
P-value		0.48	0.21	0.05	0.17	0.37	0.45	0.15
TOTAL SF36 Score								
Control	Mean(SE)	52.7(1.9)	71.8(1.8)	71.2(2.1)	70.0(2.0)	74.2(2.4)	75.6(1.8)	73.7(2.1)
Trial	Mean(SE)	53.1(1.7)	74.2(1.8)	76.3(2.0)	75.1(2.1)	75(2.2)	78.5(1.8)	79.4(1.9)
Number		92	86	88	86	80	80	79
P-value		0.93	0.52	0.21	0.22	0.86	0.42	0.15

Note: OPH = overall physical health; OMH = overall mental health; SE = standard error; * = $P \leq 0.05$

4.3.2. Annual questionnaire data

Intent-to-treat analysis of the annual questionnaire data was not statistically significant with the exception of three results (Table 4.4). These results were related to doctor visits and levels of medication. A smaller proportion of participants in the trial group compared with the control group, used medication for their low back in the second year of follow-up (P = 0.03) (Table 4.4). A smaller proportion of trial group participants visited their doctor compared with the control group during the combined three year follow-up period (p = 0.01) (*fig 4.3*).

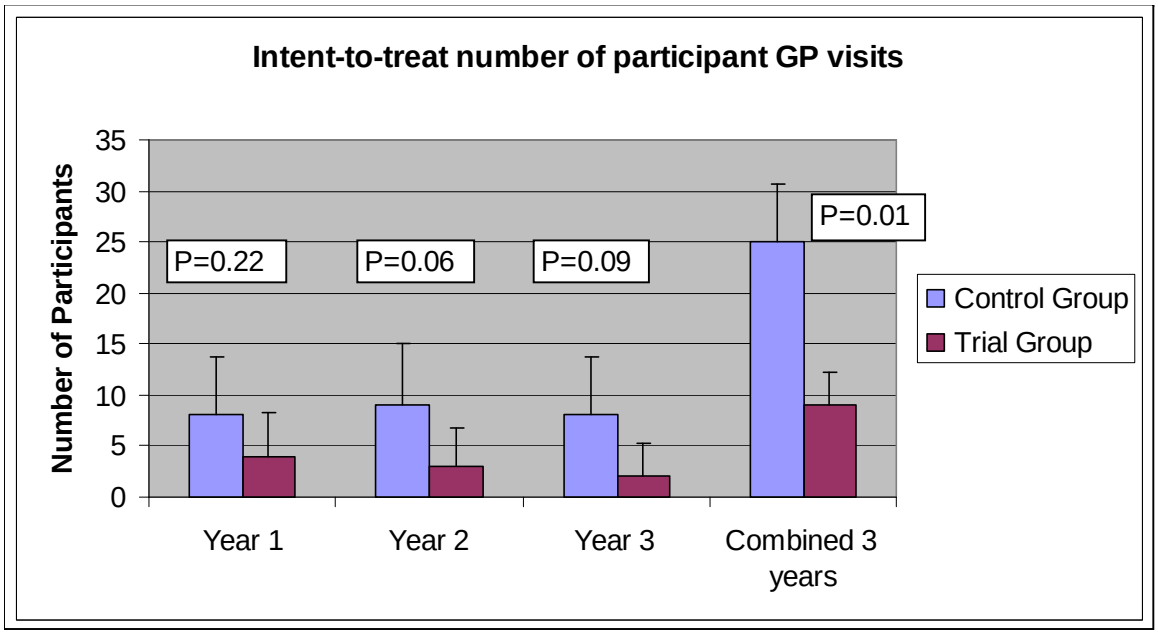


Figure 4.3. The number of participants who visited their GPs in years 1, 2 and 3, and the combined three-year period.

The second aspect of 'doctor visits' was the 'total' number of times that participants visited their doctors. For example, some participants in both groups visited their doctors on more than one occasion in any one year. Although results for each individual year were not statistically significant, the combined three year result indicated that trial group participants had fewer mean numbers of doctor visits during the three year combined follow-up (P = 0.03)(*fig 4.4*).

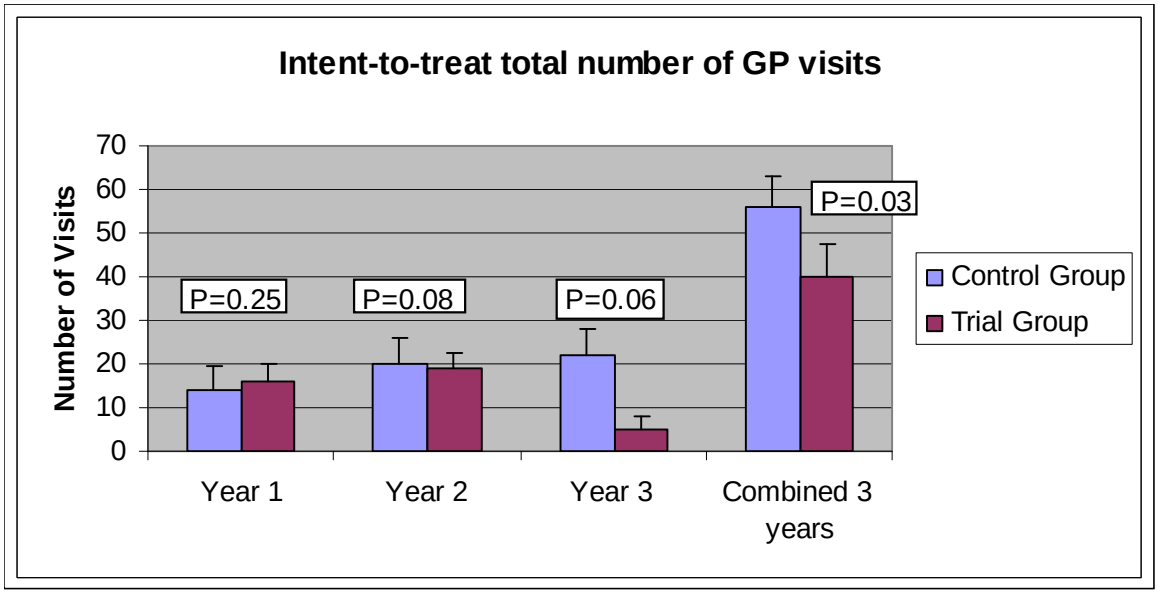


Figure 4.4. The total number of GP visits in each group for each year and the three-year combined results for the intent-to-treat subgroup.

The mean time to return to work immediately post-surgery for the control group was 37.1 days, and for the trial group 31.7 days, (P = 0.65) (Table 4.1). Therefore on average, the trial group participants returned to work approximately five days earlier than the control group participants. The categories of ‘other treatment providers’ and ‘satisfaction’ did not statistically differentiate the control and trial groups.

Table 4.4

Results of the intent-to-treat annual questionnaire years one, two and three.

Year one	Participant GP visits	Other therapist visits	Med levels	Episodes off work		Total No of GP visits	Days off work
Count within control grp (%)	8 (17.8)	6 (13.3)	22 (48.9)	8 (17.8)	Mean (SE)	0.31(0.11)	1.66(0.73)
Count within trial grp (%)	4 (8.9)	7 (15.6)	21 (46.7)	4 (8.9)	Mean (SE)	0.36(0.23)	3.84(3.73)
Total Number	90.00	90.00	90.00	90.00		90.00	90.00
P-value	0.22	0.76	0.83	0.22		0.25	0.12
Year two	Participant GP visits	Other therapist visits	Med levels	Episodes off work		Total No of GP visits	Days off work
Count within control grp (%)	9 (20)	7 (15.6)	22 (48.9)	8 (17.8)	Mean (SE)	0.44(0.19)	6.22(3.91)
Count within trial grp (%)	3 (6.7)	6 (13.3)	12 (26.7)	4 (8.9)	Mean (SE)	0.42(0.23)	4.84(3.10)
Total Number	90.00	90.00	90.00	90.00		90.00	90.00
P-value	0.06	0.76	0.03*	0.22		0.08	0.25
Year three	Participant GP visits	Other therapist visits	Med levels	Episodes off work		Total No of GP visits	Days off work
Count within control grp (%)	8 (18.2)	3 (6.8)	18 (40.9)	6 (13.6)	Mean (SE)	0.50(0.28)	2.70(2.06)
Count within trial grp (%)	2 (4.7)	3 (7.0)	11 (25.6)	3 (7.0)	Mean (SE)	0.12(0.08)	6.12(6.05)
Total Number	87.00	87.00	87.00	87.00		87.00	87.00
P-value	0.09	1.00	0.13	0.48		0.06	0.29

Note: grp = group; GP = General Practitioner; Med = Medication; No = number; SE = standard error; * = $P \leq 0.05$

4.4 Per-protocol analysis

4.4.1 Roland-Morris, Oswestry Low Back Index and SF36

Analysis of this data set, outlined in Table 4.5, provides no statistically significant results for the RM or OLBI outcome measures.

Table 4.5
Per-protocol: means, standard errors, number of completed outcome measures and p-values for the Roland-Morris and Oswestry Low Back Index at all time points.

Group RM		6 wks	14 wks	23 wks	32 wks	58 wks	84 wks	110 wks	136 wks	156 wks
Control	Mean (SE)	10.29 (0.80)	7.13 (0.81)	4.53 (0.71)	4.02 (0.64)	4.53 (0.74)	4.23 (0.76)	4.98 (0.92)	3.23 (0.62)	3.97 (0.71)
Trial	Mean (SE)	9.55 (0.96)	6.31 (0.84)	4.15 (0.65)	3.35 (0.61)	3.06 (0.75)	3.03 (0.78)	2.82 (0.68)	2.25 (0.53)	2.33 (0.64)
Number		83	84	82	81	81	82	81	79	78
P-value		0.55	0.48	0.70	0.45	0.17	0.27	0.07	0.23	0.09
Group OLBI										
Control	Mean (SE)	26.48 (2.54)	15.76 (1.90)	12.77 (1.76)	11.11 (1.80)	12.00 (1.84)	12.50 (1.77)	12.28 (2.10)	10.67 (1.62)	12.62 (1.93)
Trial	Mean (SE)	22.76 (2.24)	13.84 (1.70)	10.47 (1.29)	9.14 (1.43)	9.41 (1.93)	8.00 (1.69)	8.16 (1.69)	7.15 (1.37)	8.19 (1.68)
Number		82	83	82	82	82	82	81	79	78
P-value		0.28	0.46	0.31	0.41	0.34	0.07	0.14	0.10	0.08

Note: RM = Roland-Morris; OLBI = Oswestry low back index, wks = weeks; SE = standard error

The SF36 results for this population show a small cluster of statistically significant results around the 56 and 84 weeks post-surgery period, and four of the 11 categories at 156 weeks post-surgery also reached statistical significance (Table 4.6). Although there were a small number of statistically significant results, overall this data did not reveal a consistent significant difference between the groups. A category of note however, is that of ‘vitality’, which achieved a statistically significant difference at every assessment point throughout the follow-up period.

Table 4.6
Per-protocol: means, standard errors, number of completed outcome measures and p-values for the SF36 outcome measure at all time points.

Physical function		6 wks	32 wks	58 wks	84 wks	110 wks	136 wks	156 wks
Control	Mean(SE)	57.6 (2.4)	78.9 (2.1)	80.6(2.1)	81.0 (2.3)	82.0 (2.2)	83.3 (2.1)	82.6 (2.1)
Trial	Mean(SE)	54.9 (2.8)	79.5 (2.5)	83.1(2.1)	85.5 (2.0)	81.8 (2.7)	87.4 (1.7)	87.3 (1.9)
Number		76	73	75	72	68	69	68
P-value		0.60	0.89	0.56	0.31	0.96	0.29	0.25
Role physical								
Control	Mean(SE)	18.3 (4.1)	73.8 (4.3)	72 (4.6)	64.7 (4.8)	85.8 (3.6)	76.4 (4.1)	75.0 (4.3)
Trial	Mean(SE)	10.0 (3.3)	71.9 (4.6)	79.4 (3.5)	78 (4.0)	72.6 (4.6)	81.1 (4.3)	82.8 (3.9)
Number		76	73	75	72	68	69	68
P-value		0.27	0.83	0.37	0.14	0.11	0.56	0.35
Bodily pain								
Control	Mean(SE)	46.6 (2.5)	64.9 (2.1)	67.5 (2.5)	65.6 (2.3)	73.5 (2.5)	71.4 (2.4)	66.6 (2.9)
Trial	Mean(SE)	45.8 (1.9)	70.4 (2.2)	70.3 (2.1)	74.0 (1.9)	69.0 (2.6)	77.2 (2.0)	75.9 (2.2)
Number		76	73	75	72	68	69	68
P-value		0.86	0.20	0.55	0.53	0.38	0.19	0.08
General health								
Control	Mean(SE)	75.6 (1.9)	74.8 (2.1)	74.3 (2.4)	71.8 (2.0)	73.3 (2.4)	74.3 (2.6)	70.8 (2.7)
Trial	Mean(SE)	81.0 (1.6)	82.7 (1.8)	80.6 (1.8)	79.1 (2.0)	79.3 (2.4)	79.6 (2.1)	81.0 (2.1)
Number		76	73	75	72	68	69	68
P-value		0.14	0.05*	0.14	0.80	0.21	0.27	0.04*
Vitality								
Control	Mean(SE)	43.7 (1.9)	57.4 (1.7)	58.8 (2.1)	56.5 (2.0)	56.2 (2.2)	57.6 (2.0)	55.7 (2.8)
Trial	Mean(SE)	53.3 (2.2)	65.5 (1.8)	68.8 (1.7)	68 (1.8)	66.0 (2.1)	68.2 (1.8)	69.1 (1.9)
Number		76	73	75	72	68	69	68
P-value		0.02*	0.02*	0.01*	0.01*	0.03*	0.01*	0.01*
Social function								
Control	Mean(SE)	58.8 (2.9)	81.1 (2.2)	80.2 (2.3)	82.1 (2.4)	88.9 (2.1)	81.9 (2.6)	83.3 (2.5)
Trial	Mean(SE)	58.6 (3.2)	87.1 (1.8)	89.0 (1.9)	91.7 (1.5)	89.1 (2.2)	84.5 (2.3)	89.8 (2.1)
Number		76	73	75	72	68	69	68
P-value		0.97	0.14	0.04*	0.02	0.95	0.61	0.17
Role emotional								
Control	Mean(SE)	58.5 (5.4)	82.9 (4.4)	76.4 (3.9)	76.9 (4.7)	90.1 (4.0)	82.4 (3.9)	83.3 (3.4)
Trial	Mean(SE)	62.9 (4.9)	85.4 (3.1)	90.2 (2.2)	85.9 (3.7)	87.1 (3.5)	85.9 (3.5)	90.6 (2.9)
Number		76	73	75	72	68	69	68
P-value		0.68	0.75	0.04*	0.30	0.70	0.64	0.26
Mental health								
Control	Mean(SE)	68.0 (2.0)	74.9 (2.0)	75.4 (1.8)	75.6 (1.8)	77.0 (1.8)	76.9 (2.0)	76.7 (2.0)
Trial	Mean(SE)	74.9 (1.8)	81.5 (1.5)	82.7 (1.1)	84.1 (1.6)	83.4 (1.4)	82.3 (1.4)	82.9 (1.4)
Number		76	73	75	72	68	69	68
P-value		0.08	0.07	0.02*	0.02*	0.06	0.12	0.08
OPH								
Control	Mean(SE)	48.4 (1.9)	70.0 (1.9)	70.6 (2.4)	67.9 (2.1)	74.2 (2.2)	72.6 (2.1)	70.2 (2.3)
Trial	Mean(SE)	49.0 (1.7)	74.0 (1.9)	76.4 (1.8)	76.9 (1.8)	73.7 (2.3)	78.7 (1.50	79.2 (2.1)
Number		76	73	75	72	68	69	68
P-value		0.86	0.30	0.18	0.03*	0.92	0.13	0.05*
OMH								
Control	Mean(SE)	60.9 (2.1)	74.2 (1.9)	73.0 (2.0)	72.6 (2.0)	77.1 (1.9)	74.6 (2.1)	74 (2.2)
Trial	Mean(SE)	66.1 (2.1)	80.4 (1.5)	82.3 (1.3)	81.7 (1.8)	81 (1.9)	80.1 (1.8)	82.7 (1.8)
Number		76	73	75	72	68	69	68
P-value		0.22	0.08	0.01*	0.02*	0.32	0.16	0.04*
Total score								
Control	Mean(SE)	53.4 (2.0)	73.6 (1.9)	73.1 (2.1)	71.8 (2.0)	78.4 (2.1)	75.5 (2.0)	74.3 (2.2)
Trial	Mean(SE)	55.2 (1.9)	78.0 (1.7)	80.5 (1.5)	80.8 (1.8)	78.5 (2.1)	80.8 (1.8)	82.4 (1.9)
Number		76	73	75	72	68	69	68
P-value		0.66	0.23	0.06	0.02*	0.97	0.17	0.05*

Note: OPH = overall physical health; OMH overall mental health; SE = standard error; * = P ≤ 0.05

4.4.2 Annual questionnaire data

The annual questionnaire data for the per-protocol group resulted in three statistically significant results. The three year combined results for participant doctor visits (p = 0.03) (*fig 4.5*), the total number of doctor visits (p = 0.05) (*fig 4.6*), and medication levels in year two (p = 0.04) (*fig 4.7*). These results indicated that during the three year follow-up, fewer trial group participants required doctor visits and the total number of doctor visits used by the trial group was less than that used by the control group; and that in year two of the study, trial group participants required less medication. A summary of all results for this population is in Table 4.7.

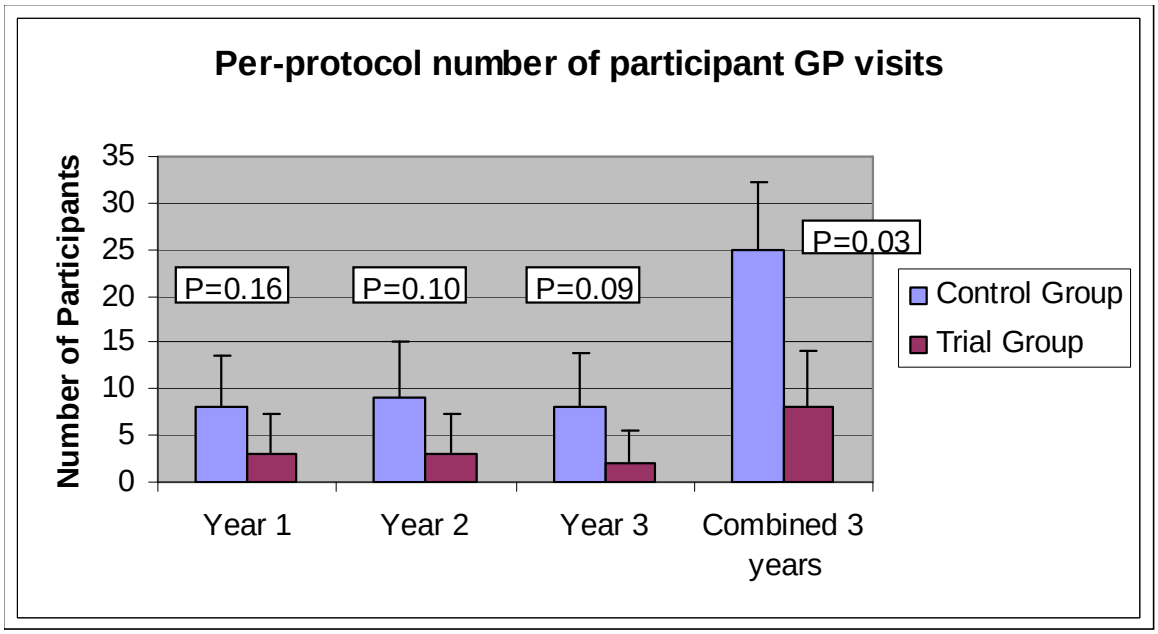


Figure 4.5. The number of participants in the per-protocol group who visited their GPs in years 1, 2, and 3 with the combined three-year result.

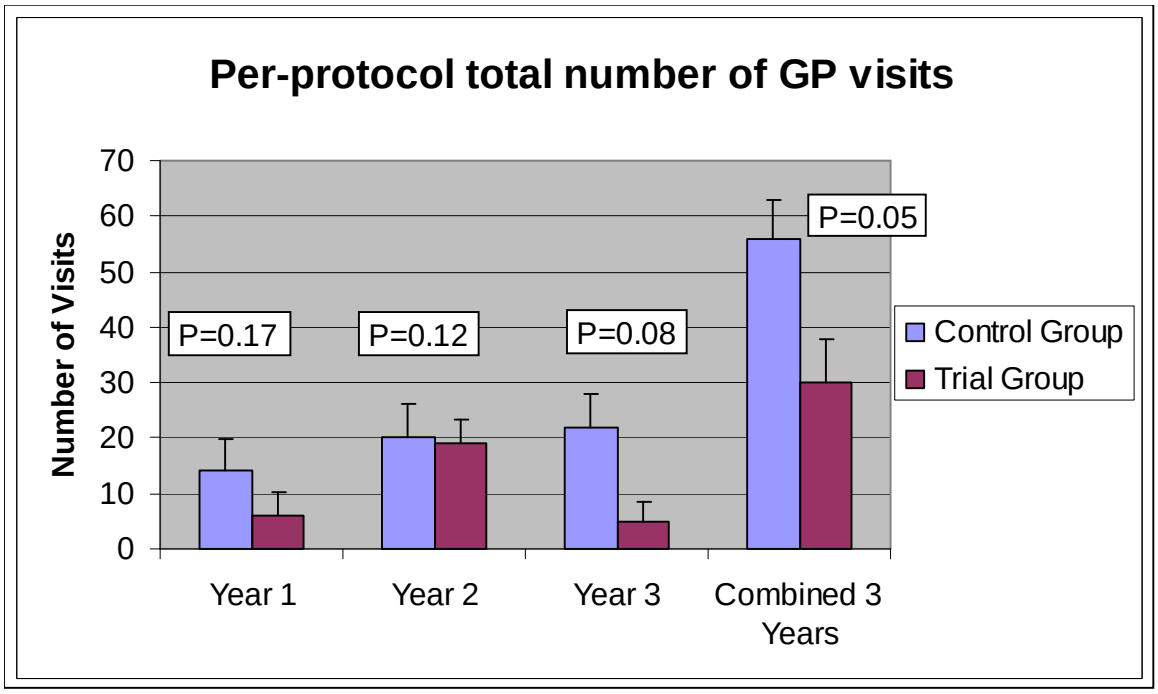


Figure 4.6. The total number of GP visits in each year for the per-protocol group with the combined three-year result.

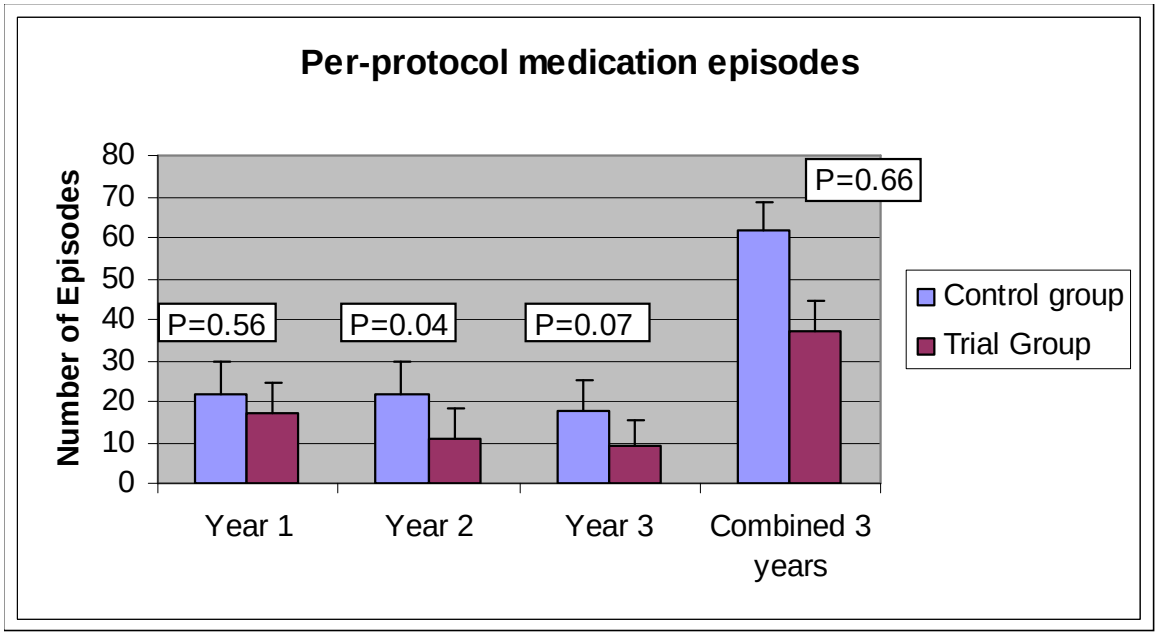


Figure 4.7. Episodes of medication use in the per-protocol sub-group for years 1, 2, and 3 as well as the combined three-year data.

Table 4.7

Per-protocol: results of the annual questionnaire, years one, two and three.

Year one	Participant GP visits	Other therapist visits	Med levels	Episodes off work		Total No of GP visits	Days off work
Count within control grp (%)	8 (17.8)	6.(13.3)	22 (48.9)	8 (17.8)	Mean (SE)	0.31(0.11)	1.67(0.73)
Count within trial grp (%)	3 (7.5)	6 (15.0)	17 (42.5)	2 (5.0)	Mean (SE)	0.15(0.09)	0.08(0.06)
Total Number	85.00	85.00	85.00	85.00		85.00	85.00
P-value	0.16	0.83	0.56	0.10		0.17	0.06
Year two	Participant GP visits	Other therapist visits	Med levels	Episodes off work		Total No of GP visits	Days off work
Count within control grp (%)	9 (20)	7 (15.6)	22 (48.9)	8 (17.8)	Mean (SE)	0.44(0.19)	6.22(3.19)
Count within trial grp (%)	3 (7.5)	6 (15)	11 (27.5)	3 (7.5)	Mean (SE)	0.48(0.33)	3.95(3.19)
Total Number	85.00	85.00	85.00	85.00		85.00	85.00
P-value	0.10	0.94	0.04*	0.16		0.12	0.18
Year three	Participant GP visits	Other therapist visits	Med levels	Episodes off work		Total No of GP visits	Days off work
Count within control grp (%)	8 (18.20)	3 (6.8)	18 (40.9)	6 (13.6)	Mean (SE)	0.50(0.28)	2.71(2.06)
Count within trial grp (%)	2 (5.0)	2 (5.0)	9 (22.5)	2 (5.0)	Mean (SE)	0.13(0.09)	6.55(6.05)
Total Number	84.00	84.00	84.00	84.00		84.00	84.00
P-value	0.09	1.00	0.07	0.27		0.08	0.18

Note:grp=group;GP = General Practitioner;Med = Medication; No = number; SE = standard error; * = $P \leq 0.05$

4.5 Per-protocol minus operation analysis

4.5.1 Roland-Morris, Oswestry Low Back Index and SF36

There were no statistically significant results in the RM and OLBI data (Table 4.8).

Table 4.8

Per-protocol minus operation: means, standard errors, number of completed outcome measures and p-values for the Roland-Morris and Oswestry Low Back Index at all time points.

Group RM		6 wks	14 wks	23 wks	32 wks	58 wks	84 wks	110 wks	136 wks	156 wks
Control	Mean (SE)	9.73 (0.82)	6.49 (0.79)	3.92 (0.64)	3.45 (0.59)	3.85 (0.67)	3.53 (0.65)	3.85 (0.78)	3.17 (0.66)	3.69 (0.65)
Trial	Mean (SE)	9.67 (1.08)	6.38 (0.92)	4.29 (0.72)	2.75 (0.63)	2.59 (0.68)	2.55 (0.69)	2.30 (0.66)	2.14 (0.58)	2.26 (0.70)
Number		74	75	73	72	73	73	72	71	71
P-value		0.96	0.93	0.70	0.42	0.20	0.31	0.14	0.25	0.14
Group OLBI										
Control	Mean (SE)	26.03 (2.76)	14.27 (1.82)	11.85 (1.8)	9.66 (1.71)	9.90 (1.58)	11.35 (1.74)	9.64 (1.76)	9.94 (1.62)	11.89 (1.89)
Trial	Mean (SE)	23.12 (2.51)	13.88 (1.88)	10.48 (1.42)	8.00 (1.45)	8.36 (1.60)	7.09 (1.73)	6.91 (1.66)	6.69 (1.43)	7.53 (1.72)
Number		73	74	73	73	73	73	72	71	70
P-value		0.45	0.88	0.57	0.48	0.5	0.09	0.27	0.14	0.09

Note: RM = Roland-Morris; OLBI = Oswestry low back index, wks = weeks; SE = standard error

The SF36 data demonstrated a small number of statistically significant results. Similar to the per-protocol population for this outcome measure, there was a small cluster of significant results at 58 and 84 weeks. Some pertinent results were found in the categories of vitality and mental health. ‘Vitality’ demonstrated a statistically significant difference between the groups at every time point during the three year follow-up, and ‘mental health’ provided statistically significant results from six weeks through to 110 weeks post-surgery. Despite these results, overall analysis of the SF36 data did not demonstrate a consistent significant difference between the groups.

