

DOCTOR OF PHILOSOPHY

A pragmatic RCT comparing specific spinal stabilisation exercises and conventional physiotherapy in the management of recurrent low back pain

Cairns, Melinda

Award date:
2002

Awarding institution:
Coventry University

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of this thesis for personal non-commercial research or study
- This thesis cannot be reproduced or quoted extensively from without first obtaining permission from the copyright holder(s)
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

**A Pragmatic RCT Comparing Specific
Spinal Stabilisation Exercises And
Conventional Physiotherapy In The
Management Of Recurrent
Low Back Pain**

Melinda Claire Cairns

**A thesis submitted in partial fulfilment of the
University's requirements for the Degree of Doctor of
Philosophy**

MARCH 2002

**Coventry University in collaboration with
The Royal Orthopaedic Hospital, Birmingham, UK**

TABLE OF CONTENTS	PAGE
ABSTRACT.....	I
PUBLICATIONS AND PRESENTATIONS	II
ACKNOWLEDGMENTS	III
1. CHAPTER 1: INTRODUCTION.....	1
1.1 TITLE	1
1.2 RATIONALE FOR STUDY	1
2. CHAPTER 2: LITERATURE REVIEW	3
2.1 INTRODUCTION	3
2.1.1 <i>Outline of spinal stabilisation training concept</i>	3
2.2 THE PROBLEM OF LOW BACK PAIN.....	4
2.3 CLASSIFICATION OF LOW BACK PAIN	6
2.4 RECURRENCE.....	10
2.5 OUTCOME MEASURES	14
2.6 MEASURING FUNCTIONAL STATUS IN LOW BACK PAIN	15
2.6.1 <i>Physical measures versus self-reported measures</i>	15
2.6.2 <i>Roland Morris Disability Questionnaire</i>	19
2.6.3 <i>Oswestry Disability Index</i>	23
2.6.4 <i>Meaningful change</i>	24
2.7 MEASURING PAIN STATUS IN LOW BACK PAIN	27
2.7.1.1 Visual Analogue Scale and Numerical Rating Scale	28
2.7.1.2 McGill Pain Questionnaire.....	29
2.8 SOCIAL AND PSYCHOSOCIAL FACTORS AND LOW BACK PAIN.....	31
2.8.1 <i>Measuring psychological variables in low back pain</i>	33
2.9 GENERIC MEASURES IN LOW BACK PAIN	40
2.10 COMPLIANCE, MOTIVATION AND ADVICE.....	44
2.11 EXERCISE AND LOW BACK PAIN	46
2.12 CURRENT PHYSIOTHERAPY PRACTICE	48
2.13 THEORY BEHIND THE SPINAL STABILISATION APPROACH.....	50

2.13.1	<i>Properties of stabilising muscles</i>	53
2.14	EVIDENCE SUPPORTING THE SPINAL STABILISATION APPROACH	58
2.14.1	<i>Neuromotor control and recruitment</i>	59
2.14.2	<i>Mechanical effect</i>	62
2.14.3	<i>Reaction to predictable perturbations</i>	63
2.14.4	<i>Relationship to load</i>	64
2.14.5	<i>Exercises used in the recruitment of transversus abdominis in clinical practice</i>	65
2.15	CLINICAL STUDIES	71
2.16	LIMITATIONS OF STUDIES TO DATE AND CURRENT RESEARCH NEEDS.....	78
2.17	SUMMARY	79
3.	CHAPTER 3: METHODOLOGY	83
3.1	RESEARCH QUESTION/HYPOTHESIS	83
3.2	DESIGN	83
3.3	TREATMENT INTERVENTIONS.....	84
3.3.1	<i>Development of guidelines for stabilisation interventions</i>	87
3.3.2	<i>Categories of treatment modalities</i>	87
3.3.3	<i>Standardized advice and educational literature</i>	88
3.3.4	<i>Diagnostic real-time ultrasound</i>	88
3.4	SAMPLE SELECTION CRITERIA	89
3.4.1	<i>Inclusion and exclusion criteria</i>	89
3.5	ETHICAL APPROVAL AND CONSENTING PROCESS	93
3.6	RECRUITMENT OF PATIENTS	94
3.6.1	<i>Randomisation</i>	98
3.7	OUTCOME ASSESSMENT	101
3.8	DEVELOPMENT OF PROTOCOL.....	101
3.8.1	<i>Background fieldwork</i>	102
3.8.2	<i>Profile of clinicians</i>	104
3.8.3	<i>Recruitment of clinicians involved in the RCT</i>	105
3.9	PILOT STUDY.....	107
3.9.1	<i>Sample size calculation</i>	108
3.9.2	<i>Assessment of inclusion and exclusion criteria</i>	109

3.9.3	<i>Assessment of the randomisation strategy</i>	110
3.9.4	<i>Feasibility</i>	110
3.10	DATA ANALYSIS	112
3.10.1	<i>Baseline comparisons</i>	113
3.10.2	<i>Analysis of change</i>	113
3.10.3	<i>Average change over time/ Area Under the Curve</i>	114
3.10.4	<i>Intention-to-treat analysis and handling of missing data</i>	115
3.10.5	<i>Scoring of the SF-36</i>	117
3.10.6	<i>Data monitoring and accuracy</i>	117
3.10.7	<i>Achieved Power</i>	118
4.	CHAPTER 4: RESULTS	119
4.1	INTRODUCTION TO RESULTS	119
4.2	BASELINE DEMOGRAPHICAL AND BIOGRAPHICAL DATA	119
4.3	STUDY VARIABLES AT BASELINE	121
4.3.1	<i>Treatment completers versus treatment non-completers</i>	122
4.3.2	<i>Responders versus non-responders at 12-month follow-up</i>	123
4.4	FOLLOW-UP RATES	124
4.5	CHANGE OVER TIME.....	125
4.5.1	<i>Summary of mean responses at each follow-up point</i>	126
4.6	FUNCTION	127
4.6.1	<i>Roland Morris Disability Questionnaire (Primary Outcome)</i>	128
4.6.1.1	Results of the intention-to-treat analysis with imputation.....	130
4.6.2	<i>Covariate analysis</i>	131
4.6.3	<i>Oswestry Disability Index (ODI)</i>	134
4.6.4	<i>Summary of Functional Assessments</i>	134
4.7	PAIN OUTCOME.....	135
4.8	PSYCHOLOGICAL VARIABLES.....	138
4.9	GENERIC HEALTH MEASURE.....	142
4.10	TREATMENT DETAILS	145
4.11	SUMMARY OF RESULTS	148
5.	CHAPTER 5: DISCUSSION	150

5.1	SUMMARY	150
5.2	INTRODUCTION	151
5.3	EVALUATION OF RESULTS.....	151
5.4	IMPLICATIONS FOR CLINICAL PRACTICE	153
5.4.1	<i>Pain and function</i>	153
5.4.2	<i>Measure of distress</i>	169
5.5	METHODOLOGICAL QUALITY.....	173
5.5.1	<i>Pragmatic designs</i>	174
5.5.2	<i>Randomisation</i>	177
5.5.3	<i>Blinding</i>	179
5.5.4	<i>Sample size</i>	180
5.5.5	<i>Patient management and withdrawal</i>	183
5.5.6	<i>End-points, outcomes and statistical analyses</i>	184
5.6	LIMITATIONS	185
5.7	FUTURE RESEARCH.....	189
5.8	CONCLUSION	191
	REFERENCES	193
	APPENDICES.....	230

TABLES AND FIGURES	PAGE
Table 2-1: Classifications one to four of the Quebec Task Force Classifications of activity related spinal disorders.....	10
Table 2-2: Minimal clinically important difference for Roland Morris Disability Questionnaire scores.....	21
Table 2-3: Activities covered by Oswestry Disability Index	23
Table 2-4: Suggested change scores discriminating between improved and non-improved patient outcome	24
Table 2-5: Summary of the changing emphasis of psychosocial assessment within back pain management guidelines	35
Table 2-6: Summary of categories within Distress Risk Assessment Method.....	39
Table 2-7: Summary of physiological, anatomical and biochemical basis for function of the local and global muscle systems.....	54
Table 3-1: Description of treatment modalities used.....	86
Table 3-2: Summary of exclusion criteria.....	91
Table 3-3: Properties of Outcome measures used in the RCT.....	103
Table 3-4: Grades of treating clinicians.....	104
Table 3-5: Demographics and study variables at baseline for the pilot trial	108
Table 3-6: Strategy for imputation of missing cases.....	116
Table 4-1: Baseline demographical and biographical profile of patients in each group	120
Table 4-2: Summary of study variables at baseline	123
Table 4-3: Study variables at baseline for responders and non-responders.....	124
Table 4-4: Summary of follow-up rates in both groups.....	125
Table 4-5: Summary of analyses undertaken and presentation of results	126
Table 4-6: Summary of mean change scores from baseline to 12-months for outcome measures for each group.....	127
Table 4-7: Summary of mean RMDQ change scores by baseline grouping.....	128
Table 4-8: Summary of results of intention-to-treat analyses for RMDQ scores	130
Table 4-9: Summary of adjusted mean difference in change scores for the two groups, and 95% (or 99%) CI, and results of the analysis of covariance	133
Table 4-10: DRAM classifications at baseline and 12-month follow-up.....	142
Table 4-11: Summary of treatment details for each group	146
Table 5-1: Methodological Quality Criteria	174

Figure 2-1: Summary of coloured flags concepts in assisting with the management of low back pain.....	32
Figure 2-2: Summary of SF-36 measurement model: contribution of items to scales and summary scales.....	42
Figure 2-3 - Diagrammatic representation of the Neutral Zone	51
Figure 2-4: Spinal stability system	52
Figure 2-5: Summary of movement system balance.....	53
Figure 2-6: Anatomical position and relationship of transversus abdominis	56
Figure 2-7: Gross Anatomy of Lumbar Multifidus	57
Figure 2-8: Detailed Anatomy of Lumbar Multifidus	58
Figure 2-9: Basic set-up for motor control studies.....	61
Figure 2-10: Diagrammatic representation of suggested neuromusculoskeletal factors contributing influencing spinal control and development of low back dysfunction	81
Figure 3-1: Summary of recruitment procedure.....	95
Figure 3-2: Progress of Patients Recruited to the RCT	97
Figure 3-3: Minimization procedure and resulting stratification categories.....	100
Figure 4-1: Baseline Quebec Task Force Classification of Symptoms for each group .	121
Figure 4-2: 95% CI for RMDQ for 12-month responders	129
Figure 4-3: 99% CI for mean ODI for 12-month responders.....	134
Figure 4-4: 99% CI for mean SFMPQ for 12-month responders.....	135
Figure 4-5: 99% CI for mean SFMPQ (Sensory) for 12-month responders	136
Figure 4-6: 99% CI for mean SFMPQ (Affective) for 12-month responders.....	137
Figure 4-7: 99% CI for mean VAS for 12-month responders.....	137
Figure 4-8: 99% CI for mean NRS for 12-month responders.....	138
Figure 4-9: DRAM categories of patients screened for inclusion to RCT.....	139
Figure 4-10: Distribution of the DRAM categories at baseline per group.....	140
Figure 4-11: 99% CI for mean MZ score for 12-month responders.....	141
Figure 4-12: 99% CI for mean MSPQ score for 12-month responders.....	142
Figure 4-13: 99% CI for mean PCS and MCS at baseline and 12-months	143
Figure 4-14: Physical Component Summary at Baseline and 12-months	144
Figure 4-15: Mental Component Summary at Baseline and 12-months	144
Figure 4-16: Percentage response to U.K. SF-36 change in health status.....	145

Figure 4-17: Summary of total number of treatment modalities used in each group.... 146

Figure 4-18: Percentage usage of first choice modality for each group..... 147

Abstract

Background

Altered muscular function of the deep abdominal and back muscles has been implicated as a factor in the development and continuation of low back pain (LBP) and small-scale studies, on specific sub-groups of LBP patients, have reported favourable outcomes when these dysfunctions are addressed using specific exercise training. However, these techniques are increasingly being incorporated into treatment packages for non-specific LBP in the UK despite little evidence of their effectiveness in this patient group.

A multi-centered, pragmatic, randomized clinical trial, with 12-month follow-up, was therefore designed to investigate the effectiveness of incorporating specific spinal stabilisation exercises within a physiotherapy treatment package in the management of recurrent LBP patients.

Methods

Following ethical approval, consenting patients with recurrent LBP, without significant levels of distress (as measured by the distress risk assessment method {DRAM}), were randomized to two groups; 'conventional' physiotherapy and the provision of an advice booklet (CT) and 'conventional' physiotherapy, the provision of an advice booklet with the addition of specific spinal stabilisation exercises (SSSE).

Randomisation was stratified for laterality, duration of symptoms and initial functional disability level {Roland Morris Disability Questionnaire-RMDQ} using a minimization procedure. Functional disability (RMDQ) was the main outcome, and generic, disease-specific and psychological measures were also collected. The trial was powered to detect a 5-point difference between groups using 90% power. A total of 221 patients were screened for entry into the trial and 97 were recruited from three metropolitan physiotherapy departments within the UK between May 1999 and September 2000.

Results

All patients were between the ages of 19 and 60 years (mean 38.6, SD: 10.5) and had an average duration of symptoms of 8.7 (8.1) months. Over 30% of the patients screened for entry to the trial were excluded as they showed evidence of psychological distress. Both groups demonstrated improved functioning, reduced pain intensity and an improvement in the physical component of quality of life. Mean change (95% CI) for RMDQ scores between baseline to 12-month follow-up were -4.5 (-6.2 to -3.6) for the SSSE group and -5.2 (-6.7 to -3.6) for the CT group. No statistically significant differences between the two groups were demonstrated for any of the outcome variables. Patients in the spinal stabilisation group received a slightly greater mean number of treatment sessions over a longer period than the conventional physiotherapy treatment group (7.5 (2.5) over 11 weeks compared to 5.9 (2.3) over 8 weeks respectively). Exploration of the content of each treatment package revealed a combination of treatments was used, most frequently active exercise and manual therapy, with little use of electrotherapy or mechanical lumbar traction.

Discussion and Conclusion

This trial represents the largest to date investigating the effects of specific spinal stabilisation exercises, and the first examining their use in a recurrent LBP population. Results indicate that physiotherapy is effective in reducing functional disability and to a lesser extent pain intensity, with improvements maintained at one year following completion of treatment, but that the addition of spinal stabilisation exercises to conventional physiotherapy and an advice booklet, does not provide any obvious additional benefit in terms of functional disability or pain intensity. These findings are of importance as they support the ongoing use of physiotherapy treatment packages in the management of recurrent LBP patients, without significant levels of distress, but challenge the assumption that stabilisation training provides an additional benefit in this particular group of LBP patients.

Publications and presentations

Presentations

Cairns, MC, Foster, NE, Wright, C (2001) A pragmatic randomised controlled trial of specific spinal stabilization exercises and conventional physiotherapy in the management of recurrent lumbar spine pain and dysfunction- 6-month follow-up. 4th Interdisciplinary World Congress on Low Back and Pelvic Pain, Montreal Canada

Cairns, MC, Foster, NE, Wright, CC (2000) Level Of Distress Measured Using The DRAM In A Recurrent LBP Population Referred For Physiotherapy. Society for Back Pain Research, Manchester, UK.

Cairns, MC, Foster, NE, Wright, C (2000) A pragmatic randomised controlled trial of specific spinal stabilization exercises and conventional physiotherapy in the management of recurrent lumbar spine pain and dysfunction-Preliminary Results. International Federation of Orthopaedic Manipulative Therapists (IFOMT), Perth, Western Australia

Yelland, NE & Cairns, MC (2000) Outcome of triaging from a collaborative physiotherapy/consultant low back pain triaging clinic. Chartered Society of Physiotherapy Congress, ICC Birmingham, UK

Peer Reviewed Papers

Cairns, MC, Foster, NE, Wright, CC, Pennington, D (2002) Level Of Distress In A Recurrent Low Back Pain Population Referred For Physical Therapy. Spine (In Press)

Cairns, MC, Harrison, K, Wright, C (2000) Pressure Biofeedback: A useful tool in the quantification of deep abdominal muscle dysfunction? Physiotherapy 86(3): 127-138

Published Abstract

Cairns, MC, Foster, N, Wright, C (1999) A pragmatic randomised controlled trial of specific spinal stabilization Exercises in the management of recurrent lumbar spine pain and dysfunction-Pilot Results. Presented at Physiotherapy Research Society Conference, November 1999. Published in: Physiotherapy 86(1): 36

Cairns, MC & Harrison, K (1998) Pressure Biofeedback: A useful tool in the quantification in deep abdominal muscle dysfunction? Society for Back Pain Research, London., UK

Posters

Cairns, MC, Harrison, K & Wright, C (1998) Pressure biofeedback: A useful tool in the quantification of muscular abdominal dysfunction? Manipulation Association of Chartered Physiotherapists, AGM and Annual Conference, London, UK

Yelland, NE & Cairns, MC (1999) The outcome of a collaborative physiotherapy/consultant low back pain triage clinic. Chartered Society of Physiotherapy Congress, Birmingham, UK

Acknowledgments

I would like to acknowledge the support and assistance of the following people and extend my grateful thanks;

To all my supervisors for their support and encouragement throughout in addition to constructive criticism;

- Dr Nadine Foster, Lecturer, Department of Physiotherapy Studies, Keele University, UK
- Christine Wright, Principal Lecturer and Deputy Director of Psychosocial Centre, Coventry University, UK
- Professor Donald Pennington, Pro-Vice Chancellor, Coventry University, UK

To Associate Professor Gwen Jull and colleagues, University of Queensland, Brisbane for their support and advice during the development of the trial protocol and acting as an external expert for this project.

To Professor Gordon Waddell and Tracey Bury (CSP) for their constructive criticism as external referees for the methodology

To the clinical and administrative staff at The Royal Orthopaedic Hospital, Queen Elizabeth Hospital and Selly Oak Hospital, Birmingham for fully supporting this trial from initial development to completion; especially those involved in screening and treating patients within the trial. Thanks to all the patients who participated in the trial and follow-up.

To the League of Friend volunteers at The Royal Orthopaedic Hospital, especially Christine, who willingly stuffed envelopes so skilfully.

To Judy Dawson, Medical Librarian at the Cadbury Library, Research and Teaching Centre at Royal Orthopaedic Hospital for her unstinting support over the last four years, always rising to the challenge and never saying no.

To Carolyn Edwards and Nicoli Yelland (MCSP), for their clinical input into the design of the trial and all their support during recruitment

To all my friends and family who have provided the support and perspective needed to undertake this work, but most especially to my husband Nigel.

1. Chapter 1: Introduction

1.1 Title

A Pragmatic Randomised Controlled Trial Comparing Specific Spinal Stabilisation Exercises And Conventional Physiotherapy In The Management Of Recurrent Low Back Pain.

1.2 Rationale for study

The current study was stimulated by the increasing clinical use of endurance training of the trunk muscles for the treatment of hypothesised 'clinical instability'; associated with various type of low back pain (LBP). Early experimental and observational studies have found that muscle function alters in the presence of pain, and this work has been developed and expanded into the area of LBP. The resultant findings that spinal muscles and abdominal function alters in the presence of LBP led to the development of clinically based programmes, specifically by physiotherapists, to address this muscle dysfunction. Small-scale studies and anecdotal evidence have shown beneficial outcomes from using exercise techniques to address this muscle dysfunction.

These techniques have gained an extensive following and some authors have reported widespread implementation within clinical treatments provided by physiotherapists (Jull and Richardson, 1994; O'Sullivan et al., 1998). Possibly due to the lack of a research tradition and evidence-based practice, their use initially appeared to be indiscriminately applied and not critically appraised. Therefore, with the increasing emphasis on evidence-based health care and the need to critically assess both the efficacy and effectiveness of physiotherapeutic interventions, these

concepts and techniques were identified as research priorities by the Chartered Society Of Physiotherapy¹ (CSP, 1997).

As a clinician, the motivation for the design of the study was to ensure that it closely represented clinical practice and that the results could be generalized to daily practice. As a researcher, the motivation was to ensure the incorporation of best practice from available research guidelines and a robust methodological design. Therefore, following extensive review of the literature, it was apparent that a pragmatic design examining the use of spinal stabilization exercises, as part of an overall physiotherapy treatment package, was appropriate in order to provide clinically applicable results and inform future research in a meaningful manner.

¹ **Topic area:** Low back pain (Page 25). **Specific research area:** “Are spinal stabilisation exercises effective in reducing pain and disability?” (Page 26)

2. Chapter 2: Literature Review

2.1 Introduction

This chapter presents a review of the literature supporting the development and practice of specific exercises designed to re-train the neuromuscular control and endurance of the deep abdominal and trunk muscles. The empirical basis for the techniques and the previous clinical studies are critically appraised. In order to facilitate this, a brief outline of the concept of stabilisation training will be presented next, with greater detail in section 2.1.3 onwards. The chapter also reviews the current state of knowledge regarding treatment of LBP, specifically recurrent LBP, and the outcome measures used in both research and clinical practice.

2.1.1 Outline of spinal stabilisation training concept

In contrast to general aerobic or conditioning training, rehabilitation of active stabilisation is concerned with the coordination of optimal patterns of muscle activity, which provide postural stability in order to allow activities to be undertaken safely and effectively (Jull and Richardson, 1994). The concept is based on the premise of different functional groupings of muscles (Bergmark, 1989) and that the individual or specific contribution of each requires specific rehabilitation. Bergmark's classification introduced the concept of a local muscle system, generally muscles which have their origin or insertion on the vertebrae and which are primarily used to control spinal stiffness through curvature control and thus provide mechanical stiffness, and the global muscle system which are non-segmental, linking the pelvis and the thorax, and are primarily concerned with torque production and movement. It is the local system that stabilisation training initially targets. A new model of functional classification has been reported (Mottram and Comerford, 1998; Comerford and Mottram, 2001b) which expands the original classifications to local stability muscles, global stability muscles and global mobiliser muscles. The training of local stability muscles is based on their physiological properties and function and involves low load, high numbers of repetitions, an emphasis on accuracy and specificity, volitional input (Miller and Medeiros, 1987), individualization of recruitment patterns and facilitation strategies

and finally incorporation into functional activities (Jull and Richardson, 1994; Richardson and Jull, 1995; Norris, 1995; Richardson, 1995; Comerford and Mottram, 2001a). The concept has an increasing evidence base, with both experimental and clinical research undertaken primarily by physiotherapists, much of which is based at the University of Queensland, Brisbane, Australia. A basis of a training programme is presented in Appendix A but the reader is directed to a number of articles for greater detail (Jull and Richardson, 1994; Richardson and Jull, 1995; Norris, 1995; Comerford and Mottram, 2001a).

2.2 The problem of low back pain

The continuing epidemic caused by increasing levels of disability due to low back pain (LBP) has vast socio-economic as well as personal implications (CSAG, 1994a). The current drive for increased evidence based practice and research into treatment efficacy has highlighted the lack of substantive evidence regarding effective treatment interventions for LBP. Although generally accepted that acute LBP has a relatively short natural history and should be a benign, self-limiting condition, the huge problem associated with the high recurrence rate is often underestimated (Waddell, 1987). With one-year recurrence rates estimated between 20-86% (Bergquist-Ullman and Larsson, 1977; Troup et al., 1981; Abenhaim and Suissa, 1987; Von Korff et al., 1993) the need to address this problem and prevent the subsequent disability associated with chronicity is pressing (Klaber Moffett et al., 1995). There has been increasing international attention in the exponential growth, compared to the relatively stable prevalence, in the socio-economic and medical costs associated with LBP, which has inevitably led to a marked increase in the amount and scope of research in the area.

Physiotherapy is often used in the treatment of LBP, with an estimated 9% of patients with back pain visiting physiotherapists in both the NHS and private settings (OCPS, 1997). Despite this, there is little conclusive evidence supporting any specific physiotherapeutic intervention into acute or recurrent LBP (Evans and Richards, 1996). The economic burden caused by LBP within Britain makes the need to provide efficient, cost-effective interventions a research priority. An estimated £36 million was spent on physiotherapy within both the UK public and private sectors in the year 1992/3 (Klaber Moffett et al., 1995). More recently, an

analysis of the economic burden of LBP in the UK estimated that of the £1632 million direct health care costs, 37% related to physiotherapy and care from allied specialists (Maniadakis and Gray, 2000).

The scope of physiotherapy in LBP encompasses many common treatment modalities that can be broadly split into “passive” modalities; heat, mobilization and manipulation, massage, traction and electrotherapy and “active” modalities; various exercise regimes and education (Foster et al., 1999). So-called passive treatment modalities of mobilization and manipulation (often termed as ‘manual therapy’) encompass a number of different approaches such as Maitland mobilizations (Maitland, 1986), McKenzie (McKenzie, 1981), Cyriax (Cyriax, 1984), and Kaltenborn (Kaltenborn, 1970). One recent development in the active treatment modalities has been the increased use of low load, high repetition training of the abdominal and trunk muscles, so called specific or segmental stabilisation training or muscle imbalance techniques. This has been a response to the recent research that has demonstrated that neuromuscular dysfunction and fatigue of the back and abdominal muscles exists within the back pain population (Richardson et al., 1995). Early small scale clinical trials have suggested that addressing this dysfunction improves both objective and subjective outcomes of treatment in certain sub-groups of the LBP population i.e. prolapsed intervertebral disc, patients with radiological evidence of instability and acute, first-episode LBP (Saal and Saal, 1989; O'Sullivan et al., 1997c; Hides et al., 2001). Despite these initial favourable results, the effect of incorporating these exercises within treatment packages that are currently available, and in a population with recurrent LBP has not been investigated.

The CSAG guidelines stressed the need to prevent chronic disability and advocated an active rehabilitative approach for simple LBP (CSAG, 1994a; CSAG, 1994b). Although clinically the wide variety of active exercise regimes available are often assumed to alter the natural history of LBP, the high recurrence rate possibly belies an underlying aspect of dysfunction that has not, as yet, been addressed. It is hypothesised that spinal stabilisation exercises may influence this muscular dysfunction. The specific nature of ‘traditional’ prescribed exercise regimes and their empirical and physiological basis has been called into question (Richardson and

Jull, 1995). The issue of recruitment and timing of onset of the deep stabilizing abdominal muscles is now being investigated and explored in addition to pure strength indexes (Hodges and Richardson, 1998; Gardner-Morse and Stokes, 1998). Increasingly, the stabilizing function of deep-seated muscles is seen as essential for the correct functioning of specific joint complexes. Similarly it has been suggested that this more centrally controlled component of stabilisation appears to play a pivotal role in the dysfunctional state within LBP subjects (Hodges and Richardson, 1997b). The evidence base and research in this area will be discussed in detail in Section 2.13 onwards.

The next two sections introduce the concepts of classification systems and the associated problems with their development and use. They also cover the issues of recurrence in LBP, definitions and the problems inherent in its measurement.

2.3 Classification of low back pain

Despite the obvious benefits of a universal classification system for LBP that is equally applicable to both the clinical and research setting, to date no such system is available. The recurrent, varying nature of LBP makes the formation of any universally accepted, validated classification system a continuing goal within back pain treatment and research (Foster et al., 1999). Numerous classification systems exist based on many variables such as extent of pain referral (Abenhaim and Suissa, 1987), suspected pathology, functional or physical limitation and behaviour of symptoms (McKenzie, 1981).

Support varies greatly for the different classification systems. Rothstein (1993) reports that classification based on pathology alone is an inadequate guide for health care professionals and that, in order to achieve effective interventions, patients must be classified according to relevant variables i.e. extent of pain or movement restriction (Rothstein, 1993). Similarly it has been suggested that classification systems based on presumed pathological processes are of limited use (Riddle, 1998). Certainly, as there is evidence that clinicians have demonstrated poor inter-tester reliability in diagnosing spinal pathologies (Donahue et al., 1996), presumed pathology would seem a poor basis on which to base a classification system.

It has been suggested that classifying patients with conditions such as LBP aids the development of appropriate treatment guidelines and promotes clinical effectiveness. In addition, the use of a standardised, universally accepted classification scheme would facilitate both intra- and inter-disciplinary communication. Development of classification systems is undertaken using both statistical and judgemental approaches. The former identifies variables that can distinguish sub-groups of patients and identify homogeneous groups with similar risk of poor outcome, psychological involvement or levels of physical impairment (Main et al., 1992; Coste et al., 1992). Any classification system must demonstrate clinical utility through which clinically useful inferences are possible, based on patient subgroups. This has been attempted in national and international guidelines on the management of LBP identifying a gross classification of LBP along triaging lines, namely simple or non-specific LBP, nerve root pain and suspected serious spinal pathology (Page 54-55 CSAG, 1994a). The latter, termed 'Red Flags', were a group of diagnostic indicators that were suggestive of possible serious spinal pathology of a non-mechanical nature. Within the diagnostic triage and management system suggested within the CSAG and subsequent guidelines, the presence of red flags should trigger an urgent/emergency referral for specialist investigations. It is recommended that patients with nerve root problems, if not resolved within 4-6 weeks, with or without physical therapy, are referred for a surgical opinion. But the suggested management of the broadest, and realistically most difficult group to manage i.e. simple, non-specific LBP, may be considered somewhat generic in recommending referral for psychosocial and vocational assessment and rehabilitation for return to work if not resolved within 4-6 weeks. Although the CSAG classification system has the requirements of simplicity, facilitates the identification of patients with potentially serious causes of LBP, and assists with management decisions, it does not allow or aid with the sub-division of the largest group, namely patients with 'simple, non-mechanical LBP'.

Some specific problems exist regarding classification systems for LBP in relation to research. It has been suggested that classification of patients with non-specific LBP is favourable prior to randomisation within randomised controlled trials

(RCT) in order to maximise possible treatment effects (Deyo and Phillips, 1996). Along with many other authors, Moffroid et al. (1994) maintain that the appropriate classification of LBP is the preliminary step in defining and managing the condition. In addition they identify the detrimental effect that the inability to randomize to truly homogeneous groups has on research into LBP treatment (Moffroid et al., 1994). Although one universally accepted system is not currently available, preliminary evidence suggests that patients classified using a system designed to guide treatment may be more effectively treated than those who are not classified (Erhard et al., 1994; Delitto et al., 1995).