Table 4.9
Per-protocol minus operation: means, standard errors, and number of completed outcome measures and p-values for the SF36 outcome measure at all time points

Physical function		6 wks	32 wks	58 wks	84 wks	110 wks	136 wks	156 wks
Control	Mean(SE)	55.0 (2.4)	77.8 (2.0)	78.6 (2.1)	79.1 (2.2)	77.0 (2.8)	82.6 (2.1)	81.5 (2.2)
Trial	Mean(SE)	54.9 (2.5)	79.2 (2.2)	82.2 (2.0)	84.0 (1.9)	80.6 (2.5)	87.0 (1.6)	87.2 (1.7)
Number		85	82	84	81	77	77	76
P-value		0.98	0.73	0.39	0.24	0.50	0.23	0.16
Role physical								
Control	Mean(SE)	16.7 (3.7)	70.0 (4.3)	68.3 (4.5)	63.4 (4.6)	79.3 (4.1)	75.6 (3.8)	71.8 (4.2)
Trial	Mean(SE)	8.8 (2.9)	66.9 (4.6)	76.9 (3.5)	74.3 (4.1)	68.1 (4.6)	80.3 (3.8)	81.8 (3.9)
Number		85	82	84	81	77	77	76
P-value		0.24	0.73	0.30	0.22	0.20	0.55	0.22
Bodily pain								
Control	Mean(SE)	45.1 (2.3)	62.6 (2.1)	64.5 (2.5)	63.6 (2.2)	68.2 (2.9)	69.9 (2.3)	65 5 (2.7)
Trial	Mean(SE)	45.4 (1.9)	68.2 (2.1)	69.2 (2.3)	71.7 (2.0)	68.1 (2.6)	77.0 (2.0)	73.8 (2.4)
Number		85	82	84	81	77	77	76
P-value		0.94	0.18	0.33	0.06	0.97	0.11	0.11
General health								
Control	Mean(SE)	75.7 (1.7)	74.3 (2.0)	73.5 (2.2)	71.2 (2.0)	71.9 (2.3)	73.9 (2.5)	71.6 (2.5)
Trial	Mean(SE)	79.9 (1.6)	80.1 (1.8)	78.7 (1.7)	76.1 (2.0)	75.8 (2.6)	77.9 (2.2)	71.6 (2.7)
Number		85	82	84	81	77	77	76
P-value		0.21	0.13	0.20	0.22	0.42	0.40	0.15
Vitality								
Control	Mean(SE)	42.6 (1.9)	55.4 (1.8)	57.2 (2.2)	55.2 (2.2)	54.2 (2.3)	57.6 (2.1)	56 (2.7)
Trial	Mean(SE)	52.9 (2.0)	63.2 (1.8)	67.3 (1.8)	65.3 (2.0)	63.6 (2.3)	67.2 (1.7)	67 (2.1)
Number		85	82	84	81	77	77	76
P-value		0.01*	0.03*	0.02*	0.02*	0.04*	0.01*	0.03
Social function								
Control	Mean(SE)	56.4 (2.8)	79.4 (2.1)	78.3 (2.3)	80.2 (2.5)	83.8 (2.8)	83.3 (2.4)	83.3 (2.4)
Trial	Mean(SE)	58.4 (3.0)	85.5 (1.7)	87.2 (2.2)	88.8 (1.8)	85.8 (2.6)	84.9 (2.1)	88.9 (2.0)
Number		85	82	84	81	77	77	76
P-value		0.72	0.12	0.05*	0.05*	0.72	0.73	0.22
Role emotional								
Control	Mean(SE)	53.3 (5.2)	78.5 (4.4)	74.1 (3.9)	72.9 (4.6)	83.7 (4.5)	83.8 (3.6)	82.9 (3.3)
Trial	Mean(SE)	61.7 (4.5)	82 9 (3.3)	88.0 (2.4)	81.6 (3.9)	83.3 (3.9)	85.1 (3.5)	87.4 (3.3)
Number		85	82	84	81	77	77	76
P-value		0.40	0.58	0.04*	0.32	0.96	0.85	0.50
Mental health								
Control	Mean(SE)	67.1 (1.9)	74.0 (1.8)	74.8 (1.8)	74.5 (1.9)	75.3 (1.9)	78.2 (1.8)	76.8 (1.9)
Trial	Mean(SE)	74.9 (1.6)	81.1 (1.4)	82.2 (1.2)	82.6 (1.5)	82.7 (1.4)	82.0 (1.3)	81.4 (1.8)
Number		85	82	84	81	77	77	76
P-value		0.03*	0.04*	0.02*	0.02*	0.04*	0.23	0.22
OPH								
Control	Mean(SE)	47.0 (1.8)	68.0 (1.9)	68.4 (2.3)	66.5 (2.1)	70.1 (2.5)	71.9 (2.1)	69.3 (2.3)
Trial	Mean(SE)	48.4 (1.6)	71.5 (1.8)	74.9 (1.9)	74.3 (2.0)	71.2 (2.4)	77.9 (1.8)	77.6 (2.0)
Number		85	82	84	81	77	77	76
P-value		0.69	0.36	0.14	0.06	0.82	0.13	56
OMH								
Control	Mean(SE)	59 (2.0)	72.3 (1.9)	71.6 (2.0)	70.8 (2.1)	73.8 (2.2)	75.3 (1.9)	74.1 (2.2)
Trial	Mean(SE)	65.6 (1.9)	78.6 (1.6)	80.7 (1.5)	78.9 (1.9)	78.2 (2.1)	79.4 (1.7)	80.6 (1.9)
Number		85	82	84	81	77	77	76
P-value		0.01*	0.08	0.01*	0.05*	0.32	0.27	0.11
Total score								
Control	Mean(SE)	51.5 (2.0)	71.5 (1.9)	71.2 (2.2)	70.0 (2.1)	74.2 (2.5)	75.6 (1.8)	73.7 (2.2)
Trial	Mean(SE)	54.6 (1.7)	75.9 (1.6)	79.0 (1.7)	78.1 (2.0)	76.0 (2.2)	80.2 (1.7)	80.7 (1.9)
Number		85	82	84	81	77	77	76
P-value		0.40	0.22	0.05*	0.05*	0.70	0.20	0.09

Note: OPH = overall physical health; OMH = overall mental health; SE = standard error; * = P ≤ 0.05

4.5.2 Annual questionnaire data

There were four statistically significant results in the annual questionnaire data for this population. They were participant doctor visits in the combined three year period, total number of doctor visits in year three and the combined three year period, and medication levels in year three.

A smaller proportion of participants in the trial group compared with the control group visited their doctor during the combined three year follow-up period (P = 0.01) (*fig 4.8*).

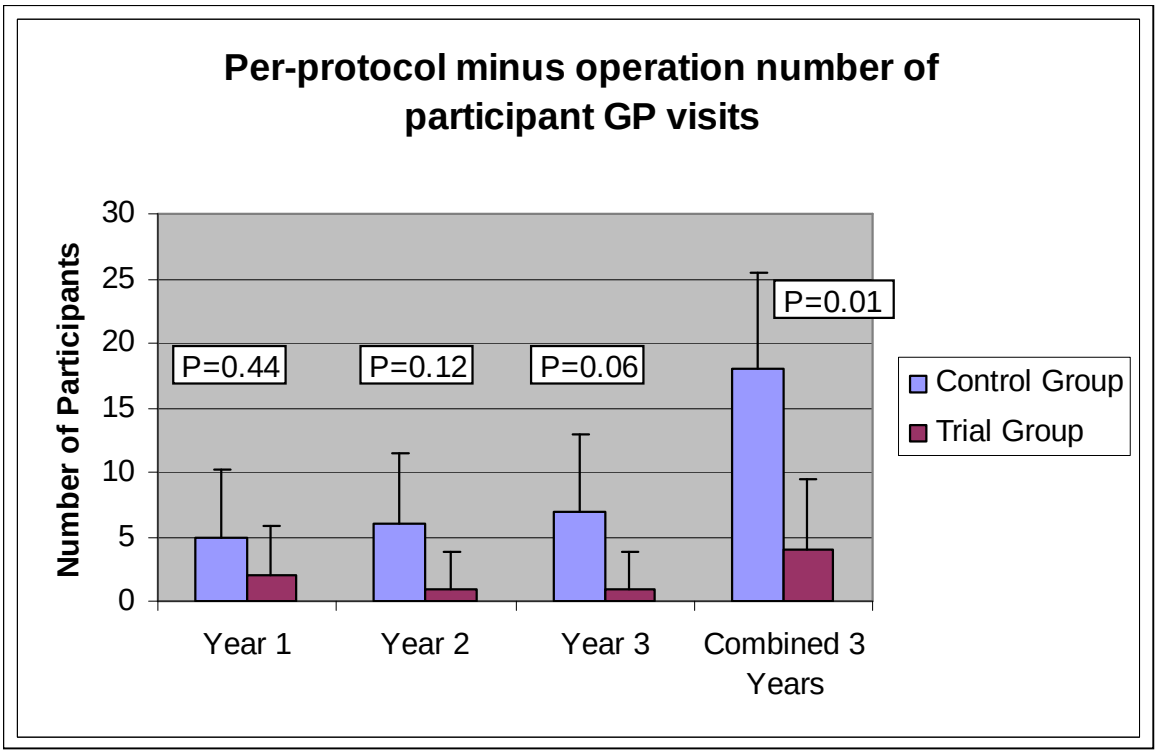


Figure 4.8. The number of participants in the per-protocol minus operation sub-group, who visited their GPs in each year, as well as the combined three-year result.

For the category of 'total number of doctor visits', the results demonstrated a gradual differentiation between the groups resulting in a statistically significant difference in year three (P = 0.05) (*fig 4.9*). This category was also statistically significant in the three year combined follow-up period (P = 0.02) (*fig 4.9*). These results demonstrated that trial group participants had fewer doctor visits in year three and during the combined three year follow-up period (*fig 4.9*).

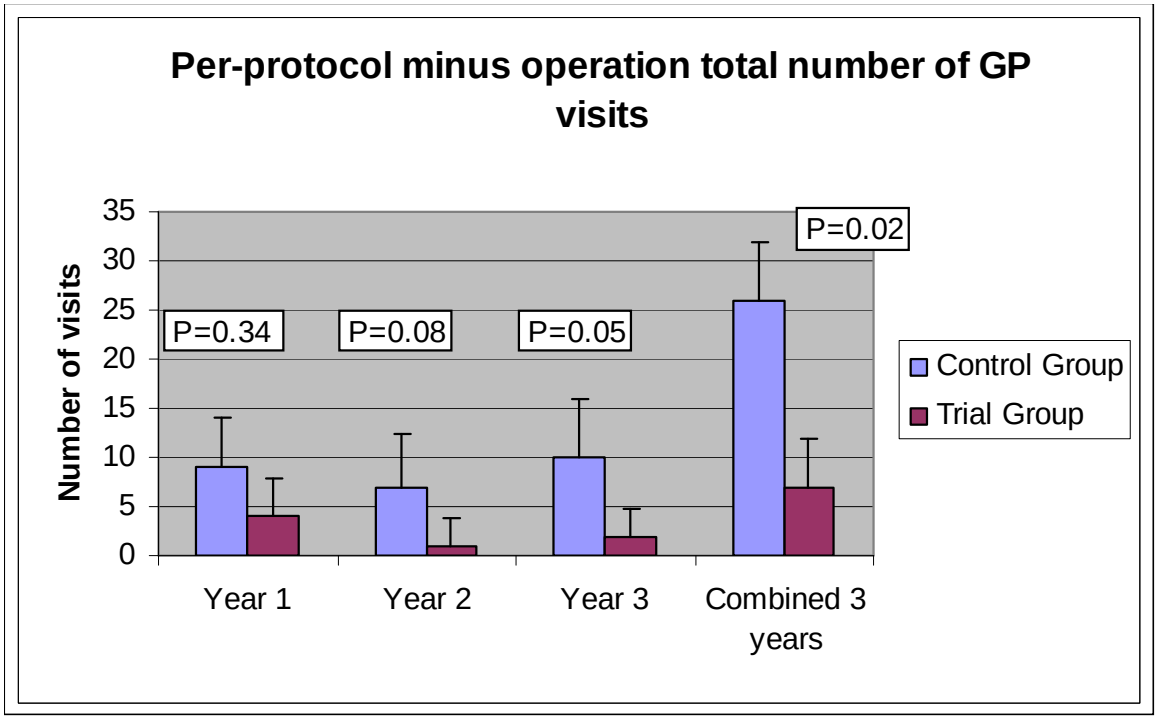


Figure 4.9. Total number of GP visits for the per-protocol minus operation sub-group for each year, as well as the combined three-year result.

In the category of 'medication use', participants in the trial group used less medication compared with the control group in year three of the follow-up period (p =0.05) (*fig 4.10*).

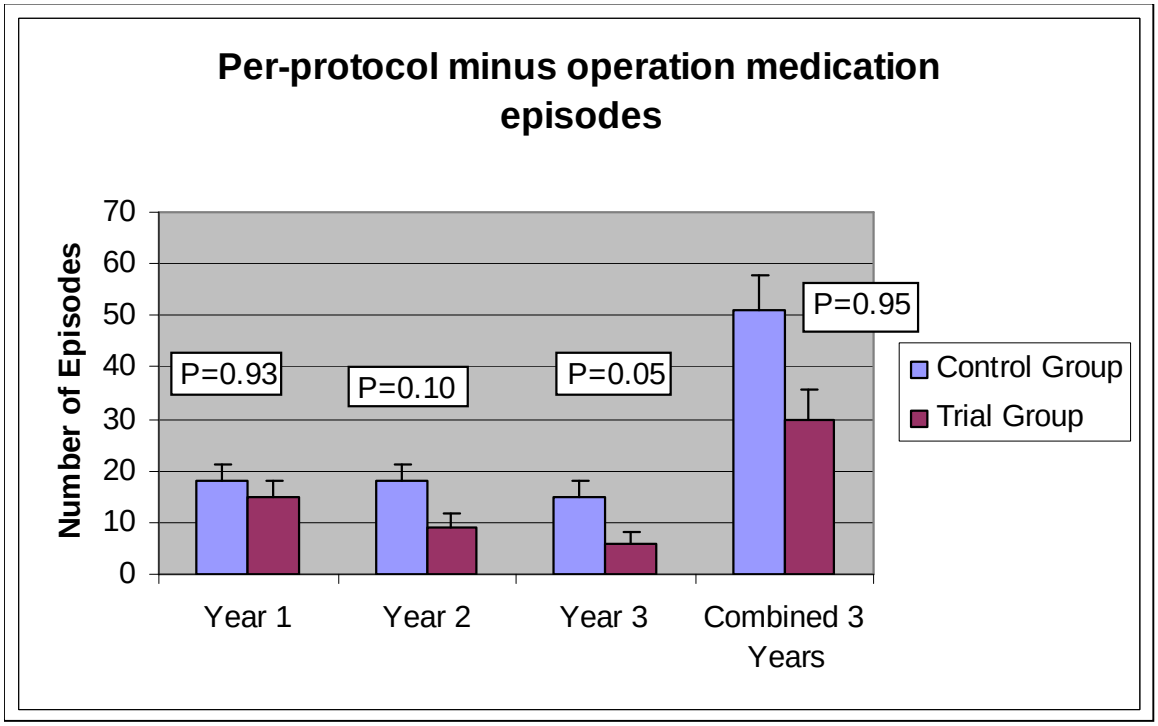


Figure 4.10. Number of medication episodes in the per-protocol minus operation sub-group for each year, as well as the combined three-year result.

The combined three year result for the ‘days off work’ category in this population exhibits a difference of 122 days for the control group compared with 8 days for the trial group; this represents a difference of 22.8 weeks (5 day working weeks) between the groups in terms of paid weekly compensation. Although this is a difference in terms of actual ‘days off’, the result is not statistically significant due to the low numbers of participants from either group taking time off work (*fig 4.11*).

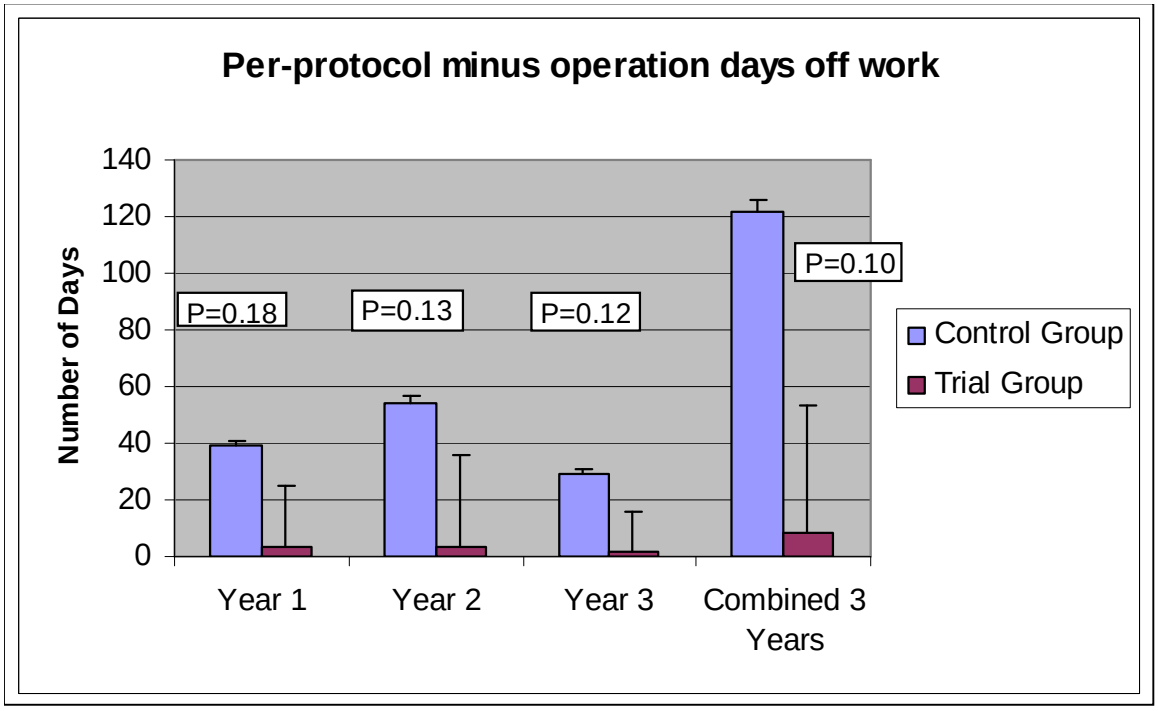


Figure 4.11. The number of days off work for each group in the per-protocol minus operation sub-group for years 1, 2 and 3, as well as the combined three-year result.

The categories of ‘satisfaction’ and ‘other treatment provider’ did not achieve statistical significance in this population.

Table 4.10

Per-protocol minus operation: results of the annual questionnaire, years one, two and three.

Year one	Participant GP visits	Other therapist visits	Med levels	Episodes off work		Total No of GP visits	Days off work
Count within control grp (%)	5 (12.2)	2 (4.9)	18 (43.9)	6 (14.6)	Mean(SE)	0.21(0.10)	0.95(0.54)
Count within trial grp (%)	2 (5.7)	5 (14.3)	15 (42.9)	2 (5.7)	Mean(SE)	0.11(0.09)	0.09(0.06)
Total Number	76.00	76.00	76.00	76.00		76.00	76.00
P-value	0.44	0.24	0.93	0.28		0.34	0.18
Year two	**Participant GP visits**	**Other therapist visits**	**Med levels**	**Episodes off work**		**Total No of GP visits**	**Days off work**
Count within control grp (%)	6 (14.6)	4 (9.8)	18 (43.9)	5 (12.2)	Mean(SE)	0.17(0.07)	1.32(0.08)
Count within trial grp (%)	1 (2.9)	4 (11.4)	9 (25.7)	1 (2.9)	Mean(SE)	0.03(0.03)	0.09(0.09)
Total Number	76.00	76.00	76.00	76.00		76.00	76.00
P-value	0.12	1.00	0.10	0.21		0.08	0.13
Year three	**Participant GP visits**	**Other therapist visits**	**Med levels**	**Episodes off work**		**Total No of GP visits**	**Days off work**
Count within control grp (%)	7 (17.5)	2 (5.0)	15 (37.5)	5 (12.5)	Mean(SE)	0.25(0.09)	0.73(0.35)
Count within trial grp (%)	1 (2.9)	1 (2.9)	6 (17.1)	1 (2.9)	Mean(SE)	0.06(0.06)	0.06(0.06)
Total Number	76.00	76.00	76.00	76.00		76.00	76.00
P-value	0.06	1.00	0.05*	0.21		0.05*	0.12

Note: grp = group; GP = General Practitioner; Med = Medication; No = number; SE = standard error; * = $P \leq 0.05$

The results from the annual questionnaire in all populations (intent-to-treat, per-protocol and per-protocol minus operation) for the three year combined analysis, demonstrated a statistical advantage for the trial group in the categories of ‘participant GP visits’ and ‘total number of GP visits’. This indicates that compared to the control group, a smaller proportion of participants in the trial group visited their doctors, and also, the total number of doctor visits in the trial group was lower than in the control group. In total, 17 control group participants visited their doctor for LBP and six trial group participants did so during the three year follow-up period. These results are demonstrated in *Figures 4.12* (combined three year participant GP visits, all sub-groups) and *4.13* (combined three year total number of GP visits, all sub-groups).

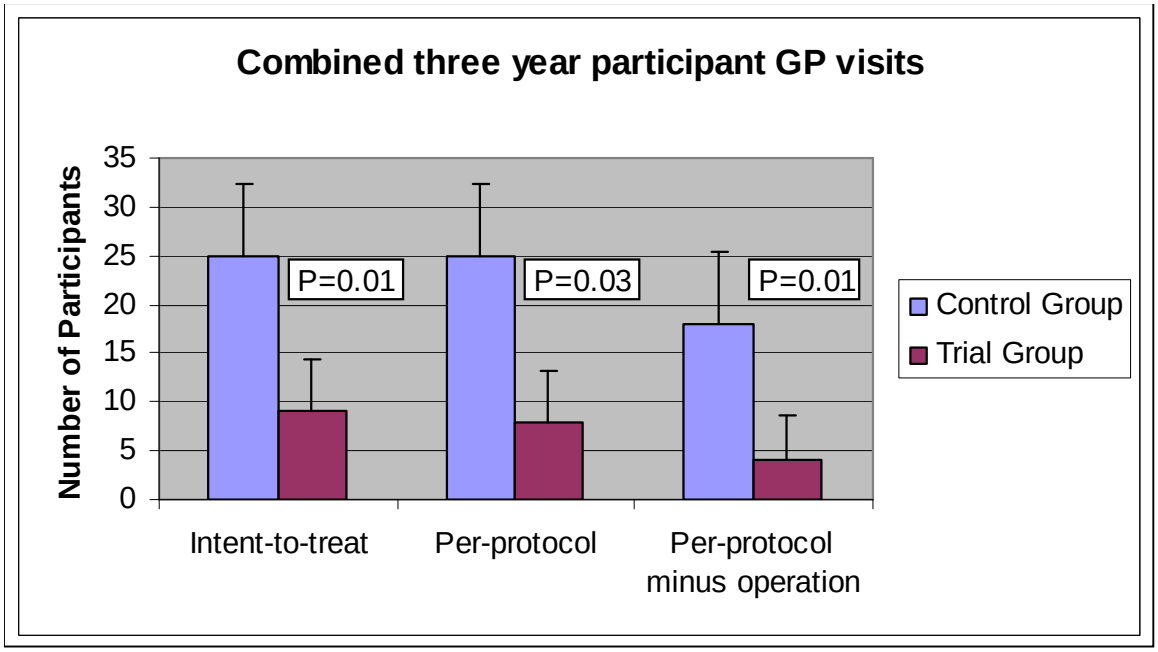


Figure 4.12. The data for 'participant GP visits' for the combined three-year follow-up in each sub-group: intent-to-treat, per-protocol and per-protocol minus operation

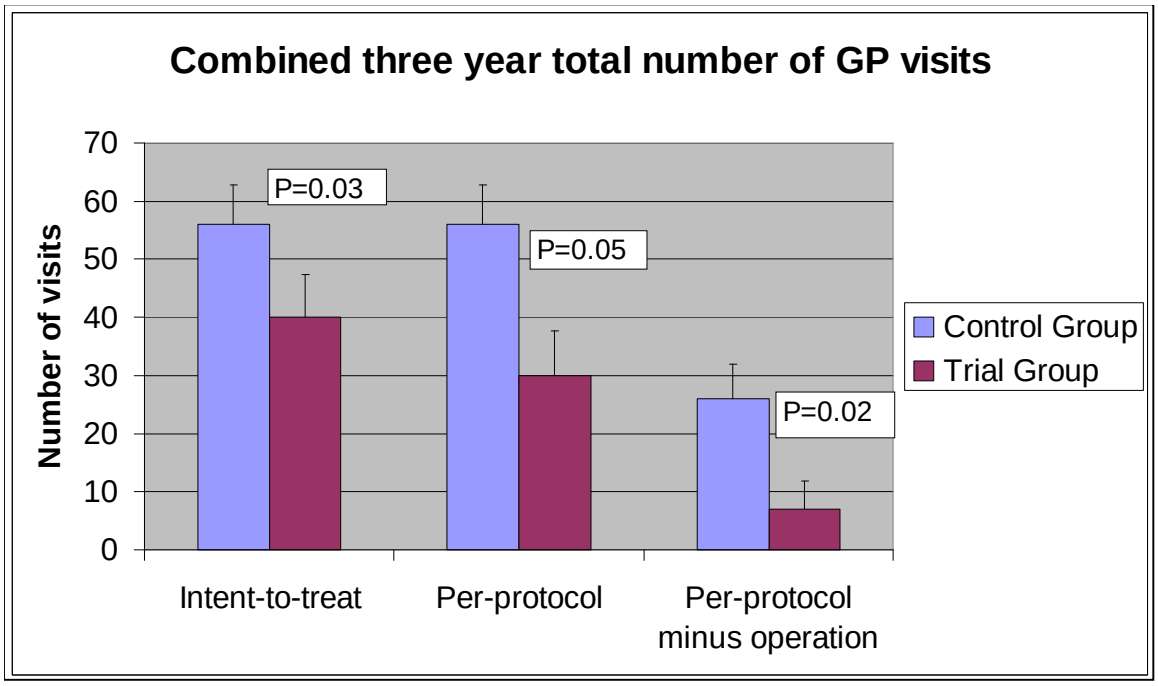


Figure 4.13. The data for the 'total number of GP visits' for the combined three-year follow-up in each sub-group: intent-to-treat, per-protocol and per-protocol minus operation.

4.6 The 'no-exercise participants in the control group' versus the trial group

4.6.1 Intent-to-treat, per-protocol and per-protocol minus operation analysis

The reason for analysing this data was to assess whether participants in the control group performed exercise of their own volition after surgery. This data, collected from the annual questionnaire, revealed that the majority of participants did perform some form of exercise after surgery, this was a confounding factor in this study. Collection of this data provided an indication as to the size or effect of this confounding factor. The no-exercise population in the control group were those participants who did not report participating in any exercise during the three year period of the study. This issue will be further discussed in Chapter Five.

The results from this population should be considered with caution due to the low number of participants in the 'no exercise' group (N = 9). For the same reason, results from the annual questionnaire data were not considered.

In the intent-to-treat population, no statistical significance was achieved with this data set in terms of the RM and OLBI outcome measures. Some categories in the SF36 outcome scores achieved statistical significance and these are outlined in Table 4.11.

Table 4.11
Intent-to-treat: 'no exercise group' versus 'trial group' mean, standard error, number and p-value for the SF36 outcome measure.

	No exercise group			Trial group			
Weeks & Category	Mean	SE	Number	Mean	SE	Number	P-value
58 weeks							
Social function 58	65.3	9.1	9	85.2	11.8	43	0.02
Role emotional 58	63.0	8.8	9	85.3	11.9	43	0.04
OMH 58	65.7	9.1	9	78.8	10.9	43	0.03
136 weeks							
Social function 136	60.4	8.8	6	83.1	12.0	41	0.02
OMH 136	62.5	9.1	6	78.1	11.3	41	0.04
Total score 136	62.6	9.1	6	78.5	11.4	41	0.03

Note: SE = standard error; OMH = Overall Mental Health

In the per-protocol population of this cohort, again, the RM and OLBI did not reach statistical significance at any time points. The SF36 results provided a greater number of statistically significant scores compared with the intent-to-treat data set but overall did not

demonstrate a significant difference between the groups. These results are summarised in Table 4.12.

Table 4.12

Per-protocol: 'no exercise group' versus 'trial group' mean, standard error, number and p-value for the SF36, Roland-Morris and Oswestry low back index outcome measures.

	No exercise group			Trial group			
Weeks & Category	Mean	SE	Number	Mean	SE	Number	P-value
32 weeks							
Social function 32	69.4	10.2	9	85.5	12.6	37	0.02
58 weeks							
Role physical 58	50.0	7.3	9	71.9	10.4	39	0.04
Social function 58	65.3	9.5	9	87.2	12.6	39	0.01
Role emotional 58	63.0	9.1	9	88.0	12.8	39	0.01
OPH 58	60.6	8.8	9	74.9	10.9	39	0.05
OMH 58	65.7	9.5	9	74.9	10.9	39	0.01
Total score 58	63.3	9.2	9	79.0	11.5	39	0.02
136 weeks							
Vitality 136	52.5	8.0	6	67.3	10.2	38	0.04
Social function 136	60.4	9.2	6	84.9	12.9	38	0.01
OPH 136	60.9	9.2	6	77.9	11.8	38	0.03
OMH 136	62.5	9.5	6	79.4	12.0	38	0.02
Total score 136	62.3	9.4	6	80.2	12.2	38	0.01

Note: SE = standard error; OPH = Overall Physical Health; OMH = Overall Mental Health

In the per-protocol minus operation population, the results are similar to those of the previous two groups analysed, with slightly more significant scores in the SF36 outcome measure and one statistically significant score in the OLBI, which is probably a chance result. The statistically significant results for the per-protocol minus operation population are outlined in Table 4.13.

Table 4.13

Per-protocol minus operation; 'no exercise group' versus 'trial group', mean, standard error, number and p-value for the SF36, Roland-Morris and Oswestry low back index outcome measures.

	No exercise group			Trial group			
Weeks & Category	Mean	SE	Number	Mean	SE	Number	P-value
32 weeks							
Social function 32	69.4	10.8	9	87.1	13.6	32	0.01
58 weeks							
Role physical 58	50.0	7.6	9	79.4	12.0	34	0.03
General health 58	66.1	10.0	9	80.6	12.2	34	0.03
Social function 58	65.3	9.9	9	89.0	13.5	34	0.00
Role emotional 58	63.0	9.6	9	90.2	13.6	34	0.00
OPH 58	60.6	9.2	9	76.4	11.6	34	0.03
OMH 58	65.7	10.0	9	82.3	12.5	34	0.00
Total score 58	63.3	9.6	9	80.5	12.2	34	0.00
136 weeks							
Bodily pain 136	61.0	9.7	6	77.2	12.3	33	0.04
Vitality 136	52.5	8.3	6	68.2	10.8	33	0.03
Social function 136	60.4	9.6	6	84.5	13.4	33	0.01
OPH 136	60.9	9.6	6	78.7	3.0	33	0.02
OMH 136	62.5	9.9	6	80.1	12.7	33	0.02
Total score 136	62.6	9.9	6	80.8	12.8	33	0.01
156 weeks							
General health 156	62.5	10.1	6	81.0	13.1	32	0.03
136 weeks							
OLBI 136	14.7	1.6	6	6.7	1.3	35	0.04

Note: OLBI = Oswestry low back index; SE = standard error; OPH = Overall Physical Health; OMH = Overall Mental Health

4.7 Trainer A (principal researcher) versus trainer B

This comparison was only made for the intent-to-treat cohort due to the low numbers in each group. Trainer A supervised 28 participants and trainer B supervised 14. Analysis of the OLBI and RM outcome measures provided no statistically significant results. This means that trainer A and B produced similar outcomes. This evidence demonstrated that the outcomes achieved by the trial group were related to the rehabilitation programme rather than to the person who supervised the programme.

There were a small number of significant results in the analysis of the SF36 data. These results are summarised in Table 4.14. All of these values were in favour of trainer B.

However, from a statistical viewpoint, it is likely that these results have occurred by chance, or are a consequence of chance differences between the groups prior to the intervention.

Table 4.14
Intent-to-treat results of Trainer A versus Trainer B: mean, standard error, number and p-value for the relevant categories in the SF36 outcome measure.

	Trainer A			Trainer B			
Category & weeks	Mean	SE	Number	Mean	SE	Number	P-value
Vitality 58	63.9	2.8	27	75.0	1.6	12	0.05
Vitality 110	59.4	3.6	26	74.5	1.9	10	0.04
Vitality 136	63.3	2.2	27	76.8	2.3	11	0.01
Mental Health 156	77.9	2.7	26	89.8	1.3	11	0.03

Note: SE = standard error

4.8 Analysis of trends within the data

The key questions that require answers from the data are:

- Did the subjective functional outcome measures correlate?
- What trends, if any, were found within each subjective functional outcome measure, considering all three sub-groups; intent-to-treat, per-protocol and per-protocol minus operation?
- Was there a correlation of trends within the subjective functional outcome measures?
- Were there any correlations between the subjective functional outcome measures and the objective data from the annual questionnaire?
- Was there a correlation between poor cognitive results from the SF36 outcome measure and the Roland-Morris, Oswestry Low Back Pain Index and the objective findings from the annual questionnaire?

These questions will be discussed in relation to the results of this study in the following sections.

4.8.1 Did the subjective functional outcome measures correlate?

Using a Pearson's correlation test and comparing the OLBI with the RM (intent-to-treat), this analysis demonstrates a correlation of 0.8 and above for all time points from 23 weeks to 156 weeks post-surgery. The only time point that fell below the critical threshold of 0.7 (Roland & Fairbank, 2000) was at base line, six weeks post-surgery (0.67). However, all but one correlation is greater than 0.7, and hence, this indicates a reliable and robust correlation between the RM and OLBI outcome measures. This data is summarised in Table 4.15.

Table 4.15

Intent-to-treat; Pearson correlations comparing relevant categories of the SF36, Oswestry low back index and Roland-Morris outcome measures.

		OLB6	OPH6	OMH6	TOT SCORE 6
RM6	Pearson correlation	0.67	-0.71	-0.60	-0.70
	N	86	87	87	87
OLB6	Pearson correlation		-0.68	-0.54	-0.65
	N	86	86	86	86
		OLB32	OPH32	OMH32	TOT SCORE 32
RM32	Pearson correlation	0.84	-0.79	-0.56	-0.74
	N	84	84	84	84
OLB32	Pearson correlation		-0.77	-0.59	-0.75
	N	85	85	85	85
		OLB58	OPH58	OMH58	TOT SCORE 58
RM58	Pearson correlation	0.90	-0.83	-0.76	-0.82
	N	85	85	85	85
OLB58	Pearson correlation		-0.80	-0.72	-0.80
	N	86	86	86	86
		OLB84	OPH84	OMH84	TOT SCORE 84
RM84	Pearson correlation	0.86	-0.76	-0.62	-0.72
	N	87	86	86	86
OLB84	Pearson correlation		-0.77	-0.65	-0.75
	N	87	86	86	86
		OLB108	OPH110	OMH110	TOT SCORE 110
RM108	Pearson correlation	0.87	-0.82	-0.81	-0.86
	N	84	79	79	79
OLB108	Pearson correlation		-0.81	-0.75	-0.81
	N	84	79	79	79
		OLB132	OPH136	OMH136	TOT SCORE 136
RM132	Pearson correlation	0.83	-0.78	-0.58	-0.75
	N	81	79	79	79
OLB132	Pearson correlation		-0.83	-0.65	-0.79
	N	81	79	79	79
		OLB156	OPH156	OMH156	TOT SCORE 156
RM156	Pearson correlation	0.86	-0.76	-0.55	-0.70
	N	81	79	79	79
OLB156	Pearson correlation		-0.84	-0.64	-0.79
	N	81	79	79	79

Note: RM = Roland-Morris; OLB = Oswestry low back index; OPH = overall physical health; OMH = overall mental health; TOT SCORE = total score; N = number

Significant correlations were also found between some categories of the SF36, RM and OLBI outcome measures. The SF36 categories of Overall Physical Function, Total Score and to a lesser extent Overall Mental Health, correlated with the RM and OLBI outcome measures. These specific SF36 categories generated correlations greater than 0.7 from six weeks to 156 weeks. The Overall Mental Health category displayed a weaker correlation, generating fewer scores greater than 0.7, compared with the Overall Physical Function and Total Score categories. This is, as expected, due to physical function being more closely related to the RM and OLBI outcome measures than is mental health. This data is also summarised in Table 4.15.

4.8.2 What trends, if any, were found within each subjective functional outcome measure considering all three sub-groups: intent-to-treat, per-protocol and per-protocol minus operation?

Section 4.8.1 outlined the correlations between each outcome measure at each time point. This section, 4.8.2, outlines the trends generated within each individual outcome measure for the whole three year time period; these trends are then compared.

Due to there being no statistically significant results found in the RM and OLBI data sets for any of the separate populations (intent-to-treat, per-protocol, or per-protocol minus operation) there were no trends to be found in these populations.

In the SF36 data set the category of vitality demonstrated a statistical difference between the groups in each population; in the intent-to-treat population, five of the seven ‘time points’ were statistically significant; in the per-protocol and per-protocol minus operation populations all time points for this category reached statistical significance (Tables 4.3, 4.6, & 4.9). However, this category of vitality was the only category in the SF36 outcome measure that revealed a trend in the results.

Two other points of interest in the per-protocol SF36 data set included a cluster of statistically significant results at the 56 and 84 week time points (Table 4.6), and in the category of mental health in the per-protocol minus operation population (Table 4.9) that demonstrated a statistical difference between the groups in the first two years post-surgery. However, due to these results not being supported in the other populations at these same time points, these findings may be attributable to chance.

4.8.3 Is there a correlation of trends within the subjective functional outcome measures?

Although there was a correlation of trends between the RM, OLBI, and some categories of the SF36 including the ‘total score’ category, as reported in section 4.8.1, most categories in the SF36 outcome measure did not correlate well with either the RM or OLBI.

However, if a subset of SF36 questions that related directly to the participants’ cognitive/emotional state of mind was compared with the RM and OLBI results, a direct relationship between poor cognitive/emotional scores in the SF36 and poor functional scores in both the RM and OLBI was found. This was a strong correlation across all populations and related to scores of greater than 6 for the RM outcome measure and greater than 20 for the OLBI, compared with questions 5 (a, b, c), Q6, Q9 (b, c, d, f, h) and Q10 of the SF36 outcome measure. The intent-to-treat results for this population are reported in Table 4.16. The results for the per-protocol and per-protocol minus operation populations mirror the results of the intent-to-treat group for these outcome measures. This finding was not reliant on whether participants performed the programme or not, therefore there was no benefit in publishing the results for the per-protocol and per-protocol minus operation groups. This was a secondary finding of this thesis.

Table 4.16

The correlation between poor cognitive scores from the SF36 and high levels of dysfunction reported in the Roland-Morris and Oswestry low back index; intent-to-treat.

RM 58wks		Q5a	V20 5b	V21 5c	Q6	V26 9b	V27 9c	V28 9d	V30 9f	V32 9h	Q10
RM ≤ 6	Mean	1.94	1.86	1.90	1.27	5.46	5.67	2.84	5.43	2.35	4.41
	(SE)	(0.03)	(0.04)	(0.03)	(0.05)	(0.08)	(0.08)	(0.14)	(0.07)	(0.010)	(0.11)
RM > 6	Mean	1.55	1.36	1.64	2.64	5.05	5.05	3.68	4.59	3.05	3.36
	(SE)	(0.06)	(0.05)	(0.05)	(0.11)	(0.11)	(0.11)	(0.14)	(0.13)	(0.11)	(0.11)
	Number	85	85	85	85	85	85	85	85	85	85
	P-value	0.00	0.00	0.00	0.00	0.05	0.00	0.01	0.00	0.00	0.00
OLBI 58wks											
OLBI ≤ 20	Mean	1.91	1.81	1.88	1.34	5.4	5.66	2.96	5.37	2.43	4.36
	(SE)	(0.03)	(0.04)	(0.04)	(0.05)	(0.08)	(0.07)	(0.14)	(0.07)	(0.10)	(0.10)
OLBI > 20	Mean	1.58	1.42	1.68	2.63	5.16	5.0	3.47	4.63	2.95	3.37
	(SE)	(0.05)	(0.05)	(0.05)	(0.12)	(0.12)	(0.12)	(0.12)	(0.14)	(0.12)	(0.13)
	Number	86	86	86	86	86	86	86	86	86	86
	P-value	0.00	0.00	0.04	0.00	0.26	0.00	0.14	0.00	0.05	0.00
RM 110 wks											
RM ≤ 6	Mean	2.00	1.95	1.96	1.32	5.54	5.77	2.70	5.46	2.26	4.72
	(SE)	(0.02)	(0.03)	(0.03)	(0.07)	(0.09)	(0.05)	(0.11)	(0.08)	(0.09)	(0.07)
RM > 6	Mean	1.59	1.36	1.59	2.64	4.86	5.09	3.73	4.68	3.09	3.59
	(SE)	(0.05)	(0.05)	(0.05)	(0.12)	(0.11)	(0.13)	(0.13)	(0.13)	(0.13)	(0.11)
	Number	79	79	79	79	79	79	79	79	79	79
	P-value	0.00.	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
OLBI 110wks											
OLBI ≤ 20	Mean	2.00	1.90	1.95	1.33	5.49	5.71	2.79	5.44	2.27	4.65
	(SE)	(0.02)	(0.04)	(0.03)	(0.07)	(0.08)	(0.07)	(0.12)	(0.08)	(0.10)	(0.07)
OLBI > 20	Mean	1.44	1.31	1.50	3.06	4.81	5.06	3.75	4.44	3.38	3.44
	(SE)	(0.06)	(0.05)	(0.05)	(0.16)	(0.12)	(0.13)	(0.14)	(0.12)	(0.12)	(0.12)
	Number	79	79	79	79	79	79	79	79	79	79
	P-value	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00
RM 156 wks											
RM ≤ 6	Mean	1.95	1.92	1.90	1.34	5.54	5.63	2.64	5.36	2.27	4.63
	(SE)	(0.02)	(0.03)	(0.03)	(0.08)	(0.10)	(0.10)	(0.13)	(0.13)	(0.09)	(0.09)
RM > 6	Mean	1.75	1.50	1.75	2.25	4.75	4.95	3.80	4.90	3.05	3.75
	(SE)	(0.05)	(0.06)	(0.05)	(0.10)	(0.15)	(0.13)	(0.12)	(0.09)	(0.11)	(0.10)
	Number	79	79	79	79	79	79	79	79	79	79
	P-value	0.01	0.00	0.10	0.00	0.00	0.01	0.00	0.10	0.00	0.00
OLBI 156wks											
OLBI ≤ 20	Mean	1.95	1.89	1.92	1.34	5.51	5.66	2.69	5.40	2.31	4.62
	(SE)	(0.02)	(0.03)	(0.03)	(0.08)	(0.10)	(0.09)	(0.13)	(0.11)	(0.09)	(0.09)
OLBI > 20	Mean	1.64	1.43	1.57	2.64	4.57	4.50	4.07	4.50	3.21	3.43
	(SE)	(0.06)	(0.06)	(0.06)	(0.10)	(0.17)	(0.16)	(0.12)	(0.12)	(0.10)	(0.09)
	Number	79	79	79	79	79	79	79	79	79	79
	P-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Note: RM = Roland-Morris; OLBI = Oswestry low back index; wks = weeks; SE = standard error; Q = Question

4.8.4 Were there any correlations between the RM, OLBI and SF36 outcome measures and data from the annual questionnaire?

The answer to this question was essentially no. The RM and OLBI provided no statistically significant results in any populations and therefore, were unable to reveal any trends. The SF36 outcome measure did demonstrate some statistically significant results but these results represented no trend or relationship with the results found in the QoL annual questionnaire data. The only trends in the QoL data set were found in the the three year combined results for ‘participant doctor visits’ and ‘total number of doctor visits’ categories and these findings have no correlation with the statistically significant results found in the SF36 outcome measure.

Both data sets provided a number of results that were close to statistical significance but did not quite reach the threshold ($p < 0.05$). The rational for selecting the original sample size of 94 participants will be discussed in Chapter Five.

CHAPTER FIVE DISCUSSION

5.1 Introduction

The study that contributes to this thesis is the first randomised controlled trial investigating the post-surgical management of lumbar discectomy patients that compared a physiotherapy rehabilitation programme with a control group that did not participate in any formal post-surgical rehabilitation, and followed participants for three years. The results of the study did not demonstrate a clear benefit of this rehabilitation programme for lumbar discectomy patients.

Two other studies, using shorter follow-up periods, have compared a post-surgical physiotherapy exercise group with a control group that did not include any formal post-surgical rehabilitation (Dolan et al. (2000) & Erdogmus et al. (2007)). Dolan et al. conducted a pilot study that followed participants for one year and Erdogmus et al. followed participants for 18 months. The Dolan study reported a statistical advantage for the physiotherapy rehabilitation programme at the end point of the study. Erdogmus et al. reported an advantage for the physiotherapy programme at the end point of the intervention (12 weeks post-surgery) but then reported that this advantage had been lost by the end point of follow-up at 18 months post-surgery.

In the following sections the key findings from the results of this study are discussed. The study investigated in this thesis is compared with previous studies that have been published in the literature on the topic of post-surgical management of lumbar discectomy. The key points of the rehabilitation programme are discussed with reference to the literature. The outcome measures used in this study and the results found are compared with other studies published on this topic. Further to these findings, issues of clinical implications, limitations of this study, recommendations and finally the conclusions are reported.

5.2 Summary of the results of the study investigated in this thesis

The key findings from the results of this thesis are:

- The results of the RM, OLBI and SF36 did not distinguish a difference between the control and trial groups.

- The category of ‘vitality’ in the SF36 outcome measure demonstrated a consistent trend in favour of the trial group in all three populations analysed (intent-to-treat, per-protocol and per-protocol minus operation). This suggests that trial group participants felt on average more ‘vital’ compared with control group participants.
- With the exception of the ‘doctor visits’ category in the combined three year data, the QoL annual questionnaire did not distinguish a difference between the control and trial groups.
- There was a strong correlation between the RM and OLBI outcome measures.
- There was a strong correlation between poor cognitive state of mind and high levels of physical dysfunction.
- There was a strong correlation between the RM, OLBI and some of the categories (overall physical health, overall mental health and total score) in the SF36 outcome measures.
- The results that compared trainer A with trainer B, supported in a limited manner the point that the rehabilitation programme was not therapist dependent.
- A significant number of control group participants chose to take part in some form of exercise after surgery.

The lack of statistically significant results generated from the subjective functional questionnaires used in this study were similar to the findings of previous research on this topic. Of the 16 studies reviewed in Tables 2.2 and 2.3, ten reported no statistical difference between the groups at the end of their respective follow-up periods. The outcome measures used in these studies were subjective functional questionnaires and time off work. Six studies reported a statistical advantage for the trial group; one of these studies used return-to-work data only (Donceel et al., 1999), while the other five studies used subjective functional questionnaires and time off work.

The category of ‘vitality’ in the SF36 results was noteworthy, as this was the only category in all of the subjective functional questionnaires that demonstrated a trend in all populations (intent-to-treat, per-protocol, and per-protocol minus operation). Due to the consistency of significant results in all populations during the complete three year follow-up, this trend was not likely to be due to chance. This result demonstrated that the trial group

participants continued to feel more vital compared with the control group participants, long after the completion of the programme.

This feeling of vitality may be attributed to trial group participants continuing to exercise, as they were encouraged to continue some form of exercise at the completion of the rehabilitation programme. However, a review of the final annual questionnaire that asked participants whether they were continuing to exercise revealed that of the 45 completed questionnaires in the trial group, 26 were exercising and 19 not. In the control group, out of 42 completed questionnaires, 20 were exercising and 21 not. Approximately half the participants in each group were exercising of their own accord. This finding nullifies the argument that trial participants felt more vital because they were exercising. Hence, it is difficult to explain the difference between the groups in relation to the 'vitality' category found in the SF36 measure.

The only statistically significant differences found in the QoL questionnaire, in all populations, were the numbers of participants who visited their doctor, and the total number of visits they had with their doctors during the combined three year follow-up. Seventeen control group participants visited their doctors during the three year follow-up period, compared with six trial group participants.

A potential confounding issue related to these data was the practice of the participants' doctors. Some doctors may have monitored their patients more closely than others, which would have led to increased doctor visits. Randomisation would usually balance this factor between groups. If the 'doctor visit' category was affected by some doctors seeing participants for purposes of monitoring rather than direct issues related to pain, it is assumed that this monitoring would have been in the first year post-surgery. It would be unusual for a doctor to monitor a patient after lumbar discectomy surgery for more than one year. The results show that in year one, eight control and four trial group participants visited their doctors. In year two, nine control and three trial participants visited their doctors. In year three, eight control and two trial participants visited their doctors. These data indicated that some participants visited their doctors more than once during the follow-up period. These results suggest that 'doctor monitoring' was not a significant confounding factor in this category.

The category of 'other treatment providers' in the QoL annual questionnaire, asked if participants had sought treatment for their low back from treatment providers other than their

doctor. The results for this category generated no statistical difference between the groups. Considering that participants in the trial group had easy access to a supervising physiotherapist, it might be expected that they would be more likely to complain of problems and seek treatment for any ailments relating to their low back. Control group participants had to make a specific effort to contact their doctor or another treatment provider. The fact that trial group participants saw another treatment provider no more or less than the control group participants, despite the easy availability for them to do so, potentially indicated that either trial group participants had fewer problems that required attention or they were able to self-manage more effectively.

An important underlying philosophy of the rehabilitation programme was to reduce reliance on medical support and promote independence. This was achieved by answering any queries that participants had regarding their low back, and by reassuring them that what they were experiencing was normal and would eventually settle. This philosophy is supported in the literature (Burton et al., 1999; Pfingsten et al., 1997; Poulter, 1999). This was particularly important when participants were still aware of significant low back or leg pain in the early stages after surgery. This reassurance provided the participant with positive cognitive reinforcement which was an integral component of the rehabilitation programme. The findings in the ‘other treatment provider’ and the ‘doctor visit’ categories suggested that the rehabilitation programme, and the method in which it was supervised, encouraged self-management and independence.

There was a strong correlation between the RM and OLBI outcome measures. With the exception of baseline scores at six weeks, Pearson’s correlations for the RM and OLBI were greater than 0.8 for all time points. Roland and Fairbank (2000) reported that a correlation of greater than 0.7 was regarded as strong. Both the RM and OLBI primarily measure physical function. Physical function was a key outcome of this study. This is why two instruments that both measure physical function were used. It was then possible to validate the scores of one against the other and hence obtain a valid judgement of physical function. The strong correlation between the RM and OLBI in this study confirmed that the findings of physical function were valid.

The Pearson correlation also demonstrated a strong correlation between the SF36 categories of ‘overall physical health’, ‘total score’ and, to a lesser extent, ‘overall mental health’ with both the RM and OLBI. As would be expected the category of ‘overall physical

health' correlated well with the RM and OLBI. These latter two instruments both measure physical health. A correlation between the RM, OLBI and total health was also understandable, as five of the categories measured by the SF36 'total score' related to physical function.

The correlation between the RM, OLBI and 'overall mental health' category, was an unexpected finding considering the different determinents measured by these three instruments. The RM and OLBI primarily measure physical function, the OLBI also incorporates a measure of pain (Bombardier et al., 2001) whereas the 'overall mental health' category of the SF36 is a psychological measure. A possible explanation may be found in the correlation between poor/low scores in the cognitive categories of the SF36 and the poor/high scores in the RM and OLBI (Table 4.16). This correlation demonstrated that participants' mental health was closely associated with the reporting of their physical health status.

The strong correlation between poor cognitive scores in the SF36 outcome measure and poor functional scores in the RM and OLBI outcome measures resulted in a linear correlation between these variables (Table 4.16). These results are consistent with previously published literature that has reported a correlation between levels of activity and levels of well-being (North et al., 1990; Ruuskanen & Ruoppila, 1995; Tsutsumi et al., 1998; Watanabe et al., 2001). Higher levels of activity are associated with higher levels of well-being and low levels of activity are associated with low levels of well-being.

Two physiotherapists (trainers A and B) supervised the participants; trainer A is the PI and trainer B is a physiotherapist familiar with the rehabilitation programme used in this study. The purpose of using two different physiotherapists in the study was to provide some support for the generalisation of results. A small number of statistically significant results in the SF36 outcome measure were found in the comparative analysis of the two groups that were supervised by different physiotherapists. Although these results may have occurred by chance due to the large number of statistical analyses performed (Type I error), they all favour trainer B.

A possible explanation of this is that all participants supervised by trainer B completed their rehabilitation programmes in the physiotherapy clinic in which trainer B works. This meant that the study participants maintained regular informal contact with trainer B. Both trainers A and B consulted nine times with the participants, on a formal basis, throughout the whole programme. However, all trainer A's participants completed their programmes at

gymnasiums across the city, therefore, they had contact with trainer A only on occasions when their programmes were being progressed. This confounding factor might explain the few results found in favour of trainer B.

The use of a second physiotherapist to perform the programme was an attempt to demonstrate the generalisability of the rehabilitation programme. This has been limited by the fact that trainer B was an experienced clinician already familiar with the programme. The ideal method would have been to have two independent supervising physiotherapists who were not already familiar with the rehabilitation programme but who had been taught how to administer the programme. Logistically, due to lack of funding to train and supervise different physiotherapists to teach and monitor the rehabilitation programme this was not possible.

The final issue relating to the results found in this thesis was the number of control group participants who took part in some form of exercise after their surgery. Only nine control group participants did not exercise after their surgery during the three year follow-up, meaning the majority undertook some exercise after surgery. This was a confounding factor in this study, as it meant that the rehabilitation programme was being measured against other exercise programmes rather than against none. This has the potential to change the hypothesis of this study; in comparing a formal physiotherapy rehabilitation programme with a patient-centered self-managed programme. Erdogmus et al.(2007) also found a significant number of participants undertook exercise after surgery; these authors comment that this was a confounding factor in their study.

Although this is a confounding factor for studies of this nature, it raises an important question. Why did so many participants exercise after surgery? Was it because they realised they were being compared with a population that was performing exercise? Did this lead them into thinking that exercise was an appropriate option after lumbar discectomy surgery? Or was their own form of self-managed post-surgical exercise a normal occurrence that happens anyway, and that goes unnoticed by surgeons, doctors and physiotherapists? Further research is required to answer this.

5.3 A comparison of this thesis study with previous research

In the Cochrane review of Rehabilitation After Lumbar Disc Surgery, Ostelo et al. (2004) stated that there was strong evidence (level 1) “that intensive exercise programmes are more effective on functional status and faster return to work (short term follow-up)” for programmes that commenced between 4 - 6 weeks post-surgery (p. 2). However, the authors also reported level 1 evidence that concluded no advantage for intensive exercise regimes on long term follow-up.

The review by Ostelo et al. (2004) compared the results of 13 papers, six of which were considered to be of high quality. The other seven papers (Alaranta et al., 1986; Brennan et al., 1994; Burke et al., 1994; Johannsen et al., 1994; Kitteringham, 1996; Rothhaupt et al., 1997; Timm, 1994) had methodological issues that limited their ability to provide conclusive outcomes and recommendations. Ostello et al. grouped these papers together in a ‘low quality’ category. Another four papers were not included in this review, as they were published after the review process had been completed. Three of these papers fitted the ‘low quality’ category due to their short follow-up time period (Filiz et al., 2005; Ostelo et al., 2003; Yilmaz et al., 2003). The fourth paper, by Erdogmus et al. (2007), satisfied the criteria for high quality set by Ostello et al.. This left seven papers, including that by Erdogmus et al. (2007), on the topic of post-surgical management of lumbar discectomy that were considered to be high quality (Danielsen et al., 2000; Dolan et al., 2000; Donceel et al., 1999; Kjellby-Wendt & Styf, 1998; Manniche, Asmussen et al., 1993; Manniche, Skall et al., 1993).

There are several points that differentiate this thesis study from previous research. They are, 1) a large cohort (N=94), 2) the rehabilitation programme was periodized, 3) the control group did not follow a formal post-surgical exercise programme, and 4) the follow-up time period was of three years duration.

The study by Manniche et al. (1993a) used 96 participants; Erdogmus (2007) used 120 participants but these 120 participants were separated into three groups (the number for each group was 40). The study by Donceel et al. (1999) involved 710 individuals. However, these 710 individuals were managed by 60 case managers, and therefore, statistically for this study, the number per group was 30. The other four studies that satisfied the Cochrane review ‘high quality’ category had cohorts of less than 63 participants.

The exercise programme in this thesis consisted of a fully periodized six-month intervention. The six-month programme included three phases: conditioning, hypertrophy and strength. None of the previous papers that have studied the topic of post-surgical management of lumbar discectomy have utilised periodization, and the longest intervention periods were twelve weeks. According to the literature, periodized exercise programmes provide superior results compared with non-periodized programmes (Garhammer, 1991; Kraemer, 1997; Kraemer et al., 2003; Kraemer et al., 1997; Marx et al., 2001; O'Bryant et al., 1988; Stone et al., 1997; Stowers et al., 1983; Willoughby, 1993). Manniche et al. in their 1993(b) paper, reported that post-surgical exercise programmes would benefit from being longer than 12 weeks. Kraemer et al. (2004) report that 16 sessions of resistance exercise are required before hypertrophy of muscle begins to occur. Prior to this, substantial strength gains are possible but these gains are from increased neural activity only and have been shown to quickly come and quickly go (Hakkinen et al., 2003 & 1991; Hakkinen & Hakkinen, 1995; Judge, Moreau, & Burke, 2003; Kraemer et al., 1995; Phillips, 2000; Ploutz et al., 1994; Staron et al., 1994).

The six month intervention used in this thesis is twice the length of time of any other programme studied previously. Whilst a longer period of post-surgical rehabilitation may have some support in the literature (Manniche et al., 1993(b); Kraemer et al 2004), the results in this current study did not support an extended post-surgical rehabilitation time frame. However, these results pertain only to the rehabilitation programme used in this study; therefore, a different long-term programme design might generate superior results. This is a potential topic for future research.

This thesis study was a randomised controlled trial which utilised a control group that followed no formal post-surgical rehabilitation programme, compared with a group that undertook a post-surgical exercise rehabilitation programme. The control group in this study followed usual conservative surgical advice, that is, to return to normal activity as soon as pain allowed. Therefore, this study provided the opportunity to test whether a post-surgical physiotherapy rehabilitation programme would provide any functional long-term benefit, when compared to a no-treatment control group. There have only been two previous studies published on this topic that have been recognised for the quality of their method, and that compared an exercise group with a no-treatment control group (Dolan et al. 2000; Erdogmus et al., 2007).

This thesis study followed participants for three years post-lumbar discectomy. Its results did not significantly differentiate between the groups during this time; hence, the long-term follow-up offered no advantage compared with a one year or 18 month period. The combined three year results for ‘participant doctor visits’ and ‘total number of doctor visits’ were however, statistically significant in this study in all populations (intent-to-treat, per-protocol and per-protocol minus operation). However, only the per-protocol minus operation population generated a statistical difference between the groups in the third year of the study to further support the combined three year data. These key differences in ‘doctor use’ were not anticipated when designing the study and would generally require larger sample sizes than were recruited. These outcomes are important but require further confirmation in future studies. Erdogmus et al. (2007) followed participants for 18 months and also found no statistical difference between the groups.

Weber (1983) reported that after four years, the difference between conservatively treated patients and those that underwent discectomy was not significant. Weber stated that any difference after four years was probably due to aging and degeneration, rather than intervention-related. Weber found the two comparison groups to be most differentiated in the first year of follow-up; these differences gradually disappeared by the fourth year of follow-up. Bessette et al. (1996) commented that two years was an optimal follow-up time frame for lumbar-discectomy patients.

The longest follow-up time frame for any previously published randomised controlled trial on this topic is 18 months, in the study by Erdogmus et al. (2007). All other studies followed participants for only one year or less (Alaranta et al., 1986; Brennan et al., 1994; Burke et al., 1994; Danielsen et al., 2000; Dolan et al., 2000; Donceel & Du Bois, 1999; Filiz et al., 2005; Johannsen et al., 1994; Kitteringham, 1996; Kjellby-Wendt & Styf, 1998; Manniche, Asmussen et al., 1993; Manniche, Skall et al., 1993; Ostelo et al., 2003; Timm, 1994; Yilmaz et al., 2003). The literature suggests that between one to two years of follow-up post-lumbar-discectomy is adequate to test for differences between groups and that further follow-up may not be justified.

In summary, the majority of studies that have investigated the post-surgical management of lumbar discectomy have compared one rehabilitation programme with another. These studies all reported a benefit for post-surgical rehabilitation programmes. Some countries have accepted that post-surgical rehabilitation programmes for lumbar

discectomy are an integral part of the surgical experience. Erdogmus et al. (2007) stated that this concept has been accepted by both patients and physicians to the point whereby withholding programmes of this nature would be considered unethical. However, this philosophy does not have universal acceptance.

A review of the literature found only two studies that have compared a post-surgical rehabilitation programme with no treatment. These studies reported contrasting results. One supported post-surgical rehabilitation, and the other did not. The results from this thesis study did not significantly differentiate the control and trial groups, although there were a small number of results in the SF36 and QoL data that did reach statistical significance. The results from these three studies raise two important points. Firstly, that further research is required to clarify whether post-surgical rehabilitation programmes provide an advantage for lumbar discectomy patients compared with conservative surgical advice; that is, to return to normal function as soon as pain allows, and secondly, what type of exercises should be included in a post-lumbar-discectomy rehabilitation programme, that aims to optimise patient function.

5.4 Key points of the exercise programme used in this thesis in relation to previous research

The design of the exercise programme used in this thesis was based on the physiology and principles of exercise prescription as described in section 2.12.6. The key training parameters of the programme were 1) utilisation of a periodized gymnasium based graded strengthening programme, utilising evidence-based time frames, rest periods and multiple set design, and 2) a lumbar spine stabilising programme focussed on an abdominal training regime using specific deep abdominal and Swiss ball exercises to activate both abdominal muscles and multifidus.

The programme was periodized, and consisted of three phases; conditioning, hypertrophy and strength. In keeping with the principle that beginner trainers do not require excessive variation (Stone et al., 1998), the conditioning phase was of eight weeks duration. The hypertrophy and strength phases were both of nine weeks duration and consisted of three, 3-week programmes for each phase. The structure of three weeks per programme maximises variation, which has been shown to optimise progression (Kraemer et al., 2003; Kraemer & Ratamess, 2004; Potteiger, Judge, Cerny, & Potteiger, 1995; Signorile, Carmel, Lai, & Roos, 2005; Smith, 2003). The exercise programmes were performed on three days per week with at

least one day of rest between programmes. This was recommended by Kraemer and Ratamess (2004) and Ratamess et al. (2009) for whole body training sessions. During each phase of the programme, the rest periods between sets were based on micro-cycle guidelines as reported by Kraemer and Ratamess (2004).