A classification system based on expert consensus opinion and utilising extent of referral of symptoms, The Quebec Task Force on Spinal Disorder (QTF) (Spitzer et al., 1987) provided a universal classification for LBP (See Table 2-1). This group proposed an *a priori* classification, based on simple clinical criteria, which provided mutually exclusive categories. Expert consensus and a review of the literature were used in the development of the system, with two additional axes incorporating symptom duration and work status, as these factors had previously been identified as influencing the outcome of treatment. Designed as a hierarchical scale, the mutuality of the categories in practice fails i.e. patients with chronic pain syndrome (category 10) also reporting radiating leg pain (category 3) can be placed in two possible categories (Riddle, 1998).

In a review of eight classification systems, including the QTF, developed to classify sub-acute LBP and guide choice of physiotherapeutic treatments, Peterson et al. (1999) examined the content, face and construct validity², feasibility and generalizability of use (Petersen et al., 1999). They concluded that the QTF only partially met the content validity and feasibility criteria and failed to meet the face and construct validity criteria. They did, however, acknowledge the generalizability of the classification system, citing its use as a standard by which other classification

² **Content validity** relates to the extent to which a tool taps the full domain of content

Face validity relates to the extent to which a tool appears to be a valid representation of what it purports to represent

Construct validity relates to how a scale or index correlates with measures of other variables that are predicted by a theory of how the variables are related

systems are measured and suggesting its use in various settings. In agreement with most other authors in the field, Petersen et al. (1999) concluded that no classification system fulfilled all requirements, suggesting that future work should concentrate on evaluating existing classification systems capable of meeting the basic measurement criteria outlined above but also developing new classification systems (Riddle, 1998; Petersen et al., 1999).

Table 2-1: Classifications one to four of the Quebec Task Force Classifications of activity related spinal disorders

1	Pain in the lumbar, dorsal or cervical areas, without radiation below the gluteal fold or beyond the shoulder respectively, and in the absence of neurological signs Category believed to represents most cases. The pain is intermittent or constant, its intensity varying with the patient's tolerance, and is almost always aggravated by mechanical factors
2	Pain in the lumbar, dorsal or cervical areas, with radiation distally (i.e. to the upper or lower limb but not beyond the knee or elbow, respectively) and not accompanied by neurological signs In this category, the pain that radiates to the distal part of the limb can be neurogenic, but it originates most often from the deep structures of the rachis.
3	Pain in the lumbar, dorsal or cervical areas, with radiation distally (i.e. beyond the knee or elbow respectively) and without neurological signs Pain that radiates to whole limb may occupy a specific dermatome, suggesting a radicular origin, or it may be more diffuse suggesting vascular or somatic
4	Pain in the lumbar, dorsal or cervical areas, with radiation to a limb with the presence of neurological signs (e.g. focal muscular weakness, asymmetry of reflexes, sensory loss in a dermatome, or specific loss of intestinal, bladder or sexual function) Includes radicular syndromes
Taken from: Quebec Task Force on Spinal Disorders (1987)	

In summary, although a plethora of classification systems exist, there is little consensus regarding their utility. In practice, it may be that, different classification systems are suited to different settings and applications. However, what is essential is that any system used yields meaningful classifications.

2.4 Recurrence

Chronic and recurrent LBP is a pervasive and costly problem with recurrence rates reported at between 20-83% (Bergquist-Ullman and Larsson, 1977; Troup et al., 1981; Abenhaim and Suissa, 1987; Von Korff et al., 1993). Although there is evidence to suggest that acute LBP runs a natural course, the high recurrence rate suggests that targeting effective interventions to this group may reduce the potentially high costs associated with recurrent back pain (Klaber Moffett et al., 1995; Waddell, 1998). Controversy exists regarding both the recurrence rate, and the natural history of recurrent LBP, although some of the variance within the

reported recurrence rates may be an artefact caused by differing definitions of 'recurrence'. CSAG report (1994 p 11) suggests that 70% of people who experience back pain will suffer three or more recurrences with recurrences tending to settle over several years (CSAG, 1994a). Similarly, Biering-Sorensen (1983) demonstrated that the likelihood of recurrence decreases with increasing time since the last attack (Biering-Sorensen, 1983). Conversely, Troup et al. (1981) reported that the risk of recurrence increases with the number of previous attacks (Troup et al., 1981) whilst other authors have reported 89% of patients experience at least one recurrence of LBP (Caldwell and Galanville, 1993). Some of these discrepancies may be explained by the differing definitions of recurrence, where stated, that were used e.g. Biering-Sorensen (1983) used a broad definition of any 'pain' or 'trouble' in the lower back whilst Troup (1981) used sickness-absence or treatment. Regardless, little consensus exists regarding the factors that accurately predict future recurrences of LBP. The QTF (1987) identified that recurrent episodes do not justify a different treatment approach from acute episodes, but that relapses and recurrences should suggest continuing exposure to risk factors or untreated components of dysfunction. This concurs with the hypothesis suggested by Sahrman (1993) that repetitive micro-trauma can be causative of much musculoskeletal pain (Sahrman, 1993) (See Figure 2-5). It also fits well with the concepts underlying the muscle imbalance approach that disruption with the normal functioning of supportive muscles can lead to the continuance and recurrence of pre-existing problems.

The difficulties associated with establishing universally accepted definitions for comparative purposes are widely accepted (Von Korff, 1994; Bouter et al., 1998; Leboeuf-Yde and Kyvik, 1998; Loney and Stratford, 1999) and defining recurrent LBP remains problematic within the research literature. A number of criteria have been suggested such as the level of symptoms, which prevents a person from undertaking their normal activities (Richardson et al., 1992; Jull et al., 1993; Hodges et al., 1996), number of days in pain is present in a set timescale (Von Korff, 1994; Bouter et al., 1998), treatment-seeking behaviour or previous episodes of back pain (Waddell, 1998). The advantages and disadvantages can be argued over for all definitions but methods relying on the number of days in pain or numbers of previous attacks are subject to recall bias, in addition to blurring the distinction

between pain and function (Von Korff, 1994; Bouter et al., 1998). Treatment seeking behaviour is intrinsically linked with factors other than simple pathology and function and therefore may cast doubt on this as a universal definition for recurrence (Waxman et al., 1998).

A three-level classification scheme for LBP chronicity was proposed by Von Korff (1994) using transient, recurrent or chronic, based on reported number of days of pain (Von Korff, 1994). Recurrent is defined as pain on fewer than half of the reporting days and chronic as pain on more than half of the reporting days. Bouter et al. (1998) proposed a similar scheme, suggesting that recurrent pain is defined as multiple episodes where pain is present on less than half of the days in a 12-month period (Bouter et al., 1998). In a more traditional, time-based concept, Waddell (1998 p 73) suggests the definition of recurrence as having experienced previous attacks with attacks lasting less than 3 months (Waddell, 1998). The problem with all of these definitions, along with the obvious issues surrounding recall bias, is the over-simplification of complex presentations into mutually exclusive categories of 'best-fit', leading to heterogeneous groupings with all the associated problems such groups bring to the design of a randomised trial.

The potential hazards of using care seeking as a measure of recurrence are outlined clearly by a study undertaken by Croft and co-workers (1998). For many years LBP has been considered a benign, self-limiting condition where typically 90% of cases resolve within 6 weeks (Waddell, 1987), but more recent research has strongly challenged this traditional view (Croft et al., 1998). In a study of 490 adults consulting their general practitioner (GP) for LBP over 12 months, Croft et al. (1998) reported that of the 463 consulting with a new episode, 59% had a single consultation while 32% had repeat consultations within 3 months (Croft et al., 1998). Using a visual analogue score for 'pain'³ and the Hanover back pain daily activity schedule which represents ease of performing 12 activities of daily living for 'disability'⁴, three categories were used to classify patients i.e. No pain or disability, pain or disability (but not both) and both pain and disability. At 3 month

³ Pain = VAS of 0 or 1

⁴ Hanover score 0 -100% (full to no restriction). Disability >90%. No disability ≤90%

interview of 218 patients who consented to participate in a cross-sectional survey (212 complete data sets), undertaken by a research nurse, 21% (39/188) had no residual pain or disability and 25% (42/170) at 12 months. Although based on GP consultations, and therefore only identifying treatment seeking, and using arbitrary cut off points for classification, these findings clearly refute the idea of a 90% resolution rate, suggesting that the majority of patients are not symptom free, but simply ceasing to seek treatment from their GP. Potential bias was identified by Croft et al. (1998) as only patients prospectively consenting to participate in the cross-sectional survey prior to the onset of any LBP were interviewed at 3 months, however, similar results were obtained from a sample of non-participants with an overall recovery rate of 27%, therefore the results can be viewed with a reasonably high degree of confidence. These findings, and those from other researchers demonstrating that reason for consultation varies with duration of symptoms (Waxman et al., 1998), support the case for not using treatment seeking alone as a measure of recurrence or prevalence.

The problem in defining recurrence is a major methodological issue in research. The York report (Klaber Moffett et al., 1995) on the economic cost of LBP strongly supported secondary prevention aimed at reducing recurrences and the need to establish which programmes and specifically which components are effective in the management of LBP by physiotherapy, a sentiment re-enforced by the 2nd International Forum for Low Back Pain Research (Deyo et al., 1998). However, a number of problems are associated with this type of research. Questionnaires and surveys suffer from problems associated with recall of recurrence and pain states, as discussed previously, and therefore introduces potential sources of inaccuracy into the research process. Power analysis of the necessary sample size is difficult when there is little evidence available on which to base these calculations. With the increased emphasis on longer-term follow-up (Von Korff, 1994), longer periods of recall are required, potentially increasing inaccuracies. Thus the dividing line between new episodes, recurrences and continuing dysfunction also causes methodological problems. One potential solution to this is prospective studies with baseline measurements and repeated follow-ups at short regular intervals over longer periods.

In summary, as with classification systems for LBP, at present there is no one accepted definition of recurrence. By necessity, studies to date specifically looking at recurrent LBP, tend to adopt one of the many definitions and relate findings to other studies using the same definition. In the absence of consensus therefore, the pragmatic view may be to follow the definition suggested of alteration of activity or treatment seeking, which is both easily applied and clinically meaningful.

The next section covers issues surrounding the measurement of outcome of treatment interventions for LBP and reviews and evaluates a number of systems and measures currently in use.

2.5 Outcome measures

The ultimate aim of any healthcare intervention is an improvement in the condition or symptoms but, traditionally, the only “measure” of success or failure has been subjective reports or arbitrary pronouncements by healthcare professionals (Deyo et al., 1994). However, increasing pressure to demonstrate ‘objective’ improvement of interventions and satisfy growing patient expectations has led to the development of numerous outcome measures, many of which have subsequently been validated for use in specific conditions. Outcomes can assess many different facets of the healthcare process and associated disease status such as quality of life, physical, social or mental functioning, satisfaction with health care interventions or the process of healthcare. They can be disease-specific or generic and measure a single aspect of the disease process (uni-dimensional) or multiple aspects (multi-dimensional). For an outcome measure to be useful, it must possess certain fundamental properties i.e. reliability, validity, responsiveness and practicality of use (Deyo, 1988). It should be meaningful to patients, essentially measuring health-related quality of life issues (Beattie and Maher, 1997). Specific socio-economic indicators of successful treatment such as return to work are certainly useful and provide a quantifiable endpoint. However problems are inherent with such measures as achievement of such endpoints e.g. return to work, are often multifactorial (Linton and Hallden, 1998). The often obscure aetiology of LBP contributes to the numerous problems with outcomes designed to measure LBP. The diversity of outcome measures utilised within the back pain literature has been identified as problematic in comparison of results and extrapolation of findings and

has led to the recommendation of a core set of measures (Deyo et al., 1998). These include bothersomeness of symptoms, back-related function, generic well-being, disability (social role) and satisfaction with care. The aim of this core set was not to limit the areas that individual researchers can examine but to provide a basis on which to build so that all research can be more usefully compared. These properties and evidence for use of the proposed measures will be discussed in the following sections.

2.6 Measuring functional status in low back pain

Traditionally physical measures which have been utilised in the clinical setting, such as the modified Schober's test, have been used as an indirect assessment of function but increasingly evidence from the literature suggests physical measures correlate very poorly with physical functioning (Waddell et al., 1980; Ohnmeiss et al., 2000; Cox et al., 2000) and consequently there has been a move toward more functionally based measures of back pain related disability. This section discusses and summarizes this situation.

2.6.1 Physical measures versus self-reported measures

Despite extensive research examining the use of reliable, simple objective clinical measures in LBP subjects, these are still the source of much debate. One reason for the increasing rejection of physical measures as outcomes in the treatment and research of LBP is the underlying basis for their use. The conventional biomedical model of disease, relying on the identification of a specific pathology on which to base subsequent treatments and assessment of limitation/function, are fundamentally flawed in light of the evidence that the majority of LBP cannot be definitively diagnosed (Waddell, 1998). This convention, maintained that the worse the underlying pathology, the worse the associated impairment, limitation and ultimately disability. However, as the majority of LBP is considered non-specific and does not receive an exact diagnosis (Evans and Richards, 1996), this biomedical model appears to be inappropriate. It is now widely accepted that patients' perspective of level of functioning and results of treatment is essential in the assessment of health care interventions (Guyatt et al., 1993).

Certainly there is evidence that 'objective' measurements suitable for use in the clinical setting can be relatively reliable as evidenced by Tillotson and Burton (1991). These authors examined the use of a flexicurve to measure sagittal spinal mobility and found it to be relatively reliable and a close approximation to radiographs of the upper and lower lumbar spine (Tillotson and Burton, 1991). Although this work identified the reliability of the flexicurve as a physical measure, no support was given for its clinical usefulness. Later, investigating the use of physical measures in the clinical setting, 344 subjects complaining of LBP and 118 subjects who denied ever having experienced LBP were assessed (Thomas et al., 1998). The authors reported a statistically significant reduction in all planes of movement in LBP subjects measuring side flexion, flexion (using a Modified Schober's method and fingers-to-floor measurement), spinal extension and knee extension. More clinically relevant, however, was that restriction in three or more planes occurred in 50% of back pain subjects as compared with 3% of those with no back pain. When attempting to translate these findings into useful outcome measures however, a number of difficulties arise. Firstly, reproducibility of physical measures is often problematic and thus can be a source of error. Secondly, physical measurements do not necessarily correlate with functional ability and thirdly, physical measures can be highly influenced by a multitude of non-physical factors such as psychology, fear, motivation and mood (Cox et al., 2000). Cox et al (2000) examined the correlations between patient self-assessment and simple functional measures such as ROM and higher order spinal coordination patterns, such as range of lordosis and estimated segmental mobility, in 91 non-acute (pain for at least 10 weeks) LBP subjects. They reported a high correlation between self-assessment and simple functional measures and no correlation between self-assessment and complex spinal tasks. Cox et al. (2000) concluded that higher-order coordination patterns of the spine are predominantly independent of self-assessment and may be a useful adjunct to clinical assessment, whereas simple parameters of functional examination are strongly correlated with cognitive state and hence are not a useful indication of functional status in LBP patients. Accordingly, in light of their findings, Cox et al. (2000) questioned the role of self-assessment questionnaires for LBP. However, measurements of this type require specialist, expensive equipment rarely available in the typical clinical environment and therefore the inclusion of such measurements in a pragmatic examination of

packages of care would be inappropriate and make generalizability of results to the clinical setting difficult.

Some attempts to distinguish specific groups of LBP on the basis of physical measures in combination with self-reported pain, functional levels and demographic data, have indicated that symmetry, flexibility, strength and dynamic mobility are distinguishing characteristics of specific homogeneous groups (Moffroid et al., 1994). Moffroid et al. (1994) studied 115 LBP patients and 112 matched controls using the 53 component items (25 test items) comprising the National Institute of Occupational Safety and Health (NIOSH) Low Back Atlas (Moffroid et al., 1992) and self-assessment questionnaires, including measures of job satisfaction, physical activity level, pain, functional capacity and psychological health. Using the 24 composite variables and cluster analysis 4 patient categories (fit, unfit, flexible and inflexible) and 5 clusters in the control group. They reported that the physical tests used were reliable but development of this work has not been undertaken and other authors have suggested that the clinical utility of the clusters is limited due to the lack of inclusion of the influence of chronicity and pain behaviour (Riddle, 1998) and overlap between the clusters (Waddell, 1998). The relationship between self-report and physical measures has recently been explored in a study that assessed both these factors to examine the effects of active therapy for chronic LBP (Mannion et al., 2001b; Mannion et al., 2001a). Using stepwise linear regression, lumbar range of motion was shown to be a poor predictor of baseline disability with pain and psychological distress at baseline being the strongest indicators. However, the use of both types of measures, do allow scope to analyse how the two co-vary which has potential to inform both future research and clinical decision-making.

Similarly, strong arguments can be made for the use of both mechanical measures, such as the Biering-Sørensen fatigue test (Beiring-Sorensen, 1984), and EMG measures such as the median frequency change, of back muscle function and endurance as a physical parameters to assess the outcome of LBP treatment. The Biering-Sørensen fatigue test is well accepted by patients and clinicians and endurance times have been shown to correlate well with EMG median frequency decline in healthy subjects (Mannion and Dolan, 1994) and have good test-retest

reliability (Dolan et al., 1995). However the Biering-Sørensen fatigue test is also influenced by psychological factors and motivation (Mannion et al., 1996; Mannion et al., 2001b) and therefore clinical use in isolation may be of limited use. More encouragingly, is the finding that a number of studies have reported that these measures reflect closely self-reported measures of function, and this may be a useful way forward in future clinical trials.

Recently, there has been increased interest in the use of the shuttle-walking test (SWT) as an adjunct to outcome measures both in the clinical setting and in studies of LBP patients (Frost et al., 1995; Fairbank, 2000; Goldby et al., 2000). Initially developed for use with respiratory patients (Singh et al., 1992), the SWT has been used as an indicator of physical functioning in patients with spinal stenosis and found to be acceptable for group analysis, although the limitation of providing a 'snapshot' only was acknowledged with the suggestion that multiple testing would improve sensitivity (Fairbank, 2000). Suggested to be a more valid measure of functional ability than range of movement, Frost et al. (1995) found the SWT to be responsive to change following a progressive fitness programme for chronic LBP patients. However, the SWT measures only one aspect of function, albeit an important one, and therefore requires supplementary information regarding functioning from other sources. Additionally, in the planning of follow-up studies, it should be considered that any outcome that requires additional attendances of participants will have an impact on the follow-up rates and consequently sample size and therefore, the inclusion of such outcomes must be carefully considered.

In summary, although traditionally physical measurements (e.g. spinal range of motion) have been considered a good gauge of improvement or deterioration in LBP, there is increasing evidence that their clinical utility is limited and correlates poorly with functional disability levels. More recently investigations into various EMG measures of spinal muscle parameters have shown encouraging results but are primarily laboratory based at present and are not as yet available routinely in the clinical setting. Consequently, the development of disease-specific, functional, self-report measures designed to quantify the impact of LBP has occurred. Issues surrounding the use of some of these measurement tools will be discussed in the

next section in relation to both clinical practice and research methodology. Issues of responsiveness to change and the identification and meaning of clinically meaningful change will also be discussed.

There is now a plethora of self-report generic and disease or condition specific questionnaires available, which purport to 'measure' functional status in LBP patients. These are self-completed and can be used in the clinical and research settings and include for example, the Oswestry Disability Index (ODI) (Fairbank et al., 1980), the Roland-Morris Disability Questionnaire (RMDQ) (Roland and Morris, 1983b; Roland and Morris, 1983a), Low Back Pain Outcome Score (Greenough and Fraser, 1992), Aberdeen Low Back Pain Questionnaire (Ruta et al., 1994) and the North American Spine Society Lumbar Spine Questionnaire (NASS LSQ) (Daltroy et al., 1996).

2.6.2 Roland Morris Disability Questionnaire

A popular back-pain specific outcome measure, used for both research and clinical decision-making, is the RMDQ (Roland and Morris, 1983a) (See Appendix B). The RMDQ consists of 24 questions abstracted from the Sickness Impact Profile (SIP) (Bergner et al., 1981) with the addition of "because of my back" to each question to improve specificity (Roland and Morris, 1983a). The questionnaire is scored from 0, representing no back-pain related disability, to 24, representing maximum disability. It takes approximately 5 minutes to complete and has a high degree of face validity. With any outcome measure used to assess change over time, one essential property, which should be assessed, is responsiveness. Responsiveness, or sensitivity to change over time, is the ability of an instrument to detect a minimal clinically important change in function as the result of treatment (Cohen, 1977; Deyo, 1988). This can be measured in a number of ways, including statistical tests, effect sizes, Guyatt's responsiveness statistic, receiver operator curves (ROC) and sensitivity co-efficient (Kopac and Esdaile, 1995). Two distinctive properties are important when considering change in any outcome measure;

- **minimal detectable change (MDC)** is reported as the level at which any change in a score truly represents a change in patients and is calculated statistically using the standard error (SE) of the change (Stratford et al., 1998).
- **minimal clinically important difference (MCID)** is the smallest difference in score that a patient perceives as important and is a value judgement (Jaeschke et al., 1989; Stratford et al., 1998).

Responsiveness of the RMDQ has been well examined in numerous studies (Kopeck and Esdaile, 1995; Patrick et al., 1995; Beurskens et al., 1996; Stratford et al., 1996a; Stratford et al., 1998; Riddle et al., 1998; Stratford and Binkley, 1999). Although initially a change score of 2-3 on the RMDQ was suggested as the MCID (Roland and Morris, 1983b) later studies by Stratford et al. (1998), concluded that individuals MCID was dependent on a patient's initial score. A change score of 4-5 points (90-95% CI) was identified as the MDC for initial scores between 7 and 19 calculated using the standard error of measurement, or within patient variability (Stratford et al., 1996b). However, this assumes a constant SE throughout the range of a scale when in reality the error may depend on where the score falls within that scale. Subsequently the MDC was calculated using the conditional standard error of measurement (CSEM), which acknowledges that the SE is conditional on the score of interest (Stratford et al., 1996b). This work identified a MDC of 4-5 points (90% CI) using CSEMs, for initial scores between 4 and 20, but that improvement in patients with initial scores of under 4 and deterioration in patients with scores of over 20 could not be identified with a high degree of confidence. This work and identification of MDC strengthened the use of the RMDQ as both a research and clinical outcome measure and was developed further in 1998 by two studies looking at sensitivity to change of the RMDQ and whether sensitivity to change was dependent on initial score (Stratford et al., 1998; Riddle et al., 1998).

In a study of 226 LBP patients, Stratford et al. (1998) investigated the RMDQ change score which best-classified patients who had achieved an important change

and whether the estimate of change is dependent on initial scores. Using receiver operating characteristics curves⁵ (ROC curves) as a measure of sensitivity and specificity to change and a 15-point global rating of change from deterioration to improvement, they concluded that for the entire scale, a change of 5 RMDQ points represented the MCID at the 90% CI. However, better levels of sensitivity and specificity were obtained for the MCID when the initial RMDQ score was taken into account. The finding that higher initial scores required greater change in order for a meaningful change to be presumed seemed intuitively correct and further supported the use of the RMDQ in both the clinical and research setting. The findings of Stratford et al. (1998) however, applied to changes in scores across large groups of patients and, as such, were of limited use in clinical decision-making regarding individual patients. Riddle et al. (1998) addressed this issue in an accompanying piece of research using a similar methodology but assessing sensitivity to change in individual patients against whether or not therapist generated goals were achieved. They found that slightly higher change scores were required to be confident that a clinically significant change had been achieved but identified that methodological issues might have been responsible for these differences. The specific MCIDs identified in both pieces of work for overlapping and the mutually exclusive bands of the RMDQ are summarised in Table 2-2.

Table 2-2: Minimal clinically important difference for Roland Morris Disability Questionnaire scores

Minimal clinically important difference			
Initial score	(Stratford et al., 1998)	(Riddle et al., 1998)	Combined *
0-8	2	3	2-3
5-12	4	5	
9-16	5	8	5-8
13-20	8	11	
17-24	8	13	8-13
*Represent minimum and maximum MCID as advocated by Riddle et al. (1998)			
Figures in shaded rows represent overlapping bands of initial scores used in analysis			

⁵ ROC curves are used to identify the best 'cut-off' point for a test or score. They are created by plotting sensitivity versus 1-specificity for each cut-off and joining the points.

More recently, however, in a comment regarding customary RMDQ discharge scores, Stratford and co-workers (1999) suggested a RMDQ score of 4 or less was typical in patients classified as 'successes' and following 12-weeks natural history of LBP. Successful outcomes were patients who had either achieved agreed set treatment goals or undergone a measurable important clinical change (Stratford and Binkley, 1999).

The issue surrounding what difference in RMDQ scores should be considered as an important change in group comparisons has been discussed in a number of articles. The current advice given by Roland and Fairbank (2000) is that sample size calculations should be based on change scores of 2-3 points in order to avoid under-powering trials (Roland and Fairbank, 2000). However, no justification is given for this statement and it is difficult to rationalise as a change of 2-3 points will only represent a clinically significant change in patients who's initial scores are less than 9 points (Stratford et al., 1998; Riddle et al., 1998). In addition, it would appear important to consider the types of patients involved in the study as well as the setting e.g. primary care versus secondary care and acute versus chronic LBP etc. Serious consideration is also required regarding methodological design and MCID. In the situation where a MCID is achieved within groups, a between group difference of 2-3 points may well be significant, however, as yet conclusive evidence to support this is lacking.

The identification of minimum change score that is of clinical significance allows effect size estimations and power calculations to be more easily performed. However, a recent review by Bombardier et al. (2001) has highlighted the issues surrounding the measurement of change and identified that the concept of meaningfulness is very much context based. Factors affecting what is considered an important change will include whether change between groups or individuals is measured, between or within individuals and what type of change is measured (Beaton, 2000; Bombardier et al., 2001). For use in RCT designs, Bombardier (2001) identifies three main types of change within groups that can be measured: differences in the observed changes between the groups, change found in a single cohort and change observed in those estimated to have an important change. In conclusion, identifying the varying responsiveness of the RMDQ quoted in the

literature (from 2-8 dependent on what change is being measured), she concludes that the trialists should refer to published studies to find a match when planning future trials.

2.6.3 Oswestry Disability Index

Another widely used disease-specific functional questionnaire is the Oswestry Disability Index (ODI) (Fairbank et al., 1980) (See Appendix C). The ODI consists of ten sections, with six graded responses covering a number of areas including pain intensity, hygiene, lifting, sleeping and walking (See Table 2-3). Each section has 6 graded responses from 0 to 5, with zero indicating no limitation and five the maximum limitation in each of the ten sections. The total score obtained is doubled and a percentage disability level is calculated for each respondent. Fairbank et al. (1980) found the ODI to be a valid indicator of disability caused by LBP, as it mirrors observed disability but it has been suggested that the inclusion of the section on pain intensity deviates from the true definition of a disability measure *per se* (Fisher and Johnston, 1997).

Table 2-3: Activities covered by Oswestry Disability Index

Section	Area
1	Pain intensity
2	Personal care
3	Lifting
4	Walking
5	Sitting
6	Standing
7	Sleeping
8	Sex life
9	Social activities
10	Travelling

2.6.4 Meaningful change

No definitive MCID has been identified for the ODI and it has recently been suggested that more work is needed in the area (Roland and Fairbank, 2000). However, the responsiveness of the ODI and RMDQ has been examined by a number of researchers (Beurskens et al., 1996; Stratford et al., 1996a; Bombardier et al., 2001). Beurskens (1996) examined the responsiveness of the ODI, RMDQ, and visual analogue scale (VAS) for pain (average severity during the last week) and VAS for the main complaint. The latter was defined as the main activity, identified by the patient, with which their LBP interferes, scored on a VAS. Global perceived effect of treatment was recorded by self-assessment on a 7-point scale (1= completely recovered through to 7= vastly worse). Effect size statistics as described by Cohen (1977) and ROCS were used to calculate the responsiveness of each instrument. Beurskens et al. (1996) concluded that each instrument could discriminate between improved and non-improved patients, with the RMDQ showing the best discrimination (0.93 and 0.91 ROC) and the ODI being the most specific to change according to the effect statistic. For their study, using the assumption that false negatives⁶ and false positives⁷ were of equal importance, they calculated what score changes which represented the best cut-off points between improved and non-improved patients (See Table 2-4).

Table 2-4: Suggested change scores discriminating between improved and non-improved patient outcome

Outcome measure	Cut-Off Points
ODI	2-3 points (4-6%)
RMDQ	2.5-5 points
VAS (Pain)	10-18mm
VAS (Main complaint)	18-24mm
Source: Beurskens et al. (1996)	

⁶ Failing to identify the presence of the property of interest when it is present i.e. in this case incorrectly identifying a patient as having deteriorated

Issues surrounding floor and ceiling effects⁸ in relation to both the ODI and RMDQ have been investigated (Deyo, 1988; Kopec and Esdaile, 1995). It has been suggested that at higher levels of disability, the ODI may still be responsive to change, with the RMDQ more suitable for lower levels of disability (Roland and Fairbank, 2000). The limited range of problems addressed in the RMDQ and the lack of any specific measure of psychological or social issues has been suggested as both a weakness and strength of the RMDQ (Roland and Fairbank, 2000). However, it has been strongly advocated in both clinical (RCGP, 1996b) and research situations (Deyo et al., 1998) that assessments of both functional and psychological factors are undertaken where appropriate. Although purported to be primarily disability scales, both the ODI and RMDQ include questions about pain, independent of activity and consequently it has been argued that they should be considered as measures of both pain and disability (Delitto, 1994).

Evaluation of both the ODI and RMDQ compared to other measures of function has been undertaken by a number of researchers (Triano et al., 1993; Beurskens et al., 1996; Taylor et al., 1999; Garratt et al., 2001). Triano et al. (1993) compared the reliability, validity and change in clinical status over time following treatment between six outcome questionnaires; the ODI, VAS, Modified Somatic Perception Questionnaire (MSPQ), Modified Zung (MZ), Locus of control and Pain Diagram. Outcome measures were taken, pre- and post-assessment and at 6-weeks following assessment from 168 LBP patients. The QTF classification definitions were used to classify location and chronicity of symptoms with the addition of a 'recurrent' category, defined as patients with a symptom-free period 4-6 weeks prior to the current episode, with 6 or more episodes within a year. A forced descriptive classification of back pain was used of entrapment, mechanical, muscular and unclassified. Triano et al. (1993) concluded that there were substantial differences

⁷ Identifying the presence of the property of interest when it is not present i.e. in this case identifying that a patient had improved, when they had in fact deteriorated (Bland, 2000)

⁸ **Ceiling effects** occur where several participants respond at the highest score on a scale, but on a higher scale would respond against a higher score (Sim and Wright, 2000). **Floor effects** represent the same scenario at the lowest scores of a scale.

in the validity and reliability of the questionnaires studied. Only the MZ, ODI and VAS showed stability between pre and post-assessment administration and the latter two were found to be the most reliable and responsive to clinical change. However, Triano et al. (1993) advocates that any outcome measure should be evaluated in the context of unit, staff and patient population with which they will be used.

Modifications have been suggested for both the RMDQ and ODI (Patrick et al., 1995; Daltroy et al., 1996; Underwood et al., 1999). The reliability, validity and acceptability of a modified form of the RMDQ and Von Korff scale, designed to measure function over the preceding four weeks, was investigated by Underwood et al. (1999). Modification was undertaken by rephrasing the introductory paragraph to inquire about pain over the preceding four weeks, using a graduated score for each question from 0-1 (no days or not applicable scored 0, 1 to 7 days scored 0.2, 8-14 days scored 0.4 etc) (Underwood et al., 1999). They suggested that the high level of "Not Applicable" responses on the modified RMDQ made analysis more difficult than for the Modified Von Korff with its 0-10 response options. Interestingly, the time-scale used to measure the disability and pain caused by LBP on both modified questionnaires was completed relating to the previous four weeks, possibly utilizing a more representative time-scale than point measurement for a condition that runs a recurrent relapsing, remitting course. In the study by Underwood et al. (1999), the modified RMDQ correlated better with the physical functioning scale of the SF-36 (Ware and Sherbourne, 1992) than with the pain scale of the SF-36, supporting the view that pain and disability should be measured separately.

In conclusion, there is good consensus from research studies and international expert opinion that both the ODI and RMDQ are valid, reliable, responsive measures of functional ability in patients with LBP. The identification of MCID for the RMDQ may provide an advantage in the research setting, as it allows effect size estimations and power calculations to be more readily performed than for the ODI. Although concerns can be raised regarding the validity and reliability of the extrapolation of results from self-completed questionnaires, the suitability of these measurement tools for both research and clinical purposes is clear. It may be

impossible to confirm that data obtained from questionnaires exactly and categorically represent the level of functioning at a given time. However, the use of such questionnaires is ideally suited to pragmatic research designs where representation of clinical practice is a priority.

The next section discusses the measurement of pain in LBP patients and presents a number of methods commonly used for the purpose in both the clinical and research settings.

2.7 Measuring Pain Status In Low Back Pain

Assessment of pain in all areas of musculoskeletal medicine is problematic and controversial and LBP is no exception. The difficulties associated with the multidimensional nature of the pain experience can confuse the clinical presentation and have been shown to influence the outcome of treatment (Main et al, 1992). Although for much of the 20th century pain was considered primarily as a purely sensory experience, it is now accepted to consist of three dimensions; sensory-discriminative, motivational-affective and cognitive-evaluative, and it is suggested that these are controlled by physiologically specialized systems within the brain (Melzack, 1987). This necessitates measurement tools and strategies, which are sensitive to these different domains. The reasons for attempting to measure pain are numerous including determining pain intensity, quality and duration, aiding diagnosis and choice of therapy and evaluating effectiveness of any intervention (Melzack and Katz, 1994). By definition, any pain rating is subjective and although acute pain may sometimes be proportional to the extent of injury (Melzack and Katz, 1994), in the majority of situations pain does not match any physiologic or pathologic change and controversy exists regarding whether pain scales measure pain or distress (Waddell, 1998).

Many methods exist for the assessment of pain, all of which have the common goal of accurately representing the human pain experience (Price et al., 1983). Similarly, all measurement methods aim to be reliable and valid, measure a specific dimension of pain and yield consistent results over time. The requirements for pain assessment may vary between clinical and research settings, with condition and

chronicity (Sim and Waterfield, 1997) but measurement methods should be versatile, dealing equally well with both the experimental and clinical settings. Intensity, duration, location in addition to bothersomeness and the affective aspects of pain may all be of interest when considering pain measurement.

2.7.1.1 *Visual Analogue Scale and Numerical Rating Scale*

Numerous measurement scales exist for the rating of the sensory-quantitative dimension or intensity of pain (Bolton and Wilkinson, 1998) such as a VAS (Scott and Huskisson, 1976), the 101-point numerical rating scale (NRS-101), the 11-point numerical rating scale (NRS-11) (Ekblom and Hansson, 1988) (See Appendix D), 6-point behavioural rating scale, 4-point verbal rating scale and the 5-point verbal rating scale (VRS-5) (Jensen et al., 1986). Although, controversy exists as to the optimum number of levels needed for pain intensity measurements (Jensen et al., 1994), such scales are popular in both clinical and research settings because they are easily administered, are acceptable to patients and provide an easily interpreted indication of the change in pain intensity. In a study of chronic pain patients, Jensen et al. (1994) found that patients essentially treated a 101-point scale as 21-point scales responding in multiples of 5 or 10. Jensen et al. (1994) advocated that pain intensity measures with 11 or 21-point scales provide sufficient levels for chronic pain patients to describe pain intensity and are adequate for detecting changes in pain intensity within that group.

The VAS is a 10cm horizontal line, which is anchored at each end with the left-hand anchor representing the minimum score and the right hand anchor the maximum (Scott and Huskisson, 1976). Although traditionally accepted to produce a ratio level of measurement it may more correctly be considered an ordinal measurement tool (Sim and Wright, 2000) since there is a debate whether the VAS has a meaningful zero. The use of VAS is widespread throughout clinical practice, as it is simple to use, acceptable to patients and inherently meaningful. Problems of non-completion have been identified (Scott and Huskisson, 1976; Herr and Mobily, 1993), which although not an insurmountable issue in the clinical setting may cause difficulties with research methodologies requiring remote completion of VAS. However, the widespread use of VAS within the clinical setting, makes its inclusion within research methodologies intuitively meaningful to

both clinicians and patients alike, and therefore results from such trials applicable to the clinical setting and significant to clinicians.

The measurement of 'average' pain has been suggested as a valid, practical and more representative measure than 'current' pain for patients with back pain (Bolton, 1999). In a study of 200 back pain patients, four daily recordings, using an 11-point NRS, were taken over a seven-day period (used to compute 'actual average'), followed by an estimation of the least, worst and average ('usual') pain for the previous week. Intra-class correlation coefficient for reliability showed estimates of pain 'on average' to be a reliable measure of 'actual average' pain intensity. The responsiveness of an 11-point NRS, VAS and verbal rating scale (VRS) has also been investigated in relation to 'usual' pain (Bolton and Wilkinson, 1998). Findings indicated that mean usual levels of pain were consistently higher than current pain, with the NRS being the most responsive, and that when 'usual' pain was reported the responsiveness of all measures was enhanced.

The use of both VAS and NRS is widespread in clinical practice and they offer a simple, practical method for the assessment of pain and the impact of interventions. They provide a useful indication of the intensity of pain, but do not provide an insight into the multi-dimensional aspects of pain, which other measures can.

2.7.1.2 *McGill Pain Questionnaire*

The McGill Pain Questionnaire (MPQ) and the more recent short form (SFMPQ) were developed to respond to the need for a multi-dimensional measure of pain (Melzack, 1987). The MPQ is comprised of 78 word descriptors, arranged in 20 lists, measuring sensory, affective and evaluative and miscellaneous dimensions of pain with sensory descriptors depicting the pain experience in terms of temporal, spatial, thermal and pressure properties and affective descriptors depicting pain in terms of fear, tension and autonomic properties (Jenkinson et al, 1995). The MPQ has been well studied and validated but the disadvantage is that it takes approximately 20 minutes to complete and contains complex vocabulary (Dudgeon et al, 1993). The SFMPQ was developed specifically for use in the research setting, in response for the need for increased information in a limited amount of time and,

in contrast, takes between 2-5 minutes to complete, contains 15 word descriptors covering the sensory (n=11) and affective (n=4) pain dimensions, a 100mm VAS and present pain intensity rating (PPI) (Melzack, 1987). The 15 descriptors were chosen from the MPQ on the frequency of use by patients representing a variety of conditions. The addition of 'splitting' was made as it was suggested to be key discriminator for dental pain (Grushka and Sessle, 1984). Present rating index (PRI) is rated from 0 to 3 (none, mild, moderate, severe) (See Appendix E). The PPI is a verbal rating scale of pain intensity, providing a unidimensional and relatively insensitive rank order of the degree of pain sensation, whereas the VAS, although strictly an ordinal measure, is generally considered a ratio scale and highly sensitive to variations in pain intensity (Deschamps et al., 1988) (See below). Scoring for the SFMPQ involves summing the scores, assuming the data to be continuous in nature. This however, is a matter of controversy within the literature as some research has shown this not to be the case, and suggested that the observations are ordinal in nature (Heft and Parker, 1984). Whilst, authors concede that the SFMPQ represents an ordinal scale, they concede that measurements must be interval in order to confer meaningfulness (Wright and Linacre, 1989) and this has therefore become the accepted convention.

The responsiveness of the SFMPQ has been studied in cancer pain and found that the scores reflected changes over time in a similar manner to the original longer version (Dudgeon et al., 1993). Jenkinson et al. (1995), in a study comparing the sensitivity to change of the MPQ, a 10cm VAS and a four-word scale for acute pain (general surgical and orthopaedic procedures), found all to be sensitive to change and suggested that simple measures are worthwhile indicators of the impact of interventions. The SFMPQ appears well suited for assessing pain in LBP. It is brief, easily completed and has a high degree of face validity. Despite the controversy surrounding its measurement level, its use in RCTs provides a quantification of both the different domains of pain, and an indication of intensity.

To conclude, multiple methods exist which purport to measure pain, both in simple quantitative terms and wider multi-dimensional terms. The literature suggests that simple measures of pain, such as the NRS and SFMPQ, are sensitive to change over time and appropriate indicators of the impact on pain of treatment

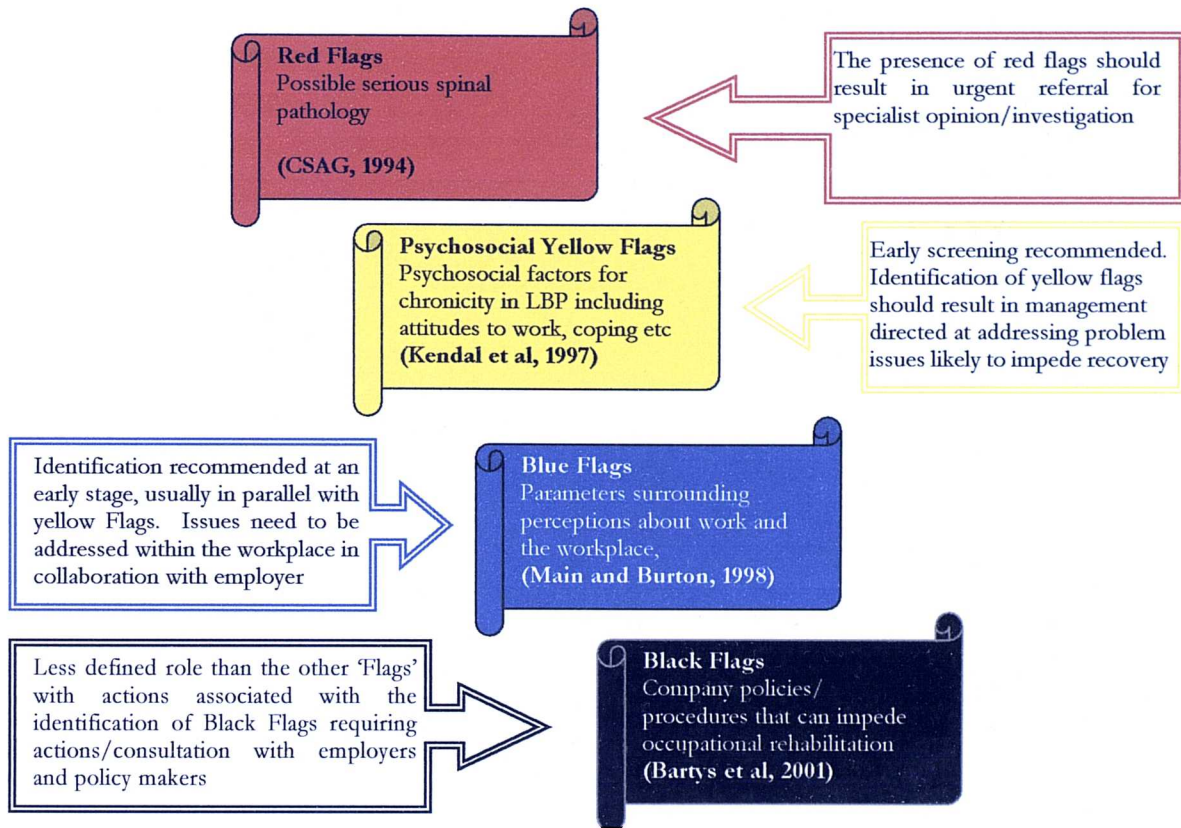
interventions in various conditions (Dudgeon et al., 1993; Jenkinson et al., 1995; Chatman et al., 1997).

2.8 Social and psychosocial factors and low back pain

It is generally accepted that social influences on back pain and disability are both extensive and complex (Waddell, 1998 p 90). After headache and tiredness, LBP is the third most common bodily symptom and therefore has a massive social impact (Waddell, 1998). Many risk factors for the development of LBP have been identified including heavy physical work, frequent bending, lifting, twisting, pulling or pushing, repetitive tasks, static postures and vibrations (van Tulder and Koes, 2001). Additionally, psychosocial risk factors include anxiety, depression, lack of job satisfaction (Papageorgiou et al., 1998) and mental stress at work (Hoogendoorn et al., 2000). In response to the increasing realisation of the importance of psychosocial factors, there was the development and introduction of psychosocial 'yellow flags' for chronicity in LBP (Kendall et al., 1997), along the lines of 'red flags' for potential serious spinal pathology. Yellow flags include a variety of maladaptive beliefs and attitudes regarding LBP, expectations and behaviour (including frequent displays of pain behaviour), reinforcement of pain behaviour by family members, heightened emotional reactivity and distress, job dissatisfaction, poor social support and compensation issues (Kendall et al., 1997). They encompassed both individual psychological parameters, and parameters related to perceptions about work and the workplace. A further concept, that of 'blue flags' arose from the suggestion that the latter i.e. parameters surrounding perceptions about work and the workplace, should be considered as separate and distinct (Main and Burton, 1998). More recently still, following a large-scale trial examining an occupational rehabilitation programme, a further category of 'flags' has been introduced (Bartys et al., 2001). These so called 'black flags' are company policies and procedures that might impede occupational rehabilitation programmes and quite clearly a complicated relationship exists between all these concepts. A summary of the concept of coloured flags in the assessment and management of LBP is presented in Figure 2-1. The development and introduction of the system of coloured flags has helped integrate the knowledge that socio-economic and psychosocial factors affect patients' responses to LBP, and also assist healthcare professionals to identify when liaison with other colleagues may be desirable or

essential. Their introduction, like the now well-established 'red flags', has been undertaken to facilitate the decision-making processes and management of LBP.

Figure 2-1: Summary of coloured flags concepts in assisting with the management of low back pain



The impact of social class on the prevalence and course of LBP has been the cause of much debate in the literature, but although there is conflicting evidence regarding the prevalence of LBP and lower social class, there is fairly consistent evidence of increasing loss of work secondary to LBP with decreasing social class. Similarly, although there is consistent evidence of a higher prevalence of LBP associated with smoking, as smoking varies greatly with social class, education and occupation, it may be that this association reflects a combination of demographic, psychological and life style factors that influence LBP. One pivotal study, which successfully controlled many of the confounding factors associated with research in this area, was that of Battie & Videman (1991), who studied pairs of identical twins with discordant smoking habits (Battie et al., 1991). Reporting an 18% greater mean disc degeneration in the lumbar spines of smokers compared to non-

smokers, this study, with the control of certain key factors such as early environment, education and social class identified a likely systemic effect of smoking on spine health if not the development of LBP. The hypothesised increased risk of LBP that cigarette smoking causes is controversial within the literature (Leboeuf-Yde, 1999; Scott et al., 1999; Goldberg et al., 2000). In a systematic review of 41 articles Leboeuf-Yde (1999) found only a weak association between smoking and LBP, citing 'clearly apparent' associations only in large sample studies. In contrast, just a year later Goldberg and colleagues (2000), following a 'structured' review of the same literature, concluded that the data supported an association between smoking and non-specific LBP. What both sets of authors agree on, however, is the lack of a definitive causal link between smoking and the prevalence or severity of non-specific LBP, the difficulties associated with research in this area and the possibility that findings to date are simply statistical artefacts.

2.8.1 Measuring psychological variables in low back pain

The failure to find an acceptable, effective, treatment based on specific pathology has led to an increasing awareness of a wider range of factors involved in the development, recurrence and associated chronic disability seen in patients with LBP (Waddell et al., 1999). The typical picture of LBP is one of recurring and increasingly disabling episodes, providing evidence that the biomedical approach to LBP has failed (Croft et al., 1998). Increasingly, the universal adoption of a biopsychosocial approach by healthcare professionals is advocated as the only way that LBP and its associated costs and disability might be managed more effectively (Kendall, Linton, Main 1997, Waddell 1998).

As it has also become accepted that psychological factors are intricately linked with the progression and outcome of LBP, the inclusion of psychological screening, either formal or informal, has been advocated within the clinical assessment of LBP (RCGP, 1996b; Kendall et al., 1997). This has led to a psychological profile of subjects entered into research trials becoming more common, as this provides information on how research cohorts compare to wider clinical populations. It has been argued that the identification of subjects at risk of psychological distress or

with pre-existing anxiety and depression should be included in any outcome evaluation and analysis (Main et al., 1992; Deyo and Phillips, 1996).

Multidisciplinary 'best-practice' guidelines for the management of acute LBP introduced the now familiar concept of 'red flags' as indicators of serious spinal pathology, and provided assistance in decision-making for appropriate and timely intervention (Waddell et al., 1999). In order to assist the identification of important biopsychosocial risk factors in acute LBP, Linton and co-workers introduced the concept of 'yellow flags' (Kendall et al., 1997) as discussed previously. Psychological factors have recently been shown to be more closely related to the progression of pain and disability than most biomedical and biomechanical variables and some authors have concluded that psychosocial factors are pivotal in the transition from acute to chronic pain (Linton, 2000; Pincus et al., 2002). The importance of these psychosocial issues has been further acknowledged by the gradual incorporation of the importance of psychosocial assessment into international guidelines, summarised in Table 2-5, and has culminated in the inclusion of 'yellow flags' into the most recent review of the acute back pain guidelines in the UK (RCGP Guidelines, Waddell et al. 1999). Prior to the introduction of the concept of 'yellow-flags', psychosocial issues had been identified as issues in LBP, but the implications of psychosocial and psychological elements of LBP had not been fully explored. Indeed, even with the current increased awareness and emphasis of psychosocial issues in recent clinical guidelines, these mostly relate to the management of acute and sub-acute, rather than chronic, LBP and therefore further implications may yet be revealed.

Table 2-5: Summary of the changing emphasis of psychosocial assessment within back pain management guidelines

Year	Guideline or summary	Element of Psychosocial inclusion
1987	Quebec Task Force on Spinal Diseases, Canada	Psychological issues recognised only as a secondary issue, not relevant in management
1993	Work Cover, South Australia	Psychological assessment appended, untested scale for work loss included
1994	Agency for Health Care Policy and Research (AHCPR), USA (Bigos et al., 1994)	Psychological issues acknowledged and emphasised
1994	Clinical Standards Advisory Group (CSAG), UK (CSAG, 1994a; CSAG, 1994b)	Recommendation to adopt biopsychosocial model Comprehensive assessment recommended at 6 weeks
1995	Pain in the Workplace Task Force (PIW) IASP	
1995	Quebec Task Force on Whiplash Associated Disorders (QWAD), Canada	Multidisciplinary assessment to include Psychological expertise
1996	Accident Rehabilitation & Compensation Insurance Corporation (ACC) and National Health Commission (NHC), New Zealand	Task force to develop NZ guidelines including psychological factors
1996	Royal College General Practitioners (RCGP), UK	Revised CSAG guidelines. Strong recognition that psychological factors are important in chronic LBP and disability. Psychological factors important at an earlier stage than previously considered
1997	ACC and NHC, New Zealand (ACC and the National Health Committee)	Publication of Guide to Assessing Psychosocial Yellow-Flags: Risk Factors for Long-Term Disability and Work Loss
1999	Royal College General Practitioners	Adoption of Psychosocial Yellow-Flags
1999	ACC and NHC, NZ	Updated version based on systematic review of literature since January 1997
2001	SUMMARY REVIEW- An International Comparison: Clinical Guidelines in the Management of LBP in Primary Care (Koes et al., 2001b)	Identified generally good consensus between guidelines from different countries. Consistent recognition of psychosocial factors as risk for chronicity but varying detail and emphasis
Adapted from; (Kendall, 1999)		

Research regarding the prevalence of psychiatric and psychological disorders in musculoskeletal pain provides an important insight into the extent of co-existence of these problems, but fails to provide answers about cause and effect. Coste et al. (1992) in a sample of 330 outpatients with non-specific 'mixed' LBP of varying duration, found that 41% demonstrated a classifiable psychiatric disorder as defined by the DSM-III (Diagnostic and Statistical manual of Mental Disorders). Affective disorders were found in 110 (33.3%) patients, confirmed depression in 91 (27.6%) and an anxiety disorder in 131 (39.7%) patients. Although concluding that the presence of psychiatric disorders did not demonstrate a clear cause and effect, this work is nevertheless informative in the body of knowledge regarding psychological disturbance associated with LBP. Sullivan et al. (1992) found similar results to those of Coste et al. (1992), reporting that major depression was three to four times greater in patients with chronic LBP than in the general population. Subsequently, these authors argued that specifically targeting depressive symptoms may be an integral component of pain management programmes (Sullivan et al., 1992).

The increasing awareness that social and psychological aspects of LBP are intrinsically linked with progression and outcome have led to calls for the early identification of these factors as an important part of the assessment of a LBP patient (Hoogendoorn et al., 2000; Pincus et al., 2002). Although it is now generally accepted that the assessment of both social and psychological factors are important in LBP, what is less clear is how, where, when and by whom this process is best undertaken.

One method for assessing psychosocial and psychological issues in LBP patients is the use of simple questionnaires administered in the clinical setting. Screening and outcome questionnaires have been advocated as an aid to the detection of anxiety and depression and the clinical decision-making process (Wright, 1994). Multiple questionnaires are available for the assessment of psychosocial factors in LBP patients e.g. Fear-avoidance belief questionnaire (Waddell et al., 1993b), the coping strategies questionnaire (Rosenstiel and Keefe, 1983), modified Zung depression index (Zung et al., 1965) and the back belief questionnaire (Symonds et al., 1996). Although all these questionnaires measure varying social and psychological

elements, it is clear from reviews of the literature that there is some difficulty in differentiating between psychological distress, depressive symptoms and depressive mood because of the properties of measurement instruments used. Therefore the term 'distress' tends to be adopted to represent a composite of these parameters. Heightened autonomic, somatic awareness or somatic anxiety has traditionally been understood as a form of distress. In assessing patients with LBP in an orthopaedic clinic, Main (1983) observed that they routinely described symptoms of increased sympathetic activity and concluded that they were demonstrating an increased awareness of bodily symptoms and function. Subsequently the modified somatic perception questionnaire (MSPQ) was developed to measure heightened somatic and autonomic awareness within the clinical setting (Main, 1983). The concept of somatic anxiety, or increased somatic awareness, is explicitly different from somatisation, which is considered a psychiatric disorder characterized by multiple, recurrent, changing physical symptoms in the absence of a physical disorder. The terminology, however does appear to be interchangeable within much of the literature. Despite recent literature suggesting that the routine administration of questionnaires for depression and anxiety does not necessarily influence clinical decision-making, there remains a need to identify distress and anxiety associated with LBP (Gilbody et al., 2001). At present there is no accepted method for the objective assessment for distress and anxiety in the LBP population, despite the widespread acceptance that these factors are important in the prognosis of LBP patients.

In summary, distress, increased somatic awareness and depressive symptoms have all been shown to impact on the outcome of treatment for LBP and therefore should be assessed and reported in research projects. Indeed, recent work has indicated that psychological factors are more closely related to the progression of pain and disability than most biomedical and biomechanical variables (Linton, 2000) and that depression, in particular, is an important factor in the transition from acute to chronic pain (Pincus et al., 2002).

A number of well-established tools exist for assessing psychological status in LBP patients. The distress risk assessment method (DRAM) is a screening tool that has

been in use now for approximately 10 years and has been shown to perform well in LBP populations. It was designed as a simple screen to be used in the clinical setting to identify patients who might benefit from referral for further psychological assessment (Main et al., 1992; Waddell, 1998). It utilises the Modified Zung (MZ) (See Appendix F) and the MSPQ (See Appendix G) and has been validated for use with a back-pain population (Zung et al., 1965; Main, 1983). Using the combined scores from the MZ and MSPQ, the DRAM categorises patients as showing no psychological distress (or 'N'), those at risk of developing distress (or 'R') and those who are clearly distressed (either distressed depressive 'DD' or distressed somatic 'DS' see Table 2-6). Early work undertaken by Main et al. (1992) clearly demonstrated the increasing risk of poor outcome of treatment associated with increasing levels of distress. As the psychological status of patients who develop LBP is an important indicator of their future outcome, the need to try and measure this parameter is clear (Burton et al., 1997). This has been highlighted recently by a systematic review of psychological risk factors in back and neck pain (Linton, 2000) and of predictors of chronicity in LBP (Pincus et al., 2002). In a review of 37 prospective studies, Linton (2000) graded the evidence presented against the methodological guidelines for systematic reviews, which had been previously reported (van Tulder et al., 1997). He concluded that psychological factors were more closely related to the development of pain and disability than most biomedical and biomechanical variables. Specifically, he identified that psychosocial factors were pivotal in the transition from acute to chronic pain, in addition to being influential in the onset of pain, a finding re-iterated more recently by another systematic review (Pincus et al., 2002). In their review of psychological factors as predictors of chronicity/disability in chronic LBP, Pincus and co-workers identified strong evidence for the role of psychological distress/depressive mood in the transition from acute to chronic LBP and moderate evidence for the role of somatisation. They found no specific predictive value for clinical outcome of the DRAM but were only able to identify one study of acceptable standard to be included (Pincus et al., 2002). They did, however, find that the depressive symptom component (MZ) significantly discriminated between recovered and non-recovered patients. These findings support previous suggestions that the identification of subjects at risk of psychological distress or with pre-existing

anxiety and depression should be included in any outcome evaluation and analysis (Main et al., 1992).

The clinical validity of patient types identified by the DRAM has been presented by Main and co-workers (1992) in 567 mixed orthopaedic and pain clinic patients. Overall, they found that increasing levels of distress were associated with an increase in severity of most variables such as higher disability scores, increased use of medication, particularly analgesics in the DD group, and increasing work loss. The findings of Main et al. (1992) support the need for early identification and treatment of distress in a LBP population. Main et al. (1992) demonstrated a clear increase in the percentage of poor outcome with increasing levels of distress as measured by the DRAM. The relative risk of poor outcome for patients classified as 'at risk' or 'R' was approximately twice that of patients classified as 'Normal' or 'N'. This rose to an increased relative risk of 3-4 for distressed (DD and DS) patients compared to those classified as 'Normal'.

Table 2-6: Summary of categories within Distress Risk Assessment Method

Category	Scores		Description
Normal	N	MZ <17	No evidence of distress or abnormal illness behaviour
At Risk	R	MZ 17-33 MSPQ <13*	Patients show slightly higher scores than Normal patients, with the largest difference in the depressive symptomology (MZ)
Distressed- Depressive	DD	MZ >33	Clear elevation in all variables, particularly high scores on depressive symptomology (MZ)
Distressed- Somatic	DS	MZ 17-33 MSPQ >12	Elevation of all variables, comparable levels of illness behaviour with DD, but elevation in somatic awareness (MSPQ)
*Score based on Waddell (1998)			
Score range: Modified Zung (MZ) = 0-69			
Modified Somatic Perception Questionnaire (MSPQ): 0-39			

The predictive validity of the DRAM was also examined in 261 patients with chronic LBP (LBP of more than 3 months) entering a physical rehabilitation programme for chronically disabled workers by Williams and co-workers (Williams

et al., 1995). They reported higher baseline distress rates than Main et al. (1992) with 41% of patients classified 'Distressed' (either DD or DS), 14% classified as 'Normal' and 45% as 'At Risk'. Similarly to previous findings, Williams and co-workers reported an increased odds of poor outcome with increasing levels of distress; 3.3 to 8.1 times higher for distressed patients compared to those classified as 'Normal' and 2.4 to 5 times greater for 'At Risk' patients than for non-distressed patients. The higher levels of distress reported in this study compared with other studies may be explained by the chronically disabled nature of the population, with 97% in receipt of earnings related compensation for 3 months or longer and 87% off work for more than 3 months. Williams et al. (1995) suggest that the screening of patients prior to enrolment in a rehabilitation programme may identify those requiring additional psychological input.

In summary, evidence clearly exists of the importance of psychosocial factors in LBP patients. Additionally, evidence exists for the clinical utility of the DRAM in identifying patients with LBP who are distressed and for its use in research to provide a simple classification system in order to describe the psychological status of the population.

2.9 Generic Measures in Low Back Pain

Over the last decade there has been increasing interest in the measurement of subjective accounts of health and quality of life in relation to healthcare interventions (Jenkinson et al., 1993). The development of questionnaires specifically designed to measure quality of life, such as the Nottingham Health Profile (Hunt et al., 1985), EuroQol (EuroQol-Group, 1990), and Short-Form 36 (SF-36) (Ware and Sherbourne, 1992) have allowed both clinicians and researchers to assess the subjective impact of different healthcare interventions on varying dimensions of health and well-being. In respect to LBP research, the International Forum for Back Pain Research in 1998 (Deyo et al., 1998) and an expert panel in 2000 (Bombardier, 2000a) recommended the use of generic outcome measures, in addition to a back-pain specific outcome measure in back pain trials.