Multiple set programmes were used in this study because of the injured patient population for which it was designed. To achieve optimal effect with a single set programme, the individual has to 'work to failure' using between 8 - 12 repetitions (Carpinelli & Otto, 1998; Winett, 2004; Winett & Carpinelli, 1999). Working to the point of failure is not an option for an injured person, as the risk of further injury is too great (Baechle & Earle, 2000). At the same time however, muscle overloading is required to generate strength. By using multiple sets with short rest periods, muscles are forced to recruit greater numbers of motor units (Kraemer et al., 1996; Kraemer & Ratamess, 2004; Ploutz et al., 1994). This process was described by Kraemer et al.(1996) as progressive overload, which is the use of heavy, moderate, and light resistances all in the same exercise session. According to Kraemer et al.(1996) progressive overload "elicits a physiological stimulus for strength gains and tissue growth" whilst not causing an excessive stress to the muscle (p. 365). This increased motor unit recruitment is achieved with lower resistance levels, which is a safer zone for the injured patient to train in (Zatsiorsky, 1995).

Furthermore, multiple sets provide optimal results with fewer exercises that are easier for the injured population to master (Gotshalk et al., 1997; Kraemer, Adams et al., 2002; Kraemer et al., 1997; Marx et al., 2001; Stone et al., 1998). A single set regime requires the individual to perform a number of different exercises per muscle group. Winnett suggested that 10-12 exercises may be used in a rehabilitation programme but reported that often more than this number are used (personal communication October 2009). Winnett provided examples of an upper and lower body programme using the single set philosophy. These programmes consisted of 18 exercises for the lower body programme and 19 exercises for the upper body programme (personal communication October 2009). This means that athletes need to be familiar and skilled with a greater number of exercises each time they train. This is not a problem for trained athletes but may be difficult for an injured patient who has never exercised in a gymnasium environment previously.

The lumbar spine stabilisation exercises used in the programme and described by Richardson and Jull (1995) have demonstrated, in previous studies, to reduce current pain

issues for patients suffering from low back pain and also have the potential to prevent or mimimise the likelihood of recurrent episodes (O'Sullivan et al., 1997; Hides et al., 2001; MacDonald, Moseley, & Hodges, 2006). The Sahrmann method of training the deep abdominal muscles (referred to in section 3.8) has not been clinically tested in randomised clinical trials. However, the concept of abdominal indrawing is identical to that used and proven by Hides et al., the only difference being that the Sahrmann method trains the muscle in a different body position. Therefore, although not scientifically proven, it is possible that the two different exercises might provide similar results.

The swissball lumbar stabilising exercises used in the programme have no scientific evidence to demonstrate their specific effectiveness for reducing or preventing low back pain. These are global exercises and were used as a progression after the participants had learned to activate the local muscles. Global exercises have however, been shown to assist in the stabilisation of the lumbar spine during certain movements (Crisco & Panjabi, 1991: Kavcic et al., 2004a; Panjabi et al., 1989 &1992a).

Specific multifidus re-education was not included in the rehabilitation programme used in the study contributing to this thesis. The reason for this was that some authors indicated that the indrawing technique used to re-activate transverse abdominus also has a role in re-activating multifidus (O'Sullivan et al., 1997; Richardson & Jull, 1995). Hence it was thought that by using the methods reported by Richardson and Jull (1995), and Sahrmann (Diagnosis and Treatment of Movement Impairment Syndrome. Shirley A Sahrmann, pp 373-377), to re-activate the deep abdominal muscles these exercises would also re-activate multifidus. This was an assumption, and a weakness of this study, and has been acknowledged as such in the limitations of this thesis (section 5.7).

A positive cognitive approach was used when communicating with participants. This was not a formal cognitive behavioural management programme based on time, task or pain contingent activities such as that used by Ostello et al. (2003). The approach used in this study was to apply a realistic, optimistic attitude to the participants' return to function. This was achieved by providing appropriate advice regarding work and recreational activity levels, providing advice on posture, avoiding poor movement patterns, and reassuring participants if they had still had issues with pain. An integral component of this approach was educating the participant. If a participant had a question about any aspect of their surgery or how they were feeling at the time, the supervising physiotherapist took time to ensure that the participant's

questions were answered in a way that satisfied his/her query. This concept is supported in the literature (Burton et al., 1999; Cohen & Rainville, 2002; Pfingsten et al., 1997; Poulter, 1999; Rainville et al., 1997; Sollner & Doering, 1997).

Finally, the programme aimed to optimise mental health and vitality through physical activity. Patients tend to experience a sense of well-being and increased confidence as they regain their pre-injury levels of fitness and strength (McCain et al., 1988; North et al., 1990; Ruuskanen & Ruoppila, 1995; Tsutsumi et al., 1998; Watanabe et al., 2001). This decreases fear avoidance and increases levels of functional activity (Frost et al., 1995; Lindstrom et al., 1992; Lively, 2002; Mannion et al., 1999; Pfingsten, 2001; Sculco et al., 2001). This in turn assists the patients in returning to normal activity levels and normalising their daily function.

A deficit of the rehabilitation programme used in this study was that it did not specifically aim to correct pain provocative motor patterns. The literature has demonstrated that lumbar flexion and rotation are the most common movement patterns associated with disc herniation (Adams et al., 2000; Adams & Hutton, 1982, 1983; Callaghan & McGill, 2001; Gordon et al., 1991; Kelsey et al., 1984; McNally et al., 1993; Mundt et al., 1993; Simunic et al., 2004). To prevent reoccurring problems, exercises should focus on correcting pain provocative motor patterns (McGill & Karpowicz, 2009; O'Sullivan, 2000).

Rehabilitation exercise programmes for low back disability should aim to apply minimal forces and avoid excess loading of the lumbar spine (Cholewicki & McGill, 1996). Recent literature has outlined a number of exercises that are appropriate for low back rehabilitation and categorised these exercises in terms of muscle activation levels and lumbar spine compression loads (Kavcic et al., 2004b). Many of the exercises used in the rehabilitation programme for this thesis study were performed with the participant in the seated position. Although participants were taught to maintain lumbar spine neutral, this is often a difficult concept for people to self-manage. In the seated position it is easy to ‘lose form’ with the result that the lumbar spine moves into flexion. This is a high risk position for post-lumbar-discectomy patients. Issues such as these need to be considered when designing exercise programmes for low back disability patients.

5.5 Outcome measures

The measurement of outcomes for population-based research remains complex. The key issue of any instrument used is that it must be valid, reliable and reproducible (Deyo et al., 1998). Outcome measure instruments generally provide hard or soft data. The results from instruments that collect hard data do not always correlate with functional abilities (Bernard, 1990; Boden et al., 1990; Deyo et al., 1998; Jensen et al., 1994; Korres et al., 1992; Simmons et al., 1991; Yen et al., 1993). However, instruments that collect soft data are susceptible to personal bias and other confounding psychosocial factors that may not be directly related to the outcome that is being measured (Deyo et al., 1994). Therefore, to obtain optimal results, a combination of hard and soft outcome instruments are required (Deyo et al., 1994; Lee et al., 2001).

When this thesis study was initiated in 2002, evidence suggested that the Roland-Morris, Oswestry Low Back Index and SF36 provided the most reliable data on functional outcomes (Beurskens et al., 1996; Bombardier et al., 2001; Fairbank & Pynsent, 2000; Hutchinson et al., 2000; Roland & Fairbank, 2000). These measures have been further validated by other studies (Brouwer et al., 2004; Hakkinen et al., 2003; Ostelo & de Vet, 2005; Ostelo, de Vet, Knol et al., 2004; Pengel et al., 2004; Sigl et al., 2006; Wittink et al., 2004).

The results of the outcome measure data in this study, with the exception of the ‘doctor visit’ category, did not provide a clear distinction between the control and trial groups. Four possible explanations for these findings are 1) the design of the rehabilitation programme, 2) the outcome measures used, 3) the statistical power calculated for this study, and 4) potential participant memory bias.

The philosophy underpinning the rehabilitation programme was that by increasing participants’ whole body strength, this would also improve their overall function. This concept is supported in the literature in cases when negative psychosocial factors are absent (Albright et al., 2001; Brill et al., 1999; Canning et al., 2004; Hurley & Scott, 1998; O'Reilly et al., 1998; Rainville, Hartigan, Jouve, & Martinez, 2004; Rainville, Hartigan, Martinez et al., 2004; Rainville et al., 1997; Vuori, 2001).

However, in contrast, other studies have demonstrated a stronger correlation between psychosocial factors and function (Bigos et al., 1991; Burton et al., 1995; Carragee et al.,

1999; Donceel & Du Bois, 1999; Dvorak, Gauchat, & Valach, 1988; Dvorak, Valach et al., 1988; Graver et al., 1999; Hasenbring et al., 1994; Junge et al., 1996). The results found in the study investigated in this thesis support this research. These results demonstrated a strong correlation between poor psychosocial scores in the SF36 outcome measure and poor functional scores in the OLBI and RM (Table 4.16).

The rehabilitation programme used in this study was designed to increase strength although strength was not measured as an outcome. This was because the study design was influenced by research that indicated a stronger correlation between psychosocial factors and function, and a weaker correlation between biomedical assessment and function (Bigos et al., 1991; Burton et al., 1995; Carragee et al., 1999; Donceel & Du Bois, 1999; Dvorak, Gauchat et al., 1988; Dvorak, Valach et al., 1988; Graver et al., 1999; Hasenbring et al., 1994; Junge et al., 1996). In hindsight, this remains a methodological limitation.

However, measuring participants' strength levels would not have changed the functional outcomes found in this study. Although strength was not formally recorded as an outcome measure, all participants increased the levels of resistance as they progressed through the rehabilitation programme. This indirectly suggests that their strength levels increased as well.

There are four factors associated with the rehabilitation programme that may have had a significant effect on the outcome results. These factors are, 1) the trial group participants' may not have increased their strength levels, 2) the trial group participants' might have been non-compliant with the requirements of the rehabilitation programme, 3) the trial group had greater numbers of participants with higher levels of negative psychosocial factors, or 4) the rehabilitation programme was not effective.

The first factor can be eliminated if it is accepted that the participants did increase their strength levels on the basis that they were all able to work at much higher levels of resistance at the end of the programme compared with the beginning. Psychosocial factors were measured using the SF36 outcome measure; analysis of these results found no difference between the groups. Hence, these data did not show a bias towards poor psychosocial scores in the trial group. These arguments indicate that the rehabilitation programme used in this thesis provided no advantage compared with usual surgical advice.

Participant compliance to the programme was critical in relation to whether or not participants increased strength. Although participant compliance was audited via the marking

of their gymnasium attendance cards, this information was not recorded as an outcome; this was a limitation of the study and is reported as such. Prior to every change in the programme, the participants' gymnasium attendance cards were checked to ensure that they had completed the required number of sessions. This process relied on the honesty of the participants to mark their cards accurately. If participants attended the gymnasium three times per week for six months and worked at the levels set by the supervising physiotherapists, it is reasonable to assume that they would have increased their strength levels. However, because strength levels were not measured, this is only an assumption.

Participant compliance in the trial group was very high. Occasionally a participant missed a session or two due to being ill or a family issue that needed attending to. This was acceptable and if required, the programme time frame was extended to accommodate a catch-up period so that the required number of sessions were completed per programme. The supervising physiotherapists noted that trial group participants were diligent in their programme attendance. This may have been because the participants were in a study where they knew they were being closely monitored via the supervision provided and the outcome measures being completed. The only other study among the 16 that have been reviewed in this thesis (outlined in Tables 2.2 and 2.3) which reported participant compliance, is that of Erdogmus et al. (2007). These authors report an 85% compliance rate to the exercise programmes taught in this study. Erdogmus et al. comment that they were surprised with this level of compliance.

Whilst the programme design may have contributed to the inability to distinguish between the groups in this study, other confounders also need consideration. Both the subjective functional questionnaires and the Quality of Life annual questionnaire provided soft data. The questions asked in the annual questionnaire were subject to interpretation by participants and reliant on participants' memory. Participants' memory was the reason that the time frame of the 'previous three months' was chosen, and related to doctor visits, medication levels and other therapist visits. Participants' memory of events beyond three months has been shown to be unreliable (Brouwer et al., 2004; Burdorf et al., 1996; Davidson & Keating, 2002; Linton & Melin, 1982; Ostelo & de Vet, 2005; Patrick et al., 1995; Severens et al., 2000; Von Korff et al., 2000).

Doctor visits, medication levels and other therapist visits, were not independently audited due to difficulties in obtaining accurate data. Participants may have seen a doctor for

their LBP during follow-up but may not have seen their usual general practitioner. They may have seen an after-hours doctor or another doctor in the clinic they attended. In these cases a follow-up audit of doctor notes may not necessarily have confirmed the doctor visits. For ‘doctor visits’ and ‘other therapist visits’ to be accurately audited, the questionnaire should have requested information on the names of the practitioners that the participants visited. This was not the case, and was a limitation of the annual questionnaire. As participants could have been self-medicating, there was no way of confirming medication use other than the participants’ self-reporting.

Time off work was a key outcome measure. According to the literature, workers were accurate in remembering time off work after one year but had difficulty remembering the number of days off work (Severens et al., 2000). Therefore, when participants in this thesis study reported an episode of time off work, the amount of time was confirmed via an audit of their doctor notes.

A potential confounding argument in this process was if a participant saw different doctors and had more time off work than was indicated by the audited doctor notes. If this had occurred, there would have been a discrepancy between the amount of time off work reported by the participant and the amount of time off recorded in the doctor notes. This was not found to be the case. The time off work reported by the participants and that recorded in doctor notes was identical in all but two cases. In one case the participant had over-estimated by one day and in the other, the participant under-estimated by one day.

The accuracy of memory recall, by the participants in this study is at odds with the work by Severens et al. (2000). A possible explanation for this is that the Severens study was retrospective. Therefore, when the participants studied by Severens et al. had time off work during the year, they had no idea they were going to be asked to recall exactly how much time off work they had in one year’s time. Whereas, the participants in this thesis study knew they were going to be asked about this and may have recorded it themselves or at least highlighted it as something to remember. The accuracy between the time off work reported by the participants and that recorded by the doctor notes indicated that the participants in this thesis study had not visited different doctors for different episodes off work. Time off work is the only hard data recorded in the outcome measures of this thesis study.

A primary goal of the study investigated for this thesis was to test whether the rehabilitation programme would generate a statistical difference between the groups in time

off work. This did not occur. In hindsight an outcome measure like the Work Status Questionnaire or the WL27 may have been more appropriate to use. These questionnaires provide qualitative data, which provide for a better analysis of how the participant is coping at work and via this data a superior cost analysis is able to be obtained (Amick et al., 2000; Bombardier, 2000). These questionnaires were not commonly used at the time this study was designed and therefore were not included.

When designing the annual questionnaire for this study, a number of internationally renowned researchers who were working on outcome measures and LBP, were contacted. All reported that they had designed their own questionnaires specific to their requirements at the time, and suggested the same be done for this study. None of the researchers suggested that they knew of the Work Status Questionnaire or WL26 as it was in 2000. The first publication found on the WL26 is that of December 2000 (Amick et al., 2000). A recent Medline search for the Work Status Questionnaire did not find the questionnaire itself. Hence, the PI designed the annual questionnaire for use in this study. The limitations of this questionnaire are accepted and are reported in the section entitled 'limitations of this study'. An outcome measure such as the WL27 may have provided valuable data for this study.

A possible explanation for the lack of difference found between the groups lies in the original statistical power calculation. The statistical power for the study investigated in this thesis was based on a study by Stratford et al. (1996). In the Stratford study 40 participants per group were required to achieve 90% statistical power (to show as statistically significant two-tailed α=0.05) with a differential change between groups of five points on the Roland-Morris Scale, and 80% with a differential change of four points. The subjects that Stratford used all had low back pain, they were not post-discectomy patients and therefore the power for this study is probably somewhat less than calculated, given that there is more variability in response in this study. The study was not directly powered to show differences in the questionnaire data (e.g. GP visits and pain medications used) where it seems positive effects of such exercise programmes may exist. Post-hoc calculations suggest that sample sizes of approximately 70/group would be required to provide adequate statistical power for these outcomes.

Another explanation for these findings is that the RM, OLBI and SF36 measures, like all subjective outcome measures, are susceptible to personal bias. A primary cause of this bias is memory of pain and dysfunction. The literature reports that individuals have a poor

capacity to remember pain or levels of function after approximately six weeks (Burdorf et al., 1996; Patrick et al., 1995). However, some authors reported that memory of pain and dysfunction is reliable for up to two months (Burdorf et al., 1996; Severens et al., 2000), while Von Korff (2000) reported the reliability of pain memory is up to three months. Linton et al. (1982) stated that caution is warranted when using "post-hoc measures of pain with chronic pain patients" (p. 284). These studies demonstrated that the reliability of pain memory is short and all authors agree that after three months, memory of pain is unreliable.

A potential explanation for this unreliability is that individuals are unduly influenced by pain and dysfunction at the time it is present but tend to quickly forget it thereafter. This affects the reporting of pain and dysfunction. For example, if participants complete a subjective outcome measure while in pain, or having recently suffered problems, this is likely to be reflected in their reporting. Whereas, if they had experienced pain and problems three to four months previously, they were likely to have forgotten many of the problems experienced at the time and report less disability. Therefore, if a participant in a research study is completing a subjective functional outcome measure every six months, as were the participants in this thesis study, the results may not be reliable.

This explanation is based on a reliable correlation between pain and function. The literature reports that the correlation between pain and function is variable (Lee et al., 2001; Simmonds et al., 1998; Smeets et al., 2007; Wittink, 2005). However, these authors do report a link between pain and disability. The strength of this correlation is often dependent upon the method used to measure function or the terminology used to define 'function' (Smeets et al.; Wittink). Hence, this research suggests that how a participant feels at the time of completing a subjective outcome measure may influence what is reported.

The questionnaire relating to 'satisfaction' in this thesis study produced no meaningful results. After critiquing the literature, it is apparent that the Satisfaction Questionnaire in this study, according to the criteria outlined by Hudak and Wright (2000), failed to meet adequate standards that are required for gathering satisfaction information. Hudak and Wright report that 'satisfaction' is a topic that is often not well managed.

Based on the criteria outlined by Hudak and Wright (2000), there were two areas where the 'satisfaction' questions failed to meet adequate standards for the collection of meaningful data. A multi-dimensional measure was not used and the two questions that were asked of the participants were global questions. Because only two questions were asked, very little

information was collected. A combination of a multi-dimensional measure with a global measure would have been beneficial.

A key issue was the ambiguity of the second question, namely: ‘How satisfied have you been with your post-operation management?’ This question should have been separated into at least two components, with one referring to the post-surgical management and one referring to the overall clinical management.

The surgeon involved in this study deliberately minimised post-surgical contact with patients, in order to encourage self-management and medical independence. In speaking with participants in this study, the surgeon’s attitude or philosophy was sometimes misinterpreted as arrogance or lack of care. This led to several participants reporting low levels of satisfaction with their post-surgical management when in fact they were being managed through the rehabilitation programme and reported to their supervising physiotherapists that they were happy with their overall progress. It is possible that these participants were not happy with the rehabilitation programme or with how they were being managed. They may have stated one thing in conversation with the supervising physiotherapist but reported another on the questionnaire. Whether this was the case or not was not able to be determined because of the lack of specificity of the questions asked.

Both the ‘satisfaction’ questions used in the questionnaire were closed questions; Hudak and Wright (2000) suggested the use of a combination of open and closed questions. An example of an open question that may have revealed relevant information is: ‘What might have been done to improve your post-surgical satisfaction’. The final area of weakness relative to the satisfaction measure used in this study was that a sample of non-respondents was not measured. Hudak and Wright reported that this should be performed in all cases.

As recommended by Hudak and Wright (2000), the ‘satisfaction’ questions used did separate satisfaction of outcome and care, and as recommended by Sitzia &Wood (1997) and Williams et al. (1994, 1998) the word ‘satisfied’ was used in the question. However, only four responses were available for participants to answer; these ranged from ‘very satisfied’ to ‘very unsatisfied’. According to Ware and Hays (1988), adding the response ‘fairly satisfied’ would have given participants five options rather than four and this may have provided for superior statistical correlations with measures of behavioural intention.

A final confounding factor of the satisfaction measure used in this study was due to an administration error. The RA1 sent out the year two annual questionnaires to several participants who were completing the year three annual questionnaires. These participants were unable to answer the satisfaction questions as these questions were only included in the year three annual questionnaire. By the time that the error was noted the PI had recognised the weaknesses in the questionnaire, and did not think it appropriate to approach the participants again to fill out the final two questions. On intent-to-treat analysis 69 participants completed the satisfaction questionnaire, leaving 24 participants who did not complete the questions. The question of satisfaction was poorly measured and administered in this study, and as such, has been reported in the limitations of this thesis.

Satisfaction is recognised as an important domain to be measured. Satisfied patients generally follow treatment guidelines, and continue to access medical advice and assistance appropriately. Dissatisfied patients tend not to follow medical advice and distance themselves from their supervising clinicians. This leads to poor outcomes (Aharony & Strasser, 1993; Carr-Hill 1992; Ferris et al. 1992).

5.5.1 Comparison of outcome measures used in previous studies with those used in this study

Comparison of the outcome measures used in this thesis with other studies on the same topic highlights the issues reported by Bombardier (2000). Bombardier stated that comparison of research was difficult due to the variety of different outcome measures used.

Tables 2.2 and 2.3 in this thesis describe the outcome measures used in 16 different studies that investigated the post-surgical management of lumbar discectomy. Examination of the outcome measures used in these studies and described in these tables highlighted the variety of outcome measures available for the study of low back pain.

Some of the outcome measures used in this thesis study were also used in previous research. Time off work was used as an outcome measure in five of the seven studies in the ‘high quality’ group (Danielsen et al., 2000; Donceel et al., 1999; Erdogmus et al., 2007; Kjellby-Wendt & Styf, 1998; Manniche, Skall et al., 1993) and three of the nine studies in the low quality group (Alaranta et al., 1986; Burke et al., 1994; Johannsen et al., 1994). One study in each of the high and low quality groups used the RM Questionnaire (Danielsen et al., 2000; Ostelo et al., 2003). None of the high quality group used the OLBI, two studies in the low quality group used the OLBI (Kitteringham, 1996; Timm, 1994); two others in the low

quality group used the modified OLBI (Filiz et al., 2005; Yilmaz et al., 2003). None of the studies used the SF36; one study in the high quality group used a satisfaction scale (Kjellby-Wendt & Styf, 1998). Therefore, comparison of the results of this thesis study with these other 16 studies is limited due to a lack of comparable results.

However, some comparisons can be made. Danielson et al. (2000), and Ostelo et al. (2003) used the RM questionnaire and compared one post-surgical intervention against another. Similar to the results of this thesis study, no statistical difference was found between the groups. Two studies used the OLBI to compare post-surgical exercise programmes, and again similar to this thesis study, neither revealed a statistically significant difference between the control and trial groups at any point during the follow-up period (Kitteringham, 1996; Timm, 1994). Kjellby-Wendt and Styf (1998) compared one post-surgical exercise regime with another and reported on satisfaction. These authors found no statistical difference in satisfaction at two years post-surgery.

Of the eight studies that reported time off work as an outcome measure, two reported a statistically significant difference in favour of the trial group. These were the study of Donceel et al. (1999) that compared a pro-active management approach with usual management in insurance medicine, and the study by Burke et al. (1994) that compared a functional restoration programme with usual care. Similar to this thesis study, the other six studies found no statistical difference between the trial and control groups (Alaranta et al., 1986; Danielsen et al., 2000; Erdogmus et al., 2007; Johannsen et al., 1994; Kjellby-Wendt & Styf, 1998; Manniche, Skall et al., 1993). With the exception of Alaranta et al., all of these studies collected raw data on time off work, and statistically analysed this data. In the Alaranta study the raw data was presented but not statistically analysed. None of the studies used qualitative questionnaires in relation to time-off-work data.

5.6 Clinical Implications.

Rates of lumbar surgery have increased in recent decades (Frymoyer & Cats-Baril, 1991; Gray et al., 2006; Taylor et al., 1995; Weinstein, Lurie, Tosteson et al., 2006). This has led to a significant rise in the cost of surgery (Frymoyer & Cats-Baril; Schroth et al., 1992; van Tulder et al., 1995; Weinstein, Lurie, Tosteson et al.). Lumbar discectomy is by far the most common operation performed on the lumbar spine. According to Gray et al. 70 - 90% of all surgery performed on the lumbar spine is lumbar discectomy. Direct medical costs associated

with surgery are only a small proportion of the total cost involved with restoring the patient to full function (van Tulder et al., 1995). The bulk of the financial burden is related to indirect costs, namely; further medical supervision and investigations for complications related to the initial problem, social support, vocational retraining and support, and compensation for time off work.

Compensation for time off work is the most significant indirect cost associated with lumbar discectomy (Preventing Low Back Injuries; A Literature Review. ACC 1998). The total cost of lumbar discectomy to national economies has been measured in terms of billions of dollars per annum (Frymoyer & Cats-Baril, 1991; Schroth et al., 1992; van Tulder et al., 1995; Weinstein, Lurie, Tosteson et al., 2006). Therefore, any intervention that hastens the return to work of patients after surgery and that reduces likelihood of post-surgical complications can potentially save national health budgets millions of dollars per year.

The primary clinical implications generated from this study are: QoL data needs to be reviewed in conjunction with an audit of the clinical and administrative records so that hard data is gathered; outcome measures that gather information on employment after the intervention should include some qualitative data; statistical power should be based on QoL measures because these measures provide data that allows economic analysis of the interventions being tested; the study population size should be approximately 140 participants to provide adequate statistical power; the follow-up period should be between one to two years post-intervention; if improved function is a primary outcome measure then functional tests such as ‘lift from floor’ or ‘push above head’ should be tested; if increasing strength is a focus of the intervention being tested then strength gains should be tested also and finally if an exercise rehabilitation programme is designed to improve function, the exercises need to be functionally based.

5.7 Limitations of this study

This study utilised the patient list of one orthopaedic consultant. Therefore the results of this study relate to that surgeon only. However, the technique used in all operations for participants in this study was a standard technique commonly used in lumbar discectomy. The expertise of the operating surgeon may be a consideration. The surgeon involved in this study was a specialist orthopaedic surgeon in the area of lumbar surgery. It is therefore possible, that had the rehabilitation programme been compared with results from surgeons with less

experience in the area of lumbar surgery, that the results may have been different. However, this is hypothetical and has not been tested in this study.

Because only two physiotherapists supervised the rehabilitation programme, questions could be raised as to whether the skill and experience of these physiotherapists may be a contributing factor to the results of this study. Both physiotherapists were experienced clinicians familiar with gymnasium-based exercise rehabilitation programmes. A more rigorous method of testing the generalisability of the study would have been to train physiotherapists who were not already familiar with the rehabilitation programme used.

This study endeavoured to measure two populations post-lumbar discectomy, one that followed a structured exercise programme and one that did not perform any exercise at all. Due to the number of control group participants who chose to exercise after their surgery, this did not happen. Therefore, to some extent this study has compared two populations that performed different post-surgical exercise regimes, those regimes being a supervised structured exercise programme and a self managed exercise programme.

The Satisfaction Questionnaire was (with hindsight) poorly managed, this led to less than useful data. The questions relating to satisfaction with the surgical result and post-surgical management were not asked until the final annual questionnaire had been completed at three years post-surgery. This made it difficult for participants to accurately answer the questions. In addition the questions were too general to define exactly how the participants felt about their post-surgical management. This was compounded by an administrative error that saw 24 participants in the intent-to-treat population fail to receive the satisfaction measure.

Trial participants' strength gains were not measured and reported. This was a limitation of this study. The lack of difference demonstrated between the control and trial groups could have been due to the fact that the trial participants did not increase their strength. All trial group participants were able to work at higher levels of resistance by the end of the programme. Therefore, an increase in strength was assumed. However, as this was not measured it remains an assumption.

The primary mechanism for posterior disc herniation is repeated flexion or flexion/rotation movement patterns. If patients have suffered a disc herniation because of these movement patterns, part of the rehabilitation process should be to correct these pain

provoking movement patterns. Although the rehabilitation programme used in this study incorporated one session on correct lifting patterns, the exercises used in the programme were not specifically aimed at correcting movement patterns. The participants were taught to maintain their back in a neutral postion while performing the exercises. However, patients are not good at maintaining ideal positions and postures when they are not supervised. Most of the exercises used in this rehabilitation programme were performed in sitting and lying positions. The types of exercises and the positions used did not reinforce correction of pain provocative motor patterns.

Compliance to the programme also related directly to the effect generated by the programme. Participant compliance was regularly checked by the supervising physiotherapists, both by asking the participants how many sessions they had performed and by checking their dated attendance cards. This method of compliance depended on the honesty of the participants. It is possible that the participants had been dishonest with their adherence to the programme and therefore not completed the programme in the manner that it was intended. This could have led to a reduced effect of the programme. Neither supervising physiotherapist suspected this to be the case. Had this been the case, the participants would not have been able to increase the levels of resistance during the rehabilitation programme. However, because compliance was not independently audited there was no method of accurately testing the participants' levels of compliance.

The abdominal exercise regimes taught to trial group participants depended on the individual judgement of the supervising physiotherapist. This factor cannot be generalised; the abdominal exercises should have been more standardised. An assumption was made that by using these abdominal exercises any multifidus muscle weakness or inhibition that the participants' had was also likely to be resolved. Specific multifidus exercises should have been included as part of the rehabilitation programme, the fact that they were not is a limitation of this study. Since this study was conceptualised, a significant amount of research on the topic of appropriate abdominal training has been published. This research is valuable for future studies of this nature.

Participants in this study were not subcategorised in any way. Recent research has demonstrated that treatment results might be dependent on the type of condition the patient presents with (Fritz, Whitman, Childs, 2005; Hicks, Fritz, Delitto, McGill, 2005). This research has demonstrated that lumbar stability treatment regimes were more effective for

patients suffering from lumbar instability than lumbar hypo-mobility; conversely manipulation/mobilisation treatment regimes were more effective for patients suffering from hypo-mobility and less effective for hyper-mobile patients. The abdominal regime used in the rehabilitation programme for this thesis utilised exercises that aimed to stabilise the lumbar spine. Whether the gymnasium- based exercises used in the rehabilitation programme for this thesis study resulted in lumbar spine stabilisation or mobilisation is unknown. However, this combination of exercise was used for all the trial group participants. It would have been valuable to subcategorise participants into hyper-mobile and hypo-mobile subgroups, and then assess any difference in the outcomes between these two groups.

Most of the outcome measure data were soft data; the problems of using subjective functional questionnaires, particularly in relation to participants' memory, have been discussed. However, in the context of rehabilitation programmes many of the most relevant outcomes can only be captured by using 'soft' data. The literature reports decreasing correlations between accurate results and time intervals of more than six weeks with these outcome measures. With the exception of time off work, the QoL data was soft. The answers to all of the questions in the QoL, with the exception of time off work, were approximations, hence this information is open to bias and variability.