The SF-36 was designed as a generic indicator of health status, but used in conjunction with a disease-specific measure, it has been suggested as an appropriate measure for clinical practice and research (Ware and Sherbourne, 1992). It consists of 10 questions, with 35 individual items, designed to measure 8 different health dimensions or scales including function, distress, well-being, and self-evaluated health status (See Appendix H). One additional question assesses change of health over the past year and does not contribute to the calculation of any of the scales. Individual items are coded, summed and transformed onto a scale of 0 (worst health) to 100 (best health). Figure 2-2 summarizes the contribution of individual items to the construction of each dimension of the SF-36.

Figure 2-2: Summary of SF-36 measurement model: contribution of items to scales and summary scales

Items		Scales	Component Summary
3a	Vigorous activities	Physical Functioning (PF)	Physical Health (PCS)
3b	Moderate activities		
3c	Lift, carry groceries		
3d	Climb several flights		
3e	Climb one flight		
3f	Bend, kneel		
3g	Walk, mile		
3h	Walk several blocks		
3i	Walk one block		
3j	Bathe, dress		
4a	Cut down time	Role Physical (RP)	
4b	Accomplished less		
4c	Limited in kind	Bodily Pain (Pain)	
4d	Had difficulty		
7	Pain-magnitude	General Health (GH)	
8	Pain-Interfere		
1	General health perception		
11a	Sick Easier		
11b	As Healthy		
11c	Health to get worse	Energy/Vitality (EV)	Mental Health (MCS)
11d	Health excellent		
9a	Pep/life		
9e	Energy		
9g	Worn out		
9i	Tired		
6	Social-Extent		
10	Social-Time		
5a	Cut down time		
5b	Accomplished less		
5c	Not accomplished		
9b	Nervous		
9c	Down in dumps		
9d	Peaceful		
9f	Blue/sad		
9h	Happy		

More recently, two summary scales have been developed; the physical and mental component summaries (PCS and MCS respectively), which are summary component measures derived from the eight scales/dimensions. Factor analyses of the correlations of the eight SF-36 scales have consistently identified two

factors, which can be interpreted as representing the 'physical' and 'mental' dimension of health status. These two summary scales using norm-based scoring with means, standard deviations and factor score coefficients from a general population allow the summary scales to have a mean of 50 and standard deviation of 10. It is suggested that this allows easier comparison and intuitive interpretation of results (Ware, 1994). Population means for the UK are available from The Oxford Healthy Life Survey (OHLS) with coefficients for use in generation of the PCS and MCS (Jenkinson et al., 1993). This allows meaningful comparisons to be made between the results from varying clinical population and the population 'norms'.

Although the SF-36 has been well validated in the UK and is widely used in research, it is not without its critics. Mawson (1995) has suggested that issues with the underlying construct of the questionnaire make the SF-36 an inappropriate measure of physiotherapy intervention. Artificial 'floor effects' can be encountered by the severely disabled within the physical functioning section and, as such, Mawson suggested that the SF-36 may lack validity and sensitivity as a measure of therapeutic intervention when the level of function is likely to remain low (Mawson, 1995). However, in subjects likely to return to full function then the SF-36 may provide a useful measure. Responsiveness in detecting change over time is an important property for any outcome measure. Patrick et al. (1995) investigated the responsiveness of the SF-36 in detecting small changes in 427 patients with sciatica over a 3-month period. They concluded that most scales of the SF-36 changed in the expected direction in response to treatment and suggested that a 7-point change in the physical functioning scale identified a MCID. Taylor et al. (1999) investigated the responsiveness of the eight scales and two summary scales of the SF-36 compared with two functional measures, the ODI and Low Back Outcome Score (LBOS) (Greenough and Fraser, 1992). Data were collected pre- and post- treatment, at 6, 12 and 24-month follow-ups from 318 back pain patients and compared with patients' subjective global ratings of change. Using three different methods for calculating effect size, all scales except general health, showed moderate or large effect sizes for those patients who rated themselves as better or much better after treatment. None of the three analyses showed consistently higher or lower values, but the magnitude of effect did differ according

to the effect calculation utilized. Negative effect size values were obtained for patients who reported a worsening condition, indicating that the questionnaires were able to identify both deterioration and improvement.

The recommendation that back pain research should include generic health measures, in addition to disease-specific measures reflects the fact that although both will often co-vary i.e. both show increases or decreases, crucially they measure different aspects of health status. Taylor et al. (1999) examined the relationship of the back pain specific measure, the ODI, and the generic measure, the SF-36. In all of the subgroups of LBP patients studied (better, worse and unchanged) the sensitivity to change between the ODI and SF-36 co-varied, suggesting that the SF-36 addresses health concepts that are important to patients with LBP. The use of generic measures of health is now well accepted as a mainstay of healthcare research in general and also within LBP research. Generic health measures generally have good face validity for patients, which should help facilitate response and follow-up rates. They provide invaluable information about baseline and subsequent changes in perceived health following healthcare interventions and, as such, should be considered an essential component in the design of research projects.

2.10 Compliance, motivation and advice

Compliance with any treatment or healthcare intervention is a vital component for a successful outcome. Both attitude and behaviour can influence compliance. A one to two third non-compliance rate with therapeutic exercise regimes has been reported (Sluijs et al., 1993) and compliance with physiotherapy regimes have been reported as consistently worse than with medication (Deyo, 1982). It has been suggested that there is a dose-response dependent outcome with regimes used to treat LBP, however the exact nature of such a relationship is unknown. Authors have identified that established exercise habits are a predictor variable for compliance and that in coronary rehabilitation, an inactive lifestyle is often associated with non-compliance with exercise regimes (Oldridge, 1982).

Expectations and motivation both play important roles in compliance with exercise regimes with subjects who expect exercise to form part of their treatment possibly being more compliant. Similarly, motivational issues concerned with the activation and persistence of behaviour influence compliance and it has been suggested that these are cognitively based and can also influence outcome. Friedrich et al. (1998) examined the effects, on compliance and level of disability, of combining a motivational programme with an exercise regime. Using subjects with chronic and recurrent LBP, they found that the addition of a motivational programme was significantly more effective in reducing disability and pain than a standard exercise programme alone (Friedrich et al., 1998). However, although both groups were advised to continue exercising after termination of treatment, no difference was found between the two groups with regards to long-term exercise compliance (Friedrich et al., 1998). Evidence exists that verbal instruction alone facilitates only low levels of compliance, supporting the need for written instructions to accompany any exercise regime (Schneiders et al, 1998). It is widely acknowledged that there are problems of a lack of valid, reliable measures of patient compliance (Friedrick et al., 1998). Inconsistencies regarding a universally accepted definition of compliance make assessment and measurement virtually meaningless, leading Friedrich et al. (1998) to conclude that a reliable, longitudinal measure to objectively validate adherence with exercise therapy does not exist. Certainly, this is highlighted by a study by Hagins et al. (1999). In an attempt to study compliance to a 4-week stabilisation regime of 22 non-symptomatic subjects, a 'compliance' sheet was given to each subject who was asked to return the form having indicated the dates on which they had performed their exercises. Unsurprisingly, these forms reported a 100% compliance rate prompting the authors to acknowledge the possibility that the reported compliance did not accurately reflect the true compliance rate. In doing so, these authors highlight the difficulty in accurately measuring compliance in both clinical practice and research and cast doubts on the reliability of subjective reports of compliance. Interestingly, Bergquist-Ullman & Larson (1977), who undertook one of the earliest trials utilizing specific spinal exercise training, did not monitor compliance with the prescribed exercise regime, citing the lack of a reliable method, whereas two later studies (Hides et al., 1996; O'Sullivan et al., 1997c), which will be discussed in detail in Section 2.15, both attempted to monitor compliance via a diary sheet. This highlights the

inconsistencies within the literature regarding the monitoring of compliance and suggests that, as yet, no universally acceptable, reliable system is available.

The preceding sections have discussed the various outcome measures used within LBP, both in the clinical and research settings and discussed their relative merits. The next sections summarise the use of exercise as a treatment intervention for LBP and the role of physiotherapists both in the prescribing of exercise and the treatment of LBP patients.

2.11 Exercise and Low Back Pain

Exercise regimes and exercises in relation to LBP have been the source of much debate in the search for the answer to the perennial questions of “Which ones are worth trying, for which patients and when?” (Faas, 1996). Consequently, exercise in the management of LBP has been the subject of many good quality systematic reviews with varying results (Koes et al., 1995; Maher et al., 1999; van Tulder et al., 2000b; van Tulder et al., 2000c; Abenhaim et al., 2000; Linton and van Tulder, 2001). The debate regarding the relative merits of aerobic training (Moffett et al., 1999) and graded activity programmes or more specific regimes such as McKenzie therapy (McKenzie, 1981) and spinal stabilisation training (Richardson and Jull, 1995; Richardson et al., 1998) is central to much of the debate. There are many stakeholders in this debate, not least the physiotherapy profession who, with exercise as one of three core skills, feel strongly about their continuing involvement in exercise provision.

The case for active rehabilitation and exercise as a generic term was highlighted by the CSAG (1994a) guidelines and in subsequent clinical guidelines (RCGP, 1996b; Waddell et al., 1999). More recently van Tulder and co-workers (2000) undertook a systematic review, as part of the Cochrane Collaboration⁹, of exercise therapy for LBP, concluding that exercise was not effective for acute LBP but more effective than GP care for chronic LBP. Although this review included one trial which studied specific stabilising spinal exercises, this trial was incorporated within a generic 'exercise' category and not a specific sub-category. By using this strategy, the individual nature of this type of exercise is not reviewed, however, methodological considerations make basing any recommendation on one RCT unwise. As the number and quality of RCTs studying spinal stabilisation training increases this will hopefully be remedied.

The evidence for the use of exercise in LBP was strengthened further by the recent international comparison of clinical guidelines for the management of LBP in primary care (Koes et al., 2001b). Although identifying that many guidelines do not extend recommendation beyond the acute stages of an episode, those that did (Netherlands, Germany, Denmark and the UK) all recommended exercise as an important intervention, but still with no consistency regarding type and intensity of exercise. The importance given to the use of exercise in the management of LBP is therefore clear, if the exact nature of that exercise remains obscure. However, what is certain is the key role that physiotherapy as a profession has to play both in direct patient contact, and also assisting in the development and revision of clinical guidelines. This latter point was evidenced by the active involvement of the physiotherapy profession in the development of the original UK guidelines (CSAG, 1994) and the former point by the evidence of the extensive use of exercise therapy by physiotherapists in the management of LBP. This point will be discussed in detail in the next session.

⁹ Cochrane Collaboration is an international not-for-profit organisation. It aims to make up-to-date accurate information about the effects of healthcare readily available worldwide. The major product of the Collaboration is the Cochrane Database of Systematic Reviews, which is part of The Cochrane Library. (Source: www.cochrane.org)

2.12 Current Physiotherapy Practice

An audit undertaken by the South East Thames Health Authority in 1997/8 highlighted the frequent use of exercise as a treatment by physiotherapists (Moore, 1997/8). The aim of this audit was to establish a tool to measure the effects of physiotherapy interventions in a general outpatient setting and to gather data on current clinical practice. Despite a planned sample size of 2000, only 564 patients were recruited, with only 331 receiving a normal discharge¹⁰. For patients who were normally discharged, mobilisations, active exercises and advice were the most frequently employed treatment strategies, with 70.2% of patients receiving some kind of active exercise regime. Muscle imbalance re-education techniques were utilised little in the rehabilitation process of patients normally discharged. As the first choice modality, they accounted for only 3.0% (n=10), rising to 8.2% (n=4) as third choice. The authors noted surprise at the low use of muscle imbalance techniques, identifying recent research work that provided evidence on which to base such strategies. Moore et al. (1998) identified a limitation, that muscle imbalance re-education may have been rated as *active exercises* thereby giving an inaccurate indication of usage. The authors concluded that the exact nature of active exercise strategies required further investigation and that the theories of muscle imbalance re-education should be incorporated into these strategies. Foster (1999) investigated current clinical practice in the treatment of LBP by physiotherapists in Britain and Ireland. In a survey undertaken over an 18-month period from Autumn 1994 to Spring 1996, a response rate of 58.3% was achieved (n=1548), with 813 therapists reporting that they were regularly treating LBP patients. Results indicated the most frequently used methods for the management of LBP were mobilisation techniques namely the Maitland (Maitland, 1986) and McKenzie concepts (McKenzie, 1981). However, it also exposed a reliance on passive electrotherapeutic techniques. This led the authors to suggest that there was a poor uptake of research information regarding the efficacy of physiotherapy techniques for the treatment of LBP.

¹⁰ Normal discharge defined as those patients who completed a course of treatment agreed between themselves and the treating therapist

Increasingly, therapists have been encouraged to view and treat LBP holistically, looking to the prevention of chronicity and long-term disability in preference to purely symptomatic relief of symptoms (CSAG, 1994a; RCGP, 1996b). Recent investigation into the functioning of the intrinsic stabilising system of the lumbar spine has led to a major re-think regarding the rehabilitation of LBP patients (Richardson and Jull, 1995; Panjabi, 1995; Gardner-Morse and Stokes, 1998). The finding that the deep abdominal and spinal muscles, transversus abdominis (TrA) and lumbar multifidus, purported to provide a major stabilising role to the spine become dysfunctional in the presence of LBP has caused a fundamental shift in the paradigms underlying the physiotherapeutic management of LBP. The basis for the re-training of these muscles as part of a holistic rehabilitation programme has been established for a specific sub-group of the LBP population such as non-operative prolapsed intervertebral discs (Saal and Saal, 1988) and patients with radiological evidence of spondylolisthesis (O'Sullivan and Twomey, 1997). However, extrapolation of such findings to a wider LBP population cannot be made. Despite this, in Britain, courses providing instruction in these techniques are increasingly popular, and anecdotal evidence suggests that they are in wide-spread use in clinical practice. This therefore indicates that there is a need for high quality, pragmatic trials that represent current clinical practice that will allow results to be extrapolated to the widest clinical population.

As outlined in section 2.1.1, spinal stabilisation exercises in contrast to general aerobic or conditioning training, are concerned with the coordination of optimal patterns of muscle activity, which provide postural stability in order to allow activities to be undertaken safely and effectively (Jull and Richardson, 1994). The concept of a local muscle system (Bergmark, 1989), controlling spinal stiffness and thus providing mechanical stiffness, is linked closely with a concept of spinal stability put forward by Panjabi (1992a & b). This forms one of the bases of stabilisation training and will be discussed in detail in the next section. For clarity, the following section will be divided up into the theory behind the spinal stabilisation approach, the evidence supporting it from experimental studies and clinical studies, the limitations of studies to date and current research needs.

2.13 Theory behind the spinal stabilisation approach

Although generally accepted that mechanical stability is essential for both static and dynamic tasks of everyday living, what is more controversial is the concept of instability in relation to spinal mechanics (Grieve, 1982; Vaccaro et al., 1997; Bogduk, 1997; Floman, 2000). ‘True’ instability such as that arising from spondylolysis¹¹, spondylolisthesis¹² or fracture is generally demonstrable by reliable, valid investigations such as radiographs and MRI scans and hence not a subject of such debate and discussion. Minor instabilities, or those termed ‘clinical instabilities’ are much more difficult to diagnose definitively and therefore are much more controversial.

Panjabi (1992a) defines clinical instability as:

“A significant decrease in the capacity of the stabilizing system of the spine to maintain the intervertebral neutral zones within physiological limits which results in pain and disability”

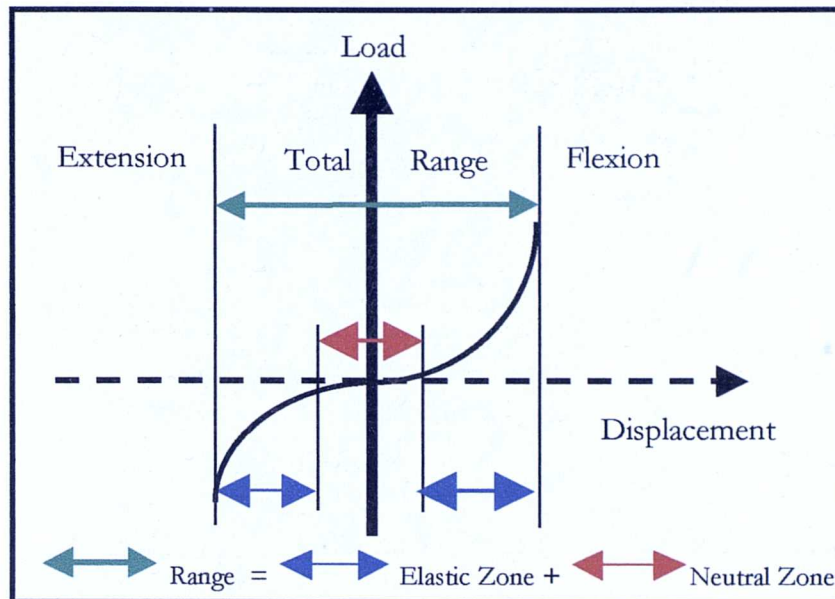
The hypothesis of the ‘neutral zone’ proposed by Panjabi (1992a & b) is key to the spinal stabilisation training approach (Panjabi, 1992b; Panjabi, 1992a). The neutral zone is hypothesised to be a region of intervertebral motion around a neutral position, where little or no resistance from the supporting spinal structures occurs. The neutral and elastic zone, where resistance is encountered, together form the total range of movement at any spinal motion segment (Figure 2.3).

¹¹ Pars interarticularis defect

¹² Refers to a mechanical consequence of vertebral slippage in an anterior direction and may be i) congenital ii) isthmic (bilateral spondylolysis) iii) degenerative iv) traumatic or v) pathological (Grieve, 1982)

Figure 2-3 - Diagrammatic representation of the Neutral Zone

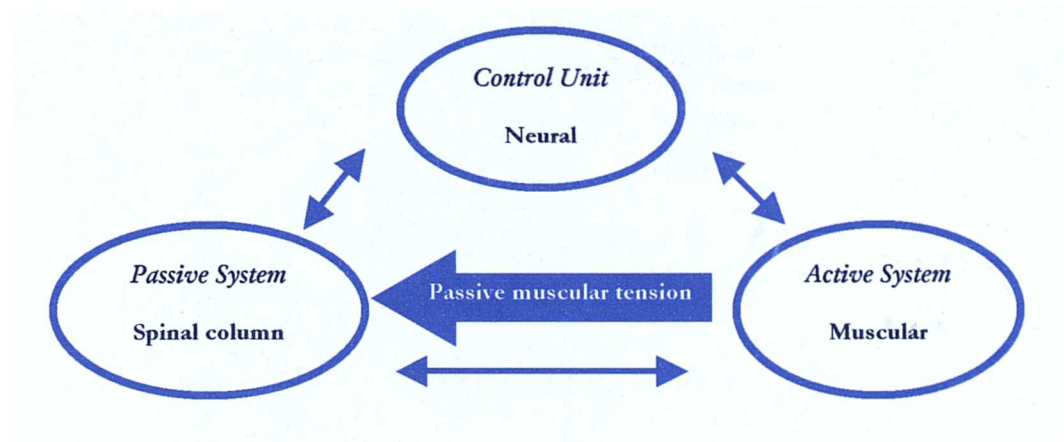
(Adapted from Panjabi, 1995)



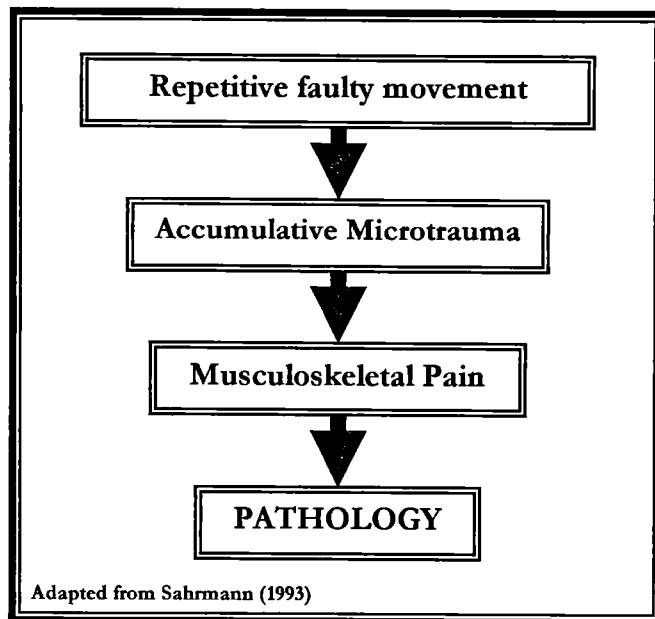
Control of this hypothesised neutral zone is at the heart of the stabilisation training approach. Panjabi (1992b) maintained that an increase in the neutral zone can be associated with degenerative changes and is important clinically as it is a measure of spinal 'stability'. In suggesting a model by which spinal stability is achieved, incorporating a combination of three sub-systems, muscular, neural and passive, Panjabi presented a key concept of stabilisation training. This interconnected system is summarized in Figure 2-4. In citing the stabilising influence of the muscles of the lumbar spine as a method of returning the neutral zone to within normal physiological limits, Panjabi and others presented a potential method by which spinal stability could be accessed and, more importantly, influenced (Panjabi et al, 1989; Wilke et al, 1995). The model of the functional stability unit consisting of the abdominal and trunk muscles anteriorly and posteriorly and the diaphragm and pelvic floor in a cephalad and caudad position respectively has been put forward as central to spinal stability and retraining (Richardson et al., 1998).

Figure 2-4: Spinal stability system

(Adapted from Panjabi, 1992)



Alongside the concept of a neutral zone, a challenge to the traditional thinking of pathology as the cause of faulty movement was made by Janda (1977) and later by Sahrman (1993). It was suggested that reflex changes and functional impairments of the motor system may manifest not only as pain but also influence the results of any motor re-education programme (Janda, 1977). The role of muscle dysfunction in the pathogenesis of musculoskeletal disorders and postural defects was argued to support the hypothesis that certain muscles respond to certain situations, for example pain, by either tightening and overactivity (hamstrings and trunk erectors), or inhibition, atrophy and weakness (abdominals and gluteals). The maintenance of balanced muscle co-ordination in order to provide protection for the osteoarticular system (Janda, 1977; Sahrman, 1993) mirrors the concept of the neutral zone and maintenance of it within normal physiological limits maintaining spinal stability (Panjabi, 1992b). The concept suggested by Sahrman (1993) of repetitive faulty movement and accumulative microtrauma being the source of much musculoskeletal pain is summarised in Figure 2-5.

Figure 2-5: Summary of movement system balance

The aim behind spinal stability training is the management of lumbar spine symptoms through maintenance of spinal stability. It is suggested that maintaining the hypothesised neutral zone within normal physiological limits, provides a protective environment for the spine to allow healing of damaged tissues and therefore, ultimately, improvement in spinal pain and dysfunction (Richardson, 1995). As such, preservation of a 'neutral' lumbar spine position during muscle re-education has been recommended (Richardson, 1995).

2.13.1 Properties of stabilising muscles

The physiological, anatomical and biochemical basis of muscle functioning summarised by Bergmark (1989) and refined more recently by a number of other authors (Bergmark, 1989; Comerford and Mottram, 2001b) describes two distinct muscle systems; *local or static system* and *global or dynamic system*. These are at the core of the spinal stability approach to rehabilitation and provide some scientific basis for the type, amount and frequency of exercises used in the retraining. It is suggested that the distinct physiological, anatomical and biochemical properties (See Table 2-7) these muscles possess make them ideal for their specific functions.

However more recent research has cast doubts over the accuracy of certain parts of this classification system (Mannion et al., 1997).

Table 2-7: Summary of physiological, anatomical and biochemical basis for function of the local and global muscle systems

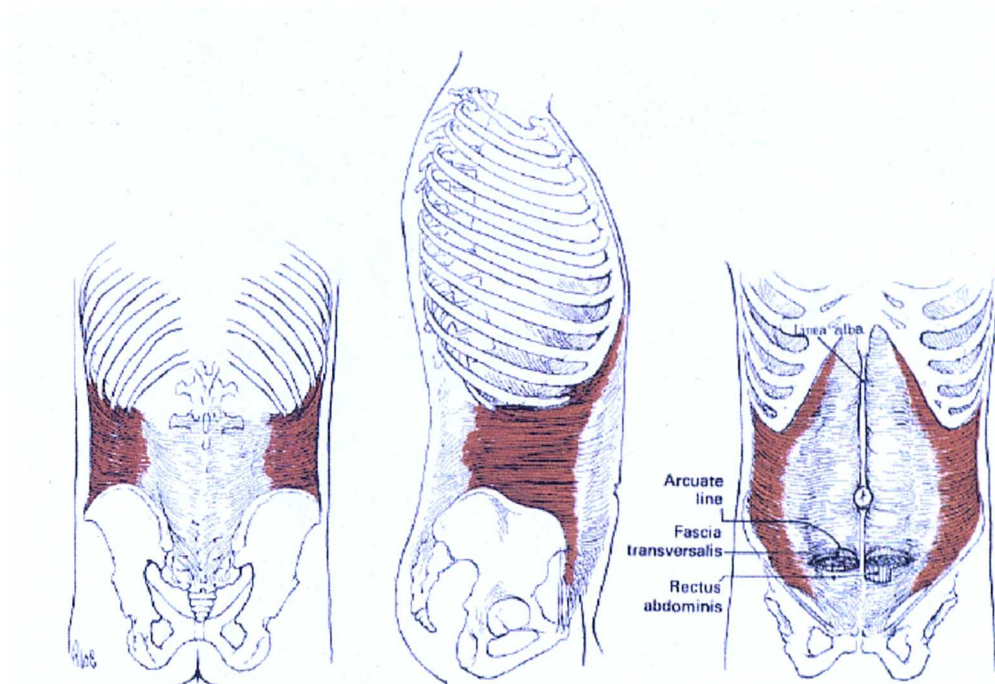
LOCAL SYSTEM (STATIC)	GLOBAL SYSTEM (DYNAMIC)
Physiological	
Tonic (slow), Type I fibres	Phasic (fast), Type II (IIB)
Slow contraction time	Fast contraction time
Low threshold- 20-30% MVC*	High threshold-40% + MVC*
Sustained discharge pattern	Erratic, uneven activity
Fatigue resistant	Fatigues quickly
Feedback afferent system important	Pre-programmed
Proprioception, fine tuning	Torque producing
Anatomical	
Deep, cross one joint, aponeurotic	Superficial, multiarthrodial, tendinous
Origins & insertions on vertebrae	Non-segmental
Control curvature -provide stiffness	Link thorax & pelvis-torque producing
Oppose gravity	Distance from joint
Continuous contraction-postural holding	Burst/phasic-repetitive/rapid
Biochemical	
Low ATPase activity	High ATPase activity
High oxidative capacity	Low oxidative capacity
Increased post-synaptic potential	Low post-synaptic pot.
Long refractory time	Short refractory time
Based on Bergmark, 1989	
*MVC= maximal voluntary contraction	
ATPase=adenosine triphosphate. Energy source for enzyme reactions in muscle fibres	

The erector spinae muscle, as a multiarthrodial muscle linking the thorax and pelvis, would be classified within the 'global muscle system' using the characteristics suggested by Bergman (1989), and latter Comerford and Mottram (2001b), and as such should possess phasic, torque producing qualities and fatigue quickly due to a high proportion of Type II muscle fibres. However, it has been demonstrated that the erector spinae has a predominance of relatively large Type I (slow twitch) muscle fibres which indicates a strong postural function (Mannion et al., 1997). Despite these findings however and the growing body of literature supporting erector spinae as a key muscle in postural control, the teaching and

literature surrounding stabilisation training continues to focus, rightly or wrongly, more on specific muscles, such as Lumbar Multifidus which will be discussed in the next section and the classification originally proposed by Bergmark (1989).

The main muscles involved in the early phases of stabilisation training are the Transversus abdominis (TrA) and Lumbar Multifidus (LM). As indicated in Table 2-7, it has been suggested that both muscles are anatomically well suited to provide a key role in muscular spinal stabilisation. The deepest of the abdominal muscles, the TrA has extensive aponeurotic attachments to the inguinal ligament, iliac crest and thoraco-lumbar fascia.

Innervated by the anterior primary rami of T7-T12 and L1 (Gray, 1980) the anatomical relationship of TrA to the other abdominal muscles is shown in Figure 2-6. A more detailed anatomical description is given in Appendix I but the reader is directed to one of the numerous texts with highly detailed descriptions of the anatomy and function of these muscles (Gray, 1980; Bogduk, 1997; Richardson et al., 1998; Hodges and Richardson, 1999).

Figure 2-6: Anatomical position and relationship of transversus abdominis

Source: Kendall, FP, McCreary, E and Provance, PG (1993) *Muscle Testing and Function*. Page 151

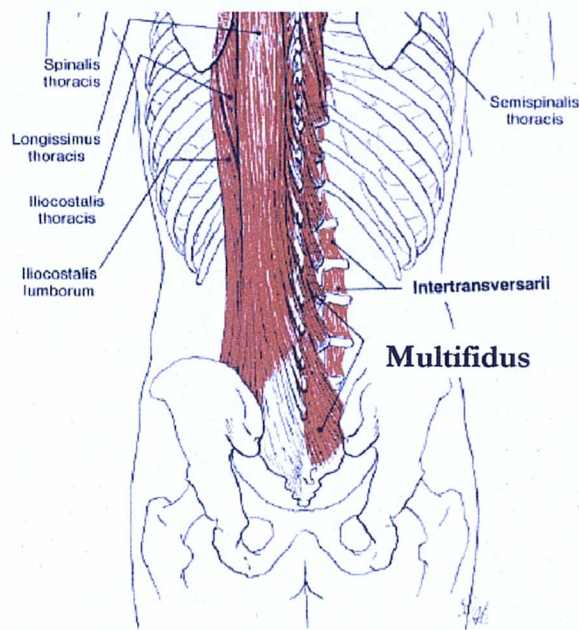
Reproduced with permission from © Lippincott Williams & Wilkins.

The architecture and attachments of TrA have led to the hypothesis that tensioning the middle layers of the thoraco-lumbar fascia might produce 40% of trunk stability in the coronal plane (Tesh et al., 1987) and that the circumferential nature of the TrA has been identified as pivotal to its stabilising function (De Troyer et al., 1990). Detroyer et al. (1990) suggested that the circumferential nature of TrA increased the effectiveness for raising intra-abdominal pressure (IAP), which was demonstrated experimentally, with other researchers demonstrating that TrA was the abdominal muscle consistently related to changes in IAP (Cresswell et al., 1992). They identified that the orientation of TrA fibres prevented the muscle from being a major flexion or extension torque generator and hypothesised that TrA, either with or without increased IAP, provided trunk stabilisation during maximal isometric flexion.

Similarly, the LM as the deepest of the lumbar extensors, with the deeper fibres attaching to the collagenous fibres of the joint capsule of the zygapophyseal joints next to the mamillary processes, and segmental innervation (Bogduk, 1997),

possesses the ideal properties to function as a deep spinal stabiliser and provide proprioceptive feedback to influence the neuromotor element within this system (McGill, 1991). Figure 2-7 shows the gross anatomy and position of LM with the detailed anatomy shown in Figure 2-8.

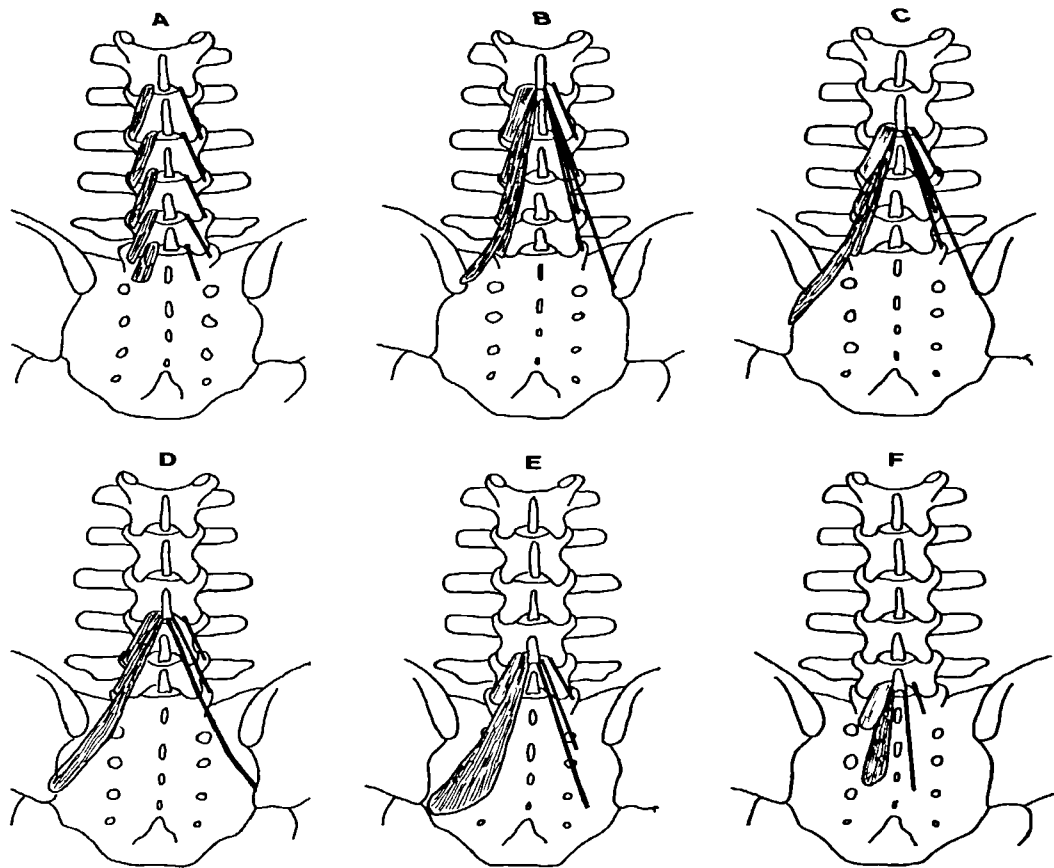
Figure 2-7: Gross Anatomy of Lumbar Multifidus



Source: Kendall, FP, McCreary, E & Provance, PG (1993) *Muscle Testing and Function*. Page 139

Reproduced with permission from © Lippincott Williams & Wilkins.

In summary, therefore, it is clear that anatomically both muscles are well suited to a stability role and as such are key to the concept of spinal stabilisation. The next section will discuss the current knowledge in respect to the neuromotor control, activation and recruitment of these main two muscles.

Figure 2-8: Detailed Anatomy of Lumbar Multifidus

The component fascicles of the lumbar multifidus muscles.

A: The laminar fibres. B-F The fascicles from the L1 to L5 spinous processes respectively

Source: Bogduk, N (1997) *Clinical Anatomy of the Lumbar Spine and Sacrum* Page 106

Reproduced with permission from © Churchill Livingstone

2.14 Evidence supporting the spinal stabilisation approach

The evidence base of spinal stabilisation training is built on a wide range of research, both experimental and clinical. It includes work that evaluates the ability of LM and in particular TrA to contribute to spinal forces and therefore stability and indirect evidence of how the central nervous system (CNS) utilises TrA during the varying tasks of everyday living (Hodges, 1999). A recent review article summarized the experimental evidence supporting the role of TrA as a key muscle in spinal stabilisation training (Hodges, 1999) namely regarding:

- i. Neuromotor control and recruitment
- ii. Mechanical effects via intra-abdominal pressure (IAP)
- iii. Reaction to predictable perturbations
- iv. Relationship with the diaphragm
- v. Relationship to load.

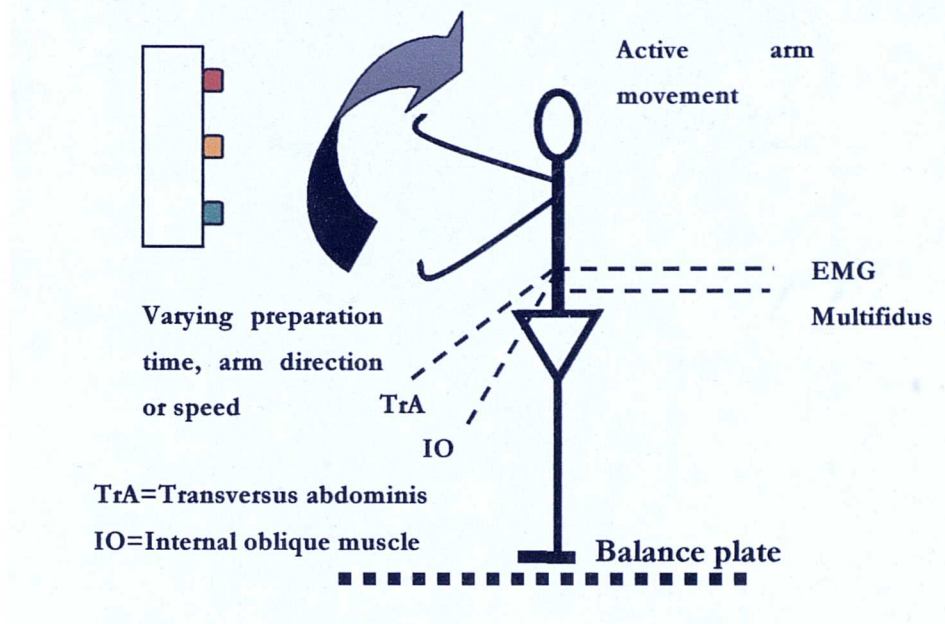
Early work investigating changes in muscle activity during fast alternating flexion/extension of the knee provided the basis for much of the work regarding the stabilising function of the abdominal and trunk muscles (Richardson and Bullock, 1986). Findings indicated that increasing velocity during training resulted in selective training of phasic muscles and subsequent muscular imbalances. Similarly, studies examining the effect of progressive loading on overall strength of the human calf muscle found that, while power of gastrocnemius increased, there was a decrease in the isometric strength of the stabilising muscle of the calf, namely soleus (Ng and Richardson, 1990). This identified that specificity of training had differing effects on different parameters of muscle function and provided a stimulus for similar work regarding the trunk and spinal muscles.

2.14.1 Neuromotor control and recruitment

The nature and neuromotor control of activation of trunk muscles to provide appropriate stability has been extensively researched and is fundamental to the concept of stabilisation exercises. Thorstensson et al. (1985) described the performance of 'appropriate' trunk movements as the result of specific patterns of muscle co-ordination, and hypothesised that interruption in such patterns may result in dysfunctional mechanics. Thorstensson and co-workers concluded that patterns of activity in trunk muscles, during voluntary trunk activity, were related to direction of movement, amplitude and velocity, citing the 'stabilising' role of the external oblique muscle, and suggested that the mechanical advantages afforded to rectus made its function more suited to a prime mover role. They also hypothesised that passive internal forces, such as tension in ligaments and viscoelastic properties, could be utilised by the CNS to minimise the need for voluntary muscle activity, suggesting a more multifactorial nature of the complex task of motor control.

This work was developed further by a series of studies undertaken by Hodges and co-workers (1996 and 1997 a,b,c) looking at the recruitment and timing of TrA. This and other work has led to the conclusion that TrA is recruited prior to all other abdominal muscles in response to sudden perturbations of the trunk (Richardson et al., 1995; Hodges and Richardson, 1996; Hodges and Richardson, 1997b). Using the experimental set-up similar to that shown in Figure 2-9, Hodges and Richardson (1996) recorded EMG activity of abdominals and multifidus, using fine-wire and surface electrodes, in response to rapid shoulder abduction, flexion and extension. They demonstrated that TrA was the first muscle to be activated in normal subjects (n=15) in all movements and was not influenced by movement direction i.e. TrA activation was direction independent. This automatic response of TrA was anticipatory exhibiting feed-forward characteristics rather than acting on a biofeedback system. A similar result was found using response to lower limb movement (Hodges and Richardson, 1997a; Hodges and Richardson, 1998). Conversely, a small patient population (n=15) with chronic LBP (minimum 18 month duration but with minimal pain at the time of testing), demonstrated a loss of this anticipatory effect, with TrA activation following the onset of the prime mover in all directions tested (Hodges and Richardson, 1996). These findings echo the assertions made by Thortensson et al. (1985), who suggested that there was a temporal element to the patterns of activity in trunk muscles during voluntary activity. Additionally, the loss of the direction-independent response of TrA demonstrated in the back pain subjects indicated that TrA was responding in a similar fashion to the other abdominal muscles, consequently reducing the protection of the spinal structures to loading. It was hypothesised that the CNS utilises contraction of the TrA to stabilise the lumbar spine in anticipation of the activation of the prime movers of the limbs, supporting the influence of a temporal element suggested by Thorstensson et al. (1985).

Figure 2-9: Basic set-up for motor control studies



Although fundamental in providing the groundwork for further experimental research, the work undertaken by Hodges and co-workers is not without its limitations. Very limited numbers, both of control and back pain subjects, were used and as such, care should be exercised with extrapolation of these findings until the results have been replicated on larger groups. The use of a chronic LBP population with minimal or absent symptoms at the time of testing limits the generalizability of these findings to a painful population, although work by the same research group is underway to address this last point (Hodges, 2001).

In summary, these experiments suggest that, in a healthy population, the deep-seated abdominal muscles, specifically TrA, show anticipatory activation independent of direction of any perturbation force and thus are ideally suited to the task of assisting spinal stability. The clinical implication of these findings is that potentially, patients with LBP lack this fundamental protective action due to alterations in the activation of TrA. It is hypothesised that as a result of the loss of this anticipatory, forward-feed action, the spine in LBP patients is left at risk of further damage because of this low-grade instability. It is this that spinal stabilisation exercise training aims to address.

2.14.2 Mechanical effect

The effect of IAP as a stabilizing force for the spine has been debated for many years. Cresswell and colleagues (Cresswell et al., 1992; Cresswell et al., 1994a; Cresswell et al., 1994b) identified a delay between the onset of muscular activity and IAP, and between peak EMG activity and peak IAP. They suggested that these delays might be due to 'sluggish' co-ordination of the large flat muscles of the abdominal wall and the compression of the elastic abdominal cavity. Conversely, Hodges and Richardson (1997a) later concluded that TrA activity occurred prior to any activity of the lower limb prime movers. Although both investigations were undertaken in an upright standing position, Cresswell et al. (1992) utilised a pulsed Valsalva manoeuvre (coughing and straining), whereas Hodges and Richardson (1997a) recorded automatic abdominal EMG activity secondary to limb movement. Hodges et al. (1997) demonstrated that the diaphragm contracted prior to upper limb movement suggesting that this helped generate IAP, useful as short-term assistance but not longer-term postural stability (Hodges et al., 1997). Allison et al. (1998) further examined the role of the diaphragm during abdominal hollowing exercises in 26 healthy subjects and how activation patterns vary with differing levels of IAP (Allison et al., 1998). In the crook supine position, the normalised EMG activity of the anterolateral abdominals, rectus abdominis and diaphragm were all monitored and a pressure sensor (Pressure Biofeedback Unit-PBU) used to monitor pressure changes. With a pressure change of 5mmHg, all EMG amplitudes increased but with sustained loading at 15mmHg pressure change, both the diaphragm and rectus abdominis showed significantly elevated levels. As it is maintained that specific activation of certain muscles such as TrA is desirable in the early stages of stabilisation retraining (Richardson and Jull, 1995), rather than a more generalised group activation, these findings support the diaphragm playing a significant role in different levels of abdominal co-activation, and the concept of exercises designed to isolate the suggested 'core stabilising' muscles such as TrA being undertaken at lower intensity to assist with specificity rather than at higher intensity levels that activate all abdominal muscles. Similarly, recent research has indicated the pelvic floor makes an important contribution to the functional stability unit consisting of the abdominal, trunk, diaphragm and pelvic floor. EMG tracings from pubococcygenus showed similar activity to that of the diaphragm and TrA (Sapsford et al., 1997). Further research in healthy female subjects with no

history of stress incontinence showed that a voluntary abdominal contraction resulted in increased pelvic floor EMG activity and that there was an increase in pelvic floor pressure prior to an increase in the abdominal pressure (Sapsford and Hodges, 2001). It was suggested that this was a normal reaction and was pre-programmed.

This is an ongoing, important area of research as pelvic floor dysfunction is a common problem and the emerging role of the pelvic floor in spinal stability may indicate areas for potential interventions to address this dysfunction in relation to both urinary and spinal symptoms (Sapsford, 2001).

2.14.3 Reaction to predictable perturbations

The concept of reactive forces acting on the body as a result of limb movement is neither new, nor controversial (Bouisset and Zattara, 1987). It is widely accepted that the CNS prepares for predictable challenges to posture by alteration in muscle activity (Bouisset and Zattara, 1987; Friedli et al., 1988). However, the effect of both predictable and unpredictable challenges and changes that might occur in the presence of pain were less well documented, but of vital importance in relation to a spinal stabilisation concept. These have been studied by a number of authors, and the following provides a summary of the results, a critical appraisal, and discussion of the relevance of the findings to the clinical application of spinal stabilisation training.

Using a biomechanical model, Gardener-Morse and Stokes (1998) examined the effects of different co-activation patterns of spinal loading. They hypothesised that the neuromuscular control system 'pre-sets' muscular co-activation to ensure stability and reduce the need for active neuromuscular control responses, thus reducing the inherent time delays associated with any feedback system. Their findings indicated that antagonistic activation of abdominal muscles increased spinal stability, however this was associated with an increased muscle fatigue rate and a small increase in spinal compression (Gardner-Morse and Stokes, 1998). Although supporting the role of muscle activation in spinal stability, these conclusions contrast to those of Hodges et al. (1997 a, b, c) who maintain that control occurs primarily through a feedforward rather than feedback mechanism.

However Gardener-Morse and Stokes (1998) did maintain that dynamic stability is important under changing loads and postures, and that 'pre-set' of co-activation is probably only useful in response to small perturbations. The examination of the responses of the abdominal musculature to perturbation was investigated further in a series of studies by Hodges and co-workers. Hodges and Cresswell (1999), using rapid, bilateral arm movements in healthy subjects ($n=8$), found that small, consistent preparatory motions of the trunk, in the opposite direction to the movement, occur as a result of superficial trunk muscle activity. They suggested that the CNS utilises this as a means of dampening down or absorbing the forces imposed on the spine during movement, and not solely as a method of 'stiffening' the spine (Hodges et al., 1999). This muscle activity occurred in short, phasic bursts in contrast to the sustained, low-level activity of TrA, which continued to have an anticipatory, feedforward activation. These findings were replicated using a three-dimensional trunk motion and asymmetrical upper limb motion ($n=8$ healthy males) (Hodges et al., 2000) indicating that activation of the superficial trunk muscles resulted in anticipatory postural adjustments rather than simple rigidification of the spine.

Again, although innovative and informative in regards to future research directions, these experiments are limited by the very small number of subjects and they lack immediate clinical applicability.

2.14.4 Relationship to load

The effects of speed of perturbation, and therefore resultant load, have been investigated in relation to activation of abdominal muscles by many authors (Thorstensson et al., 1985; Cresswell and Thorstensson, 1994; Hodges and Richardson, 1997c; Hodges and Richardson, 1997a; Hodges and Richardson, 1997b). Thorstensson (1985) found that increasing speed of trunk movement resulting in increased activity in rectus abdominis (RA) but with no change in the anterolateral abdominal function. Similarly, Hodges & Richardson in a piece of unpublished work, cited in Richardson et al. (1999), reported an association between speed and TrA activation. Previous work had reported that in healthy subjects anticipatory TrA activity occurred with movements at fast speeds (Hodges and Richardson, 1997c) and although TrA failed to be active before the prime

mover, at natural or functional speeds its action fell within an anticipatory criteria (Aruin and Latash, 1995). Thus, they argued that TrA could be considered to have an anticipatory effect associated with fast and to a lesser extent natural speed arm movements whereas no trunk muscles were active prior to the prime mover in slow speed arm flexion ($30^{\circ}/\text{sec}$). It was extrapolated that trunk muscles response was pre-programmed with self-paced and rapid arm movements and not influenced by speed of movement or magnitude of force, above a certain threshold. In a piece of unpublished work however, Hodges & Richardson reported that in LBP subjects ($n=15$) the response of TrA and other abdominal muscles was altered, contracting prior to the prime mover, only in response to limb movement at fast speeds. They hypothesised that the threshold for activation of TrA had been raised and that the loss of the proposed stabilizing effect of TrA placed the spine at an increased risk of damage. In addition to the obvious limitations of this study, small numbers and lack of publication in a peer-reviewed journal, the significance of the results needs close scrutiny, as TrA activation did not occur prior to prime mover activation in either group at natural speeds, but was found to be within the 50ms range for anticipatory muscle activation in the non-LBP group. Despite these limitations, this work has been seen as important in adding to the evidence-base for the spinal stabilisation concept and providing an initial basis for the provision of exercises that graduate through increasing load.

The next section discusses issues surrounding the clinical application of stabilisation techniques. It briefly reviews the literature regarding the specific contraction of the TrA and LM, facilitation strategies to assist in the teaching of these contractions and the reliability of physiotherapists' applications of these techniques. It also reviews the literature regarding spinal proprioception and joint position sense as they relate to the 'neutral' lumbar spine position used in specific spinal stabilisation training.

2.14.5 Exercises used in the recruitment of transversus abdominis in clinical practice

With the increasing awareness of the importance of muscle activation on spinal stability there has been increasing use of these techniques within clinical practice.

Additionally the motor-control changes that have been shown to exist in deep stabilising muscles necessitates re-training of these muscles to be viewed as a motor re-training programme. This highlights a number of issues of clinical importance of how, where and when these muscles should be re-educated.

The type of exercise which most appropriately activates deep abdominal muscles was examined by Richardson et al. (1992) using EMG activity recorded during three commonly used training exercises, namely abdominal hollowing, bracing and pelvic tilt. The aim was to investigate the most suitable method of facilitating appropriate stabilisation patterns using an indirect measure of spinal stability, an air filled tri-sectional bag with a pressure transducer (PBU), to register changes in movement hypothesised to correspond to decreases in dynamic stability. Isometric resisted rotation of the trunk had previously been shown to promote a suitable co-contraction stability pattern and this was used as a standard (Richardson et al, 1990). Pressure sensor and EMG tracings were utilised and stability patterns were measured in reclined sitting and crook supine lying. The findings suggested that abdominal bracing and hollowing demonstrated similar activity levels, but abdominal hollowing showed activation patterns closest to that of trunk rotation. Conversely, backward or posterior pelvic tilting, which are both commonly prescribed exercise for patients with LBP, showed marked deviation from the ideal stability pattern. Although there are acknowledged methodological problems associated with EMG recording (Soderberg and Cook, 1984; Turker, 1993) these results lead to the adoption of the abdominal hollowing manoeuvre in the facilitation of a TrA contraction within a motor retraining programme.

The ability to maintain a selective isometric contraction of the anterolateral abdominals i.e. Internal Oblique (IO) and TrA, has been suggested to discriminate between LBP patients and those without LBP (Richardson et al., 1995). Tests of abdominal displacement or abdominal 'drawing-in', which is a more refined form of abdominal hollowing, purported to more selectively activate TrA and IO, have been used as an indirect measure of abdominal muscle function in a number of studies (Richardson et al., 1992; Richardson et al., 1995; Hodges et al., 1996; Cairns et al., 2000). A diagrammatic representation of an abdominal drawing-in manoeuvre is shown in Figure A 1 in Appendix A.

One fundamental problem with the re-training of deep-seated muscles is the problems associated with identification of a correct contraction (Richardson et al, 1998). It has been suggested that visual assessment does not give reliable information regarding the performance of TrA function, and measurement tools such as pressure biofeedback (PBU) provide only an indirect, albeit clinically useful, indication of function. Increasingly, real-time ultrasound is being suggested as the method of choice to provide feedback to both clinicians and patients (Hides et al, 1995a; Hides et al, 1998). However at present, this is not routinely available in physiotherapy clinical practice. Certainly, it is well-accepted that feedback is important during the teaching of any new skill and the effects of training utilising multisensory cueing¹³ on the recruitment of IO and TrA have shown positive outcomes (Miller and Medeiros, 1987). Integrated EMG values of IO and TrA after training (n=20 healthy males), consisting of a variety of multisensory cueing strategies used during five paced eccentric curl backs, were compared to those from a control group. The control group (n=20 healthy males) undertook abdominal muscle strengthening exercise that they frequently undertook with no cueing. EMG activity was tested in a crook supine position during the curl back phase of a curl-up exercise. The findings showed a statistically significant increase in the activity of IO oblique and TrA of the experimental group ($p < 0.001$). Clinically, these results are of interest as they indicate that multisensory cueing could positively affect the recruitment of the deep stabilising muscles and thus potentially improve spinal and pelvic stability. Their conclusion that sensory factors are important in the correct recruitment of TrA and other deep stabilisers concurs with other later investigations (Jull et al, 1993).

It is accepted that exercise and motor skill learning are not synonymous (Stevens and Hall, 1998). The need to be aware of the various motor learning strategies when incorporating specific exercises within the rehabilitation of LBP patients is of importance as identified by Miller and Medeiros (1987). As spinal stabilisation training involves motor retraining, the key points of motor learning theories should be acknowledged and applied (Stevens and Hall, 1998). *Transfer-appropriate processing*

¹³ Multisensory cueing including auditory, tactile, visual and kinaesthetic input

involves accurately identifying and simulating potentially harmful activities and then using the treatment technique within these situations i.e. using co-contraction of trunk and abdominal muscles during provocative activities. Secondly, *contextual interference* involving practice sessions, which are cognitively challenging, i.e. random sequencing of practice and finally, *self-evaluation*, involves the patient giving an evaluation of practice before external feedback. In forcing the learner to focus on motor patterns and repetitive self-evaluation, awareness of inappropriate habitual motor patterns is increased and it has been suggested that use of these techniques may facilitate transfer of motor re-learning skills into functional settings. When this is considered in relation to the proposed cause of pathology as repetitive faulty motion (Sahrmann, 1993), it becomes clear that the use of spinal stability re-education needs to be placed in a motor relearning context and, subsequently the exercises need to be undertaken regularly and accurately if any change is to be achieved.

In the same way that exercise and motor skill learning are not synonymous, teaching and learning are also two separate concepts. There is an obvious need for physiotherapists to be able to accurately teach a relatively isolated contraction of TrA if segmental spinal stabilisation training is to be clinically usable. The sensitivity and specificity of 52 therapists, who regularly used stabilisation techniques, to identify a correct TrA contraction and different substitution strategies¹⁴ was tested by Glass et al. (1997). Six exercises, on four separate occasions, over an eight-day period undertaken by three simulated patients were assessed. Overall, sensitivity for correct identification was 80% and specificity 92% with slightly better results in four-point kneeling than in prone (83% compared with 76% sensitivity and 93% compared with 90% specificity respectively) (Glass et al., 1997). Although a well-designed study with a high degree of clinical applicability, limitations lie in the reliance on EMG, video analysis and expert-opinion rather than a more quantifiable method such as real-time ultrasound. However the overall conclusion was that physiotherapists familiar with the concept of spinal stabilisation training and teaching of isolated anterolateral abdominal contraction could differentiate accurately between correct and incorrect patterns of

¹⁴ Substitution strategy- aberrant or unwanted muscle action

TrA activation. This suggests that in planning trials incorporating stabilisation training, clinicians involved should be adequately trained and familiar with the techniques in order that the results can be confidently applied to the clinical setting.

Similarly, proprioception and postural awareness are both important elements in the rehabilitation and motor re-learning in subjects with LBP. Protection against injurious insults to the spine requires proper anticipation of events, appropriate sensation of body position (proprioception including joint position sense), and appropriate muscular responses. Delayed lumbar muscle responses to sudden loads are known to occur with lumbar fatigue (Wilder et al., 1996; Taimela et al., 1999). It may be a combination of failure in the sensation of position and output of the response that is responsible for this delay, but either way it is of great importance clinically in the rehabilitation of LBP using muscle re-education training. Numerous studies have shown that the ability to accurately re-position spinal postures in a healthy population is high (Jakobs et al., 1985; Taylor and McCloskey, 1990; McGlashen et al., 1991; Maffey-Ward et al., 1996; Swinkels and Dolan, 1998; Swinkels and Dolan, 2000) but impaired in the presence of LBP (Maffey-Ward et al., 1996; Gill and Callaghan, 1998; Brumagne et al., 1999; Taimela et al., 1999). These findings are of importance in relation to segmental stabilisation training as the ability to maintain a neutral lumbar spine during endurance and motor-control re-training of the deep abdominal and trunk muscles is fundamental to the underlying concept.

Examinations of joint position sense in healthy subjects ($n=20$) have reported reproducibility of spinal postures in the upright position of 3.79° (2.56°) in the sagittal plane and 2.26° (1.59°) in the coronal plane over a two-week period (Swinkels and Dolan, 1998) and 5.27° (3.47°) and 3.70° (2.62°) respectively in flexed postures. Very similar results were obtained by Gill and Callaghan (1998) for standing flexion in the sagittal plane, which reported a mean deviation of 4.45° for a group of 20 non-LBP subjects and statistically significant differences in the lumbar proprioception of individuals with and without LBP. However, a number of demographic and selection differences did exist, with the patient group having a significantly higher mean age (43.3 years; range 21-74 compared to 32.9 years;

range 24-53) and the healthy subjects were drawn from hospital physiotherapy staff compared to the LBP subjects drawn from a patient population (Gill and Callaghan, 1998).

As one of the fundamentals on which segmental stabilisation training is based is the increase in fatigue of the core stabilizing muscles in the presence of LBP (Nicolaisen and Jorgensen, 1985; Roy et al., 1989; Roy et al., 1990; Cooper et al., 1993) the effects of such fatigue are important in relation to spinal joint position sense and proprioception. This has been tested in 106 subjects (57 with LBP, 49 healthy controls) in relation to the ability to react to rotational lumbar movement at 1°/sec before and after a fatiguing procedure (Taimela et al., 1999). They reported that patients with chronic LBP had a significantly poorer ability in sensing position change, concluding that lumbar fatigue induced significant impairment to sense of position change in both groups and supporting previous findings reporting diminishing control and co-ordination of the neuromuscular system associated with fatiguing muscle activity (Parnianpour et al., 1988). This work supports the anecdotal clinical evidence that patients with LBP tend to have poor proprioceptive awareness and also supports the suggestions that early stabilisation training may be considered more appropriately as 'awareness' training than specifically aiming for profound physiological changes (Hagins et al., 1999).

With the evidence that spinal repositioning sense plays an important role in the protection of the spine from damaging forces and positions, and that this is reduced in patients with LBP, there is an obvious need for a kinesthetic test for the lumbar spine, which could be applied relatively easily, in the clinical setting. This has been investigated by Maffey-Ward and co-workers (1996) using the 3 Space Fastrak. The error within and between days, of healthy subjects (n=10) to reproduce a neutral lumbopelvic position was examined with a small mean repositioning error for flexion/extension reported over the three repetitions (2.6° (1.2) and 2.6° (1.7) for day 1 and 2 respectively) which were well within the ranges described by other authors for various asymptomatic joint complexes. No statistically significant difference between days was demonstrated. Despite the obvious limitations due to the small number of subjects, these results provide a basis for further evaluation of this test on patients with low back pain to investigate

its ability to detect any kinesthetic deficit. This work, and that on repositioning sense described earlier, is of importance as the ability to achieve and sustain accurate spinal positioning is key to the underlying concept of specific endurance and motor-control re-training of the deep abdominal and trunk muscles. Consequently, it is important that any trials utilizing spinal stabilisation training incorporate an element of spinal positioning within the training programme.

The preceding sections have outlined the origins, research and development of the concepts behind segmental stabilisation exercises and summarized the current state of knowledge. The next section develops this further by presenting a review and critique of the clinical studies, specifically RCTs, which include spinal stabilisation training as part of the treatment interventions in the methodological design.

2.15 Clinical studies

The number of studies investigating the effects of spinal stabilisation exercises have increased slowly over the last few years, but have often suffered from poor methodological quality and reporting. However, more recent studies have shown improved quality and design with a corresponding increase in the generalizability of the results.