In relation to the categories in the QoL, the potential confounding factors of participants seeing different doctors for their LBP or seeing their doctor for post-surgical monitoring rather than complications have been mitigated by the evidence seen in the results for these categories. However, the potential for this type of confusion did exist and can not be entirely ruled out.

In the 'medication use' category, the question asked whether or not participants had used medication for LBP for a period of two days or more. Participants may have used medication every third or fourth day but over a prolonged period of time, this creates a potential confounder. In this situation they may have answered in the negative to the question, yet still been reliant on prolonged medication use.

The categories of 'doctor visits', 'medication use' and 'other treatment provider' asked whether or not participants had required any of these services or medication during the last three months. In effect the questionnaire was collecting data for just three months of each year for these categories. This limits the information that can be extracted from this data, as

participants' use of these services and use of medications may have been different during the other nine months of the year.

The 'time off work' category in the QoL collected raw time off work data. A qualitative approach to this data would have added value to this study.

5.8 Future research

Further studies on this topic are required, particularly considering the increasing rate and cost of lumbar discectomy. Rehabilitation programmes for post-surgical lumbar discectomy patients should focus on achieving functional gains. A key outcome measure for most governmental and funding agencies is cost; therefore cost effectiveness should be both targeted and measured as a primary outcome. To this end, outcome measures such as the WL26 and Work Status Questionnaire which collect qualitative data, and which allow for more detailed and accurate formulation of cost analysis, should be used.

Other important outcome measures for the study of post-surgical management of lumbar discectomy are 'satisfaction' and 'functional' measures. Satisfaction is now recognised as an important domain to measure. The design of 'satisfaction' questionnaires must include all of the critical determinants required to obtain meaningful data. Exercise programmes should measure gains in functional tasks, to ensure that there is a correlation between the gains made with the exercises and corresponding functional gains.

Studies using post-surgical exercise programmes for lumbar discectomy should be structured on evidenced-based principles of exercise rehabilitation. In particular, a focus on correcting pain proking movement patterns that may have been the cause of the injury in the first instance should be incorporated. This would add to the body of knowledge that has been published on this topic in recent years.

It is not known whether general holistic gymnasium based exercises such as those used in the rehabilitation programme for the study that contributed to this thesis, result in lumbar stabilisation or mobilisation. Future research to ascertain this information would be of interest.

Lumbar fusion is associated with higher rates of post-surgical complication when compared with lumbar discectomy. Rehabilitation programmes for lumbar fusion also need to

be trialled under experimental guidelines to assess what is best practice, for management of this type of surgery.

5.9 Summary of recommendations for clinical practice and future research

- Physiotherapy rehabilitation programmes for the post-surgical management of participants who have undergone a lumbar discectomy need to focus on physical reactivation, with an emphasis on regaining pre-injury strength levels and functional capacity.
- These rehabilitation programmes should be periodized, non-aggravating and structured according to the evidence-based principles of exercise rehabilitation.
- Further research that investigates post-surgical exercise rehabilitation programmes after lumbar discectomy should use exercises that are functionally focussed and functional outcome measures should be used to measure their effect.
- If regaining strength is a primary focus of a post-surgical rehabilitation programme, strength should be measured as an outcome.
- If a post-lumbar discectomy exercise trial group was being compared with a non-exercise control group, it would be ideal to ensure that the control group does not exercise. If the control group does exercise, the amount of exercise should be measured and considered within the design of the study.
- Outcome measures used in future research on this topic should focus on hard economic data. The WL26 or Work Status Questionnaire are examples of such measures.
- A two year follow up time frame may be the optimal follow-up period for studies of this nature.
- Satisfaction is an important domain to measure and satisfaction surveys need to be designed and administered according to appropriate criteria and method.
- Further research is required to prove or disprove the need for post-surgical rehabilitation after lumbar discectomy.

- An investigation into what patients do after lumbar discectomy in terms of their own management would be of interest.
- A comparison of a post-lumbar-discectomy cognitive education programme with a group that follows no formal post-surgical rehabilitation could provide valuable information.
- Research to investigate best-practice for the post-surgical management of lumbar fusion would be a valuable addition to the body of knowledge on this topic.

5.10 Conclusions

The results of this study did not significantly differentiate the control and trial groups with the exception of the 'doctor visit' category in the QoL questionnaire. There were however, a number of statistically significant findings in the SF36 and Qol data. These findings suggest that further research to investigate programmes similar to that used in this study, is warranted.

There are several distinctions between this study and other published literature on this topic. Firstly, this study compared an exercise group with a control group that followed no formal post-surgical rehabilitation programme. Secondly, the focus of the programme was to improve function by increasing the participants' holistic strength. Thirdly, the population sample size was one of the largest used. Fourthly, the duration of the rehabilitation programme was six months. Fifthly, the follow-up was three years and finally, the physiotherapy rehabilitation programme used was a fully periodized programme.

When this study was initially designed, the surgeon involved with the study discouraged his lumbar discectomy patients from participating in any type of formal exercise programme post-surgery. Therefore, it was assumed that participants in the control group would be unlikely to use exercise as an option post-surgery. This assumption has been proven to be incorrect with the majority of control group participants undertaking their own self-managed exercise regimes. This is an important confounding factor in this study.

To accurately measure the outcomes of different interventions, robust research method is required. Critical to this method is the selection of appropriate outcome measures and sample size. In order to obtain sound statistical results, a combination of functional and objective hard data outcome measures should be used.

Of the subjective functional measures used, the Roland-Morris, Oswestry Low Back Index and SF36 are highly recommended in the literature. However, the goal of using common outcome measures for the purpose of comparing international research, as promoted by Bombardier et al. (2000), is a worthy one. Bombardier et al. suggested the SF36 (version 2) for the category of back-specific function, the Chronic Pain Grade for the category of pain, the Work Status Questionnaire for work disability, and the Patient Satisfaction Scale or Global Questionnaire for the category of patient satisfaction.

In the modern health environment, clinicians are frequently challenged by funding institutions to prove that the interventions they are promoting are cost effective. To this end, the collection of hard economic data is critical in modern clinical research. In relation to lumbar discectomy, there are two aspects to consider, the initial direct costs of the intervention and the indirect costs of the follow-up.

This is the second largest study published on the topic of the post-surgical management of lumbar discectomy that compares an exercise group with a control group (that follows no formal post-surgical rehabilitation regime). Both this study and the previously published study by Erdogmus et al. (2007) have demonstrated no statistical difference between the treatment group and non-treatment group at the end point of the study. The study by Erdogmus et al. demonstrated an early statistical advantage but this was lost by the end point of follow-up at 18 months post-surgery.

In the study by Erdogmus et al. (2007) a sham treatment was used; no statistical difference was demonstrated between the sham treatment and physiotherapy programme. Erdogmus et al. query whether the perceived treatment advantage of any physiotherapy programme may be due to the cognitive effect of patient education and reducing fear avoidance by offering positive support. As yet, no studies have been published that compare a cognitive behavioural therapy programme with a no-treatment group. A study of this nature would be a valuable addition to the body of knowledge on this topic.

Considering that indirect costs constitute more than 90% of the total cost for lumbar discectomy, it is imperative that researchers gather information on the follow-up costs of lumbar discectomy. The bulk of these indirect costs are related to weekly compensation. work disability questionnaires such as the Work Status Questionnaire or WL26, should be considered for future research on this topic.

The literature clearly demonstrates the benefit of lumbar discectomy for patients with lumbo-sacral radicular syndrome. The literature also shows that the majority of the costs of this procedure are indirect costs incurred after surgery. The issue now, therefore, is to ascertain the most efficient method of managing patients after surgery. Considering the significant resources consumed by the procedure of lumbar discectomy and the increasing incidence of this procedure, further studies are needed to investigate the optimal post-surgical management of patients, in order to reduce the total cost of this procedure.

REFERENCES

Abbott, H. G., & Kress, J. B. (1979). Preconditioning in the prevention of knee injuries. *Archives of Physical Medicine & Rehabilitation, 50*(6), 326-333.

Accident compensation corporation, New Zealand: (1998). Preventing low back injuries: A literature review.

Accident compensation corporation newsletter: April (2008), issue 110.

Adams, M. A., & Dolan, P. (1995). Recent advances in lumbar spinal mechanics and their clinical significance. *Clinical Biomechanics, 10*(1), 3-19.

Adams, M. A., Freeman, B. J., Morrison, H. P., Nelson, I. W., & Dolan, P. (2000). Mechanical initiation of intervertebral disc degeneration. *Spine, 25*(13), 1625-1636.

Adams, M. A., & Hutton, W. C. (1982). Prolapsed intervertebral disc. A hyperflexion injury. *Spine, 7*(3), 184-191.

Adams, M. A., & Hutton, W. C. (1983). The effect of fatigue on the lumbar intervertebral disc. *Journal of Bone & Joint Surgery, 65*(2), 199-203.

Aharony, L., & Strasser, S. (1993). Patient satisfaction: what we know about and what we still need to explore. *Medical Care Review, 50*(1), 49-79.

Ahtiainen, J. P., Pakarinen, A., Alen, M., Kraemer, W. J., & Hakkinen, K. (2003). Muscle hypertrophy, hormonal adaptations and strength development during strength training in strength-trained and untrained men. *European Journal Applied Physiology, 89*(6), 555-563.

Alaranta, H., Hurme, M., Einola, S., Kallio, V., Knuts, L. R., & Torma, T. (1986). Rehabilitation after surgery for lumbar disc herniation: results of a randomized clinical trial. *International Journal of Rehabilitation Research, 9*(3), 247-257.

Albright, J., Allman, R., Bonfiglio, R. P., Conill, A., Dobkin, B., Guccione, A. et al. (2001). Philadelphia Panel evidence-based clinical practice guidelines on selected rehabilitation interventions for low back pain. *Physical Therapy, 81*(10), 1641-1674.

Allison, G., Kendle, K., Roll, S., Schupelius, J., Scott, Q., & Panizza, J. (1998). The role of the diaphragm during abdominal hollowing exercises. *Australian Journal of Physiotherapy, 44*(2), 95-102.

Amick, B. C., Lerner, D., Rogers, W. H., Rooney, T., & Katz, J. N. (2000). A review of health-related work outcome measures and their uses, and recommended measures. *Spine, 25*(24), 3152-3160.

Aprill, C., & Bogduk, N. (1992). High-intensity zone: a diagnostic sign of painful lumbar disc on magnetic resonance imaging. *British Journal of Radiology, 65*(773), 361-369.

Arrelov, B., Alexanderson, K., Hagberg, J., Lofgren, A., Nilsson, G., & Ponzer, S. (2007). Dealing with sickness certification - a survey of problems and strategies among general practitioners and orthopaedic surgeons. *Biomed Central Public Health, 7*, 273.

Atha, J. (1981). Strengthening muscle. *Exercise & Sport Science Reviews, 9*, 1-73.

Atlas, S. J., Keller, R. B., Chang, Y., Deyo, R. A., & Singer, D. E. (2001). Surgical and nonsurgical management of sciatica secondary to a lumbar disc herniation: five-year outcomes from the Maine Lumbar Spine Study. *Spine, 26*(10), 1179-1187.

Badley, E. M., Rasooly, I., & Webster, G. K. (1994). Relative importance of musculoskeletal disorders as a cause of chronic health problems, disability, and health care utilization: findings from the 1990 Ontario Health Survey. *Journal of Rheumatology, 21*(3), 505-514.

Baechle, T. R., & Earle, R. W. (2000). *Essentials of strength training and conditioning* (Second ed. Vol. One). Champaign, Ilinois: Human kinetics.

Baker, D., Nance, S., & Moore, M. (2001a). The load that maximises the average mechanical power output during explosive bench press throws in highly trained athletes. *Journal of strength and conditioning research, 15*, 20-24.

Baker, D., Nance, S., & Moore, M. (2001b). The load that maximises the average mechanical power output during jump squats in power-trained athletes. *Journal of strength and conditioning research, 15*, 92-97.

Balderston, R. A., Gilyard, G. G., Jones, A. A., Wiesel, S. W., Spengler, D. M., Bigos, S. J., et al. (1991). The treatment of lumbar disc herniation: simple fragment excision versus disc space curettage. *Journal of Spinal Disorders, 4*(1), 22-25.

Barr, J. S., Kubic, C. S., Molloy, M. K, McNeil, J. M., Riseborough, E. J., & White, J. C. (1967). Evaluation of end results in treatmentof ruptured lumbar intervertebral discs with protrusion of nucleus pulposus. *Surgical Gynecology and Obstetrics., 125,* 250-256.

Barrios, C., Ahmed, M., Arrotegui, J., Bjornsson, A., & Gillstrom, P. (1990). Microsurgery versus standard removal of the herniated lumbar disc. A 3-year comparison in 150 cases. *Acta Orthopaedica Scandinavica, 61*(5), 399-403.

Barrios, C., Ahmed, M., Arrotegui, J. I., & Bjornsson, A. (1990). Clinical factors predicting outcome after surgery for herniated lumbar disc: an epidemiological multivariate analysis. *Journal of Spinal Disorders, 3*(3), 205-209.

Battie, M. C., Haynor, D. R., Fisher, L. D., Gill, K., Gibbons, L. E., & Videman, T. (1995). Similarities in degenerative findings on magnetic resonance images of the lumbar spines of identical twins. *Journal of Bone & Joint Surgery, 77*(11), 1662-1670.

Battie, M. C., Videman, T., Gibbons, L. E., Fisher, L. D., Manninen, H., & Gill, K. (1995). 1995 Volvo Award in clinical sciences. Determinants of lumbar disc degeneration. A study relating lifetime exposures and magnetic resonance imaging findings in identical twins. *Spine, 20*(24), 2601-2612.

Beedle, B. B., & Mann, C. L. (2007). A comparison of two warm-ups on joint range of motion. *Journal of Strength & Conditioning Research, 21*(3), 776-779.

Beemsterboer, W., Stewart, R., Groothoff, J., & Nijhuis, F. (2008). The influence of sick leave frequency determinants on homogeneous groups in two socio-economically comparable, but socio-culturally different regions in The Netherlands. *Central European Journal of Public Health, 16*(4), 151-160.

Behm, D. G., & Anderson, K. G. (1995). Neuromuscular implications and applications of resistance training. *Journal of Strength & Conditioning Research, 9*, 264-274.

Behm, D. G., & Anderson, K. G. (2006). The role of instability with resistance training. *Journal of Strength & Conditioning Research, 20*(3), 716-722.

Bernard, T. N. (1990). Lumbar discography followed by computed tomography. Refining the diagnosis of low-back pain. *Spine, 15*(7), 690-707.

Bessette, L., Liang, M. H., Lew, R. A., & Weinstein, J. N. (1996). Classics in spine surgery literature revisited. *Spine, 21*(3), 259-263.

Beurskens, A. J., de Vet, H. C., & Koke, A. J. (1996). Responsiveness of functional status in low back pain: a comparison of different instruments. *Pain, 65*(1), 71-76.

Bigos, S. J., Battie, M. C., Spengler, D. M., Fisher, L. D., Fordyce, W. E., Hansson, T. H., et al. (1991). A prospective study of work perceptions and psychosocial factors affecting the report of back injury. *Spine, 16*(1), 1-6.

Biondi, J., & Greenberg, B. J. (1990). Redecompression and fusion in failed back syndrome patients. *Journal of Spinal Disorders, 3*(4), 362-369.

Blais, R. (1990). Assessing patient satisfaction with health care: did you drop somebody? *Canadian Journal of Program Evaluation., 5*, 1-13.

Boden, S. D., Davis, D. O., Dina, T. S., Patronas, N. J., & Wiesel, S. W. (1990). Abnormal magnetic-resonance scans of the lumbar spine in asymptomatic subjects. A prospective investigation. *Journal of Bone & Joint Surgery, 72*(3), 403-408.

Bogduk, N., & Twoomey, L. T. (1991). *Clinical antomy of the lumbar spine* (Second ed.). Melbourne: Churchill Livingstone.

Bogduk, N., Tynan, W., & Wilson, A. S. (1981). The nerve supply to the human lumbar intervertebral discs. *Journal of Anatomy, 132*(Pt 1), 39-56.

Bombardier, C. (2000). Outcome assessments in the evaluation of treatment of spinal disorders: summary and general recommendations. *Spine, 25*(24), 3100-3103.

Bombardier, C., Hayden, J., & Beaton, D. E. (2001). Minimal clinically important difference. Low back pain: outcome measures. *Journal of Rheumatology., 28*(2), 431-438.

Boos, N. (2003). Comment on "The clinical importance of changes in outcome scores after treatment for chronic low back pain" by O. Hagg et.al. *European Spine Journal, 12*(1).

Boos, N., Rieder, R., Schade, V., Spratt, K. F., Semmer, N., & Aebi, M. (1995). 1995 Volvo Award in clinical sciences. The diagnostic accuracy of magnetic resonance imaging, work perception, and psychosocial factors in identifying symptomatic disc herniations. *Spine, 20*(24), 2613-2625.

Brennan, G. P., Shultz, B. B., Hood, R. S., Zahniser, J. C., Johnson, S. C., & Gerber, A. H. (1994). The effects of aerobic exercise after lumbar microdiscectomy. *Spine, 19*(7), 735-739.

Brill, P. A., Cornman, C. B., Davis, D. R., Lane, M. J., Mustafa, T., Sanderson, M., et al. (1999). The value of strength training for older adults. *Home Care Provider, 4*(2), 62-66.

Brinckmann, P., Biggemann, M., & Hilweg, D. (1988). Fatigue fracture of human lumbar vertebrae. S1-23.

Brouwer, S., Kuijer, W., Dijkstra, P. U., Goeken, L. N., Groothoff, J. W., & Geertzen, J. H. (2004). Reliability and stability of the Roland Morris Disability Questionnaire: intra class correlation and limits of agreement. *Disability & Rehabilitation, 26*(3), 162-165.

Brown, S. H. M., Vera-Garcia, F. J., & McGill, S. M. (2006). Effects of abdominal muscle coactivation on the externally preloaded trunk: variations in motor control and its effect on spine stability. *Spine, 31*(13), E387-393.

Burdorf, A., Post, W., & Bruggeling, T. (1996). Reliability of a questionnaire on sickness absence with specific attention to absence due to back pain and respiratory complaints. *Occupational & Environmental Medicine, 53*(1), 58-62.

Burke, D. G., Culligan, C. J., Holt, L. E., & MacKinnon, N. C. (2000). Equipment designed to stimulate proprioceptive neuromuscular facilitation flexibility training. *Journal of Strength & Conditioning Research, 14*(2), 135-139.

Burke, S. A., Harms-Constas, C. K., & Aden, P. S. (1994). Return to work/work retention outcomes of a functional restoration program. A multi-center, prospective study with a comparison group. *Spine, 19*(17), 1880-1885.

Burnett, A., Sze, C. C., Tam, S. M., Yeung, K. M., Leong, M., Wang, W. T., et al. (2009). A cross-cultural study of the back pain beliefs of female undergraduate healthcare students. *Clinical Journal of Pain, 25*(1), 20-28.

Burton, A. K., Tillotson, K. M., Main, C. J., & Hollis, S. (1995). Psychosocial predictors of outcome in acute and subchronic low back trouble. *Spine, 20*(6), 722-728.

Burton, A. K., Waddell, G., Tillotson, K. M., & Summerton, N. (1999). Information and advice to patients with back pain can have a positive effect. A randomized controlled trial of a novel educational booklet in primary care. *Spine, 24*(23), 2484-2491.

Buttermann, G. R. (1996). Treatment of lumbar disc herniation: epidural steroid injection compared with discectomy. A prospective, randomized study. *Journal of Bone & Joint Surgery, 86-A*(4), 670-679.

Cahill, B. R., & Griffith, E. H. (1978). Effect of preseason conditioning on the incidence and severity of high school football knee injuries. *American Journal of Sports Medicine, 6*(4), 180-184.

Callaghan, J. P., & McGill, S. M. (2001). Intervertebral disc herniation: studies on a porcine model exposed to highly repetitive flexion/extension motion with compressive force. *Clinical Biomechanics, 16*(1), 28-37.

Campos, G. R., Luecke, T. J., Wendeln, H. K., Toma, K., Hagerman, F. C., Murray, T. F., et al. (2002). Muscular adaptations in response to three different resistance-training regimens: specificity of repetition maximum training zones. *European Journal of Applied Physiology, 88*(1-2), 50-60.

Canning, C. G., Ada, L., Adams, R., & O'Dwyer, N. J. (2004). Loss of strength contributes more to physical disability after stroke than loss of dexterity. *Clinical Rehabilitation, 18*(3), 300-308.

Carpinelli, R. N., & Otto, R. M. (1998). Strength training. Single versus multiple sets. *Sports Medicine, 26*(2), 73-84.

Carragee, E. J., Han, M. Y., Yang, B., Kim, D. H., Kraemer, H., & Billys, J. (1999). Activity restrictions after posterior lumbar discectomy. A prospective study of outcomes in 152 cases with no postoperative restrictions. *Spine., 24*(22), 2346-2351.

Carr-Hill, R. A. (1992). The measurement of patient satisfaction. *Journal of Public Health Medicine, 14*(3), 236-249.

Caspar, W., Campbell, B., Barbier, D. D., Kretschmmer, R., & Gotfried, Y. (1991). The Caspar microsurgical discectomy and comparison with a conventional standard lumbar disc procedure. *Neurosurgery, 28*(1), 78-86.

Cats-Baril, W. L. (1991). Economics of Spinal Disorders. In (Ed.), *The Adult Spine: Principles and Practice* (pp. 85-105). New York: Raven Press.

Cherkin, D. C., Deyo, R. A., Loeser, J. D., Bush, T., & Waddell, G. (1994). An international comparison of back surgery rates. *Spine, 19*(11), 1201-1206.

Cholewicki, J., & McGill, S. M. (1996). Mechanical stability of the in vivo lumbar spine: implications for injury and chronic low back pain. *Clinical Biomechanics, 11*(1), 1-15.

Cholewicki, J., & VanVliet, J. J. (2002). Relative contribution of trunk muscles to the stability of the lumbar spine during isometric exertions. *Clinical Biomechanics, 17*(2), 99-105.

Cleary, P. D., & Edgman-Levitan, S. (1997). Health care quality. Incorporating consumer perspectives. *JAMA, 278*(19), 1608-1612.

Cohen, I., & Rainville, J. (2002). Aggressive exercise as treatment for chronic low back pain. *Sports Medicine, 32*(1), 75-82.

Connolly, E. S. (1992). Surgery for recurrent lumbar disc herniation. *Clinical Neurosurgery, 39*, 211-216.

Cookson, R., Drummond, M., & Weatherly, H. (2009). Explicit incorporation of equity considerations into economic evaluation of public health interventions. *Health Economics, Policy, & Law, 4*(2), 231-245.

Cooper, R. G., St Clair-Forbes, W., & Jayson, M. I. (1992). Radiographic demonstration of paraspinal muscle wasting in patients with chronic low back pain. *British Journal of Rheumatology, 31*(6), 389-394.

Corning Creager, C. (1994). *Therapeutic exercises using the Swiss ball*. Minneapolis: Executive Physical Therapy Inc.

Crisco, J. J., & Panjabi, M. M. (1991). The intersegmental and multisegmental muscles of the lumbar spine. A biomechanical model comparing lateral stabilizing potential. *Spine, 16*(7), 793-799.

Cronin, J. (1997). Strength training for sport: part 4. *The New Zealand Coach*, 18-20.

Cronin, J., & Crewther, B. (2004). Training volume and strength and power development. *Journal of Science & Medicine in Sport, 7*(2), 144-155.

Danielsen, J. M., Johnsen, R., Kibsgaard, S. K., & Hellevik, E. (2000). Early aggressive exercise for postoperative rehabilitation after discectomy. *Spine., 25*(8), 1015-1020.

Dankaerts, W., O'Sullivan, P., Burnett, A., & Straker, L. (2006). Altered patterns of superficial trunk muscle activation during sitting in nonspecific chronic low back pain patients: importance of subclassification. *Spine, 31*(17), 2017-2023.

Dasinger, L. K., Krause, N., Deegan, L. J., Brand, R. J., & Rudolph, L. (1999). Duration of work disability after low back injury: a comparison of administrative and self-reported outcomes. *American Journal of Industrial Medicine, 35*(6), 619-631.

Davidson, C., & Tolich, M. (2003). *Social science research in New Zealand* (Second ed.). Auckland: Pearson education New Zealand ltd.

Davidson, M., & Keating, J. L. (2002). A comparison of five low back disability questionnaires: reliability and responsiveness. *Physical Therapy., 82*(1), 8-24.

Davis, W. J., Wood, D. T., Andrews, R. G., Elkind, L. M., & Davis, W. B. (2008). Elimination of delayed-onset muscle soreness by pre-resistance cardioacceleration before each set. *Journal of Strength and Conditioning Research, 22*(1), 212-225.

De Palma, A. F., & Prabhakar, M. (1966). Posterior-posterobilateral fusion of the lumbosacral spine. *Clinical Orthopaedics & Related Research, 47*, 165-171.

De Vaus, D. A. (2002). *Surveys in social research* (Fifth ed.). Melbourne: Allen and Unwin.

DeAnna, L. M., Sogg, S., Guarino, P., Skinner, J., Williams, D., Barkhuizen, A., et al. (2006). Predictors of exercise compliance in individuals with gulf war veterans illnesses: Department of veterans affairs cooperative study 470. *Military Medicine, 171*(9), 917-923.

den Boer, J. J., Oostendorp, R. A. B., Beems, T., Munneke, M., & Evers, A. W. M. (2006). Continued disability and pain after lumbar disc surgery: The role of cognitive-behavioral factors. *Pain, 123*(1-2), 45-52.

Der Ananian, C., Wilcox, S., Saunders, R., Watkins, K., & Evans, A. (2006). Factors that influence exercise among adults with arthritis in three activity levels. *Preventing Chronic Disease, 3*(3), A81.

Deyo, R., Battie, M., Beurskens, A., Bombardier, C., Croft, P., Koes, B., et al. (1998). Outcome measures for low back pain research. A proposal for standardized use. *Spine., 23*(18), 2003-2013.

Deyo, R. A., Andersson, G., Bombardier, C., Cherkin, D. C., Keller, R. B., Lee, C. K., et al. (1994). Outcome measures for studying patients with low back pain. *Spine, 19*(18S), 2032S-2036S.

Deyo, R. A., Ciol, M. A., Cherkin, D. C., Loeser, J. D., & Bigos, S. J. (1993). Lumbar spinal fusion. A cohort study of complications, reoperations, and resource use in the Medicare population. *Spine, 18*(11), 1463-1470.

Deyo, R. A., & Tsui-Wu, Y. J. (1987). Descriptive epidemiology of low-back pain and its related medical care in the United States. *Spine, 12*(3), 264-268.

Dolan, P., & Adams, M. A. (1998). Repetitive lifting tasks fatigue the back muscles and increase the bending moment acting on the lumbar spine. *Journal of Biomechanics, 31*(8), 713-721.

Dolan, P., Greenfield, K., Nelson, R. J., & Nelson, I. W. (2000). Can exercise therapy improve the outcome of microdiscectomy? *Spine., 25*, 1523-1532.

Donceel, P., & Du Bois, M. (1998). Fitness for work after surgery for lumbar disc herniation: a retrospective study. *European Spine Journal, 7*(1), 29-35.

Donceel, P., & Du Bois, M. (1999). Predictors for work incapacity continuing after disc surgery. *Scandinavian Journal of Work, Environment & Health, 25*(3), 264-271.

Donceel, P., Du Bois, M., & Lahaye, D. (1999). Return to work after surgery for lumbar disc herniation. *Spine, 24*(9), 872-876.

Dorgo, S., King, G. A., & Rice, C. A. (2009). The effects of manual resistance training on improving muscular strength and endurance. *Journal of Strength & Conditioning Research, 23*(1), 293-303.

Dougall, A., Russell, A., Rubin, G., & Ling, J. (2000). Rethinking patient satisfaction: Patient experiences of an open access flexible sigmoidoscopy service. *Social Science & Medicine Vol 50(1) 2000, 53-62*.

Drought, H. J. (1992). Personal training program design and periodization. *National Strength and Conditioning Association Journal, 14*(5), 31-36.

Dvorak, J., Gauchat, M. H., & Valach, L. (1988). The outcome of surgery for lumbar disc herniation. I. A 4-17 years' follow-up with emphasis on somatic aspects. *Spine, 13*(12), 1418-1422.

Dvorak, J., Valach, L., Fuhrimann, P., & Heim, E. (1988). The outcome of surgery for lumbar disc herniation. II. A 4-17 years' follow-up with emphasis on psychosocial aspects. *Spine, 13*(12), 1423-1427.

Edstrom, L., & Grimby, L. (1986). Effect of exercise on the motor unit. *Muscle Nerve, 9*(2), 104-126.

Englund, L., Tibblin, G., & Svardsudd, K. (2000). Variations in sick-listing practice among male and female physicians of different specialities based on case vignettes. *Scandinavian Journal of Primary Health Care, 18*(1), 48-52.

Enoka, R. M. (1988). Muscle strength and its development. New perspectives. *Sports Medicine, 6*(3), 146-168.

Erdogmus, C. B., Resch, K. L., Sabitzer, R., Muller, H., Nuhr, M., Schoggl, A., et al. (2007). Physiotherapy-based rehabilitation following disc herniation operation: Results of a randomized clinical trial. *Spine, 32*(19), 2041-2049.

Fairbank, J. C., & Pynsent, P. B. (2000). The Oswestry Disability Index. *Spine, 25*(22), 2940-2952.

Fathallah, F. A., Marras, W. S., & Parnianpour, M. (1998). The role of complex, simultaneous trunk motions in the risk of occupation-related low back disorders. *Spine, 23*(9), 1035-1042.

Ferreira, M. L., Ferreira, P. H., Latimer, J., Herbert, R. D., Hodges, P. W., Jennings, M. D., et al. (2007). Comparison of general exercise, motor control exercise and spinal manipulative therapy for chronic low back pain: A randomized trial. *Pain, 131*(1-2), 31-37.

Ferris, L. E., Williams, J. I., Llewellyn-Thomas, H. A., Basinski, A. S. H., Cohen, M. M., & Naylor, C. D. (1992). A guide to direct measures of patient satisfaction in clinical practice. *Canadian Medical Association Journal, 146*(10), 1727-1731.

Filiz, M., Cakmak, A., & Ozcan, E. (2005). The effectiveness of exercise programmes after lumbar disc surgery: a randomized controlled study. *Clinical Rehabilitation, 19*(1), 4-11.

Finneson, B. E., & Cooper, V. R. (1979). A lumbar disc surgery predictive score card. A retrospective evaluation. *Spine, 4*(2), 141-144.

Fitzpatrick, R. (1991). Surveys of patients satisfaction: I--Important general considerations. *British Medical Journal, 302*(6781), 887-889.

Fleck, S. J. (1999). Periodized strength training: a critical review. *The Journal of Strength and Conditioning Research, 13,* 82-89.

Fleck, S. J., & Falkel, J. E. (1986). Value of resistance training for the reduction of sports injuries. *Sports Medicine, 3*(1), 61-68.