An early study in 1977 used the concepts put forward by Janda (1977) regarding the postural effects of muscles and incorrect movement, in combination with various forms of manual therapy (Bergquist-Ullman and Larsson, 1977). In a RCT design, both 'combined therapy' and back school were found to be superior to 'placebo' short wave diathermy in respect to shorter duration of symptoms (and sick leave) during the initial episode of pain. However, no significant differences between the three groups were found for number of recurrences over a one-year period following treatment. Although well designed, including a stratified randomisation process and very detailed descriptions of the three group interventions, limitations are evident. The two active treatment interventions both included exercises, with the back school including 'isometric abdominal muscle exercises' and, therefore, the two groups may not have been sufficiently different. However this work formed the basis for future studies and set a good standard.

The effect of training the anterolateral muscles over a 10-week period in painfree subjects has been investigated with findings indicating an increase in the speed at which subjects were able to generate intra-abdominal pressure (Cresswell et al., 1994a). These results indicated that abdominal muscle functioning could be influenced by a specific training programme and were replicated by O'Sullivan et al. (1998) who reported a marked improvement in the anterolateral synergistic function and importantly the automatic recruitment of muscles during loaded tasks (O'Sullivan et al., 1998). As part of this last study, 44 patients with known spondylolysis and spondylolisthesis were randomised to either specific training of the deep abdominal muscles and co-activation of the LM, or treatment directed by their treating therapist¹⁵ (O'Sullivan et al., 1997c). The specific exercise group followed a structured 10-week programme, utilising an abdominal drawing-in manoeuvre (See Figure A 1), encouraging specific contraction of the deep abdominal muscles and co-contraction of the LM proximal to the pars defect (Richardson and Jull, 1995). In keeping with the theory of specific stabilisation training, patients were encouraged to integrate the holding patterns into functional daily activities known to provoke symptoms with daily exercise programmes designed to take approximately 15 minutes. Outcomes included measures of pain (SFMPQ), function (ODI), physical parameters (electronic inclinometry) and muscle functioning (surface EMG of abdominal muscle activation). At discharge, there was a statistically significant difference between the two groups, in favour of the stabilisation training group, for change scores on all outcomes except lumbar spine range of movement. At 3, 6 and 30 months following treatment, the improvement was maintained for pain descriptors and intensity scores, and function. Despite a number of methodological limitations, including the relatively small sample size increasing the possibility of a Type II error and no evidence of any pilot study or sample size calculations, this trial, with very long term beneficial results provided good initial evidence for the usefulness of these types of exercises with the specific sub-group of the LBP studied. In a second paper, based on the same cohort (O'Sullivan et al., 1998), the results of the surface EMG recordings indicated that the relative ratios of IO (known to co-activate with TrA) to RA activity were higher in both conscious activation (abdominal drawing in

¹⁵ Including, deep massage, local heat and ultrasound therapy

manoeuvre) and, perhaps more importantly automatic activation (double straight leg raise). Although the significant difference demonstrated between the treatment groups was only present when non-normalized data were analysed, the accepted problems associated with the normalization of EMG data may have influenced this result. Despite this, however, the finding that specific training of the deep abdominal and trunk muscles may result in changes in their automatic functioning provided important support for training in this group of LBP patients and the basis for further, larger scale trials. In a third paper, based on this cohort, the results of a comparison between 12 physically active chronic LBP (CLBP) patients¹⁶ and 10 matched controls to selectively activate their deep abdominal muscles were reported (O'Sullivan et al., 1997b). Using a crook supine position, subjects performed an abdominal drawing in manoeuvre, with surface EMG recordings of RA and IO. The control group showed significantly greater activation of IO compared to RA than the CLBP subjects during the abdominal drawing in manoeuvre, but no difference between the groups were demonstrated except for the ratio of activation (IO to RA). These findings, in accordance with other reports (O'Sullivan et al., 1997a), suggest a dysfunction in the deep abdominal muscles associated with CLBP, continues to exist despite regular, high levels of activities in these patients, supporting the role of specific training to address the dysfunction. Limitations of sample size preclude widespread application of these results. However, they provided direction for future research and potentially clinical interventions in similar patients groups.

Two studies, which have contributed to the current knowledge regarding the effects of neuromuscular dysfunction associated with LBP have reported both reductions in functional disability and, possibly more importantly, marked reductions in recurrence rates (Hides et al., 1996; Hides et al., 2000; Hides et al., 2001). In their original study, Hides and colleagues (1996) randomly allocated 39 patients with sub-acute, first-episode LBP and unilateral segmental inhibition of the

¹⁶ Defined as undertaking a minimum of three thirty-minute aerobic activity sessions per week

LM muscle, to receive either medical treatment¹⁷ or exercise therapy¹⁸ over a four-week period. At the end of the treatment period, almost total remission of symptoms and marked decreases in disability were reported, with only 4/35 patients reporting some residual painful symptoms and the mean RMDQ score of 2.3 and 0 for the medical and exercise groups respectively. Ultrasound imaging showed a more rapid and more complete recovery of the LM, in terms of percentage difference in muscle CSA between the sides at the most affected level, in patients receiving the specific exercise training compared to the medical management. This finding was replicated at the 10-week follow-up. The importance of these findings is that despite almost full resolution of symptoms and a return to pre-morbid activity levels, atrophy remained in the LM of the group who had not received the specific training. This deficit would, in theory, place these patients at a higher risk of recurrence of symptoms in light of the *in vitro* evidence indicating the importance of the muscular system in stabilizing and protecting the spine (Panjabi et al., 1989; Goel et al., 1993; Wilke et al., 1995). Obvious limitations of this trial included low sample size and the lack of reporting of power calculations but additionally, no direct causal relationship can be drawn regarding recovery of CSA and reduction of neuromuscular deficits as no functional measure of muscle function was included. Despite this, however, this study furthered the understanding of how spinal muscles react to pain and specific rehabilitation. It had been long known that patients with LBP show aberrant activity in their back muscles, but this study provided evidence that in non-surgical, acute, first episode LBP, rapid atrophy occurs in the LM muscle. The cause of this was suggested to be reflex inhibition¹⁹ and the rapid onset and persistence, despite return to normal activity, compared to the slow onset of disuse atrophy found in CLBP, indicated that interventions to address this dysfunction should be instigated early within the management of LBP. The long-term follow up to this study (Hides et al., 2000; Hides et al., 2001), reported ongoing problems in the group in

¹⁷ Medical management including advice on bedrest (1-3 days), absence from work and prescription medicine

¹⁸ A therapeutic exercise designed to re-educate the LM muscle in its stabilizing role, including isometric LM contraction with the deep abdominal muscles as based on Richardson and Jull (1994 a and b) and assisted by the use of real-time ultrasound to ensure adequate activation of the LM. Patients were treated twice a week for a four-week period.

¹⁹ Sensory stimuli impedes voluntary contraction of a muscle leading to weakness and rapid muscle atrophy. Reflex inhibition can occur in the absence of pain

whom a reduced CSA of LM was present 10-weeks following onset of symptoms and this will be discussed in the next section.

The findings of persistent rapid onset muscle atrophy despite resolution of symptoms (Hides et al., 1996) and the persistence of LM wasting despite regular high level activity in CLBP (O'Sullivan et al., 1997a), have been supplemented by recent findings indicating that the duration of symptoms has a significant effect on muscle fibre type distribution but not gross cross sectional area (CSA) (Mannion et al., 2000). In a biopsy study of the paravertebral muscles of 59 chronic LBP patients (mean duration of symptoms of 157 (124) months), when gender and age were controlled for, symptom duration showed a significant influence on fibre type distribution. Findings indicated that prolonged symptoms were associated with a significant increase in 'glycolytic' type IIX (fast twitch) fibres and a decrease in the proportion of type I (slow twitch) fibres (See Table 2-7), supporting previous studies which found increased fatigability of these muscles. Duration of pain had no significant influence on fibre size or gross CSA with the gross muscle CSA influenced most strongly by lean body mass. This contrasts with some previous studies examining CLBP and muscle atrophy, but the use of internal standards as a measure of muscle CSA e.g. vertebral end-plate size, which lack proven validity, may account for the contrary but inconclusive findings. Acknowledging the limitations of a lack of matched controls, Mannion et al. (2000) suggest that rather than selective atrophy or alteration in size of specific muscle type, fibre type transformations may occur in the back muscles of chronic LBP patients. Importantly, these suggested fibre transformations indicate that with prolonged stimuli muscle fibre type distribution is malleable (Mannion et al., 2000) and therefore, importantly for treatment interventions, presumably reversible. Additionally, as these changes are time related, these findings support the early instigation of active treatment programmes such as spinal stabilisation training in order to minimize changes and therefore resultant disability.

Using the same cohort described in the earlier section (Hides et al., 1996), a questionnaire, administered via the telephone by an independent observer, was used to assess recurrence rates one and three years after treatment (Hides et al., 2001). Recurrence rates for the exercise group were reported as 30% and 35%, and

for the medical treatment group as 84% and 75% for the one year (n=39) and 2-3²⁰ year follow-up respectively (n=36). Consequently, patients in the medical treatment group were 12.4 times more likely to experience a recurrence in the first year and 9 times more likely in year 2-3 (reduced to 5.9 using a best-case analysis) (Hides et al., 2001). Additionally, the researchers enquired about length and severity of episodes and treatment sought, reporting that recurrence episodes “as severe” as the original, were reported by 9/16 (56%) of the control group and 2/6 (33%) for the specific exercise group in year two-three, but interesting were similar for the 2-3 year-point at 2/12 (17%) and 1/7 (14.2%) respectively.

Similarly, there was little difference between the percentage of patients seeking treatment in both groups, which was 25% for the medical treatment group and 20% for the specific exercise group. Although encouraging, application of these results is severely limited for a number of key reasons. Firstly, although a highly homogenous population, the sample studied represents a very small, specific subgroup of back pain patients and therefore these results cannot be generalised to a chronic, recurrent or even sub-acute population. Secondly, as discussed in Section 2.4 many problems exist when trying to both define, and measure recurrence and reliance on recall has obvious drawbacks. Additionally, no definition of recurrence was provided, making comparison with other work difficult. Thirdly, and possibly most importantly, the very limited numbers of patients involved, make the inevitable losses, all from the medical treatment group, potentially skew the results. Although only three patients were lost to follow-up (7.5%), all were in the medical treatment group, which equates to a 15.7% (3/19) loss from this group. Additionally, the rates of further treatment sought, (25% and 20% for the medical and specific treatment groups) at years two-three introduces obvious contamination. Finally, as with the previous study, no direct measure of muscle function was undertaken so any assumption of reduced neuromuscular functioning on the basis of reduced CSA at 10-week following onset of symptoms is pure extrapolation. Therefore, in summary, the study undertaken by Hides et al. (2001) represents a good start from which to investigate specific stabilisation training using properly powered studies with adequate sample sizes, and does

²⁰ Recurrence rate in the year following treatment, and during the second and third year following treatment.

indicate that addressing deep muscle dysfunction may have a beneficial effect in certain groups of LBP patients.

Due to the very specific nature of the training, spinal stabilisation exercises are often used on an individual basis. It has been suggested that a high level of specificity of contraction is necessary, at least during the early stages of training, to facilitate the isolated contractions identified to be fundamental to the training of deep stabilizing muscles (Richardson and Jull, 1994; Comerford and Mottram, 2001a). Consequently, the use of these techniques in a group or class setting may not be the best environment for this level of specificity. Despite this, a number of studies have looked at the use of stabilization-type training in a group setting. In a retrospective cohort study of patients with non-operative lumbar disc prolapse, Saal and Saal (1989) examined the outcomes of 52 patients who underwent an aggressive treatment programme including elements of spinal stabilisation training in a back school setting. Patients had MRI evidence of herniated nucleus pulposus, a primary complaint of referred leg pain, and a positive straight leg raise test. Patient records from a 10 month period were examined, 64 patients identified and mailed a standardised questionnaire²¹, and a 91% (58/64) response rate was obtained, 52 of which were non-operative patients. The self-reported outcome for 50 patients (96%) was 'Good' or 'Excellent' with 48 patients returning to work (92%) (Saal and Saal, 1989). Interestingly, although the stabilisation training was undertaken in a group environment, initially exercises were carried out on a one-to-one basis and then patients were advanced to a class situation as the authors identified the need for a meticulous technique. The trial by Saal and Saal (1989) was vital to the development of a research base for stabilisation training and was the springboard for much future work. Despite this however, as always, results must be considered within the limitation of the study design. As a retrospective cohort review, the study relied on a very selected sample and a high degree of recall from the patients involved with a mean post-care follow-up of 31.1 months. Additionally, although the aggressive stabilisation training formed the main component of the rehabilitation, one important potential confounder was the use

of epidural and/or nerve root blocks where indicated for pain control. The rationale was that this facilitated an active, not passive, treatment, which although sound in practice, does introduce a very large variable into the treatment process.

2.16 Limitations of studies to date and current research needs

From the preceding review, it can be seen that although the studies reviewed constitute a good foundation on which to develop the evidence base behind the use of spinal stabilisation exercises, there are some obvious limitations that need addressing. The majority of studies to date have suffered from inadequate sample size with few showing evidence of power or sample size calculations. Both of these are fundamentals for an RCT design (Pocock, 1983; Jadad, 2000). Additionally, although statistical significance is invariably quoted in all studies, very few discuss whether these findings constitute clinical significance. Additionally, as it has been extensively claimed that spinal stabilisation training is already an accepted part of the management of LBP (Richardson and Jull, 1994; Richardson and Jull, 1995; O'Sullivan et al., 1997a; Jull and Richardson, 2000), this suggests a need for trials evaluating the use of the techniques as used in clinical practice. As such, pragmatic designs are the methodology of choice. In contrast to explanatory trials, which aim to answer specific questions about efficacy (Wakefield, 2000), pragmatic designs measure effectiveness or the benefits that treatment or treatment packages produce in real clinical practice (Roland and Torgerson, 1998b). This last point, regarding treatment packages of care, is of importance as evaluating techniques in isolation does not represent current clinical practice, at least in Britain and Ireland (Moore, 1997/8; Foster et al., 1999). However, it must be excepted that evaluating techniques in relative isolation does assist in identifying which specific components of packages are the most effective. The study by Mannion et al. (1999) used such a design and demonstrated that group aerobic and stretching exercises and muscle strengthening/coordination using training devices were equally efficacious in reducing pain intensity and frequency, and disability in chronic LBP (Mannion et al., 1999), and by including a combination of self-report, physical, and physiological measures was able to then attempt to analyse what

²¹ Including questions from the ODI, pain rating, work status and self-rating outcome.

specific mode of action each therapy utilised (Mannion et al., 2001b; Mannion et al., 2001a; Kaser et al., 2001). Therefore, as with any trial, a decision must be made as to which design best suits the research question. Problems also exist regarding the level of expertise of clinicians participating in the provision of treatment in trials. Similarly, although there is a need to ensure interventions are provided by clinicians with adequate skill in the technique (Koes et al., 1995), the use of expert clinicians as has been used in certain trials (Hides et al., 1996; O'Sullivan et al., 1997c) may be inappropriate as it does not represent clinical practice since it has been shown that the majority of LBP patients tend to be treated by more junior physiotherapists (Foster et al., 1999). As such, this grade of therapist would not necessarily have the level of skill or expertise required to replicate the interventions as undertaken in these trials. Equally there is a need to ascertain the optimum level of skill required to effectively apply interventions as training over and above that level may not be justified on a cost basis. Again, a pragmatic design, whilst ensuring core competency and a basic level of skill, would avoid these problems and increase the applicability of results.

Although the trials to date have reported favourable results, these have used very specific subgroups of LBP patients, which have been specifically defined and identified e.g. radiological evidence of spondylolisthesis and acute, first episode LBP with ultrasound evidence of multifidus wasting. Although this is positive, and the identification of sub-groups of patients is key to successful treatment and homogenous groups for research purposes, it is acknowledged to be difficult in the clinical setting and the blanket application of results from studies using sub-groups to all LBP patients is both inappropriate and unethical. Therefore, there is a clear need for a pragmatic trial of stabilisation training, as currently used i.e. as part of a package of care rather than in isolation, which is appropriately powered to detect clinically significant changes in outcomes deemed to be of most importance.

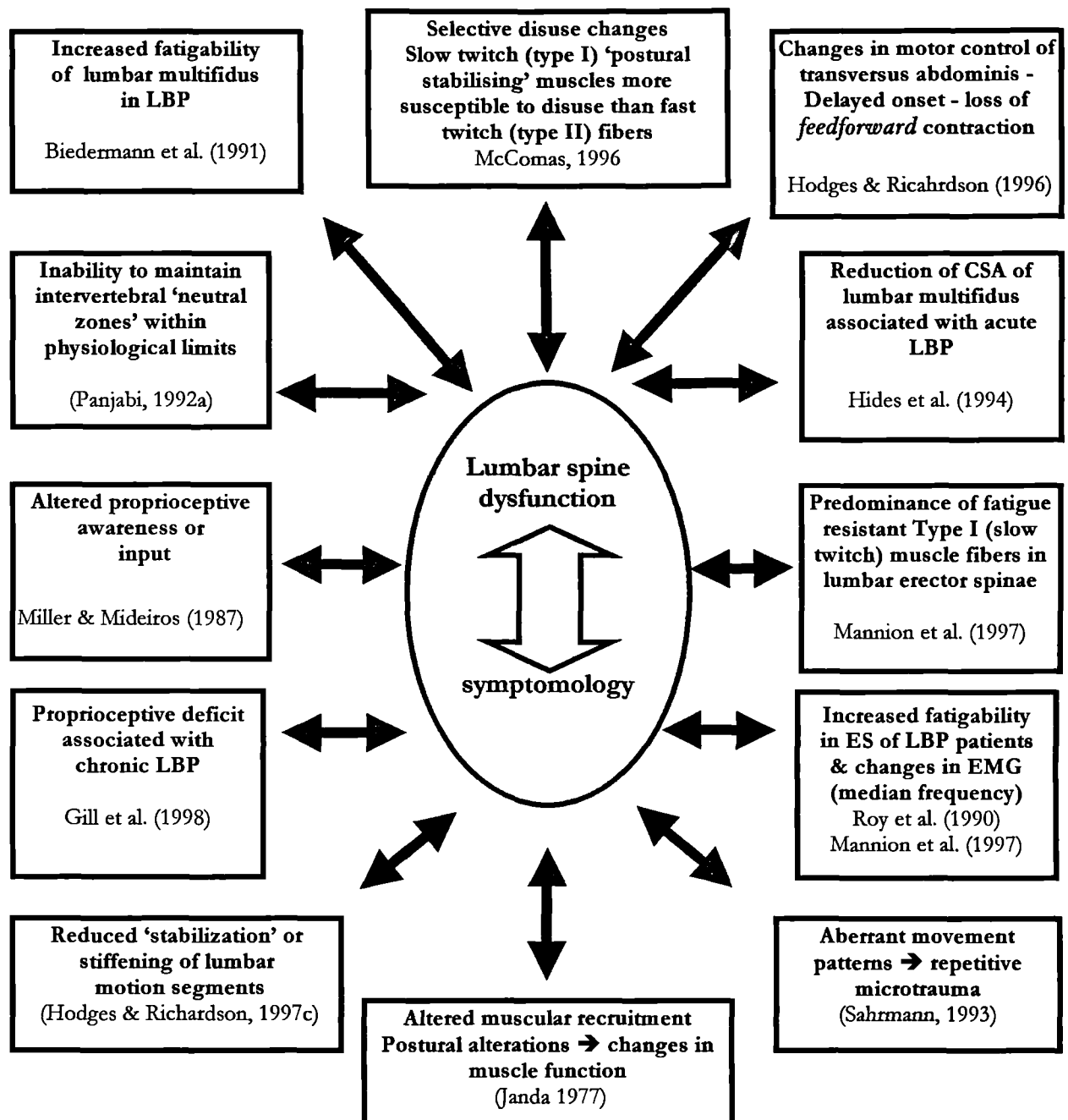
2.17 Summary

O'Sullivan (2000) maintains that successful management of LBP depends on accurate identification of the sub-group of the population who will respond to specific interventions such as spinal stabilisation training. However, observation of

current clinical practice would suggest that these techniques are being used within the general LBP population rather than just within sub-groups.

Diagnosis of clinical instability remains controversial and numerous clinical symptoms have been associated with the diagnosis of clinical instability, with varying degrees of acceptance within the medical and physiotherapeutic community. An example of such is displayed in Appendix J. Although controversy exists regarding the diagnosis, measurement and even the treatment of clinical instability, the multifactorial nature of neuromusculoskeletal control and subsequent dysfunction is unquestionable and a diagrammatic representation of some suggested contributing factors is illustrated in Figure 2-10. The multifactorial nature of low back dysfunction necessitates that trials designed to examine the effectiveness of spinal stabilisation training need to account for the multiple facets involved in the functioning, dysfunctioning and retraining of the core stabilising muscles as part of a 'spinal stabilisation' concept.

Figure 2-10: Diagrammatic representation of suggested neuromusculoskeletal factors contributing influencing spinal control and development of low back dysfunction ²²



²² Based on Cairns et al. (2000)
ES=erector spinae

Chapter 3 presents the details of the methodological design of a pragmatic, randomised controlled trial undertaken to examine the effect of incorporating specific spinal stabilisation exercise to conventional physiotherapy for a recurrent LBP population treated in a secondary healthcare setting. The development of the protocol, planned statistical analyses and data management are presented.

3. Chapter 3: Methodology

3.1 Research question/hypothesis

A number of clinical studies have examined the effect of specific stabilisation regimes on the outcome of patients with LBP. However, to date very specific subgroups of the LBP population, such as non-operative prolapsed discs, spondylolisthesis and acute first episode LBP, have been studied (Saal and Saal, 1989; Hides et al., 1996; O'Sullivan et al., 1997b; Hides et al., 2001). The need therefore existed to extend the current knowledge base and reflect the clinical application of these techniques to a general, non-specific back-pain population. The aim of the study was therefore to assess the relative effectiveness of the addition of specific spinal stabilisation training to conventional physiotherapy management of recurrent LBP compared with physiotherapy management alone. Recurrent LBP was chosen as the target population as this group have the potential to be high level consumers of healthcare resources.

3.2 Design

The design selected to answer the research question was a pragmatic, prospective, multi-centre, single blind, randomised controlled trial (RCT) with two treatment arms and six and twelve-month postal follow-up. The two treatment arms represented 'packages' of care designed to represent current clinical practice in the UK at the time the trial was designed. Both packages included conventional or current physiotherapy treatment techniques or modalities utilised at the clinicians' discretion to assist with treatment goals in the management of recurrent LBP. A full description is provided in section 3.3. The specific treatment intervention under investigation was spinal stabilisation training, designed to increase the endurance of the deep abdominal and trunk muscles with the aim of improving spinal stability and therefore assist in the management of symptoms associated with recurrent LBP. These exercises formed the basis of the second treatment arm, but in keeping with the pragmatic nature of the trial and the aim to reflect clinical practice, these were used in addition to conventional physiotherapy care directed at the discretion of the treating clinician.

Both treatment arms were delivered in a secondary healthcare setting with care provided free at the point of delivery via the National Health Service (NHS), i.e. two physiotherapy departments, of a University Teaching Hospital²³ and regional orthopaedic centre²⁴.

Multiple outcomes were chosen to reflect the wide impact that LBP can have on health and quality of life. In accordance with the recommendations of the Second International Forum on Low Back Pain Research in Primary Care, a combination of disease-specific and generic outcomes was chosen (Borkan et al., 1998). However, the primary outcome of back-related functional ability, as measured by the Roland Morris Disability Questionnaire (RMDQ), was identified (Roland and Morris, 1983a). All outcome measures used were self-completed questionnaires, which allowed for postal follow-up at six and twelve months following discharge from treatment. These two timescales were chosen to assess the short and longer-term impact of the two treatment interventions under investigation. A single-blind design was chosen as successful double-blinding requires that interventions are indistinguishable for both patients and investigators (therapists). This would obviously be impossible in the current trial and therefore single-blinding was used in which patients were blind to their treatment allocation (Jadad, 2000). Randomisation was undertaken using a minimization procedure and three stratifying variables, and was completed by an independent investigator. A full description of this procedure is given in section 3.6.1.

3.3 Treatment interventions

Although a pragmatic design had been chosen for the current study, strategies were included to reduce possible confounding factors. Treatments were undertaken on an individual basis at the discretion of the treating clinician, all of whom had undertaken appropriate training prior to the start of the trial. This is described in sections 3.8.2 and 3.8.3. Non-study clinicians did not treat any patients within the trial, and study

²³ Queen Elizabeth Hospital (QEH) and Selly Oak Hospital (SOH), Birmingham

²⁴ The Royal Orthopaedic Hospital (ROH), Birmingham

clinicians treated patients randomised to either treatment arm. In order to reduce the possible confounding impact of motivation issues associated with group work, referral to hydrotherapy²⁵ (Skinner and Thomson, 1994) and Back School²⁶ was not allowed.

A standardised education booklet, *The Back Book* (Roland et al., 1996) was provided for all participants in order to promote the evidence-based message of early return to normal activity.

Patients were randomised into two groups as follows:

- i. Conventional physiotherapy treatment alone (CT)**
- ii. Conventional physiotherapy treatment and specific spinal stabilisation training (SSSE)**

i) Conventional physiotherapy treatment alone (CT)

Interventions within the CT group consisted of any treatment method currently used within the scope of practice outlined by the Chartered Society of Physiotherapy for State Registered Physiotherapists (CSP, 1996) (See table 3-1). Any exercises utilising, low load, high repetition muscle activity were excluded as this forms the basis of the spinal stabilisation intervention. All participating departments had actively adopted the recommendations of the CSAG (CSAG, 1994a) and directed all LBP treatment accordingly.

²⁵ Therapeutic use of water e.g. water based exercise. A detailed description is provided in Skinner and Thomson (1994)

²⁶ Back School representing a group treatment session including advice on the aetiology and anatomy of LBP, aerobic circuit training +/- psychological input regarding coping strategies etc.

Table 3-1: Description of treatment modalities used

Modality	Description
1 Manual therapy	<ul style="list-style-type: none"> Any manual therapy technique applied to the musculoskeletal system such as Maitland mobilizations (Maitland, 1986), Cyriax soft-tissue mobilizations or manipulations (Cyriax, 1984), Kaltenborn (Kaltenborn, 1970) or NAGS and SNAGS (Mulligan, 1989) This included both high and low velocity manipulations and mobilisations Management of neural tissues or neural tissue tethering (Butler, 1991)
2 Active exercises	<ul style="list-style-type: none"> This included any exercise-based regime including the McKenzie approach (McKenzie, 1981) to treatment Aerobic based exercises designed to increase cardiovascular function or general musculoskeletal conditioning
3 Lumbar traction	<ul style="list-style-type: none"> Mechanically applied lumbar traction either on a static or rhythmic basis
4 Postural or ergonomic advice	<ul style="list-style-type: none"> Any advice regarding spinal positioning, seating, work station or work-place set-up Included suggested reduction of specific activities
5 Electrotherapy	<ul style="list-style-type: none"> Pain relief from any electrotherapeutic technique contained within the CSP guidelines for practice (CSP, 1991) e.g. TENS¹
The protocol excluded use of 'Back School' or hydrotherapy	
¹ Transcutaneous electrical nerve stimulation	

ii) Specific spinal stabilisation group (SSSE)

Interventions in the SSSE group consisted predominantly of endurance training for the deep abdominal (transversus abdominis) and back extensor (lumbar multifidus) muscles as outlined by Richardson and Jull (1995). A guideline was provided (See Appendix K) but treatment was individualised for each patient. A specific booklet was developed which explained this approach and was given to all patients randomised to the SSSE arm of treatment (Appendix L). The booklet outlined the anatomy and function of the muscles to be trained and the concept of endurance rather than strength training. Standard physiotherapeutic techniques as used in the CT group could be used for symptomatic relief, in order to facilitate the spinal stabilization-training regime. However, the main component of this treatment package was the specific spinal stabilisation training. Patients randomised to the SSSE group were given a copy of the *Spinal Stabilisation Exercise Booklet* (Appendix L)

designed specifically for the trial. They were instructed to read it before their next appointment as it formed the basis of their treatment.

The individual components of both treatment arms were recorded for each patient.

3.3.1 Development of guidelines for stabilisation interventions

In order to replicate current clinical practice as closely as practicable, a pragmatic RCT design was adopted. Clinicians were given an outline treatment programme against which to direct treatment, rather than a rigid protocol. This was designed to reflect the individual needs of patients and the differing activation and facilitation methods required. The outline treatment for the SSSE group is shown in Appendix K, with the study protocol requiring spinal stabilizing techniques to be started by the second treatment session. This guideline was based on the available evidence for the literature on activation patterns and progression of treatment (Richardson and Jull, 1995).

3.3.2 Categories of treatment modalities

Five groups of treatment modalities were developed following review of the literature and discussion with the clinicians involved in treatment. These are summarized and described Table 3-1. Treatment dates and techniques used were recorded for each patient and ranked on a one to five scale where one represented the most frequently used technique and five the least frequently used. Subjective patient compliance with prescribed exercises was recorded on the treatment data collection sheet by the treating clinician.

All initial assessments were one-hour in length and follow-up treatment sessions were thirty minutes long. Following discussions with other researchers and difficulties arising during the running of the pilot study (Section 3.9), the maximum number of treatment sessions was extended to 12. Similarly, the maximum timescale for treatment was extended from 10 to 12 weeks i.e. patients in both treatment groups had to receive all treatments within a period of twelve weeks and could consist of a maximum of 12 treatment sessions. This was primarily to take account of the natural

breaks that had occurred during treatment in the pilot trial such as staff or patient holidays.

3.3.3 Standardized advice and educational literature

Standardised educational information was provided for all patients in the form of *The Back Book* (Roland et al., 1996). As this book presented the latest evidenced-based advice and was designed as an adjunct to the overall management of LBP, rather than use in isolation (Burton et al., 1999), it was considered appropriate for use in the two treatment packages. Once full consent had been obtained a copy of *The Back Book* was issued to each participant with the verbal instruction:

“This booklet contains advice about back pain based on the latest research. It would be useful if you could read it before your next appointment, and your physiotherapist can answer any questions you might have.”

3.3.4 Diagnostic real-time ultrasound

In order to accurately assess the impact of the spinal stabilisation regime, it was essential that study clinicians could ensure that patients within the SSSE group could adequately activate the key necessary muscles i.e. transversus abdominis and lumbar multifidus. The development and increasing availability of real-time diagnostic ultrasound has led to an expansion of its use for more than simply evaluation of pathomorphological changes (Hides et al., 1995b). Advantages over other imaging techniques such as magnetic resonance imaging and computer tomography are clear including reduced cost, lack of exposure to ionising radiation and relatively easy accessibility (Hides et al., 1995a; Schwartz et al., 1999). Although more dependent on operator expertise than other procedures, ultrasound imaging has been suggested to be useful for both direct measurement of muscle size and visual feedback for re-education of specific muscle activity (Hides et al., 1998). Although all treating clinicians were experienced in this form of muscle training, the aim of the trial was to assess the effect of the addition of teaching specific spinal stabilisation training exercises, to ‘usual’ physiotherapy care for recurrent LBP. It became evident in the planning of the trial that in order to achieve the aim of the trial, a facility was needed

to assess the required muscle activation. Ultrasound imaging was considered the most appropriate method and this facility was arranged at the main trial centre (ROH). A training session for the primary researcher was undertaken regarding the use of the specific unit with an experienced ultrasonographer. All referrals from trial clinicians for ultrasound sessions were made directly to, and undertaken by the primary researcher. All sessions were undertaken using either a 3MHz or 5MHz head, dependent on the clinical indications. Of the total sample of 97 patients, only 4 ultrasound sessions were requested. After a review of the literature and discussion with the trial clinicians, a policy was agreed regarding referral for a single ultrasound session to assist patients who found activating transversus abdominis or lumbar multifidus difficult or impossible. This is shown in Appendix M.

3.4 Sample selection criteria

This section describes the inclusion and exclusion criteria for the trial and justification for their use. It also outlines the recruitment pathways and processes used to enrol patients into the trial. For clarity, participants in the trial are referred to as patients throughout this section. This reflects the pragmatic nature of the trial i.e. trial patients were treated in NHS secondary care setting alongside non-trial patients, and was deemed more appropriate than the use of 'subject' which may be more appropriate for experimental, non-clinical settings. Additionally, it was considered that all patients seeking healthcare and especially physiotherapy input should be active participants in their treatment.

3.4.1 Inclusion and exclusion criteria

The main inclusion criteria was that patients must have experienced at least one previous episode of LBP necessitating alteration in normal activities or for which medical care/intervention had been sought. This definition was based on previous work (Hodges and Richardson, 1996; Cairns et al., 2000) and attempted to obtain a sample of recurrent LBP patients rather than a chronic population with unvarying symptoms. The definition of LBP was based on the relevant literature with pain having to be present in the area between the lower margin of the 12th rib and the gluteal fold with or without radiating leg pain (Croft and Raspe, 1995; Waddell, 1998).

Pain had to be of a non-specific, mechanical nature (See exclusions on page 91) with or without leg pain. Patients were required to be sufficiently proficient at English in order to allow them to independently complete the questionnaires used within the trial without the assistance of a translator.

The age range of participants for the main trial was 18 to 60 years. This was reduced from 18 to 65, which was used in the pilot study. The upper limit was lowered in order to assist comparison with other work in this area, which has used a similar range (Moffett et al., 1999; Dolan et al., 2000). Additionally, there is evidence that age can exert an independent influence on muscle structure (Mannion et al., 2000) and that the proportion of muscle occupied by different fibre types is affected by age (Lexell and Downham, 1992).

Following the pilot, a minimum functional disability level as measured by the RMDQ was introduced as an inclusion criterion. Evidence from the literature has indicated that a change score of 5 points is required in order to achieve a minimal clinically important difference (MCID) across the entire RMDQ scale (0-24) (Stratford et al., 1996a). A minimum score more than 4 was therefore chosen in order for a MCID to be achievable in all patients.

The last inclusion criteria concerned distress as measured by the Distress Risk Assessment Method (Main et al., 1992). This has been described in detail in Chapter 2, and consists of two questionnaires, the Modified Somatic Perception Questionnaire (MSPQ) (Main, 1983) and the Modified Zung (Zung et al., 1965) (See Appendix G and Appendix F). The scores from these two questionnaires are used to place patients in one of four mutually exclusive categories (See Table 2.6 in Chapter 2). For the purposes of the current trial, only patients classified as showing no evidence of distress, or those at risk of becoming distressed, were included.

Exclusion criteria were kept to a minimum in order to preserve the pragmatic nature of the trial and allow results to be generalisable to the widest population and are summarised in Table 3-2. They were based on a review of the relevant literature, clinical guidelines and current clinical practice at the time the trial was designed. The

presence of red flags as defined in the CSAG guidelines (1994a) e.g. widespread neurology, unexplained weight loss etc, were routinely used clinically to screen for potentially serious spinal pathology and therefore were an obvious exclusion for the trial, as were cauda equina symptoms²⁷.

Table 3-2: Summary of exclusion criteria

Red Flags	Evidence of cauda equina compression
	Non-mechanical LBP
	Acute nerve root compression with new neurological loss
Surgical	Abdominal surgery (last 12 months), Any spinal surgery
Medical	Systemic illness Neurological or muscular degenerative disorders
Other	Pregnancy or less than one year post-partum Psychological distress (as measured by the DRAM)

Clinical presentations suggestive of acute objective motor radiculopathy or nerve root compression²⁸ with new or progressive neurological loss were excluded in line with the CSAG and RCGP guidelines (CSAG, 1994a; RCGP, 1996b) as these conditions required differential diagnosis. Objective sensory radiculopathy²⁹ was not excluded (Borenstein et al., 1995).

The surgical exclusions of abdominal surgery within the last twelve months and any previous spinal surgery are in line with other research in this area. The treatment arm undertaking specific spinal muscle retraining would be required to follow an intensive

²⁷ Difficulty with micturition, loss of anal sphincter tone or faecal incontinence or saddle anaesthesia around the anus, perineum or genitals.

²⁸ Well localized pain extending down the leg in a dermatomal distribution to below the knee, +/- paraesthesia or anaesthesia and with alteration in reflexes or myotomes. The pain is lancinating, shooting or sharp (Borenstein et al, 1995)

abdominal endurance programme, which would be contraindicated following recent abdominal surgery. Patients who had undergone previous spinal surgery were excluded, as a non-surgical LBP sample was required.

Two medical exclusions were applied to ensure that all patients entering the trial had the physical ability to actively participate in either of the treatment packages, which were being compared. Patients with neurological and muscular degenerative disorders e.g. cerebral-vascular accident or muscular dystrophy, were excluded as one arm of the trial was based on muscle retraining, and the ability to fully undertake such training would be compromised by these types of conditions. Similarly, systematic illnesses e.g. rheumatoid arthritis, were also an exclusion criterion for the same reason.

Female patients who were pregnant at the time of entry to the trial or who were less than one-year post-partum were excluded from participation. This was for a number of reasons. Firstly, as with recent abdominal surgery, pregnancy would contraindicate participation in an intensive abdominal training programme. Secondly, it is widely accepted that in the immediate post-partum period, there is increased ligamentous laxity due to hormonal changes. This would potentially act as a confounding variable and may have compromised comparability between groups.

As previously stated, no evidence of distress, as measured on the DRAM was one of the inclusion criteria for the trial, therefore any patients classified into one of two categories of distress identified by the DRAM (distressed depressed or distressed somatic) were excluded. This was based on strong evidence from the literature indicating that LBP patients who are distressed have an increased risk of poor outcome with physical treatment alone (Main et al., 1992; Hope and Forshaw, 1999). As the aim of the current trial was primarily to investigate the addition of one physically based exercise intervention, compared to a 'usual care' physical

²⁹ Paraesthesia or anaesthesia in a nerve root distribution with objective changes on sensation testing, with or without radicular leg pain. This presentation would be classified as QTF 4.

intervention, it was considered inappropriate to enrol patients with a known high risk of poor outcome with such treatment interventions.

3.5 Ethical approval and consenting process

Ethical approval was sought and obtained from the South Birmingham Local Research Ethics Committee (LREC) prior to the start of the pilot study. Ethical approval was granted dependent on further clarification of four points. The amendments made were that patients were to be given a copy of their consent form and a more detailed explanation of 'randomly allocated' was included on the patient information sheet. Further clarification of one of the exclusion criteria was required (new neurological involvement) and the committee wanted information regarding blinding of follow-up assessment. As all outcome assessments were self-administered by the patient, this last point was not applicable. These points were all clarified and ethical approval was granted for the main trial (LREC Reference number 0281).

All patients were approached regarding participation in the trial by the lead researcher at their initial assessment at one of the three centres. This process is described fully in the next section. All patients recruited to the trial provided written informed consent prior to their inclusion (Appendix N). Recruitment of patients to the trial is shown in Figure 3-1. A full verbal explanation of all aspects of the trial was given including what participation would involve and the length and amount of follow-up required. This was accompanied by full written information (Appendix O). The consent form outlined clearly the option for withdrawal at any point from the trial.

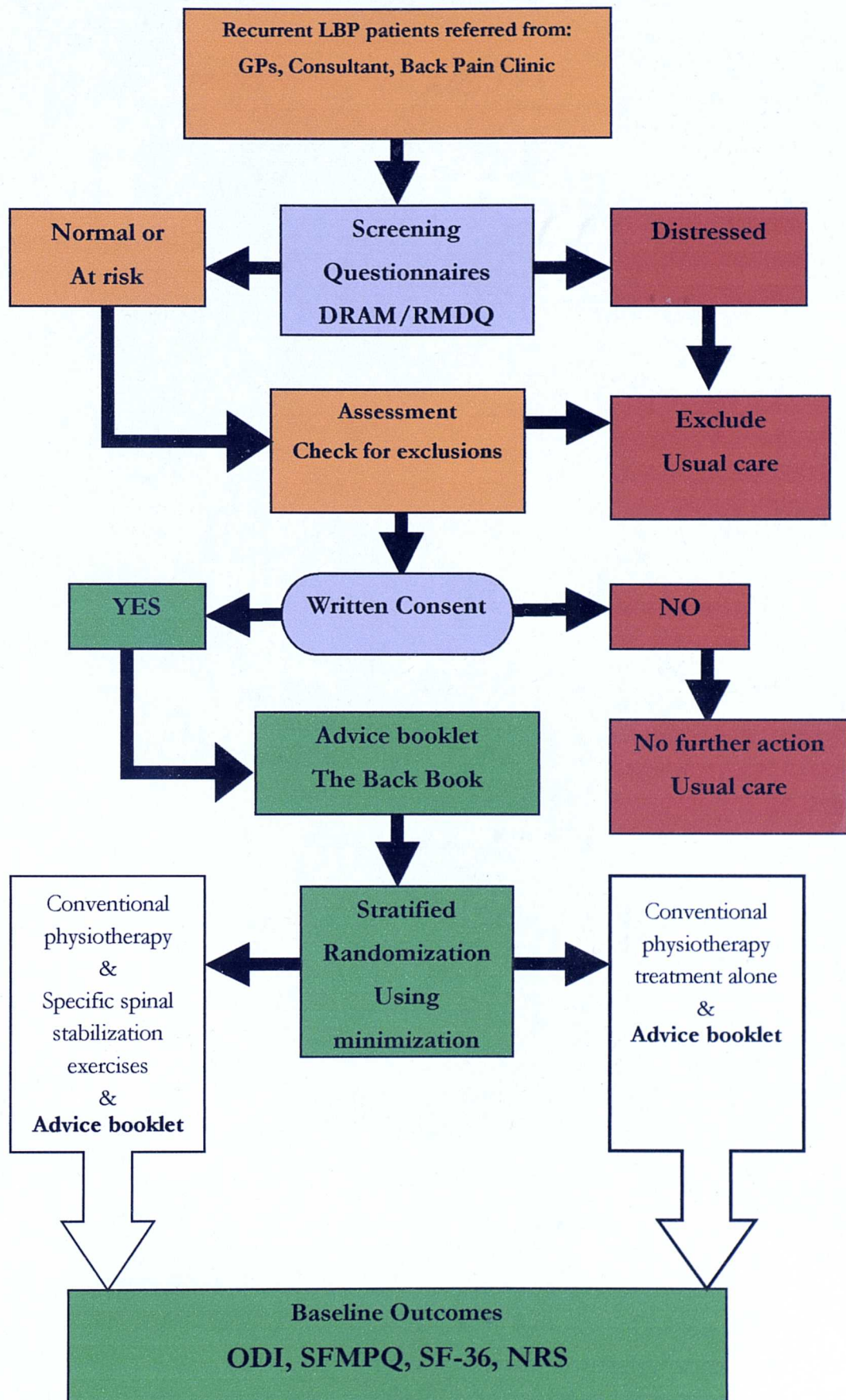
In order to facilitate recruitment, a trial was undertaken with treating clinicians consenting patients directly. This however did not prove successful due to the increased time commitment required to undertake this task. It was therefore decided to continue with the primary researcher consenting all patients.

3.6 Recruitment of patients

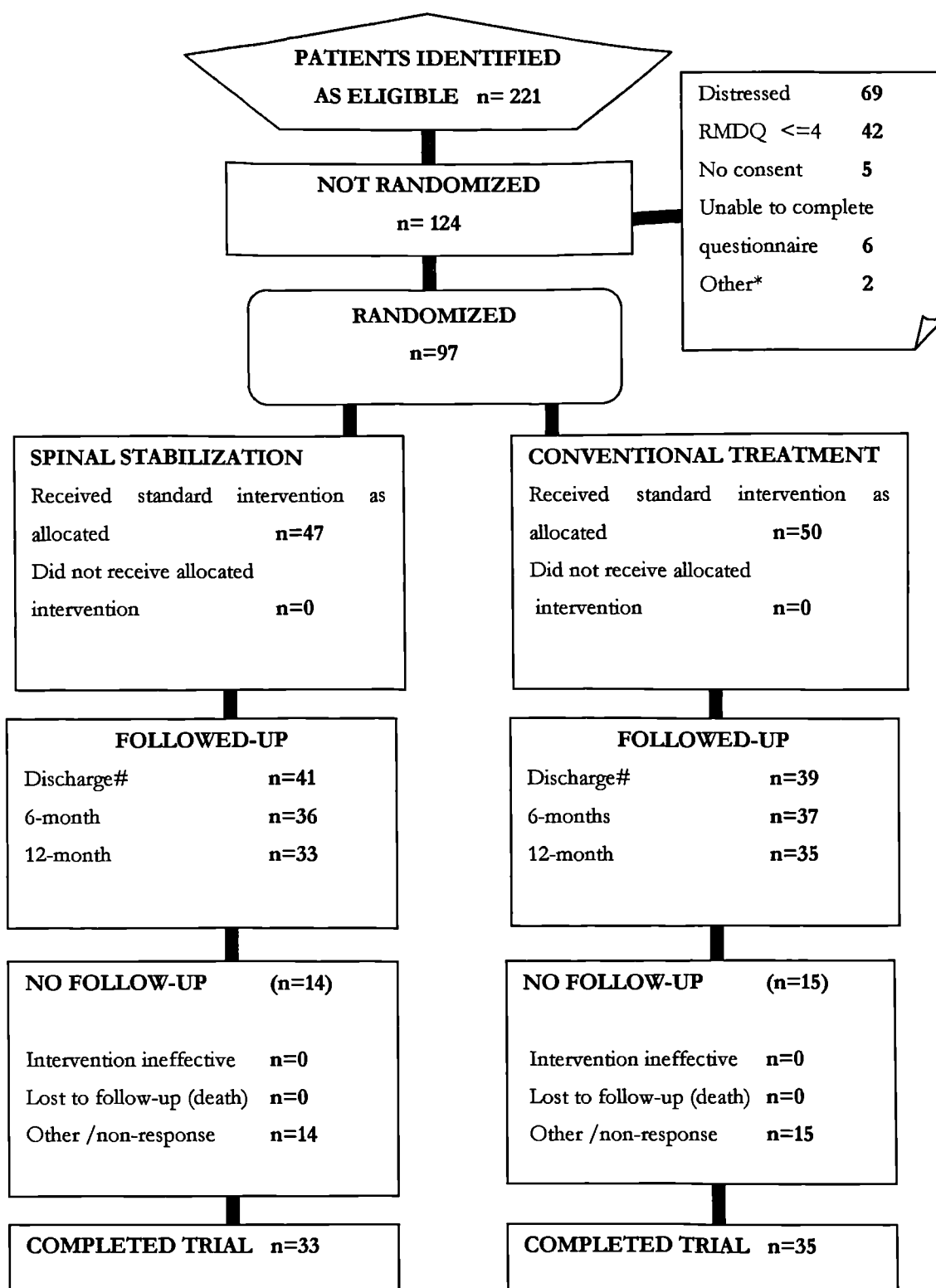
Patients were recruited to the trial from a number of sources and the process is summarized in Figure 3-1. All participants were recruited following referral from their General Practitioner (GP) or Consultant for physiotherapy at one of the participating units or for assessment at a physiotherapist-led back pain triaging clinic (BPC) at the Royal Orthopaedic Hospital. Patients attending for assessment at either physiotherapy or BPC were routinely screened using the DRAM and RMDQ. Any patients classified as distressed (distressed depressed or distressed somatic) using the DRAM were excluded from the trial. These patients received routine physiotherapy management but took no further part in the study. Similarly, patients whose RMDQ score was four or less were excluded from participation and also received usual physiotherapy management. Those meeting the inclusion criteria were identified to the primary researcher, assessed against the exclusion criteria and eligible patients invited to participate in the trial.

At all three centres, 'blitz' days were undertaken to aid recruitment to the trial. Referrals for physiotherapy of patients with LBP were identified by the main researcher at ROH and by named, treating clinicians at the other two centres. Patients were contacted either by phone or letter and booked in for assessment with a study clinician. This followed normal practice in each centre at the time, and therefore patients were not informed about the trial at the time the assessment was booked. All patients were asked to complete a DRAM and RMDQ on arrival for assessment and each clinician would assess between 3 and 4 patients per 'blitz' session. Patients not excluded on the basis of the DRAM/RMDQ scores were assessed against the exclusion criteria and any eligible patients were approached by the primary researcher and invited to participate in the trial.

Figure 3-1: Summary of recruitment procedure



Following written consent, patients were required to complete the remainder of the outcome questionnaires (See Figure 3-1). The primary researcher checked these before the patient left the department and any incorrect entries or missing items were amended. Although no formal grace period was allowed following consent, in practice a period of grace occurred as patients were on average booked to begin treatment two to three days after providing consent. No patient withdrew consent during this period but one patient was withdrawn from the trial as they reported symptoms suggestive of cauda equina involvement, thereby fulfilling an exclusion criterion and requiring a surgical opinion. A summary of the progress of all patients screened for entry to the trial is presented in Figure 3-2 in the format recommended in the Consolidated Standards of Reporting Trials (CONSORT) statement (JAMA, 1994; Altman, 1996; Begg et al., 1996).

Figure 3-2: Progress of Patients Recruited to the RCT³⁰³⁰ Based on CONSORT Statement (1996)

* Patients transferred elsewhere for treatment

Includes patients ceasing to attend for treatment but responding to postal follow-up (3-SSSE:2-CT)

3.6.1 Randomisation

An adaptive randomisation procedure was used incorporating minimization (Pocock, 1983; Friedman et al., 1998; Sim and Wright, 2000). Treatment at the start of the trial was randomly allocated to the first patient and subsequent patients were assigned using a randomisation weighted towards minimizing the imbalances in the prognostic indicators chosen (Roberts and Torgerson, 1999). Although simple randomisation should produce groups with similar baseline characteristics, it has been suggested that with smaller studies, there is no guarantee that this will occur, and imbalances between groups at baseline are possible (Friedman et al., 1998). Equally, the use of stratified randomisation using important prognostic factors can be problematic in smaller scale trials, where many strata contain few participants or remain empty. Adaptive methods of stratification avoid this problem because randomisation does not occur within strata whilst important prognostic indicators are still used (Friedman et al., 1998). The main advantage of a minimization procedure is that it protects against severe baseline imbalances for important prognostic indicators (Friedman et al., 1998). As minimization is an adaptive method of randomization, group allocation is not truly random (Sim and Wright, 2000). Despite this, Treasure & MacRae (1998) maintain that randomisation does not guarantee similarity of groups whilst minimization does, suggesting minimization as the platinum standard to the gold standard of simple randomization.

The factors used in the minimization process as strata were:

- Quebec Task Force (QTF) for Spinal Diseases Classification (Categories 1-4) (Spitzer et al., 1987)
- Total duration of symptoms
- Functional disability as measured by the RMDQ score (Roland and Morris, 1983a)

The QTF for Spinal Diseases classifications 1-4 describe the extent of LBP and referred leg symptoms. These were used to assign patients to a group depending on the extent of radiation of their symptoms. The QTF was chosen as it provides an indication of the extent of both back and referred pain. This was considered an

important prognostic factor as evidence has shown that radiating leg pain is associated with a more prolonged course and substantially more disability than LBP alone (Deyo and Tsui-Wu, 1987; Selim et al., 1998). Additionally, it has been suggested (Ren et al., 1999) that when evaluating the impact of LBP severity on functional status limitation and health-care usage, radiating leg pain should be considered although more recent research, however, has questioned this assertion (Padfield et al., 2000).

The total duration that patients had experienced symptoms, using two time-points of more than or less than 5 years, was used as a stratifying factor. Duration of pain was considered important, as evidence from the literature suggests that the reason for consultation varies with duration of pain (Waxman et al., 1998). Increasing chronicity of symptoms may be associated with changing pain behaviour, expectation and response to treatment. It was important for the current study to achieve comparable groups, therefore the use of duration of symptoms as a stratifying variable was considered justified.

Baseline functional disability (RMDQ score) was also used as a stratifying variable. Two categories were used, under 50% (0-12 points) and over 50% (13 and above). These were used as research has shown that the MCID varies depending on the initial RMDQ score (Stratford et al., 1996a). Stratford and co-workers have suggested 3 exclusive categories and associated MCID discussed in Chapter 2. Using these 3 categories would have provided an unmanageable number of stratifying categories and therefore RMDQ scores of over and less than 50% were used. The above variables produced a minimization procedure with 16 categories and are shown in Figure 3-3. After consent was obtained, patient characteristics were assessed against those shown in Figure 3-3 and one of two procedures were followed. If there were uneven numbers of patients in the same category in each treatment arm, the patient was automatically allocated to the arm with the least patients thereby equalising the distribution. If an even number or no patients had been allocated to either strata or there were even numbers in each arm, a coin was flipped by a member of the department independent of the trial and the patient allocated on the basis of the result (heads=CT, tail=SSSE).

Figure 3-3: Minimization procedure and resulting stratification categories

QTF Category	Duration (years)	RMDQ Score	Category No.
1	<5	≥12	1
		<12	2
	≥5	≥12	3
		<12	4
2	<5	≥12	5
		<12	6
	≥5	≥12	7
		<12	8
3	<5	≥12	9
		<12	10
	≥5	≥12	11
		<12	12
4	<5	≥12	13
		<12	14
	≥5	≥12	15
		<12	16

In order to enhance clinical continuity and replicate current clinical practice, randomisation of therapists was not included in the study design i.e. when therapists had assessed a patient who subsequently consented and was entered into the trial, the assessing therapist conducted the treatment following randomisation regardless of which arm the patient was allocated to. Randomisation of therapists would only have been possible in certain recruitment situations i.e. with participants assessed in BPC and those identified as suitable for the trial following assessment by members of the participating departments not involving in treating in the trial. This procedure supported the pragmatic nature of the trial design by ensuring that therapists were involved in management of patients in both treatment arms.

3.7 Outcome assessment

Assessment of all outcomes was undertaken prior to the start of physiotherapy treatment (baseline) and at completion of treatment (discharge). Six and twelve month follow-ups were undertaken by post. The outcome measures chosen incorporated both generic health measures and disease-specific outcomes and are shown in Table 3-3. The primary outcome measure was the RMDQ, which was used for calculation of the sample size. Permission was sought to use questionnaires where appropriate³¹.

A data collection sheet was completed detailing demographic data, QTF classification, work and smoking status (Appendix P). This was revised following the pilot. The main changes consisted of a more detailed identification code, to include both patient number, unit, treating clinician and group randomisation (Appendix Q). This allowed patients to be tracked and monitored more accurately.

Patients entered into the trial who failed to complete a full course of treatment or respond to follow-up questionnaires were sent tick-box, pre-paid post-cards requesting information regarding why they had stopped treatment or failed to respond to follow-up (Appendix R). An outcome sheet was manually completed for all patients and is shown in Appendix S. Clinicians completed a treatment record after each session and a discharge summary on completion of treatment (Appendix T and Appendix U respectively).

3.8 Development of protocol

This section outlines the planning and background work undertaken during the development of the study protocol. It includes details of the development of

³¹ SFMPQ: Professor Ronald Melzack, Department of Psychology, Stewart Biological Services Building, 1205 Dr Penfield Avenue, Montreal, QC, Canada, H3A 1B1 (3rd September 1999)

U.K. SF-36: Professor John Ware, Health Assessment Lab, 750 Washington Street, #345, Boston, MA 02111 (Permission gained-13th April 1999)

collaborative links with the participating treatment centres, recruitment and training of the treating clinicians and the pilot phase of the trial.

3.8.1 Background fieldwork

Spinal stabilisation exercise techniques have been developed and extensively practised and researched at the University of Queensland, Australia and therefore an educational visit was arranged to the Joint Stability Assessment Clinic to receive valuable input and advice from academic staff regarding the study protocol. This took place in September 1998 for four weeks. The input included methodological advice from and discussion with Professor Carolyn Richardson and Associate Professor Gwen Jull. Discussions regarding the clinical assessment of patients randomised to the SSSE group highlighted that there was a need to be able to identify if patients could activate the required muscles. It was therefore agreed that real-time ultrasound imagining would be available on an individual basis at the discretion of the treating clinician. The primary researcher would undertake this imaging. Practical training in teaching specific stabilisation techniques, and the use of diagnostic ultrasound for imaging was provided by Suzanne Roll during the visit to Queensland University. The specific patient protocol and training plan for clinicians involved in the study were also developed along with the patient information booklet to be used in the spinal stabilisation training groups which is Appendix L.

Table 3-3: Properties of Outcome measures used in the RCT

Outcome measure	Domains/Responses	Scores	Range of scores	Time to complete
Functional				
RMDQ	24 items	'Yes'=1	0= no disability	5 minutes
	Yes/No responses	'No' =0	24=maximum disability	(Kopec, 2000)
ODI	10 items	Percentage	0% no disability	5 minutes
	6 responses/item		100%maximum disability	(Kopec, 2000)
Quality of life				
SF-36	36 questions	0-100	0=best health	10 minutes
	8-scale health profile		100=worst possible health	(Laurie, 2000)
	2 summary scales		Both self-reported	
			perceived health	
Pain				
NRS	Pain (Usual pain)	11-point scale	0, no pain, 10, worst pain	Less than 1 minute
SFMPQ	Total	0-45	0-None, 1-Mild,	2-3 minutes
	Sensory	0-33	2-Moderate, 3-Severe	
	Affective	0-12	High scores = increase pain reporting	
	VAS	10cm	Higher scores represent	As above
	Present Pain Intensity	0-5	Higher reported pain levels	As above
Psychological distress				
DRAM	Distress		4 categories	Varies ++
MZ	Depression	0-69	Normal (N),	
MSPQ	Somatisising	0-39	At risk (R)	
			Distressed depressed (DD)	
			Distressed somatic (DS)	

3.8.2 Profile of clinicians

A breakdown of the information gained from the questionnaires (Appendix V) completed by the clinicians involved in the RCT is shown in Table 3-4. In total, 10 physiotherapists were involved in the trial. Five Senior I clinicians, two Senior II clinicians and three clinical specialists were recruited to treat the patients in the RCT. Three clinicians had been qualified between 4 to 6 years, four between 7 to 10 years, one between 11 to 14 years and two more than 14 years. The years of specialisation within musculoskeletal outpatient work ranged from 3 to over 14 years.

Table 3-4: Grades of treating clinicians

Characteristic	Category	Frequency
Clinical Grade	Senior I	5
	Senior II	2
	Clinical Specialist	3
Total number of clinicians		10
Number of years qualified	4-6 years	3
	7-10 years	4
	11-14 years	1
	> 14 years	2
Number of years specialisation	< 3 years	1
	4-6 years	1
	7-10 years	1
	11-14 years	3
	> 14 years	4
'Muscle imbalance' training	In-service training	2
	Single weekend course	2
	3 weekend course	6

In order to ensure a minimum level of competency for the clinicians who would be involved in treating patients during the trial, information was collected regarding the amount of training in spinal stabilisation techniques each clinician had undertaken. Six clinicians had completed a three weekend 'muscle imbalance' course, two had completed a single week-end course and two had only received training as part of

'In-service training'³² but were booked to complete a single week-end course before the start of the RCT.

Junior grade physiotherapists were not invited to participate in the trial as, at the time of the study, 'muscle imbalance' or spinal stabilisation training techniques were not routinely taught on undergraduate physiotherapy courses and therefore it was considered that junior grade clinicians would not have the skills required for participation. Additionally, as junior staff rotated clinical areas every four months in the hospitals involved, continuity of care for the patients involved in the trial could not be guaranteed.

3.8.3 Recruitment of clinicians involved in the RCT

Initially, the managers of two local physiotherapy outpatient departments were approached for permission to contact their staff regarding possible involvement as treating clinicians for the trial. This was followed by an open invitation to all staff involved in the treatment of LBP on an outpatient basis, outlining the proposed trial and inviting them to attend an informal meeting to discuss possible involvement. The initial meeting took place in July 1998 with a presentation of the research proposal followed by discussion regarding ethical and practical considerations. At that meeting, all clinicians expressing an interest to be actively involved in the trial completed a questionnaire, which asked details about their undergraduate and post-graduate qualifications (See Appendix V).

Following this meeting, a group of 10 clinicians were identified as possessing the appropriate skills to undergo the further training to be involved in the trial. These clinicians were required to have completed a recognised post-graduate course, or have attended, or be booked on prior to the start of the pilot study, a course covering the theory and principles of lumbar spinal stabilisation techniques. However at the time that clinicians were recruited to participate in the trial, no post-graduate courses offering to teach spinal stabilisation training techniques had been accredited or validated by either the Chartered Society of Physiotherapy or any

³² Training within work time as part of a continuing professional education programme. This can involve theoretical and practical teaching and evaluative processes e.g. critical appraisal of research.

higher educational institution. However, completion of a formal course of training in these techniques was used as a basic requirement for all study clinicians in an attempt to ensure that all clinicians had a good basic level of proficiency in these techniques prior to inclusion in the study.