Foa, E. B., & Kozak, M. J. (1986). Emotional processing of fear: exposure to corrective information. *Psychological Bulletin, 99*(1), 20-35.

Fredriksson, K., Toomingas, A., Torgen, M., Thorbjornsson, C. B., & Kilbom, A. (1998). Validity and reliability of self-reported retrospectively collected data on sick leave related to musculoskeletal diseases. *Scandinavian Journal of Work, Environment & Health, 24*(5), 425-431.

Friedrich, M., Gittler, G., Arendasy, M., & Friedrich, K. M. (2005). Long-term effect of a combined exercise and motivational program on the level of disability of patients with chronic low back pain. *Spine, 30*(9), 995-1000.

Fritz, J.M., Whitman, J. M., & Childs, J. D. (2005). Lumbar spine segmental mobility assessment: An examination of validity for determining intervention strategies in patients with low back pain. *Archive of Physical Medicine and Rehabilitation 86,* 1745-1752.

Fritz, J. M., & Irrgang, J. J. (2001). A comparison of a modified Oswestry Low Back Pain Disability Questionnaire and the Quebec Back Pain Disability Scale. *Physical Therapy, 81*(2), 776-788, 2001 Feb.

Frost, H., Klaber-Moffett, J. A., Moser, J. S., & Fairbank, J. C. (1995). Randomised controlled trial for evaluation of fitness programme for patients with chronic low back pain. *British Medical Journal, 310*(6973), 151-154.

Frymoyer, J. W. (1988). Back pain and sciatica. *New England Journal of Medicine, 318*(5), 291-300.

Frymoyer, J. W., & Cats-Baril, W. L. (1991). An overview of the incidences and costs of low back pain. *Orthopedic Clinics of North America, 22*(2), 263-271.

Frymoyer, J. W., Hanley, E., Howe, J., Kuhlmann, D., & Matteri, R. (1978). Disc excision and spine fusion in the management of lumbar disc disease. A minimum ten-year follow-up. *Spine, 3*(1), 1-6.

Fujimura, Y., Wakano, K., & Hijikata, S. (1974). Diagnostic significance of lumbar discography. *Seikeigeka, 25*, 781-790.

Garhammer, J. (1991). A comparison of maximal power outputs between elite male and female weightlifters in competition. *International Journal of Sports Biomechanics., 7*, 3-11.

Geiger, G., Todd, D. D., Clark, H. B., Miller, R. P., & Kori, S. H. (1992). The effects of feedback and contingent reinforcement on the exercise behavior of chronic pain patients. *Pain, 49*(2), 179-185.

George, S. Z., Fritz, J. M., Bialosky, J. E., & Donald, D. A. (2003). The effect of a fear-avoidance-based physical therapy intervention for patients with acute low back pain: results of a randomized clinical trial. *Spine, 28*(23), 2551-2560.

Gibson, J. N. A., & Waddell, G. (2007). Surgical interventions for lumbar disc prolapse: updated Cochrane Review. *Spine, 32*(16), 1735-1747.

Gordon, S. J., Yang, K. H., Mayer, P. J., Mace, A. H., Kish, V. L., & Radin, E. L. (1991). Mechanism of disc rupture. A preliminary report. *Spine, 16*(4), 450-456.

Gotshalk, L. A., Loebel, C. C., Nindl, B. C., Putukian, M., Sebastianelli, W. J., Newton, R. U., et al. (1997). Hormonal responses of multiset versus single-set heavy-resistance exercise protocols. *Canadian Journal of Applied Physiology, 22*(3), 244-255.

Graver, V., Haaland, A. K., Magnaes, B., & Loeb, M. (1999). Seven-year clinical follow-up after lumbar disc surgery: results and predictors of outcome. *British Journal of Neurosurgery., 13*(2), 178-184.

Gray, D. T., Deyo, R. A., Kreuter, W., Mirza, S. K., Heagerty, P. J., Comstock, B. A., et al. (2006). Population-based trends in volumes and rates of ambulatory lumbar spine surgery. *Spine, 31*(17), 1957-1963.

Grenier, S. G., & McGill, S. M. (2007). Quantification of lumbar stability by using 2 different abdominal activation strategies. *Archives of Physical Medicine & Rehabilitation, 88*(1), 54-62.

Gruber, M., & Gollhofer, A. (2004). Impact of sensorimotor training on the rate of force development and neural activation. *European Journal of Applied Physiology*.

Guyatt, G. H., Mitchell, A., Molloy, D. W., Capretta, R., Horsman, J., & Griffith, L. (1995). Measuring patient and relative satisfaction with level or aggressiveness of care and involvement in care decisions in the context of life threatening illness. *Journal of Clinical Epidemiology, 48*(10), 1215-1224.

Hagg, O., Fritzell, P., & Nordwall, A. (2003). The clinical importance of changes in outcome scores after treatment for chronic low back pain. *European Spine Journal, 12*(1), 12-20.

Hakkinen, A., Ylinen, J., Kautiainen, H., Airaksinen, O., Herno, A., & Kiviranta, I. (2003). Does the outcome 2 months after lumbar disc surgery predict the outcome 12 months later? *Disability & Rehabilitation, 25*(17), 968-972.

Hakkinen, K., Alen, M., Kraemer, W. J., Gorostiaga, E., Izquierdo, M., Rusko, H., et al. (2003). Neuromuscular adaptations during concurrent strength and endurance training versus strength training. *European Journal of Applied Physiology, 89*(1), 42-52.

Hakkinen, K., & Hakkinen, A. (1995). Neuromuscular adaptations during intensive strength training in middle-aged and elderly males and females. *Electromyography & Clinical Neurophysiology, 35*(3), 137-147.

Hakkinen, K., Kallinen, M., Komi, P. V., & Kauhanen, H. (1991). Neuromuscular adaptations during short-term "normal" and reduced training periods in strength athletes. *Electromyography & Clinical Neurophysiology., 31*(1), 35-42.

Hakkinen, K., Kallinen, M., Linnamo, V., Pastinen, U. M., Newton, R. U., & Kraemer, W. J. (1996). Neuromuscular adaptations during bilateral versus unilateral strength training in middle-aged and elderly men and women. *Acta Physiology Scandinavia, 158*(1), 77-88.

Hakkinen, K., Komi, P. V., Alen, M., & Kauhanen, H. (1987). EMG, muscle fibre and force production characteristics during a 1 year training period in elite weight-lifters. *European Journal of Applied Physiology & Occupational Physiology, 56*(4), 419-427.

Hakkinen, K., Kraemer, W. J., Kallinen, M., Linnamo, V., Pastinen, U. M., & Newton, R. U. (1996). Bilateral and unilateral neuromuscular function and muscle cross-sectional area in middle-aged and elderly men and women. *Journals of Gerontology Series A-Biological Sciences & Medical Sciences, 51*(1), 21-29.

Hakkinen, K., Pakarinen, A., Alen, M., Kauhanen, H., & Komi, P. V. (1987). Relationships between training volume, physical performance capacity, and serum hormone concentrations during prolonged training in elite weight lifters. *International Journal of Sports Medicine, 8 (Supplement 1),* 61-65.

Hakkinen, K., Pakarinen, A., Alen, M., Kauhanen, H., & Komi, P. V. (1988a). Neuromuscular and hormonal adaptations in athletes to strength training in two years. *Journal of Applied Physiology, 65*(6), 2406-2412.

Hakkinen, K., Pakarinen, A., Alen, M., Kauhanen, H., & Komi, P. V. (1988b). Neuromuscular and hormonal responses in elite athletes to two successive strength training sessions in one day. *European Journal of Applied Physiology & Occupational Physiology, 57*(2), 133-139.

Hall, J. A., & Dornan, M. C. (1988). Meta-analysis of satisfaction with medical care: description of research domain and analysis of overall satisfaction levels. *Social Science & Medicine, 27*(6), 637-644.

Hasenbring, M., Marienfeld, G., Kuhlendahl, D., & Soyka, D. (1994). Risk factors of chronicity in lumbar disc patients. A prospective investigation of biologic, psychologic, and social predictors of therapy outcome. *Spine, 19*(24), 2759-2765.

Hazard, R. G., Haugh, D., Green, P. A., & Jones, P. L. (1994). Chronic low back pain: The relationship between patient satisfaction and pain, impairment, and disability outcomes. *Spine, 19*(8), 881-887.

Hejna, W. F., & Rosenberg, A. (1982). The prevention of sports injuries in high school students through strength training. *Journal of the National Strength and Conditioning Association, 4*, 28-31.

Hermantin, F. U., Peters, T., Quartararo, L., & Kambin, P. (1999). A prospective, randomized study comparing the results of open discectomy with those of video-assisted arthroscopic microdiscectomy. *Journal of Bone & Joint Surgery, 81*(7), 958-965.

Herron, L. (1994). Recurrent lumbar disc herniation: results of repeat laminectomy and discectomy. *Journal of Spinal Disorders, 7*(2), 161-166.

Hestbaek, L., Leboeuf-Yde, C., Manniche, C. (2003). Low back pain: What is the long-term course? A review of studies of general patient populations. *European Spine Journal, 12* , 149-165.

Hestbaek, L., Leboeuf-yde, C., Endberg, M., Torsten, L., Bruun, N. H., Manniche, C. (2003). The course of low back pain in a general population. Results from a five year prospective study. *Journal of Manipulative and Physiological Therapeutics, 26*(4), 213-219.

Hicks, G.E., Fritz, J. M., Delitto, A., McGill, S. M. (2005). Preliminary development for a clinical prediction rule for determining which patients with low back pain will respond to a stabilisation exercise programme. *Archives of Physical Medicine and Rehabilitation, 86,* 1753-1762.

Hicks, A. L., Martin, K. A., Ditor, D. S., Latimer, A. E., Craven, C., Bugaresti, J., et al. (2003). Long-term exercise training in persons with spinal cord injury: effects on strength, arm ergometry performance and psychological well-being. *Spinal Cord, 41*(1), 34-43.

Hides, J. A., Richardson, C. A., & Jull, G. A. (1996). Multifidus muscle recovery is not automatic after resolution of acute, first-episode low back pain. *Spine, 21*(23), 2763-2769.

Hides, J. A., Stokes, M. J., Saide, M., Jull, G. A., & Cooper, D. H. (1994). Evidence of lumbar multifidus muscle wasting ipsilateral to symptoms in patients with acute/subacute low back pain. *Spine, 19*(2), 165-172.

Hilbe, J. M. (2005). A review of SPSS version 13.0. *The American Statistitian, 59*(2), 185, 186.

Hirsch, C., & Schajowicz, F. (1953). Studies on the structural changes in the lumbar annulus fibrosis. *Acta Orthopaedia Scandinavia, 22,* 184-231.

Hodges, P. W., & Richardson, C. A. (1997a). Contraction of the abdominal muscles associated with movement of the lower limb. *Physical Therapy, 77*(2), 132-144.

Hodges, P. W., & Richardson, C.A. (1997b). Feed forward contraction of transversus abdominus is not influenced by the direction of arm movement. *Experimental Brain Research, 114,* 362-370.

Hodges, P. W., & Richardson, C. A. (1997c). Relationship between limb movement speed and associated contraction of the trunk muscles. *Ergonomics, 40,* 1220-1230.

Hodges, P. W., & Richardson, C. A. (1998). Delayed postural contraction of transversus abdominus associated with movement of the lower limb in people with lower back pain. *Journal of Spinal Disorders, 11,* 46-56.

Hodges, P. W., & Richardson, C. A. (1999a). Altered trunk muscle recruitment in people with low back pain with limb movement at different speeds. *Archives of Physical Medicine, 80*(9), 1005-1012.

Hodges, P. W., & Richardson, C. A. (1999b). Transverse abdominus and the superficial abdominal muscles are controlled independently in a postural task. *Neuroscience letters, 265*(2), 91-94.

Hodges, P. W. (1999). Is there a role for transversus abdominis in lumbo-pelvic stability? *Manual Therapy, 4*(2), 74-86.

Hodges, P. W., Cresswell, A. G., & Thorstensson, A. (1998). Preparatory trunk motion precedes upper limb movement. *Experimental Brain Research, 124*, 69-79.

Hodges, P. W., & Richardson, C. A. (1996). Inefficient muscular stabilization of the lumbar spine associated with low back pain. A motor control evaluation of transversus abdominis. *Spine, 21*(22), 2640-2650.

Hoffman, R. M., Wheeler, K. J., & Deyo, R. A. (1993). Surgery for herniated lumbar discs: a literature synthesis. *Journal of General Internal Medicine, 8*(9), 487-496.

Holt, B. W., & Lambourne, K. (2008). The impact of different warm-up protocols on vertical jump performance in male collegiate athletes. *Journal of Strength & Conditioning Research, 22*(1), 226-229.

Huang, T. J., Hsu, R. W., Li, Y., & Cheng, C. (2005). Less systemic cytokine response in patients following microendoscopic versus open lumbar discectomy. *Journal of Orthopaedic Research, 23*(2), 406-411.

Hudak, P. L., & Wright, J. G. (2000). The characteristics of patient satisfaction measures. *Spine, 25*(24), 3167-3177.

Hurley, M. V., & Scott, D. L. (1998). Improvements in quadriceps sensorimotor function and disability of patients with knee osteoarthritis following a clinically practicable exercise regime. *British Journal of Rheumatology, 37*(11), 1181-1187.

Hutchinson, P. J., Laing, R. J., Waran, V., Hutchinson, E., & Hollingworth, W. (2000). Assessing outcome in lumbar disc surgery using patient completed measures. *British Journal of Neurosurgery., 14*(3), 195-199.

Ihlebaek, C., Hansson, T. H., Laerum, E., Brage, S., Eriksen, H. R., Holm, S. H., et al. (2006). Prevalence of low back pain and sickness absence: a "borderline" study in Norway and Sweden. *Scandinavian Journal of Public Health, 34*(5), 555-558.

Javid, M. J., Nordby, E. J., Ford, L. T., Hejna, W. J., Whisler, W. W., Burton, C., et al. (1983). Safety and efficacy of chymopapain (Chymodiactin) in herniated nucleus pulposus with sciatica. Results of a randomized, double-blind study. *Journal of the American Medical Association, 249*(18), 2489-2494.

Jensen, M. C., Brant-Zawadzki, M. N., Obuchowski, N., Modic, M. T., Malkasian, D., & Ross, J. S. (1994). Magnetic resonance imaging of the lumbar spine in people without back pain. *New England Journal of Medicine, 331*(2), 69-73.

Johannsen, F., Remvig, L., Kryger, P., Beck, P., Lybeck, K., Larsen, L. H., et al. (1994). Supervised endurance exercise training compared to home training after first lumbar diskectomy: a clinical trial. *Clinical & Experimental Rheumatology, 12*(6), 609-614.

Johns, G. (1994). Absenteeism estimates by employees and managers: Divergent perspectives and self serving perceptions. *Journal of Applied Psychology, 79*(2), 229-239.

Judge, L. W., Moreau, C., & Burke, J. R. (2003). Neural adaptations with sport-specific resistance training in highly skilled athletes. *Journal Sports Science, 21*(5), 419-427.

Jull, G. A. (1994). Physical therapy of the lumbar spine. In L. T. Twoomey & J. R. Taylor (2nd Edition), *Physical therapy of the low back.* (pp. 251-283). Michigan: Chruchill Livingstone.

Junge, A., Frohlich, M., Ahrens, S., Hasenbring, M., Sandler, A., Grob, D., et al. (1996). Predictors of bad and good outcome of lumbar spine surgery. A prospective clinical study with 2 years follow up. *Spine., 21*(9), 1056-1064.

Kahanovitz, N., Viola, K., & Gallagher, M. (1989). Long-term strength assessment of postoperative discectomy patients. *Spine, 14*(4), 402-403.

Kaigle, A. M., Holm, S. H., & Hansson, T. H. (1995). Experimental instability in the lumbar spine. *Spine, 20*(4), 421-430.

Kambin, P., & Brager, M. D. (1987). Percutaneous posterolateral discectomy. Anatomy and mechanism. *Clinical Orthopaedics & Related Research,* (223), 145-154.

Kambin, P., & Nass. (2003). Arthroscopic microdiscectomy. *Spine Journal: 3* Supplement, 60S-64S.

Karchevsky, M., Schweitzer, M. E., Carrino, J. A., Zoga, A., Montgomery, D., & Parker, L. (2005). Reactive endplate marrow changes: a systematic morphologic and epidemiologic evaluation. *Skeletal Radiology, 34*(3), 125-129.

Kavcic, N., Grenier, S., & McGill, S. M. (2004a). Determining the stabilizing role of individual torso muscles during rehabilitation exercises. *Spine, 29*(11), 1254-1265.

Kavcic, N., Grenier, S. G., & McGill, S. M. (2004b). Quantifying tissue loads and spine stability while performing commonly prescribed low back stabilization exercises. *Spine, 29*(20), 2319-2329.

Kell, R. T., & Asmundson, G. J. G. (2009). A comparison of two forms of periodized exercise rehabilitation programs in the management of chronic nonspecific low-back pain. *Journal of Strength & Conditioning Research, 23*(2), 513-523.

Kelsey, J. L., Githens, P. B., White, A. A., Holford, T. R., Walter, S. D., O'Connor, T., et al. (1984). An epidemiologic study of lifting and twisting on the job and risk for acute prolapsed lumbar intervertebral disc. *Journal of Orthopaedic Research, 2*(1), 61-66.

Kernan, T., & Rainville, J. (2007). Observed outcomes associated with a quota-based exercise approach on measures of kinesiophobia in patients with chronic low back pain. *Journal of Orthopaedic & Sports Physical Therapy, 37*(11), 679-687.

Kitteringham, C. (1996). The effect of straight leg raise exercises after lumbar decompression surgery - a pilot study. *Physiotherapy, 82*(2), 115-123.

Kjaer, P., Leboeuf-Yde, C., Korsholm, L., Sorensen, J. S., & Bendix, T. (2005). Magnetic resonance imaging and low back pain in adults: a diagnostic imaging study of 40-year-old men and women. *Spine, 30*(10), 1173-1180.

Kjellby-Wendt, G., & Styf, J. (1998). Early active training after lumbar discectomy. *Spine, 23*(21), 2345-2351.

Korres, D. S., Loupassis, G., & Stamos, K. (1992). Results of lumbar discectomy: A study using 15 different evaluation methods. *European Spine Journal, 1*(1), 20-24.

Kraemer, W. J. (1997). A Series of Studies—The Physiological Basis for Strength Training in American Football: Fact Over Philosophy. *The Journal of Strength and Conditioning Research, 11*, 131-142.

Kraemer, W. J., Adams, K., Cafarelli, E., Dudley, G. A., Dooly, C., Feigenbaum, M. S., et al. (2002). American College of Sports Medicine position stand. Progression models in resistance training for healthy adults. *Medicine & Science in Sports & Exercise, 34*(2), 364-380.

Kraemer, W. J., Fleck, S. J., & Evans, W. J. (1996). Strength and power training: physiological mechanisms of adaptation. *Exercise and Sport Science Reviews, 24*, 363-397.

Kraemer, W. J., Hakkinen, K., Triplett-Mcbride, N. T., Fry, A. C., Koziris, L. P., Ratamess, N. A., et al. (2003). Physiological changes with periodized resistance training in women tennis players. *Medicine & Science in Sports & Exercise., 35*(1), 157-168.

Kraemer, W. J., Patton, J. F., Gordon, S. E., Harman, E. A., Deschenes, M. R., Reynolds, K., et al. (1995). Compatibility of high-intensity strength and endurance training on hormonal and skeletal muscle adaptations. *Journal of Applied Physiology, 78*(3), 976-989.

Kraemer, W. J., & Ratamess, N. A. (2000). Physiology of resistance training: current issues. *Orthopedic Physical Therapy Clinics. North America, 9*(4), 467-513.

Kraemer, W. J., & Ratamess, N. A. (2004). Fundamentals of resistance training: progression and exercise prescription. *Medicine and Science in Sports and Exercise, 36*(4), 674-688.

Kraemer, W. J., Ratamess, N. A., & French, D. N. (2002). Resistance training for health and performance. *Current Sports Medicine Reports, 1*(3), 165-171.

Kraemer, W. J., Stone, M. H., O'Bryant, H. S., Conley, M. S., Johnson, R. L., Nieman, D. C., et al. (1997). Effects of single vs. multiple sets of weight training: impact of volume, intensity, and variation. *The Journal of Strength and Conditioning Research, 11*, 143-147.

Kuisma, M., Karppinen, J., Niinimaki, J., Ojala, R., Haapea, M., Heliovaara, M., et al. (2007). Modic changes in endplates of lumbar vertebral bodies: prevalence and association with low back and sciatic pain among middle-aged male workers. *Spine, 32*(10), 1116-1122.

Kuslich, S. D., Ulstrom, C. L., & Michael, C. J. (1991). The tissue origin of low back pain and sciatica. *Orthopaedic Clinics of North America, 22*, 181-187.

Larsen, D. L., Attkisson, C. C., Hargreaves, W. A., & Nguyen, T. D. (1979). Assessment of client/patient satisfaction: development of a general scale. *Evaluation & Program Planning, 2*(3), 197-207.

Lebow, J. L. (1974). Consumer assessments of the quality of medical care. *Medical Care, 12*(4), 328-337.

Lee, C. E., Simmonds, M. J., Novy, D. M., & Jones, S. (2001). Self-reports and clinician-measured physical function among patients with low back pain: a comparison. *Archives of Physical Medicine & Rehabilitation, 82*(2), 227-231.

LeVois, M., Nguyen, T. D., & Attkisson, C. C. (1981). Artifact in client satisfaction assessment: experience in community mental health settings. *Evaluation & Program Planning, 4*(2), 139-150.

Lindstrom, I., Ohlund, C., Eek, C., Wallin, L., Peterson, L. E., Fordyce, W. E., et al. (1992). The effect of graded activity on patients with subacute low back pain: a randomized prospective clinical study with an operant-conditioning behavioral approach. *Physical Therapy, 72*(4), 279-293.

Linton, S. J., & Melin, L. (1982). The accuracy of remembering chronic pain. *Pain, 13*(3), 281-285.

Lively, M. W. (2002). Sports medicine approach to low back pain. *Southern Medical Journal, 95*(6), 642-646.

Ljunggren, A. E., Weber, H., Kogstad, O., Thom, E., & Kirkesola, G. (1997). Effect of exercise on sick leave due to low back pain. A randomized, comparative, long-term study. *Spine, 22*(14), 1610-1616.

Locker, D., & Dunt, D. (1978). Theoretical and methodological issues in sociological studies of consumer satisfaction with medical care. *Social Science & Medicine 12, 283-292*.

MacNab, I. (1971). Negative disc exploration. An analysis of the causes of nerve-root involvement in sixty-eight patients. *Journal of Bone & Joint Surgery, 53*(5), 891-903.

Maezawa, S., & Muro, T. (1992). Pain provocation at lumbar discography as analyzed by computed tomography/discography. *Spine, 17*(11), 1309-1315.

Malter, A. D., Larson, E. B., Urban, N., & Deyo, R. A. (1996). Cost-effectiveness of lumbar discectomy for the treatment of herniated intervertebral disc. *Spine, 21*(9), 1048-1054.

Malter, A. D., McNeney, B., Loeser, J. D., & Deyo, R. A. (1998). 5-year reoperation rates after different types of lumbar spine surgery. *Spine., 23*(7), 814-820.

Manniche, A. K., Asmussen, K., Lauritsen, B., Vinterberg, H., Karbo, H., Abildstrup, S., et al. (1993b). Intensive dynamic back exercises with or without hyperextension in chronic back pain after surgery for lumbar disc protrusion. A clinical trial. *Spine, 18*(5), 560-567.

Manniche, A. K., Asmussen, K., Lauritsen, B., Vinterberg, H., Karsten, H., Rose-Hansen, E., et al. (1994). Back pain, sciatica and disability following first-time conventional haemilaminectomy for lumbar disc herniation. *Danish Medical Bulletin, 41*(1), 103-106.

Manniche, A. K., Skall, H. F., Braendholt, L., Christinsen, B. H., Christophersen, L., Ellegaard, B., et al. (1993a). Clinical trial of postoperative dynamic back exercises after first lumbar discectomy. *Spine., 18*(1), 92-97.

Mannion, A. F., Muntener, M., Taimela, S., & Dvorak, J. (1999). A randomized clinical trial of three active therapies for chronic low back pain. *Spine, 24*(23), 2435-2448.

Marchand, F., & Ahmed, A. M. (1990). Investigation of the laminate structure of lumbar disc anulus fibrosus. *Spine, 15*(5), 402-410.

Martin, K. (2001). The effects of exercise training on HRQL amoung people with spinal cord injury and the mediating influence of pain. *Rehabilitation Psychology*.

Marx, J. O., Ratamess, B. C., & Nindl, B. C. (2001). The effects of single-set versus periodized multiple-set resistance training muscular performance and hormonal concentrations in women. *Medicine & Science in Sports & Exercise, 33,* 635-643.

Mayer, T. G., Smith, S. S., Keeley, J., & Mooney, V. (1985). Quantification of lumbar function. Part 2: Sagittal plane trunk strength in chronic low-back pain patients. *Spine, 10*(8), 765-772.

Mayer, T. G., Vanharanta, H., Gatchel, R. J., Mooney, V., Barnes, D., Judge, L., et al. (1989). Comparison of CT scan muscle measurements and isokinetic trunk strength in postoperative patients. *Spine, 14*(1), 33-36.

McAuley, E., Blissmer, B., Marquez, D. X., Jerome, G. J., Kramer, A. F., & Katula, J. (2000). Social relations, physical activity, and well-being in older adults. *Preventive Medicine, 31*(5), 608-617.

McAuley, E., Elavsky, S., Motl, R. W., Konopack, J. F., Hu, L., & Marquez, D. X. (2005). Physical activity, self-efficacy, and self-esteem: longitudinal relationships in older adults. *Journals of Gerontology. Series B: Psychological Sciences and Social Sciences, 75.*

Mcbride, J. M., Triplett-Mcbride, T., Davie, A., & Newton, R. U. (2002). The effect of heavy vs light-load jump squats on the development of strength, power, and speed. *Journal of Strength and Conditioning Research, 16*, 75-82.

McCain, G. A., Bell, D. A., Mai, F. M., & Halliday, P. D. (1988). A controlled study of the effects of a supervised cardiovascular fitness training program on the manifestation of primary fibromyalgia. *Arthritis & Rheumatism, 31*, 1135-1141.

McDonagh, M. J., & Davies, C. T. (1984). Adaptive response of mammalian skeletal muscle to exercise with high loads. *European Journal of Applied Physiology & Occupational Physiology, 52*(2), 139-155.

McGill, S. M. (1991). Electromyographic activity of the abdominal and low back musculature during the generation of isometric and dynamic axial trunk torque: implications for lumbar mechanics. *Journal of Orthopaedic Research, 9*(1), 91-103.

McGill, S. M. (1997). The biomechanics of low back injury: implications on current practice in industry and the clinic. *Journal of Biomechanics, 30*(5), 465-475.

McGill, S. M. (2007). *Low back disorders: Evidence-based prevention and rehabilitation.* (Second ed.). Champaign: Human Kinetics.

McGill, S. M., Grenier, S., Kavcic, N., & Cholewicki, J. (2003). Coordination of muscle activity to assure stability of the lumbar spine. *Journal of Electromyography & Kinesiology, 13*(4), 353-359.

McGill, S. M., & Karpowicz, A. (2009). Exercises for spine stabilization: motion/motor patterns, stability progressions, and clinical technique. *Archives of Physical Medicine & Rehabilitation, 90*(1), 118-126.

McGill, S. M., & Norman, R. W. (1987). Effects of an anatomically detailed erector spinae model on L4/L5 disc compression and shear. *Journal of Biomechanics, 20*(6), 591-600.

McNally, D. S., Adams, M. A., & Goodship, A. E. (1993). Can intervertebral disc prolapse be predicted by disc mechanics? *Spine, 18*(11), 1525-1530.

Milner-Brown, H. S., Stein, R. B., & Lee, R. G. (1975). Synchronization of human motor units: possible roles of exercise and supraspinal reflexes. *Electroencephalography & Clinical Neurophysiology, 38*(3), 245-254.

Mimura, M., Panjabi, M. M., Oxland, T. R., Crisco, J. J., Yamamoto, I., & Vasavada, A. (1994). Disc degeneration affects the multidirectional flexibility of the lumbar spine. *Spine, 19*(12), 1371-1380.

Mochida, J., Nishimura, K., & Nomura, T. (1994). Appropriate procedures of posterior herniotomy and percutanious nucleotomy in lumbar disc herniation. *Rinsho Seikei Geka, 29*, 423-430.

Modic, M. T., Steinberg, P. M., Ross, J. S., Masaryk, T. J., & Carter, J. R. (1988). Degenerative disk disease: assessment of changes in vertebral body marrow with MR imaging. *Radiology, 166*(1), 193-199.

Moneta, G. B., Videman, T., Kaivanto, K., Aprill, C., Spivey, M., Vanharanta, H., et al. (1994). Reported pain during lumbar discography as a function of anular ruptures and disc degeneration. A re-analysis of 833 discograms. *Spine, 19*(17), 1968-1974.

Mundt, D. J., Kelsey, J. L., Golden, A. L., Pastides, H., Berg, A. T., Sklar, J., et al. (1993). An epidemiologic study of non-occupational lifting as a risk factor for herniated lumbar intervertebral disc. *Spine, 18*(5), 595-602.

Munn, J., Herbert, R. D., Hancock, M. J., & Gandevia, S. C. (2005). Resistance training for strength: effect of number of sets and contraction speed. *Medicine & Science in Sports & Exercise, 37*(9), 1622-1626.

Naclerio, F. J., Colado, J. C., Rhea, M. R., Bunker, D., & Riplett, T. (2009). The influence of strength and power on muscle endurance test performance. *Journal of Strength and Conditioning Research, 23*(5), 1482-1488.

Naidoo, J., & Wills, J. (2000). *Health promotions: Foundations for practice* (Second ed.). London: Bailliere Tindall.

Naylor, A. (1974). The late results of laminectomy for lumbar disc prolapse. A review after 10 to 25 years. *Journal of Bone and Joint Surgery, 56*, 17-29.

Newton, R. U., Hakkinen, K., Hakkinen, A., McCormick, M., Volek, J. S., & Kraemer, W. J. (2002). Mixed-methods resistance training increases power and strength of young and older men. *Medicine & Science in Sports & Exercise, 34*(8), 1367-1375.

Nork, S. E., Hu, S. S., Workman, K. L., Glazer, P. A., & Bradford, D. S. (1999). Patient outcomes after decompression and instrumented posterior spinal fusion for degenerative spondylolisthesis. *Spine, 24*(6), 561-569.

North, R. B., Ewend, M. G., Lawton, M. T., Kidd, D. H., & Piantadosi, S. (1991). Failed back surgery syndrome: 5-year follow-up after spinal cord stimulator implantation. *Neurosurgery, 28*(5), 692-699.

North, T. C., McCullagh, P., & Tran, Z. V. (1990). Effect of exercise on depression. *Exercise & Sport Sciences Reviews, 18*, 379-415.

O'Bryant, H., Byrd, R., & Stone, M. H. (1988). Cycle ergometer and maximal leg and hip strength adaptations to two different methods of training. *Journal of Applied Sports Science, 2*(2), 27-30.

O'Reilly, S. C., Jones, A., Muir, K. R., & Doherty, M. (1998). Quadriceps weakness in knee osteoarthritis: the effect on pain and disability. *Annals of the Rheumatic Diseases, 57*(10), 588-594.

Ostelo, R. W., & de Vet, H. C. (2005). Clinically important outcomes in low back pain. *Best Practice & Research in Clinical Rheumatology, 19*(4), 593-607.

Ostelo, R. W., de Vet, H. C., Berfelo, M. W., Kerckhoffs, M. R., Vlaeyen, J. W., Wolters, P. M., et al. (2003). Effectiveness of behavioral graded activity after first-time lumbar disc surgery: short term results of a randomized controlled trial. *European Spine Journal, 12*(6), 637-644.

Ostelo, R. W., de Vet, H. C., Knol, D. L., & van den Brandt, P. A. (2004). 24-item Roland-Morris Disability Questionnaire was preferred out of six functional status questionnaires for post-lumbar disc surgery. *Journal of Clinical Epidemiology, 57*(3), 268-276.

Ostelo, R. W., de Vet, H. C., Waddell, G., Kerckhoffs, M. R., Leffers, P., & van Tulder, M. W. (2004). Rehabilitation after lumbar disc surgery. *Cochrane Database of Systematic Reviews.*(2).

Osterman, H., Seitsalo, S., Karppinen, J., & Malmivaara, A. (2006). Effectiveness of microdiscectomy for lumbar disc herniation: a randomized controlled trial with 2 years of follow-up. *Spine, 31*(21), 2409-2414.

Osti, O. L., Vernon-Roberts, B., Moore, R., & Fraser, R. D. (1992). Annular tears and disc degeneration in the lumbar spine. A post-mortem study of 135 discs. *Journal of Bone & Joint Surgery 74*(5), 678-682.

O'Sullivan, P., Twomey, L., Allison, G. (1997). Evaluation of specific stabilising exercise in the treatment of chronic low back pain with radiological diagnosis of spondylolysis or spondylolisthesis. *Spine 22*(24), 2959-2967.

O'Sullivan, P., Twomey, L., Allison, G., Sinclair, J., Miller, K., & Knox, J. (1997). Altered patterns of abdominal muscle activation in patients with chronic low back pain. *Australian Journal of Physiotherapy, 43*(2), 91-98.

O'Sullivan, P. B. (2000). Lumbar segmental 'instability': clinical presentation and specific stabilizing exercise management. *Manual Therapy, 5*(1), 2-12.

O'Sullivan, P. B., Burnett, A., Floyd, A. N., Gadsdon, K., Logiudice, J., Miller, D., et al. (2003). Lumbar repositioning deficit in a specific low back pain population. *Spine, 28*(10), 1074-1079.

Otani, K., Harris, L. E., & Tierney, W. M. (2003). A paradigm shift in patient satisfaction assessment. *Medical Care Research and Review Vol 60(3) 347-365*.

Ozgen, S., Naderi, S., Ozek, M. M., & Pamir, M. N. (1999). Findings and outcome of revision lumbar disc surgery. *Journal of Spinal Disorders., 12*(4), 287-292.

Palmgren, T., Gronblad, M., Virri, J., Kaapa, E., & Karaharju, E. (1999). An immunohistochemical study of nerve structures in the anulus fibrosus of human normal lumbar intervertebral discs. *Spine, 24*(20), 2075-2079.

Panjabi, M. M. (1992a). The stabilizing system of the spine. Part I. Function, dysfunction, adaptation, and enhancement. *Journal of Spinal Disorders, 5*(4), 383-389.

Panjabi, M. M. (1992b). The stabilizing system of the spine. Part II. Neutral zone and instability hypothesis. *Journal of Spinal Disorders, 5*(4), 390-396.

Panjabi, M. M., Abumi, K., Duranceau, J., & Oxland, T. R. (1989). Spinal stability and intersegmental muscle forces: A biomechanical model. *Spine, 14*(2), 194-200.

Pappas, C. T. E., Harrington, T., & Sonntag, V. K. H. (1992). Outcome analysis in 654 surgically treated lumbar disc herniations. *Neurosurgery, 30*, 862-866.

Pascoe, G. C., Attkisson, C. C., & Roberts, R. E. (1983). Comparison of indirect and direct approaches to measuring patient satisfaction. *Evaluation & Program Planning, 6*(3-4), 359-371.

Patrick, D. L., Deyo, R. A., Atlas, S. J., Singer, D. E., Chapin, A., & Keller, R. B. (1995). Assessing health-related quality of life in patients with sciatica. *Spine, 20*(17), 1899-1908.

Pengel, L. H., Refshauge, K. M., & Maher, C. G. (2004). Responsiveness of pain, disability, and physical impairment outcomes in patients with low back pain. *Spine., 29*(8), 879-883.

Pezowicz, C. A., Robertson, P. A., & Broom, N. D. (2005). Intralamellar relationships within the collagenous architecture of the anulus fibrosus imaged in its fully hydrated state. *Journal of Anatomy, 207*(4), 299-312.

Pezowicz, C. A., Robertson, P. A., & Broom, N. D. (2006). The structural basis of interlamellar cohesion in the intervertebral disc wall. *Journal of Anatomy, 208*(3), 317-330.

Pezowicz, C. A., Schechtman, H., Robertson, P. A., & Broom, N. D. (2006). Mechanisms of anular failure resulting from excessive intradiscal pressure: a microstructural-micromechanical investigation. *Spine, 31*(25), 2891-2903.

Pfingsten, M. (2001). Functional restoration--it depends on an adequate mixture of treatment. *Schmerz, 15*(6), 492-498.

Pfingsten, M., Hildebrandt, J., Leibing, E., Franz, C., & Saur, P. (1997). Effectiveness of a multimodal treatment program for chronic low-back pain. *Pain, 73*(1), 77-85.

Philips, H. C. (1987). Avoidance behaviour and its role in sustaining chronic pain. *Behaviour Research & Therapy, 25*(4), 273-279.

Phillips, S. M. (2000). Short-term training: when do repeated bouts of resistance exercise become training? *Canadian Journal of Applied Physiology, 25*(3), 185-193.

Picavet, H. S. J., Vlaeyen, J. W. S., & Schouten, J. S. A. G. (2002). Pain catastrophizing and kinesiophobia: predictors of chronic low back pain. *American Journal of Epidemiology, 156*(11), 1028-1034.

Pincivero, D. M., Lephart, S. M., & Karunakara, R. G. (1997). Effects of rest interval on isokinetic strength and functional performance after short-term high intensity training. *British Journal of Sports Medicine, 31*(3), 229-234.

Ploutz, L. L., Tesch, P. A., Biro, R. L., & Dudley, G. A. (1994). Effect of resistance training on muscle use during exercise. *Journal of Applied Physiology, 76*(4), 1675-1681.

Pollock, M. L., Gaesser, G. A., Butcher, J. D., Despres, J., Dishman, R. K., Franklin, B. A., et al. (1998). ACSM Position Stand: The recommended quantity and quality of exercise for developing and maintaining cardiorespiratory and muscular fitness, and flexibility in healthy adults. *Medicine and Science in Sports and Exercise, 30*(6), 975-991.

Porchet, F., Wietlisbach, V., Burnand, B., Daeppen, K., Villemure, J. G., & Vader, J. P. (2002). Relationship between severity of lumbar disc disease and disability scores in sciatica patients. *Neurosurgery., 50*(6), 1253-1259.

Postacchini, F. (1996). Results of surgery compared with conservative management for lumbar disc herniations. *Spine, 21*(11), 1383-1387.

Potteiger, J. A., Judge, L. W., Cerny, J. A., & Potteiger, V. M. (1995). Effects of altering training volume and intensity on body mass, performance, and hormonal

concentrations in weight-event athletes. *The Journal of Strength and Conditioning Research, 9*, 55-58.

Poulter, D. (1999). Empower the patient. *Journal of Orthopaedic & Sports Physical Therapy, 29*(10), 616-617.

Rainville, J., Hartigan, C., Jouve, C., & Martinez, E. (2004). The influence of intense exercise-based physical therapy program on back pain anticipated before and induced by physical activities. *Spine Journal, 4*(2), 176-183.

Rainville, J., Hartigan, C., Martinez, E., Limke, J., Jouve, C., & Finno, M. (2004). Exercise as a treatment for chronic low back pain. *Spine Journal, 4*(1), 106-115.

Rainville, J., Sobel, J., Hartigan, C., Monlux, G., & Bean, J. (1997). Decreasing disability in chronic back pain through aggressive spine rehabilitation. *Journal of Rehabilitation Research & Development, 34*(4), 383-393.

Ratamess, N. A., Alvar, B. A., Evetoch, T. K., Housch, T. J., Kibler, W. B., Kraemer, W. J., et al. (2002). Position stand: progression models in resistance training for healthy adults. *Medicine and Science in Sports and Exercise, 34*, 364-380.

Ratamess, N. A., Alvar, B. A., Evetoch, T. K., Housch, T. J., Kibler, W. B., Kraemer, W. J., et al. (2009). Progression models in resistance training for healthy adults. *Medicine and Science in Sports and exercise, 41*(3), 687-708.

Rejeski, W. J., Ettinger, W. H., Martin, K., & Morgan, T. (1998). Treating disability in knee osteoarthritis with knee exercise therapy: a central role for self efficacy and pain. *Arthritis Care Research, 11*, 94-101.

Rhea, M. R., Alvar, B. A., Burkett, L. N., & Ball, S. D. (2003). A meta-analysis to determine the dose response for strength development. *Medicine & Science in Sports & Exercise, 35*(3), 456-464.

Rhea, M. R., Oliverson, J. R., Marshall, G., Peterson, M. D., Kenn, J. G., Ayllon, F., et al. (2008). Noncompatibility of power and endurance training among college baseball players. *Journal of Strength & Conditioning Research, 22*(1), 230-234.

Richardson, C. A., & Jull, G. A. (1995). Muscle control - pain control. What exercises would you prescribe? *Manuel Therapy, 1*(1), 2-10.

Ricketson, R., Simmons, J. W., & Hauser, B. O. (1996). The prolapsed intervertebral disc. The high-intensity zone with discography correlation. *Spine, 21*(23), 2758-2762.

Riddle, D., & Stratford, P. (2002). Roland-Morris scale reliability. *Physical Therapy, 82*(5), 512-515.

Robins, D. W. (2005). Post activation potentiation and its practical application: a brief review. *Journal of strength and conditioning research, 19*, 453-458.

Robinson, J. M., Stone, M. H., Johnson, R. L., Penland, C. M., Warren, B. J., & Lewis, D. R. (1995). Effects of different weight training exercise/rest intervals on strength, power, and high intensity exercise endurance. *The Journal of Strength and Conditioning Research, 9*, 216-221.

Roland, M., & Fairbank, J. (2000). The Roland-Morris Disability Questionnaire and the Oswestry Disability Questionnaire. *Spine., 25*(24), 3115-3124.

Ross, C. K., Steward, C. A., & Sinacore, J. M. (1995). A comparative study of seven measures of patient satisfaction. *Medical Care, 33*(4), 392-406.

Rothhaupt, D., Laser, T., Ziegler, H., & Liebig, K. (1997). Orthopedic hippotherapy in postoperative rehabilitation of lumbar intervertebral disk patients. A prospective, randomized therapy study. *Sportverletzung Sportschaden., 11*(2), 63-69.

Rutherford, O. M., & Jones, D. A. (1986). The role of learning and coordination in strength training. *European Journal of Applied Physiology & Occupational Physiology, 55*(1), 100-105.

Ruuskanen, J. M., & Ruoppila, I. (1995). Physical activity and psychological well-being among people aged 65 to 84 years. *Age & Ageing, 24*(4), 292-296.

Saal, J. A. (1990). Dynamic muscular stabilization in the nonoperative treatment of lumbar pain syndromes. *Orthopaedic Review, 19*(8), 691-700, 1990 Aug.

Saal, J. A. (1996). Natural history and nonoperative treatment of lumbar disc herniation. *Spine, 21*(24 supplement), 2S-9S.

Saal, J. A., & Saal, J. S. (1989). Nonoperative treatment of herniated lumbar intervertebral disc with radiculopathy. An outcome study. *Spine, 14*(4), 431-437.

Santos-Eggimann, B., Wietlisbach, V., Rickenbach, M., Paccaud, F., & Gutzwiller, F. (2000). One-year prevalence of low back pain in two Swiss regions: estimates from the

population participating in the 1992-1993 MONICA project. *Spine., 25*(19), 2473-2479.

Schade, V., Semmer, N., Main, C. J., Hora, J., & Boos, N. (1999). The impact of clinical, morphological, psychosocial and work-related factors on the outcome of lumbar discectomy. *Pain, 80*, 239-249.

Schantz, P. (1983). Capillary supply in heavy-resistance trained non-postural human skeletal muscle. *Acta Physiologica Scandinavica, 117*(1), 153-155.

Schiphorst-Preuper, H. R., Reneman, M. F., Boonstra, A. M., Dijkstra, P. U., Versteegen, G. J., Geertzen, J. H. B., et al. (2008). Relationship between psychological factors and performance-based and self-reported disability in chronic low back pain. *European Spine Journal, 17*(11), 1448-1456.

Schneiders, A. G., Zusman, M., & Singer, K. P. (1998). Exercise therapy compliance in acute low back pain patients. *Manual Therapy, 3*(3), 147-152.

Schollum, M. L., Robertson, P. A., & Broom, N. D. (2008). Microstructure and mechanical disruption of the lumbar disc annulus: part I: a microscopic investigation of the translamellar bridging network. *Spine, 33*(25), 2702-2710.

Schroth, W. S., Schectman, J. M., Elinsky, E. G., & Panagides, J. C. (1992). Utilization of medical services for the treatment of acute low back pain: conformance with clinical guidelines. *Journal of General Internal Medicine, 7*(5), 486-491.

Sculco, A. D., Paup, D. C., Fernhall, B., & Sculco, M. J. (2001). Effects of aerobic exercise on low back pain patients in treatment. *Spine Journal, 1*(2), 95-101.

Selye, H. (1976). Forty years of stress research: principal remaining problems and misconceptions. *Canadian Medical Association Journal, 115*(1), 53-56.

Severens, J. L., Mulder, J., Laheij, R. J., & Verbeek, A. L. (2000). Precision and accuracy in measuring absence from work as a basis for calculating productivity costs in The Netherlands. *Social Science & Medicine, 51*(2), 243-249.

Sahrmann, S. A. (2002). *Diagnosis and treatment of movement impairment syndrome*: Michigan. (pp. 373-377). Mosby.

Sigl, T., Cieza, A., Brockow, T., Chatterji, S., Kostanjsek, N., & Stucki, G. (2006). Content comparison of low back pain-specific measures based on the International

Classification of Functioning, Disability and Health (ICF). *Clinical Journal of Pain, 22*(2), 147-153.

Signorile, J. F., Carmel, M. P., Lai, S., & Roos, B. A. (2005). Early plateaus of power and torque gains during high- and low-speed resistance training of older women. *Journal of Applied Physiology, 98*(4), 1213-1220.

Simmonds, M. J., Kumar, S., & Lechelt, E. (1996). Psychological factors in disabling low back pain: causes or consequences? *Disability & Rehabilitation, 18*(4), 161-168.

Simmonds, M. J., Olson, S. L., Jones, S., Hussein, T., Lee, C. E., Novy, D., et al. (1998). Psychometric characteristics and clinical usefulness of physical performance tests in patients with low back pain. *Spine, 23*(22), 2412-2421.

Simmons, J. W., Emery, S. F., McMillin, J. N., Landa, D., & Kimmich, S. J. (1991). Awake discography. A comparison study with magnetic resonance imaging. *Spine, 16*(6 Supplement), S216-221.

Simunic, D. I., Robertson, P. A., & Broom, N. D. (2004). Mechanically induced disruption of the healthy bovine intervertebral disc. *Spine, 29*(9), 972-978.

Sitzia, J., & Wood, N. (1997). Patient satisfaction: a review of issues and concepts. *Social Science & Medicine, 45*(12), 1829-1843.

Smeets, R. J. E. M., van Geel, A. C., Kester, A. D., & Knottnerus, J. A. (2007). Physical capacity tasks in chronic low back pain: what is the contributing role of cardiovascular capacity, pain and psychological factors? *Disability & Rehabilitation, 29*(7), 577-586.

Smith, D., & Bruce-Low, S. (2004). Strength training methods and the work of Arthur Jones. *Journal of Exercise Physiology online, 7*(6, December).

Smith, D. J. (2003). A framework for understanding the training process leading to elite performance. *Sports Medicine, 33*(15), 1103-1126.

Sollner, W., & Doering, S. (1997). A psychological therapy methods in chronic non-radicular backache. *Orthopade, 26*(6), 535-543.

Spennewyn, K. C. (2008). Strength outcomes in fixed versus free-form resistance equipment. *Journal of Strength & Conditioning Research, 22*(1), 75-81.

Stadnik, T. W., Lee, R. R., Coen, H. L., Neirynck, E. C., Buisseret, T. S., & Osteaux, M. J. (1998). Annular tears and disk herniation: prevalence and contrast enhancement on MR images in the absence of low back pain or sciatica. *Radiology, 206*(1), 49-55.

Staff, P. H. (1982). The effects of physical activity on joints, cartilage, tendons and ligaments. *Scandinavian Journal of Social Medicine. Supplement, 29*, 59-63.

Staron, R. S., Karapondo, D. L., Kraemer, W. J., Fry, A. C., Gordon, S. E., Falkel, J. E., et al. (1994). Skeletal muscle adaptations during early phase of heavy-resistance training in men and women. *Journal of Applied Physiology, 76*(3), 1247-1255.

Stauffer, R. N., & Coventry, M. B. (1972). Anterior interbody lumbar spine fusion. Analysis of Mayo Clinic series. *Journal of Bone & Joint Surgery, 54*(4), 756-768.

Steenstra, I. A., Anema, J. R., Bongers, P. M., de Vet, H. C. W., & van Mechelen, W. (2003). Cost effectiveness of a multi-stage return to work program for workers on sick leave due to low back pain, design of a population based controlled trial. *BioMed central Musculoskeletal Disorders, 4*, 26.

Stewart, G., & Sachs, B. L. (1996). Patient outcomes after reoperation on the lumbar spine. *Journal of Bone & Joint Surgery, 78*(5), 706-711.

Stewart, I. B., & Sleivert, G. G. (1998). The effect of warm-up intensity on range of motion and anaerobic performance. *Journal of Orthopaedic & Sports Physical Therapy, 27*(2), 154-161.

Stewart, M. A., & Wanklin, J. (1978). Direct and indirect measures of patient satisfaction with physicians' services. *Journal of Community Health, 3*(3), 195-204.

Stone, M. H. (1990). Muscle conditioning and muscle injuries. *Medicine & Science in Sports & Exercise., 22*(4), 457-462.

Stone, M. H., O'Bryant, H., & Garhammer, J. (1981). A hypothetical model for strength training. *Journal of Sports Medicine & Physical Fitness, 21*(4), 342-351.

Stone, M. H., Plisk, S., Stone, M. E., Schilling, B., O'Bryant, H., & Pierce, K. (1998). Athletic performance development: volume load- 1 set vs. multiple sets, training velocity and training variation. *Journal of the National Strength and Conditioning Association, 20*, 22-31.

Stone, M. H., Potteiger, J. C., Pierce, K. C., Proulx, C. M., O'Bryant, H. S., Johnson, R. L., et al. (1997). Comparison of the effects of three different weight-training programs on the one repetition maximum squat. *The Journal of Strength and Conditioning Research, 11*, 286.

Stowers, T., McMillan, J., Scala, D., Davis, V., Wilson, D., & Stone, M. (1983). The short-term effects of three different strength-power training methods. *National Strength & Conditioning Association Journal, 5*(3), 24-27.

Stratford, P. W., Binkley, J., Solomon, P., Finch, E., Gill, C., & Moreland, J. (1996). Defining the minimum level of detectable change for the Roland-Morris questionnaire. *Physical Therapy, 76*(4), 359-365.

Stratford, P. W., Binkley, J. M., Riddle, D. L., & Guyatt, G. H. (1998). Sensitivity to change of the Roland-Morris Back Pain Questionnaire: part 1. *Physical Therapy, 78*(11), 1186-1196.

Suni, J., Rinne, M., Natri, A., Matti, P., Parkkari, J., & Alaranta, H. (2006). Control of the lumbar neutral zone decreases low back pain and improves self-evaluated work ability: a 12-month randomized controlled study. *Spine, 31*(18), 611-620.

Swartling, M., Peterson, S., & Wahlstrom, R. (2007). Views on sick-listing practice among Swedish General Practitioners--a phenomenographic study. *BioMed Central Family Practice, 8*, 44.

Swartling, M. S., Alexanderson, K. A. E., & Wahlstrom, R. A. (2008). Barriers to good sickness certification -- an interview study with Swedish general practitioners. *Scandinavian Journal of Public Health, 36*(4), 408-414.

Taaffe, D. R., & Marcus, R. (1997). Dynamic muscle strength alterations to detraining and retraining in elderly men. *Clinical Physiology, 17*(3), 311-324.

Taylor, V. M., Deyo, R. A., Cherkin, D. C., & Kreuter, W. (1995). Low back pain hospitalization. Recent United States trends and regional variations. *Spine, 19*(11), 1207-1212.

Tessitore, A., Meeusen, R., Pagano, R., Benvenuti, C., Tiberi, M., & Capranica, L. (2008). Effectiveness of active versus passive recovery strategies after futsal games. *Journal of Strength & Conditioning Research, 22*(5), 1402-1412.

Timm, K. E. (1994). A randomized-control study of active and passive treatments for chronic low back pain following L5 laminectomy. *Journal of Orthopaedic & Sports Physical Therapy, 20*(6), 276-286.

Tipton, C. M., Matthes, R. D., Maynard, J. A., & Carey, R. A. (1975). The influence of physical activity on ligaments and tendons. *Medicine & Science in Sports, 7*(3), 165-175.

Torstensen, T. A., Ljunggren, A. E., Meen, H. D., Odland, E., Mowinckel, P., & Geijerstam, S. (1998). Efficiency and costs of medical exercise therapy, conventional physiotherapy, and self-exercise in patients with chronic low back pain. A pragmatic, randomized, single-blinded, controlled trial with 1-year follow-up. *Spine, 23*(23), 2616-2624.

Tsutsumi, T., Don, B. M., Zaichkowsky, L. D., Takenaka, K., Oka, K., & Ohno, T. (1998). Comparison of high and moderate intensity of strength training on mood and anxiety in older adults. *Perceptual & Motor Skills, 87*(3/1), 1003-1011.

Urquhart, D. M., Hodges, P. W., Allen, T. J., & Story, I. H. (2005). Abdominal muscle recruitment during a range of voluntary exercises. *Manual Therapy, 10*(2), 144-153.

van Doorn, J. W. (1995). Low back disability among self-employed dentists, veterinarians, physicians and physical therapists in The Netherlands. A retrospective study over a 13-year period (N = 1,119) and an early intervention program with 1-year follow-up (N = 134). *Acta Orthopaedica Scandinavica. Supplement, 263*, 1-64.

van Tulder, M. W., Koes, B. W., & Bouter, L. M. (1995). A cost-of-illness study of back pain in The Netherlands. *Pain, 62*(2), 233-240.

Veres, S. P., Robertson, P. A., & Broom, N. D. (2008). Microstructure and mechanical disruption of the lumbar disc annulus: Part II: How the annulus fails under hydrostatic pressure. 2711-2720.

Vernon-Roberts, B., Fazzalari, N. L., & Manthey, B. A. (1997). Pathogenesis of tears of the anulus investigated by multiple-level transaxial analysis of the T12-L1 disc. *Spine, 22*(22), 2641-2646.

Vlaeyen, J. W., Kole-Snijders, A. M., Boeren, R. G., & van Eek, H. (1995). Fear of movement/(re)injury in chronic low back pain and its relation to behavioral performance. *Pain, 62*(3), 363-372.

Von Korff, M., Jensen, M. P., & Karoly, P. (2000). Assessing global pain severity by self-report in clinical and health services research. *Spine, 25*, 3140-3151.

Vuori, I. M. (2001). Dose-response of physical activity and low back pain, osteoarthritis, and osteoporosis. *Medicine & Science in Sports & Exercise, 33*(6 Supplement), S551-586.

Waddell, G., Reilly, S., Torsney, B., Allan, D. B., Morris, E. W., Di Paola, M. P., et al. (1988). Assessment of the outcome of low back surgery. *Journal of Bone & Joint Surgery British, 70*(5), 723-727.

Warburton, D. E. R., Nicol, C. R., & Bredin, S. S. D. (2006). Health benefits of physical activity: the evidence. *Canadian Medical Association Journal, 174*(6), 801-809.

Ware, J. E. (1978). Effects of acquiescent response set on patient satisfaction ratings. *Medical Care, 16*(4), 327-336.

Ware, J. E. (1981). How to survey patient satisfaction. *Drug Intelligence & Clinical Pharmacy, 15*(11), 892-899.

Ware, J. E. (2000). SF-36 health survey update. *Spine, 25*(24), 3130-3139.

Ware, J. E., & Hays, R. D. (1988). Methods for measuring patient satisfaction with specific medical encounters. *Medical Care, 26*(4), 393-402.

Ware, J. E., Snyder, M. K., Wright, W. R., & Davies, A. R. (1983). Defining and measuring patient satisfaction with medical care. *Evaluation & Program Planning, 6*(3-4), 247-263.

Watanabe, E., Takeshima, N., Okada, A., & Inomata, K. (2001). Effects of increasing expenditure of energy during exercise on psychological well-being in older adults. *Perceptual & Motor Skills, 92*(1), 288-298.

Waxman, R., Tennant, A., & Helliwell, P. (2000). A prospective follow-up study of low back pain in the community. *Spine., 25*(16), 2085-2090.

Weber, H. (1983). Lumbar disc herniation. A controlled, prospective study with ten years of observation. *Spine, 8*(2), 131-140.

Weber, K. R., Brown, L., Coburn, J., & Zindor, S. (2008). Acute effects of heavy-load squats on consecutive squat jump performance. *Journal of Strength & Conditioning Research, 22*(3), 726-730.

Weinstein, J. N., Lurie, J. D., Olson, P. R., Bronner, K. K., & Fisher, E. S. (2006). United States' trends and regional variations in lumbar spine surgery: 1992-2003. *Spine, 31*(23), 2707-2714.

Weinstein, J. N., Lurie, J. D., Tosteson, T. D., Skinner, J. S., Hanscom, B., Tosteson, A. N. A., et al. (2006). Surgical vs nonoperative treatment for lumbar disk herniation: the Spine Patient Outcomes Research Trial (SPORT) observational cohort. *Journal of the American Medical Association, 296*(20), 2451-2459.

Weinstein, J. N., Tosteson, T. D., Lurie, J. D., Tosteson, A. N. A., Hanscom, B., Skinner, J. S., et al. (2006). Surgical vs nonoperative treatment for lumbar disk herniation: the Spine Patient Outcomes Research Trial (SPORT): a randomized trial. *Journal of the American Medical Association, 296*(20), 2441-2450.

Weishaupt, D., Zanetti, M., Hodler, J., & Boos, N. (1998). MR imaging of the lumbar spine: prevalence of intervertebral disk extrusion and sequestration, nerve root compression, end plate abnormalities, and osteoarthritis of the facet joints in asymptomatic volunteers. *Radiology, 209*(3), 661-666.

White, A. H., von Rogov, P., Zucherman, J., & Heiden, D. (1987). Lumbar laminectomy for herniated disc: a prospective controlled comparison with internal fixation fusion. *Spine, 12*(3), 305-307.

Wilke, H. J., Wolf, S., Claes, L. E., Arand, M., & Wiesend, A. (1995). Stability increase of the lumbar spine with different muscle groups. A biomechanical in vitro study. *Spine, 20*(2), 192-198.

Williams, B. (1994). Patient satisfaction: a valid concept? *Social Science & Medicine,* 38 (4),509-516.

Williams, B., Coyle, J., & Healy, D. (1998). The meaning of patient satisfaction: an explanation of high reported levels. *Socail Science & Medicine, 47(9), 1351-1359.*

Willoughby, D. S. (1993). The effects of mesocycle-length weight training programs involving periodization and partially equated volumes on upper and lower body Strength. *The Journal of Strength and Conditioning Research, 7,* 2-8.

Wilson, G. J., Newton, R. U., Murphy, A. J., & Humphries, B. J. (1993). The optimal training load for the development of dynamic athletic performance. *Medicine & Science in Sports & Exercise, 25*(11), 1279-1286.

Winett, R. A. (2004). Meta-analyses do not support performance of multiple sets or high volume resistance training. *Journal of Exercise Physiology online, 7*(5), 10-20.

Winett, R. A. (2006). Comment-resistance training for strength: effect of number of sets and contraction speed. *Medicine & Science in Sports & Exercise, 38*(2), 401.

Winett, R. A., & Carpinelli, R. N. (1999). The quantity and quality of exercise for healthy adults.[comment]. *Medicine & Science in Sports & Exercise, 31*(6), 916-920.

Winkler, J. D., Kanouse, D. E., & Ware, J. E. (1982). Controlling for acquiescence response set in scale development. *Journal of Applied Psychology Vol 67(5) 555-561*.

Wise, M. B., Uhl, T. L., Mattacola, C. G., Nitz, A. J., & Kibler, W. B. (2004). The effect of limb support on muscle activation during shoulder exercises. *Journal of Shoulder & Elbow Surgery*, *13*(6), 614-620.

Wittink, H. (2005). Functional capacity testing in patients with chronic pain. *Clinical Journal of Pain, 21*(3), 197-199.

Wittink, H., Turk, D. C., Carr, D. B., Sukiennik, A., & Rogers, W. (2004). Comparison of the redundancy, reliability, and responsiveness to change among SF-36, Oswestry Disability Index, and Multidimensional Pain Inventory. *Clinical Journal of Pain, 20*(3), 133-142.

Yen, Y. C., Eaton, S., & Maxwell, M. W. (1993). The relationship between the St Thomas and Oswestry disability scores and the severity of low back pain. *Journal of Manipulative and Physiological Therapeutics, 16*(1), 14-18.

Yilmaz, F., Yilmaz, A., Merdol, F., Parlar, D., Sahin, F., & Kuran, B. (2003). Efficacy of dynamic lumbar stabilization exercise in lumbar microdiscectomy. *Journal of Rehabilitation Medicine, 35*(4), 163-167.

Ylinen, J. J., Savolainen, S., Airaksinen, O., Kautiainen, H., Salo, P., & Hakkinen, A. (2003). Decreased strength and mobility in patients after anterior cervical diskectomy compared with healthy subjects. *Archives of Physical Medicine & Rehabilitation, 84*(7), 1043-1047.

Yoganandan, N., Maiman, D. J., Pintar, F., Ray, G., Myklebust, J. B., Sances, A., et al. (1988). Microtrauma in the lumbar spine: a cause of low back pain. *Neurosurgery*, 23(2), 162-168.

Yorimitsu, E., Chiba, K., Toyama, Y., & Hirabayashi, K. (2001). Long-term outcomes of standard discectomy for lumbar disc herniation: a follow-up study of more than 10 years. *Spine*, 26(6), 652-657.

Zatsiorsky, V. M. (1995). Science and practice of strength training. Champaign: Human Kinetics.

APPENDICES

APPENDIX 1

Confirmation of Ethical Consent

Health and Disability Ethics Committees

Upper South A Regional Ethics Committee
Ministry of Health
4th Floor, 250 Oxford Tce
PO Box 3877
Christchurch
Phone (03) 372 3037
Fax (03) 372 1015
Email: uppersouth_ethicscommittee@moh.govt.nz

23 July 2009

Mr Barry Donaldson
32A Woodchester Avenue
Christchurch 8013

Dear Mr Donaldson

Comparison of a progressive non-aggravating strengthening programme versus usual post-operative regime for lumbar spine injury
Investigators: B Donaldson, Prof D Rivett, Prof T Shipton (Supervisor)
Ethics ref: CTY/01/09/122

As requested, please find copies of the following;

- Letter confirming ethical approval of the above study by the Canterbury Ethics Committee, dated 29 January 2002
- Letter confirming an extension of ethical approval for a further 18 months by the Canterbury Ethics Committee, dated 16 September 2003
- Letter confirming an extension of ethical approval for a further 12 months by the Upper South B Regional Ethics Committee, dated 30 October 2006

Unfortunately our file contains only copies of the text of the first two letters, and these are not signed or printed on letterhead. However, I can confirm that these approvals were granted by the Canterbury Ethics Committee.

If you require further information, please do not hesitate to contact me.

Yours sincerely

Alieke Dierckx
Upper South A Ethics Committee Administrator
Alieke_dierckx@moh.govt.nz

29 January 2002

Mr B Donaldson
Redwood Physiotherapy Clinic
199 Main North Road
Christchurch 5

Dear Barry

Comparison of a progressive non-aggravating strengthening programme versus usual post-operative regime for lumbar spine injury
Investigators: B Donaldson, Prof D Rivett, Prof T Shipton (Supervisor)
Ethics reference no: 01/09/122
Information sheet/consent form version 10/28/2001

Thank you for your letter of advising that Mr Inglis has agreed to provide his patients with information about the study and that you will only use his patients.

I am pleased to advise that, using the delegated authority granted her by the Committee, the Chairperson of the Canterbury Ethics Committee has given final ethical approval for this study to proceed in Canterbury.

The Committee certifies that it is satisfied this trial is not conducted principally for the benefit of the manufacturer or distributor of the medicine or item in respect of which the trial is carried out. This certification is for the purposes of the Accident Insurance Act 1998, section 35(5).

Approval is until 1 February 2004. The Committee will review the study annually and notify you if it withdraws approval. It is your responsibility to forward a progress report in January 2003. Failure to do so may result in withdrawal of ethical approval. A final report is also required at the conclusion of the study. Report forms are available from the administrator.

It is also a condition of approval that the Committee is advised of any adverse events, if the study does not commence, or the study is altered in any way, including all documentation eg advertisements, letters to prospective participants. Please quote the above ethics committee reference number in all correspondence.

The Committee wishes you well with your research.

Yours sincerely

Sally Cook
Ethics Committee Administrator

16 September 2003

Mr B Donaldson
Redwood Physiotherapy Clinic
199 Main North Road
Christchurch 5

Dear Barry

Comparison of a progressive non-aggravating strengthening programme versus usual post-operative regime for lumbar spine injury
Investigators: B Donaldson, Prof D Rivett, Prof T Shipton (Supervisor)
Ethics reference no: 01/09/122
Information sheet/consent form version 10/28/2001

Thank you for the letter to participants and consent form for participants to continue in the study for a further 18 months and complete three more sets of outcome measures.

The letter and consent form have been given ethical approval by the Chairperson under delegated authority.

Yours sincerely

Sally Cook
Ethics Committee Administrator

APPENDIX 2

Initial consent form for participants to authorise the orthopaedic surgeon to give their contact details to the principal researcher.

CONSENT FOR MR GRAHAME INGLIS (ORTHOPAEDIC SURGEON) TO GIVE YOUR NAME, ADDRESS, AND CONTACT PHONE NUMBERS TO BARRY DONALDSON

EXPLANATION

Barry Donaldson is a physiotherapist currently undertaking a Masters of Health Science degree through the Christchurch School of Medicine. Barry's research thesis is based on post-operative rehabilitation with patients who are recovering from low back surgery

This consent allows Barry Donaldson to contact you in order that he may explain the project to you with a view to seeking your consent to participate in the project.

IMPORTANT

You are under no obligation to sign this consent. If you do consent to speak with Barry Donaldson you are under no obligation and will incur no disadvantage by not agreeing to take part in the research project.

NAME

ADDRESS

CONTACT PHONE NUMBERS **home** ……

work ……

cell phone ……..

SIGNATURE **DATE**

APPENDIX 3

The participant information sheet

Information Sheet for Participants

A comparison of usual post-operative routine compared with a gym based exercise routine.

Thank you for considering participation in this research project. This project is the research component required in order to complete a Masters of Health Science Degree. Please read this information sheet carefully before deciding whether or not to participate. If you decide to participate in this study we thank you. If you decide not to take part there will be no disadvantage to you of any kind and we thank you for considering our request.

What is the aim of the project?

This study is designed to test whether a specific gym based exercise programme achieves an improved outcome compared with normal low back post-operative routine. As yet no information is known on this topic. This study has the potential for health professionals to improve their knowledge in regard to post-operative rehabilitation and hence improve outcomes for future patients.

What type of participants are being sought?

We are seeking people between the ages of 20 yrs to 65 yrs who have no major medical problems other than the low back problem for which they have just had surgery.

The operation to the low back you have had must not have been for purposes of removal of infection, tumour or inflammation. It is preferable that you are a non-smoker, and that your low back surgery was incident / accident related. ie: your low back problem was caused by a particular event.

The purpose of the above criteria is to:

a) Safeguard you the patient and

b) Give the study an optimum chance of a defining result.

What will participants be asked to do?

If you decide to participate in the study you will then be randomly allocated to one of two groups.

Group A: will undergo usual post-operative routine and this will mean that you follow the usual instructions of your surgeon and doctor, this may or may not mean that you receive follow up treatment or rehabilitation. Whether you do depends on your post-operative progress and the opinion of your supervising doctor.

Group B: will participate in an exercise based gym rehab programme. This requires exercise sessions lasting 30 - 40 minutes 3 x per week for 26 weeks. This exercise programme is very specific, progressive and painfree. We emphasise that the program must be non-aggravating. The programmes will be performed at a gymnasium of your choice, usually one which is convenient for you to attend. The exercise programmes will be supervised by two physiotherapists experienced in this type of management, Sean Wilson and Barry Donaldson. There is no cost involved for your programme, as your gym membership will be paid out of funding provided by the study.

Patients in Group A will be offered the opportunity of completing the gym programme at the end of the trial if they wish.

Outcome Measures

During the study an independent observer will contact participants in both groups. This person will advise you on filling out the outcome measures, which are provided for you with this information leaflet. These outcome measures are to be completed by you and placed in the sealed envelope supplied, then posted to Barry Donaldson. Barry Donaldson and possibly Professor Ted Shipton (research supervisor) are the only people who will see this information. The outcome measures will be recorded at Post-op

6 weeks

14 weeks

23 weeks

32 weeks

58 weeks

84 weeks

The study lasts for 20 months post-operation. This allows for patients to complete their post-op rehab phase and return to normal life.

Compensation

In the unlikely event of a physical injury as a result of your participation in this study, you will be covered by accident compensation legislation with its limitations. If you have any questions about ACC please feel free to ask the researcher for more information before you agree to take part in this trial.

Personal information required:

Name:

Address:

Contact Phone Numbers:

Age:

Gender:

What is your ethnic background?

Employment:

Current Employment status. Ie: do you have a job to return to?

GP:

Operating Surgeon:

Brief Medical History:

Date of operation:

History of low back problems prior to this last operation:

Height / Weight

Security of Personal information:

For Group A and B, only Barry Donaldson and possibly his supervisor will see all information. For purposes of contact, the independent observer will have access to contact information. For Group B, Sean Wilson will have access to necessary clinical information of patients that he is managing. All information will be kept in a secure cabinet at Redwood Physiotherapy Clinic, 199 Main North Road, Christchurch in a limited access room. On completion of the study, all information will be disposed of via an information security firm

and shredded. You have the right to access and correct personal information held on file by the investigator that relates directly to you. Results of this project may be published but any data included will in no way be linked to any specific participant. A copy of the results shall be forwarded to you if you wish. Simply circle YES or NO on the study participation consent form.

Can participants change their mind and withdraw from the study?

If you decide to take part in the study, you must understand that your participation is entirely voluntary. No remuneration is offered to you.

You may withdraw from the study at any time without having to give a reason and without any disadvantage of any kind in future dealings you may have with us, the investigators or the University of Otago.

What if participants have any questions?

If you have any questions about our project either now or in the future, please feel free to contact either:

Barry Donaldson
Redwood Physiotherapy Clinic
199 Main North Road
CHRISTCHURCH
Ph: 03 352 9900

or

Professor Ted Shipton
Christchurch Hospital
Anaesthetics Department
Ph: 03 364 0640

Further to this if you have any queries or concerns regarding your rights as a participant in this study, you may wish to contact a Health and Disability Services Consumer Advocate, Telephone (03) 377 7501 or 0800 377 766 outside Christchurch.

Barry Donaldson is a Masters of Health Science student at the Christchurch School of Medicine, Barry is the principal investigator for this study. Professor Shipton is Professor of Anaesthetics Christchurch School of Medicine and is Barry Donaldson's University supervisor for this study.

This study has received ethical approval from the Canterbury Ethics Committee.

APPENDIX 4

Participant consent form and personal information form

CONSENT FORM FOR PARTICIPANTS

A comparison of usual post-operative management compared with a gym based exercise programme.

- I have read and I understand the Information Sheet for volunteers participating in this study. I have had the opportunity to discuss the study and to ask questions, which have been answered to my satisfaction.
- I understand that any data in questionnaires relating to this study will be destroyed at the conclusion of this study.
- I understand that taking part in this study is voluntary and that I may withdraw from the study at any time.
- I understand that my withdrawal from the study will in no way affect any future dealing I have with the investigators or the University of Otago.
- I am aware of the risks involved in the study and of the arrangements for compensation should injury or illness occur.
- I understand that my participation in this study is confidential and that I will remain anonymous in the study.
- I understand that the information I provide in this study may be published in an anonymous form.
- I wish to receive a summary of the results of the study: YES ☐ NO ☐
- If you are likely to be moving address or changing contact phone numbers in the next six months please tick YES ☐ NO ☐

I, _____, (full name) hereby consent to take part in this project.

Signature: _____ (Participant)

Signature: _____ (Principal Researcher)

Date: _____

This study has received ethical approval from the Canterbury Ethics Committee.

Personal information required:

Name:

Address:

Contact Phone Numbers:

Age:

Gender:

Height: Weight

What is your ethnic background?

Occupation:

Current Employment status. ie: do you have a job to return to?

GP:

Operating Surgeon:

Brief Medical History:

Date of operation:

History of low back problems prior to this last operation:

Time from onset of leg symptoms to surgery:

Hospital stay time:

Time off work prior to surgery:

Time to return to full time duties:

If you are likely to be moving address or changing contact phone numbers in the next six months please tick YES ☐ NO ☐

APPENDIX 5

The subjective functional outcome measures:

- The Oswestry Low Back Index

- The Roland-Morris 24 questionaire

- The Short Form 36

Oswestry Low Back Index

Instructions

Mark in each section only the one box which applies to you. We realise you may consider that two of the statements in any one section relate to you, but please just mark the box which most closely describes your problem.

Section 1 – Pain Intensity

☐ I can tolerate the pain I have without having to use pain killers.

☐ The pain is bad but I manage without taking pain killers.

☐ Pain killers give complete relief from pain.

☐ Pain killers give moderate relief from pain.

☐ Pain killers give very little relief from pain.

☐ Pain killers have no effect on the pain and I do not use them.

Section 2 – Personal Care (washing, dressing etc.)

☐ I can look after myself normally without causing extra pain.

☐ I can look after myself normally but it causes extra pain.

☐ It is painful to look after myself and I am slow and careful.

☐ I need some help but manage most of my personal care.

☐ I need help every day in most aspects of self care.

☐ I do not get dressed, wash with difficulty and stay in bed.

Section 3 – Lifting

☐ I can lift heavy weights without extra pain.

☐ I can lift heavy weight but it gives extra pain.

☐ Pain prevents me from lifting weights off the floor, but I can manage if they are conveniently positioned. eg: on a table.

☐ Pain prevents me from lifting heavy weights but I can manage light to medium weights if they are conveniently positioned.

☐ I can lift only very light weights.

☐ I cannot lift or carry anything at all.

Section 4 – Walking

☐ Pain does not prevent me walking any distance.

☐ Pain prevents me walking more than 1 mile.

☐ Pain prevents me walking more than ½ a mile.

☐ Pain prevents me walking more than ¼ a mile.

☐ I can only walk using a stick or crutches.

☐ I am in bed most of the time and have to crawl to the toilet.

Section 5 – Sitting

☐ I can sit in a chair as long as I like.

☐ I can only sit in my favourite chair as long as I like.

☐ Pain prevents me from sitting more than 1 hour.

☐ Pain prevents me from sitting more than ½ hour.

☐ Pain prevents me from sitting more than 10 mins.

☐ Pain prevents me from sitting at all

Section 6 – Standing

☐ I can stand as long as I want without extra pain.

☐ I can stand as long as I want but it gives me extra pain.

☐ Pain prevents me from standing for more than 1 hour.

☐ Pain prevents me from standing for more than 30 mins.

☐ Pain prevents me from standing for more than 10mins.

☐ Pain prevents me from standing at all.

Section 7 - Sleeping

☐ Pain does not prevent me from sleeping well.

☐ I can sleep well only by using tablets

☐ Even when I take tablets I have less than 6 hours sleep.

☐ Even when I take tablets I have less than 4 hours sleep.

☐ Even when I take tablets I have less than 2 hours sleep.

☐ Pain prevents me from sleeping at all.

Section 8 – Sex Life

☐ My sex life is normal and causes no extra pain.

☐ My sex life is normal but causes some extra pain..

☐ My sex life is nearly normal but very painful.

☐ My sex life is severely restricted by pain.

☐ My sex life is nearly absent because of pain.

☐ Pain prevents any sex life at all.

Section 9 – Social life

☐ My social life is normal and gives me no extra pain.

☐ My social life is normal but increases the degree of pain.

☐ Pain has no significant effect on my social life apart from limiting my more energetic interests, eg; dancing.

☐ Pain has restricted my social life and I do not go out as often.

☐ Pain has restricted my social life to my home.

☐ I have no social life because of pain.

Section 10 – Travelling

☐ I can travel anywhere without extra pain.

☐ I can travel anywhere but it gives me extra pain.

☐ Pain is bad but I manage journeys over 2 hours.

☐ Pain restricts me to journeys of less than 1 hour.

☐ Pain restricts me to short necessary journeys under 30 minutes.

☐ Pain prevents me from travelling except to the doctor or hospital.

Roland and Morris Disability Questionnaire

When your back hurts, you may find it difficult to do some of the things you normally do.

These are some sentences that people have used to describe themselves when they have back pain. When you read them, you may find that some stand out because they describe you today. As you read the list, think of yourself today. When you read a sentence that describes you today circle YES. If that sentence does not describe you today circle NO. Remember, only answer YES if you are sure that sentence describes you today.

a) I stay at home most of the time because of my back. YES ☐ NO ☐

b) I change position frequently to try and get my back comfortable. YES ☐ NO ☐

c) I walk more slowly than usual because of my back YES ☐ NO ☐

d) Because of my back I am not doing any of the jobs I usually do around the house. YES ☐ NO ☐

e) Because of my back, I use a handrail to get upstairs. YES ☐ NO ☐

f) Because of my back, I lie down to rest more often. YES ☐ NO ☐

g) Because of my back, I have to hold on to something to get out of any easy chair. YES ☐ NO ☐

h) Because of my back, I try to get other people to do things for me. YES ☐ NO ☐

i) I get dressed more slowly than usual because of my back. YES ☐ NO ☐

j) I only stand up for short periods of time because of my back. YES ☐ NO ☐

k) Because of my back I try not to bend or kneel down YES ☐ NO ☐

l) I find it difficult to get out of a chair because of my back. YES ☐ NO ☐

m) My back is painful almost all the time. YES ☐ NO ☐

n) I find it difficult to turn over in bed because of my back. YES ☐ NO ☐

o) My appetite is not very good because of my back pain. YES ☐ NO ☐

p) I have trouble putting on my socks (or stockings) because of the pain in my back. YES ☐ NO ☐

q) I only walk short distances because of back pain. YES ☐ NO ☐

r) I sleep less well because of my back. YES ☐ NO ☐

s) Because of my back pain, I get dressed with help from someone else. YES ☐ NO ☐

t) I sit down for most of the day because of my back. YES ☐ NO ☐

u) I avoid heavy jobs around the house because of my back. YES ☐ NO ☐

v) Because of my back pain, I am more irritable and bad tempered with people than usual. YES ☐ NO ☐

w) Because of my back, I go upstairs more slowly than usual. YES ☐ NO ☐

x) I stay in bed most of the time because of my back. YES ☐ NO ☐

Score Total of all items answered **YES:** ☐

The Short Form – 36 Health Survey

Instructions

This survey asks for your views about your health. This information will help keep track of how you feel and how well you are able to do your usual activities.

Answer every question by marking the answer as indicated. If you are unsure about how to answer a question, please give the best answer you can.

1. In general would you say your health is:

Excellent……. 1
Very Good…….. 2
Good…… 3
Fair……. 4
Poor…… 5

2. Compared to a year ago, how would you rate your health in general now?

Much better now than one year ago…… 1
Somewhat better now than on year ago…….. 2
About the same now as one year ago……. 3
Somewhat worse now than one year ago…….. 4
Much worse now than one year ago…… 5

3. The following items are about activities you might do during a typical day. Does your health now limit you in these activities? If so how much ?

(circle one number on each line)

Activities	Yes, Limited A lot	Yes, Limited A Little	No, not Limited At all
a. Vigorous activities, such as running, lifting heavy objects, participating in strenuous sports.	1	2	3
b. Moderate activities, such as moving a table, pushing a vacuum cleaner, bowling, or playing golf.	1	2	3
c. Lifting or carrying groceries	1	2	3
d. Climbing several flights of stairs	1	2	3
e. Climbing one flight of stairs	1	2	3
f. Bending, kneeling or stooping..	1	2	3
g. Walking more than one mile	1	2	3
h. Walking several blocks	1	2	3
i. Walking one block	1	2	3
j. Bathing or dressing yourself	1	2	3

4. During the past 4 weeks ,have you had any of the following problems with your work or other regular daily activities as a result of your physical health?

(circle one number on each line)

	Yes	No
a. Cut down on the amount of time you spent on work or other activities.	1	2
b. Accomplished less than you would like	1	2
c. Were limited in the kind of work or other activities.	1	2
d. Had difficulty performing the work or other activities (eg it took extra effort)	1	2

5. During the past 4 weeks, have you had any of the following problems with your work or other regular activities as a result of any emotional problems (such as feeling depressed or anxious)?

(circle one number on each line)

	Yes	No
a. Cut down the amount of time you spent on work or other activities.	1	2
b. Accomplished less than you would like	1	2
c. Didn’t do work or other activities as carefully as usual	1	2

6. During the past 4 weeks, to what extent has your physical healthy or emotional problems interfered with your normal social activities with family, friends, neighbours or groups?

(Circle one)

Not at all……	1
Slightly……	2
Moderately…….	3
Quite a bit……	4
Extremely……	5

7. How much bodily pain have you had during the past 4 weeks?

(Circle one)

None…….	1
Very Mild…….	2
Mild……..	3
Moderate……..	4
Severe……...	5
Very severe……...	6

8. During the past 4 weeks, how much did pain interfere with your normal work (including both work outside the home and housework)?

(Circle one)

Not at all……	1
Slightly……	2
Moderately…….	3
Quite a bit……	4
Extremely……	5

9. These questions are about how you feel and how things have been with you during the past 4 weeks. For each question, please give the on answer that comes closest to the way you have been feeling. How much of the time during the past 4 weeks.

(circle one number on each line)

	All of the Time	Most of the time	A Good bit of the time	Some of the time	A Little of the time	None of the Time
a. Did you feel full of pep	1	2	3	4	5	6
b. Have you been a very nervous person	1	2	3	4	5	6
c. Have you felt so down in the dumps that nothing could cheer you up?	1	2	3	4	5	6
d. Have you felt calm and peaceful?	1	2	3	4	5	6
e. Did you have a lot of energy?	1	2	3	4	5	6
f. Have you felt down hearted and blue?	1	2	3	4	5	6
g. Did you feel worn out?	1	2	3	4	5	6
h. Have you been a happy person?	1	2	3	4	5	6
i. Did you feel tired?	1	2	3	4	5	6

10. During the past 4 weeks, how much of the time has your physical health or emotional problems interfered with your social activities (like visiting with friends, relatives, etc.)?

(Circle one)

All of the time…… 1

Most of the time……. 2

Some of the time…… 3

A little of the time…….. 4

None of the time…… 5

11. How TRUE or FALSE is each of the following statements for you?

(circle one number on each line)

	Definitely True	Mostly True	Don't know	Mostly false	Definitely false
a. I seem to get sick a little easier than most people	1	2	3	4	5
b. I am as healthy as anybody I know.	1	2	3	4	5
c. I expect my health to get worse	1	2	3	4	5
d. My health is excellent	1	2	3	4	5

APPENDIX 6

The annual questionnaires, version I for years one and two, and version II (including Satisfaction) for year three.

Questionnaire For Participants

Dear Participant,

This is an additional quick five questions, which I would greatly appreciate you answering. These questions provide additional information that the assessments we have already made don't cover.

1) During the last three months, have you seen your GP in relation to your low back?
YES ☐ NO ☐

If your answer is **yes**, approximately how many visits? ☐

2) During the last three months have you seen a Physiotherapist, Chiropractor, Osteopath or any other therapist for your low back? YES ☐ NO ☐

If your answer is **yes**, approximately how many visits? ☐

Please tick the appropriate box.

NO OF TREATMENTS	1 - 5	6 - 10	11 - 15	16 - 20	20 plus
PHYSIOTHERAPIST					
CHIROPRACTOR					
OSTEOPATH					
ACCUPUNCTURIST					
OTHER: please name ……					

3) During the last three months have you needed to take any form of medication (e.g. painkillers, anti-inflammatories) for your low back? YES ☐ NO ☐

If your answer is **yes**, approximately how many times would you have taken medication for two days or more? ☐

Please tick below the type of medication you have taken, you may tick more than one type of medication. If you have taken something different for your back please indicate what this may have been. Remember this information is completely confidential and coded for your further privacy.

Medication	Please tick
Anti-inflammatory	
Pain killer	
Sleeping pill	
Anti-depressant	
Other, name if you wish ……	

4) Since returning to full time normal work duties have you had to take time off work again due to your low back problems? YES ☐ NO ☐

If your answer is **yes**, approximately how many days off have you had? ☐

5) Since your discharge from Mr Inglis have you been participating in any form of regular exercise?

If you have please give a brief description of type and time eg walking 30 minutes per day 3 times per week etc.

Dear,

This is an additional quick five questions, which I would greatly appreciate you answering. These questions provide information that is not covered by the other outcome measures.

1) During the last three months have you seen your GP in relation to your low back?

YES ☐ NO ☐

If your answer is **yes**, approximately how many visits? ☐

2) During the last three months have you seen a Physiotherapist, Chiropractor, Osteopath or any other therapist for your low back. YES ☐ NO ☐

If your answer is **yes**, approximately how many visits? ☐

If your answer is yes approximately how many visits.

NO OF TREATMENTS	1 - 5	6 - 10	11 - 15	16 - 20	20 plus
PHYSIOTHERAPIST					
CHIROPRACTOR					
OSTEOPATH					
ACCUPUNCTURIST					
OTHER, please name …….					

3) During the last three months have you needed to take any form of medication (painkillers, anti-inflammatories) for your low back? YES ☐ NO ☐

If your answer is **yes**, approximately how many times would you have taken medication for two days or more? ☐

Please tick below the type of medication you have taken, you may tick more than one type of medication. If you have taken something different for your back please indicate what this may have been. Remember this information is completely confidential and coded for your further privacy.

Medication	Please tick
Anti-inflammatory	
Pain killer	
Sleeping pill	
Anti-depressant	
Other, name if you wish ……	

4) During the past **year** have you had to take time off work again due to low back problems? YES ☐ NO ☐

If your answer is **yes**, approximately how many days off have you had? ☐

5) Since your discharge from Mr Inglis have you been participating in any form of regular exercise? If you have please give a brief description of type and time eg walking 30 minutes per day 3 times per week etc.

6) How satisfied are you with the result of your surgery?

A / Very satisfied B / Satisfied C / Unsatisfied D / Very unsatisfied

7) How satisfied have you been with your post-operation management?

A / Very satisfied B / Satisfied C / Unsatisfied D / Very unsatisfied

APPENDIX 7

The gym programme exercise sheets

- Conditioning

- Hypertrophy A, B and C

- Strength A, B and C

Name: Date:

CONDITIONING PROGRAMME

Exercise	Sets and Reps	Weights				
Warm Up						
Seated Bench Press						
Prone Pull						
Prone Fly						
Front Pull Downs						
Leg Press						
Pulley Hamstring Curls (standing)						
DB Bicep Curl						
Tricep Push Down						
Warm Down						
Brace Abdominals						

Progressions 3x10,3x12,3x15 (20 - 30sec rest between sets)

Increase weights after 3x15 performed comfortably

HYPERTROPHY PROGRAMME A				
Warm up 5 mins rower				
Seated Bench Press	15	12	10	8
Prone Pull	15	12	10	8
Prone Fly				
	15	12	10	8
Front Pull Down	15	12	10	8
Leg Press	15	12	10	8
Pulley hamstring curls (standing)	15	12	10	8
DB Bicep Curl	15	12	10	8
Tricep Push Down	15	12	10	8
Warm down 5 minutes rower				

HYPERTROPHY PROGRAMME B				
Warm Up (5 mins Cardio glide)				
Flat DB Flys / Pec Dec	15	12	10	8
Seated Row	15	12	10	8
Inclined Prone flys	15	12	10	8
Lat Pull Down (Close Grip)	15	12	10	8
Lunges	15	12	10	8
Butt kicks	15	12	10	8
Bicep Curl with bar	15	12	10	8
Elbow extension	15	12	10	8
Warm Down (10 mins Cardio glide)				

HYPERTROPHY PROGRAMME C				
Warm up (5 mins Slide Board / Exercycle)				
Inclined DB Press	15	12	10	8
One Arm Pulls	15	12	10	8
Reverse Pec Deck	15	12	10	8
Upright Rows	15	12	10	8
Backward Lunge	15	12	10	8
Hamstring Curl (lying)	15	12	10	8
Inclined DB bicep curl	15	12	10	8
DB Press	15	12	10	8
Warm Down (5 mins Slide Board / Exercycle)				

STRENGTH PROGRAMME **A**				
Warm up 5 mins rower				
Seated Bench Press	15	10	8	6
Prone Pull	15	10	8	6
Prone Flys				
	15	10	8	6
Front Pull Down	15	10	8	6
Leg Press	15	10	8	6
Pulley hamstring curls	15	10	8	6
DB Bicep Curl	15	10	8	6
Tricep Push Down	15	10	8	6
Warm down 5 minutes rower				

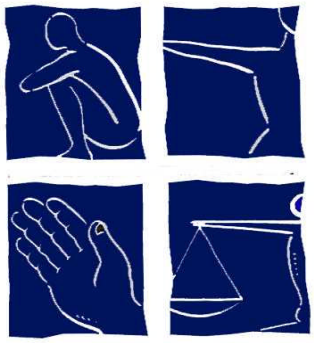
STRENGTH PROGRAMME **B**				
Warm Up (5 mins Cardio glide)				
Flat DB Flys / Pec Deck	15	10	8	6
Seated Rows	15	10	8	6
Inclined Prone flys	15	10	8	6
Lat Pull Down (Close Grip)	15	10	8	6
Lunges	15	10	8	6
Butt kicks	15	10	8	6
Bicep Curl with bar	15	10	8	6
Elbow extension	15	10	8	6
Warm Down (10 mins Cardio glide)				

STRENGTH PROGRAMME **C**				
Warm up (5 mins Slide Board / Exercycle)				
Inclined DB Press	15	10	8	6
One Arm Pulls	15	10	8	6
Reverse Pec Deck	15	10	8	6
Upright Rows	15	10	8	6
Backward Lunge	15	10	8	6
Hamstring Curl (lying)	15	10	8	6
Inclined DB bicep curl	15	10	8	6
DB Press	15	10	8	6
Warm Down (5 mins Slide Board / Exercycle)				

APPENDIX 8

Consent to check medical records for time off work

REDWOOD PHYSIOTHERAPY CLINIC
Barry L. Donaldson. M.N.Z.S.P. Dip M.T.
Manipulative Therapy * Muscle Balancing* Postural Dynamics



Dear

Once again my thanks for your assistance with the study over the last three years; the study is now complete and I am in the process of collecting and collating all of the final data. Once I have completed the final data analysis I will send you a copy of the results; I am hopeful of completing this process early in the New Year.

To this end according to the questionnaires that you filled out; you took some time off work in the last two years of the study. I have an estimation of how much time you took; I now need to validate this information, as you may have been remembering back quite a time, up to one year in some cases and it's a researched fact that our memories aren't that good over that period of time.

For me to validate this information I need to confirm the details of your time off with your GP. To obtain this information I need you to consent that you are happy for me to gather this information. I emphasize that the only information that I am seeking is the time off that you have already advised me of; I simply need to confirm this with your GP's notes.

If you are happy for me to confirm this information with your GP please sign the consent that is outlined below and return this letter in the stamped self addressed envelope. Once again, my thanks and appreciation of your ongoing support for the study. I look forward to sending you the final summary of the study early next year. If you have any queries regarding this process or this consent form please do not hesitate to contact me via the contact details on this letter head.

Yours sincerely

Barry Donaldson.

I consent to Barry Donaldson obtaining information regarding time off work that I took for my low back between the years of 2003-2006.

Name (please print**) …….**

Signature ……... **Date ……**

199 Main North Road (opp St Bedes College) Christchurch. Telephone: (03) 352 9900 Fax: (03) 3543342

APPENDIX 9

Papers published from this thesis