Ten clinicians, equating to 5.5 whole time equivalents, were identified, but time available to the trial was markedly reduced from this by pre-existing commitments to specialist clinics, clinical and teaching activities. Four clinicians were recruited to treat patients from the Royal Orthopaedic Hospital and Queen Elizabeth Hospital respectively, and two from Selly Oak Hospital, in Birmingham, UK.

A series of three training days was organised for therapists who were involved in the trial. The aim of these days was to ensure that all clinicians had an up-to-date knowledge of the current research regarding spinal stabilisation training and that they were fully competent at assessing and teaching the techniques to patients.

The first of these took place on Wednesday 14th October 1998 and consisted of an overview of the background, development and theory of spinal stabilisation including current research findings. The second session (Thursday 7th January 1999) included a teaching session utilising diagnostic ultrasound imaging of transversus abdominis and lumbar multifidus. This allowed examination of various recruitment patterns and facilitation techniques and was received well as a useful learning tool. The final study day (Tuesday 12th January 1999) was directed more at the clinical application of techniques and included three differing patient presentations³³. Clinicians divided into groups to use a problem-solving approach to identify the most appropriate recruitment strategies and treatment programme for each individual patient presentation. Feedback and discussion following these sessions allowed clinicians to identify any specific learning or training needs which could be addressed prior to the start of the main trial.

³³ Three patients currently undergoing stabilization training and represented varying recruitment patterns and subsequently requiring individualized facilitation techniques kindly agreed to attend as teaching subjects for the training day

Meetings took place at regular intervals with clinicians and support staff who were involved in the trial. These allowed feedback regarding practicalities surrounding the study protocol, updates on numbers recruited and support for any difficulties. Patients recruited to the trial were booked for treatment sessions in accordance with the system used at each participating site. They were treated in an outpatient setting alongside non-trial patients and on discharge, communication with the referring source undertaken in the usual way.

In summary, the twelve months from February 1998 to February 1999 were spent in completing the background fieldwork for the trial. This included planning the methodology, seeking and gaining ethical approval for the study, recruitment and training of clinicians and consultation with acknowledged external experts in the field and organising a pilot phase for the trial.

3.9 Pilot study

A pilot was undertaken prior to the main trial with four main aims, each of which is addressed below:

- i. collect data on the variability of the primary outcome measure, the RMDQ, on which to base sample size/power calculation for the main trial
- ii. assess the appropriateness of the inclusion/exclusion criteria for LBP patients
- iii. assess the success of the randomisation strategy to produce comparable groups
- iv. assess the feasibility of running the main trial across the three centres

In total, 21 patients (11 male, 10 female) were recruited to the pilot trial with a mean age of 44.4 years (SD: 16.2), with descriptive data presented in Table 3-5. Of the 21 patients, 11 were discharged with 9 completed follow-ups at discharge. Six-month follow-up data were available on 9 (43%) patients and one-year on 7 (33%) patients. A summary of four of the study variables at baseline is shown in Table 3-5. The distribution of patients lost to follow-up was equally distributed between both groups with 7 (58.3%) lost from the SSSE group and 5 (41.7%) from the CT group.

Table 3-5: Demographics and study variables at baseline for the pilot trial

Variable	SSSE n= 9		CT n=12	
	Mean (SD)	Range	Mean (SD)	Range
Age (years)	37 (14)	19-68	37 (14)	20-64
Height (cm)	176.6 (8.4)	166-190	171.5 (11)	155-192
Weight (Kg)	81 (19)	60-122	74 (13)	59-108
Outcome measure	Mean (SD)	n	Mean (SD)	n
RMDQ	10.4 (6.4)	12	8 (5)	9
ODI	25.8 (15.4)	12	25 (14.8)	9
SFMPQ	9.7 (7.8)	11	6.6 (4.9)	8
NRS	5.4 (2.8)	10	5 (3.1)	9

3.9.1 Sample size calculation

Data for variability in the primary outcome measure (RMDQ) for the target population i.e. recurrent LBP patients in a secondary healthcare setting, was not available at the time of the pilot study therefore it was necessary to collect this data from patients recruited to the pilot from the three clinical sites to be used for the main trial. These data were then used to calculate the appropriate sample size for the main trial.

The variability in RMDQ scores from the pilot trial allowed a power calculation to be undertaken to ensure that sufficient patients were recruited to the main study. The standard deviation for the mean RMDQ score obtained from the 21 patients recruited to the pilot trial was 6, and this was used in subsequent sample size calculations. Prior to performing the sample size calculation, a number of factors were considered, as recommended by Altman (1992). The sample size was calculated to detect a true difference between two means for an analysis of variance design (Sokal and Rolf, 1969) and is shown in Appendix W. In order to detect a 5-point difference between groups at 12-month follow-up from discharge, with 90% power at the 5% significance level the total sample size required was estimated to be 64, assuming no attrition. In order to ensure an adequate sample size was maintained, 10% attrition at each of three follow-up points i.e. at discharge, 6 and 12

months, was calculated as recommended by Altman (1992). The addition of attrition resulted in a total sample size requirement of 92 patients.

Therefore, given the standard deviation of 6, 90% power at the 5% significance level to detect a 5-point or more difference between the two groups in the mean change on the RMDQ 12-months following discharge from treatment, a total sample size of 92 was required, allowing 10% attrition at each follow-up point.

3.9.2 Assessment of inclusion and exclusion criteria

A set of inclusion and exclusion criteria was developed based on previous research in this area. As with the design of most pragmatic trials, these were kept to a minimum in order to increase the application of the results to the wider population, whilst still recruiting as homogeneous a population as possible. The pilot was used to test the appropriateness of these criteria, identify any shortcomings and allow any amendments to be made prior to the main trial.

Four patients ceased to attend (CTA) and no discharge data are available. Of the 21 patients randomised, 5 patients were withdrawn. For the purpose of clarity the definitions of *exclusions*³⁴ and *withdrawals*³⁵ suggested by Friedman et al. (1998 p 285) were used throughout the reporting of the trial.

The five withdrawals were unsuitable to continue with treatment for the following reasons; fast deterioration and referral for specialist opinion, referral for hydrotherapy (protocol violation), abnormal radiological finding requiring surgical consultation, unable to attend for treatment and pregnancy (not known at the time of randomization). As a result of the high percentage of withdrawals, the process for consenting patients was revised and checking of inclusion and exclusion criteria was undertaken more rigorously. A number of modifications were made to the study

³⁴ **Exclusions** – defined as “people who are screened as potential participants for a randomized controlled trial but who do not meet all the entry criteria and, are therefore, not randomized”

³⁵ **Withdrawals** – defined as “participants who have been randomized but are deliberately not included in the analysis”

protocol and inclusion criteria. The main two were the introduction of a minimum RMDQ score of more than 4 and the reduction of the upper age limit from 65 to 60 years. In the pilot trial 33% (7/21) of patients had RMDQ score 4 or less. This made detecting change difficult as evidence from the literature at the time indicated that a reduction of 5 RMDQ points was the MCID for the whole scale (Stratford et al., 1996a). Additionally, a reduction in 5 RMDQ points was set as the MCID for the sample size calculation.

The upper age limit was lowered from 65 to 60 for a number of reasons. This would facilitate comparison with other work in this area. Additionally, muscle physiology changes associated with increasing age may have affected the re-training required in the SSSE group and therefore confounded the results (Lexell and Downham, 1992).

3.9.3 Assessment of the randomisation strategy

As can be clearly seen from Table 3-5, the stratified randomisation procedure using a minimization technique, resulting in the strata summarized in Figure 3-3, produced comparable groups at baseline both in respect to demographic and study variables. Obviously a limitation was the small numbers of patients recruited to the pilot trial. However the method was easy to administer and as it produced groups, which were comparable on prognostic indicators that were deemed to be important, the strategy was adopted for the main trial.

3.9.4 Feasibility

As the trial was based over three geographically separate, clinical sites within the Birmingham area, it was important that problems were identified and rectified prior to the start of the main trial. Similarly, the practicalities associated with recruitment, consent and enrolment of participants needed to be examined. A full 6 and 12-month follow-up of patients recruited to the pilot study was undertaken to identify common problems or difficulties that could be addressed prior to the main trial. The pilot study also allowed the clinicians involved in the trial to familiarise themselves with the study protocol and the delivery of the treatment interventions.

A few practical difficulties were encountered during the pilot phase. A high number of patients failed to attend for assessment appointments, which resulted in wasted clinical time and slow recruitment. This was addressed in the main trial by administration staff telephoning patients the day before their appointment to check that they were going to attend. A number of patients were unable to complete the screening questionnaires as they failed to bring reading glasses to the assessment. In the main trial this was addressed by informing the patient in their appointment letter that they would be required to complete a number of questionnaires and therefore to bring any reading glasses if required. Although all referrals were checked for a diagnosis of LBP, at assessment many patients identified other musculoskeletal problems as their main concern requiring treatment and, therefore, were not eligible for inclusion into the trial. Additionally, many patients presented with acute first episode LBP or that of a chronic nature without episodes of remission and therefore did not fit the criteria for entrance.

Difficulties were experienced in the pilot study with patients not completing a full course of treatment. The maximum number of treatment sessions allowed in the pilot was 8 sessions over a 10-week period. The mean number of treatment sessions was 5.9 (SD: 4.4, range 2-12). Of the 16 patients who started treatment, 25% (4/16) received more treatment sessions than the study protocol specified (1 patient received 9 sessions, 2 received 10 and 1 received 12 sessions). Following discussion with clinicians and other researchers involved in running a smaller scale RCT, the maximum number of treatment sessions was increased to 12 sessions over a maximum of 12-week period. The length of treatment was increased to 12 weeks in order to allow for planned holidays, courses and mandatory training during the treatment period. Additionally, other researchers who had used a 10-week period of retraining suggested that the addition of a further 2 weeks would allow consolidation of the muscle re-education used as part of the SSSE treatment arm (Personal Communication³⁶).

³⁶ "Back to the Future" October 2000, course presented by Dr Peter O'Sullivan at the Hammersmith Hospital, London

Few problems were experienced with regard to the geographical separation of the three clinical sites. However, the co-ordination of clinicians' caseloads to allow recruitment days to be arranged was difficult. This was addressed by standardising, as far as possible, the days on which each unit recruited patients to the study with the primary researcher travelling to the unit in order to assist in recruitment and consent patients.

In summary, the pilot study was invaluable in informing the direction of the main trial and identifying potential practicalities with the daily running of the multi-centre RCT. Amendments made prior to recruitment to the main trial included the introduction of an additional inclusion criteria i.e. a minimum RMDQ score, stronger emphasis during the consenting procedure regarding the importance of completion of treatment and follow-up, standardisation of recruitment days at each unit and the more rigorous checking of inclusion and exclusion criteria prior to recruitment.

3.10 Data analysis

The following summarises the data analyses undertaken, justifications for the selection of statistical tests and assumptions underlying their use. All questionnaires were scored manually using the scoring systems indicated by the designers or authors of each. This produced either a single summary score³⁷ for each outcome, which was entered into a Statistical Packages for Social Sciences (SPSS) (Version 10) datafile or a number of component scores³⁸, which were individually entered into the database. For the DRAM, the scores for the MZ and MSPQ were entered into the datafile and syntax³⁹ files were written and used to code the scores into one of the four DRAM categories (N, R, DD or DS). The U.K. SF36 data were entered verbatim into a separate SPSS database and syntax files used to rcode reversed scored items as indicated in the scoring manual (Jenkinson et al., 1996).

³⁷ RMDQ, ODI, NRS

³⁸ SFMPQ-consisted of a total, sensory and affective component, PPI and VAS score

³⁹ Syntax windows are text file windows used to enter, edit, and run commands in SPSS.

3.10.1 Baseline comparisons

Descriptive data in the form of means, modes, ranges and variance (standard deviation) were calculated for all demographic, biographic and study variables prior to inferential analysis.

The aim of any randomisation procedure is to avoid bias by balancing characteristics that might influence outcome evenly across groups. These factors can be biographic such as age, height, weight, or prognostic such as severity or clinical history. However, it is well recognised that random allocation cannot guarantee comparability for all baseline characteristics or important prognostic indicators (Altman, 1985). The use of a minimization technique has been advocated as one method to reduce baseline imbalances (Pocock, 1983; Birkett, 1985; Altman, 1992) but this can only be undertaken on a limited number of variables and only for those factors known to be of clinical importance. Although hypothesis testing is often undertaken and reported for baseline characteristics in an attempt to identify any imbalances, it is now widely challenged by statisticians (Altman, 1985; Altman and Dore, 1990; Altman, 1992). Altman has argued convincingly that significance testing at baseline only assesses the success of the randomisation process and that the decision of whether baseline differences have the potential to affect the outcome of a trial is a clinical, not statistical, decision. Additionally, examination of the data demonstrated that the minimization process had been successful in producing comparable groups at baseline, in terms of the key variables selected. However, in order to acknowledge that baseline measures of the RMDQ have an effect on subsequent change, a compromise was conceded that, with the exception of the primary outcome measure, baseline demographic data were not tested for statistical differences.

3.10.2 Analysis of change

The main aim of the analysis was to examine the difference between the mean change scores of the two treatment groups, adjusted for baseline values. Hence an analysis of covariance was used to examine the difference in mean change score at 12-months between the two groups. For comparative purposes, the non-adjusted means were calculated but are only presented for the primary outcome variable, the

RMDQ score. Analysing change scores alone (such as the 12-months outcome minus the baseline score) does not take into account the increased likelihood of regression to the mean that can occur when there are baseline imbalances (Vickers and Altman, 2001). By using the change scores in a combined ANOVA/regression model, in the form of an analysis of covariance (ANCOVA), with mean change in scores adjusted for baseline scores⁴⁰, a conservative, robust analysis is achieved (Sim and Wright, 2000). Additionally, this analysis was clinically relevant as change score from baseline is frequently used as a measure of treatment effect in the clinical setting and the analysis of covariance equation is as follows:

$$\text{Change in outcome} = \text{Constant} + a * \text{baseline score} + b * \text{gp}$$

Where;

- a =** is an estimated coefficient that represents a change in score at follow-up for a unit increase in baseline score
- b =** is the estimated treatment effect (mean difference between groups)
- gp =** is the binary code for group⁴¹ (0=SSSE or 1=CT)

The use of an analysis of covariance design also provides a higher statistical power than using simply change between baseline and post-treatment or percentage change (Vickers, 2001). Additionally it has been suggested that even when there are minimal differences at baseline, covariate analysis can improve the sensitivity by reducing the variance in the test statistic attributable to baseline variables (Norman and Streiner, 1998; Friedman et al., 1998).

3.10.3 Average change over time/Area Under the Curve

In addition to the main analyses and in order to illustrate change over time for responders, error plots using confidence intervals were constructed for each outcome. As recommended when multiple outcomes are measured, 95% CI were used for the primary outcome, the RMDQ, and 99% for all secondary outcomes (Davis, 1997; Prentice, 1997). An analysis incorporating 'average' change from baseline (as defined by area under the curve or AUC) was undertaken in order to illustrate change over time. As an *a priori* decision, the AUC analysis was calculated

⁴⁰ Adjusted means used for ALL outcomes

⁴¹ Used in the SPSS calculation of the covariate analysis

only for the RMDQ as this was the primary outcome on which the sample size calculation had been based, as recommended by Pocock (1997) and Prentice (1997). Firstly, individual response curves were constructed for each patient in order that important variation could be identified and curve shapes classified based on response (Matthews et al., 1990).

A summary measure was calculated to represent change over the 12-month follow-up using AUC (Matthews et al., 1990; Altman, 1992). This provided a single summary measure, which was then used in further analysis using an analysis of covariance. The use of AUC has a number of reported advantages, including the use of individuals as a basic unit and single number representing aspects of the individual's response curve (Matthews et al., 1990). AUC analysis complimented the original analysis of covariance on adjusted mean change scores as it reflected the average levels of individual functioning over the whole year following discharge from treatment, in contrast to the mean change score at 12-months follow-up, which examined only a single arbitrary end-point.

AUC were calculated using an equation in a Syntax file of SPSS incorporating difference from baseline scores in order to reduce the effect of varying baseline scores and duration of treatment times. It was only calculated for patients with 12-month follow-up data as an end-point is required for the calculations. Interim data were estimated using linear interpolation. A full description of this process is given in Appendix DD.

3.10.4 Intention-to-treat analysis and handling of missing data

In order to deal with the inevitable problem of missing data, a specific strategy was developed. It has been suggested that one way of dealing with the issue of missing data is to perform a number of analyses; optimistic or 'best case' scenario, pessimistic or 'worst-case' scenario and 'responders'⁴² only, and compare the results of all three analyses. If similar results from each analysis are achieved, the findings can be viewed with more confidence (Altman, 1992).

⁴² Responders only - patients who complete the necessary follow-up points i.e. 12-month follow-up in the present case

Analyses can be undertaken using a number of strategies such as ‘as treated’, ‘adherers’ or ‘intention-to-treat’ where patients are analysed in the groups to which they were randomised, regardless of compliance or whether they received the specific treatment. Intention-to-treat or ITT analyses have been suggested as the analysis of choice in RCTs and have been suggested as essential for pragmatic trials (Hollis and Campbell, 1999). Other methods of dealing with protocol violations involve subjective decisions and therefore have the potential for bias (Altman, 1992). Although retaining treatment dropouts and protocol deviations preserves between-group comparability, it is accepted that analysing treatment allocation rather than treatment received may underestimate between group differences (Altman, 1992; Lewis and Machin, 1993; Sim and Wright, 2002).

ITT analysis including a worst-case accounting was undertaken (JAMA, 1994; Koes et al., 1995; Begg et al., 1996; Friedman et al., 1998). Best-case scenario was defined as ‘no-change’, ‘worst-case’ as return to baseline. In order to undertake this, imputation of missing data was necessary (Sim and Wright, 2002). Carry forward imputation of last available data was used to represent no change for the best-case scenario and baseline value imputed for the ‘worst’ case scenario. This scheme is illustrated in Table 3-6.

Table 3-6: Strategy for imputation of missing cases

Missing data	Imputation for 12-month value	
	Last value carried-forward ¹	Return to baseline ²
12 months only	6-months	Baseline
6 & 12-month	Discharge	Baseline
Discharge, 6 & 12-months	Baseline	Baseline
¹ ‘Best’ case		
² ‘Worst’ case		

3.10.5 Scoring of the SF-36

The UK SF-36-Version I, which had been adapted for use with a UK population, was used in the study following permission from the administering body⁴³. Item responses were entered directly into an SPSS (version 10) spreadsheet and recoded in accordance with the manual. Data were analysed according to the scoring system and algorithms in the UK SF-36 Interpretation Manual (Jenkinson et al, 1996). It was intended to calculate the eight dimension scores if an adequate number of item responses were available. This provided a 0-100 score where 0 represents the least desirable health state and 100 the most desirable health state. In order to calculate the two summary scales, the physical and mental component scales (PCS and MCS), the eight dimensions were standardized (z-scores) using means and standard deviations from a UK population (OHLS, 1992). They were then aggregated using weights from a general UK population (Jenkinson et al., 1996).

The PCS and MCS were then standardized to have a mean of 50 and a SD of 10 using a linear T-score transformation in SPSS (Jenkinson et al., 1996). The use of norm-based scoring allows easy interpretation of the summary measures and comparisons to a normal population as recommended in a recent review article (Ware, 2000).

3.10.6 Data monitoring and accuracy

In order to ensure the accuracy of scoring and inputting of data, a random 20% check was undertaken by a second researcher. This consisted of 52 questionnaires, covering all follow-up points. SPSS syntax files had been set up to ensure that no values exceeding the score range could be inputted for any outcomes. Checks of the RMDQ scoring (primary outcome) highlighted two inputting errors. Following the initial 20% check, a further check of the RMDQ scoring and inputting for all time points was undertaken. This revealed very few problems with either scoring or inputting and the results are summarised in Table A 2. Any discrepancies found were corrected in the main database.

⁴³ Health Assessment Lab, 750 Washington Street, #345, Boston, MA 02111 (13th April 1999)

A summary of the checks for the RMDQ and all other outcomes is presented in Appendix X.

3.10.7 Achieved Power

The current study used a power of 90% rather than 80% in order to reduce the risk of a type II error. The trial was powered to detect a standardized difference of 0.8 (MCID/variance of the sample i.e. $5/6$; where 6 was the variance in the RMDQ score from the pilot trial) at the 12-month follow-up point (Altman, 1992). This resulted in a relatively small sample size because a large treatment effect was being sought. The actual power achieved was slightly higher than the planned 90%. Given that the actual standardised difference of the trial was 1.20 ($5/3.9=1.20$: where 5 is the MCID and 3.9 is standard deviation of the baseline mean RMDQ score of 12-month responders), the power to detect this difference was over 95% (Altman, 1992). Papers published after the trial was underway have suggested that a change score of 3-points on the RMDQ should be used, rather than larger changes when sample sizes are calculated for trials (Roland and Fairbank, 2000; Bombardier et al., 2001). Although powered to detect a larger change between groups, the current study still had 89% power to detect a 3-point change ($3/3.9 = 0.76$ where 3 is the MCID and 3.9 is the standard deviation of the baseline mean RMDQ score of 12-month responders).

In summary, the covariate analysis was chosen in order to provide information on self-reported change using clinically meaningful measures, which can be easily applied to the clinical setting and interpreted, by researchers and clinicians alike. The additional use of the AUC complimented the covariate analysis by reflecting the average level of self-reported functioning over the 12-months following treatment compared to specifically at the 12-month point. The use of an ITT analysis ensured that patients were analysed in the groups to which they were randomised, and *a priori* imputation decisions regarding missing data ensured a robust and reliable analysis was undertaken.

4. Chapter 4: Results

4.1 Introduction to results

Using the methodology described in Chapter 3, a pragmatic, randomised controlled trial (RCT) was undertaken with the aim of investigating the effects of the addition of specific stabilizing exercises in the management of recurrent LBP compared with conventional physiotherapy. This section presents the baseline comparisons, follow-up results and between group analyses.

4.2 Baseline demographical and biographical data

In total, 97 patients were recruited to this multicentre RCT, between May 1999 and September 2000; 50 were randomised to the CT group and 47 to the SSSE group. Of the 97 patients recruited, 47 (49%) were male and 50 (51%) were female. The mean age of patients in the SSSE group was 38 years compared to 40 years in the CT group. Corresponding mean heights were 169cm and 170cm, and weights were 75kg and 78kg, respectively. Table 4-1 displays the breakdown of demographic data for each group, including total duration of symptoms, which was used as one of the stratifying variables in the randomisation process.

It can be seen that the demographical and biographical characteristics of the patients in the two groups are fairly well balanced at baseline with the exception of the total duration of symptoms which was longer in the SSSE group than the CT group. These variables are summarised in Table 4-2. A full summary of the distribution of all these demographic and biographical variables at baseline is given in Appendix Y.

In order to fully describe the population, supplementary information was gathered at recruitment from all patients with regard to employment status and smoking status; and, for female patients, the number of previous pregnancies and caesarean sections. At the start of treatment, the majority of patients, 77 (79%), were in employment with 67% still working and 12% off work (citing their LBP as the reason for not working). Five patients were unemployed because of their back pain, giving a total

of 17 (18%) of the sample that reported interference with employment because of their LBP.

Table 4-1: Baseline demographical and biographical profile of patients in each group

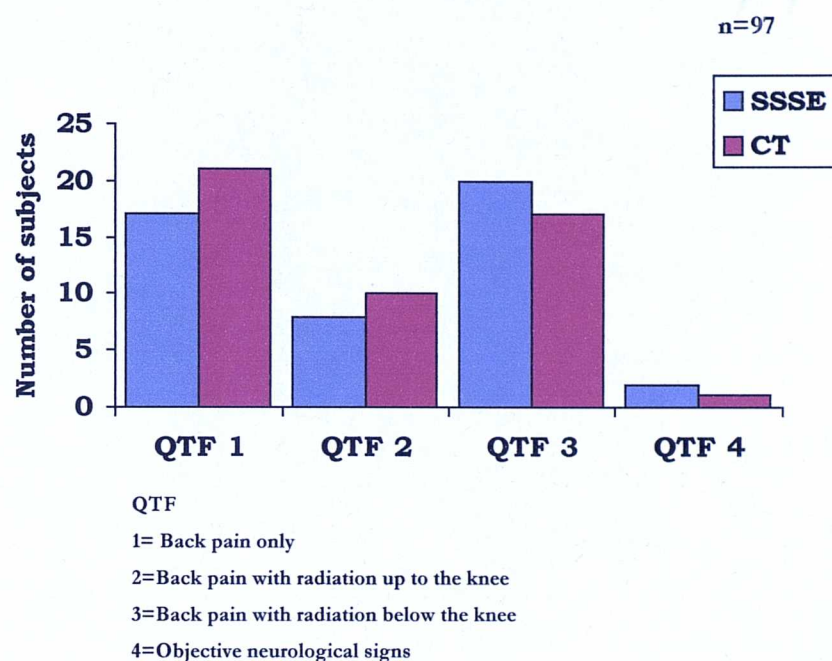
Characteristic	Group			
	SSSE (n=47)		CT (n=50)	
	Mean (SD)	Range	Mean (SD)	Range
Age (years)	37.5 (9.5)	19-56	39.9 (11.3)	21-60
Height (cm)	169.0 (10.5)	149-188	170.4 (10.7)	137-187
Weight (kg)	75.4 (15.2)	47-108	78.3 (15.2)	52-131
Current duration (months)	9.6 (8.5)	1-36	7.9 (7.6)	1-36
Total duration (months)#	118.2 (86.3)	9-300	82.0 (69.0)	2-240
	n (%)		n (%)	
Female	25 (53)		25 (50)	
Pregnancy ≥ 1	23 (49)		21 (42)	
Caesarean sections: yes	6 (13)		1 (2)	
Not working due to LBP*	10 (21)		7 (14)	
Smoker: yes	16 (34)		17 (34)	
# Indicates differing baseline variable				
*=Includes those employed as well as those not working, and those unemployed because of LBP				

The distribution of patients reporting interference with work was comparable between the two groups. A full summary and additional information is available in Appendix Z. Of the 50 female patients recruited to the trial, 88% reported at least one previous pregnancy with 16% of that number having undergone at least one caesarean section. The number of pregnancies was comparable across the two groups but more patients had undergone caesarean sections in the SSSE group than the CT group. Full details are provided in Appendix Z.

In total, 73 patients were treated at The Royal Orthopaedic Hospital and 12 each at The Queen Elizabeth and Selly Oak Hospitals. Classification of LBP was recorded using the Quebec Task Force for Spinal Diseases (QTF) (Spitzer et al., 1987) and was used as one of the stratifying variables. The distribution across both groups is

displayed in Figure 4-1 and indicates that the majority of patients reported either back pain only (QTF 1, n=38) or back pain with leg pain extending below the knee (QTF 3, n=37) with the least number of patients exhibiting neurological signs (QTF 4, n=3) e.g. objective sensation changes.

Figure 4-1: Baseline Quebec Task Force Classification of Symptoms for each group



4.3 Study Variables at Baseline

All patients completed two screening tools, the RMDQ and DRAM to assess eligibility prior to consent and randomization. The remaining 4 tools (ODI, SFMPQ, NRS and SF-36) were completed following informed consent and after randomization. A summary of study variables for each group at baseline for all patients entered into the trial is displayed in Table 4-2. The baseline RMDQ score was used as one of the stratifying variables for randomization.

The distribution of the data for each outcome at baseline was examined for normality. A full summary is provided in Appendix AA. Examination of study variables at baseline revealed that successful randomisation had achieved groups that were comparable e.g. primary outcome measure 10.36 for the SSSE group and 10.32

for the CT groups, VAS for each group of 4.18cm and 4.22 respectively. As previously indicated in Chapter 3, no significance testing was undertaken on study variables at baseline, except in the case of the primary outcome measure, RMDQ, which showed no significant differences between the groups.

4.3.1 Treatment completers versus treatment non-completers

Out of the total sample of 97, 9 (9%) failed to complete a course of treatment. In the SSSE group 5/47 (11%) failed to complete treatment and 4/50 (8%) in the CT group. The demographic and biographic data for patients who completed a full course of treatment⁴⁴ and treatment non-completers⁴⁵ are shown in Appendix BB. In summary, this shows good comparability across treatment completers and treatment non-completers in their individual groups with the exception of the 'Total duration of symptoms', which was higher in the SSSE completers group, but this did not reach statistical significance. There were no significant differences ($p < 0.05$) on study variables at baseline across treatment completers and treatment non-completers.

⁴⁴ Treatment completers=patients discharged by mutual agreement between the therapist and patient

⁴⁵ Treatment non-completers=patients who ceased to attend (CTA) for treatment

Table 4-2: Summary of study variables at baseline

	SSSE (n=47)		CT (n=50)	
	Mean (SD)	Range	Mean (SD)	Range
RMDQ	10.4 (4.3)	5-20	10.3 (4.1)	5-21
ODI	24.81 (9.4)	4-52	24.64 (9.7)	0-50
SFMPQ				
Total	11.2 (6.8)	0-29	10.6 (7.1)	1-29
Sensory	9.0 (5.0)	0-21	8.8 (5.3)	1-24
Affective	2.2 (2.5)	0-11	2.0 (2.5)	1-9
VAS	4.2 (2.00)	0-7.5	4.22 (2.3)	0.3-8.6
PPI	2.0 (0.8)	0-4	2.2 (1.0)	1-5
NRS	5.7 (1.8)	2-9	5.3 (2.3)	2-10
U.K. SF-36				
PCS	31.2 (9.5)		32.2 (9.2)	
MCS	55.5 (6.5)		55.0 (7.7)	
MZ	19 (8)	3-33	17 (8)	4-36 *
MSPQ	4.3 (2.8)	0-10	4.5 (4.3)	0-22 *
DRAM category	Number (%)		Number (%)	
Normal	19 (40)		23 (46)	
At Risk	28 (60)		25 (50)	
DD	N/A		2 (4)*	
* Includes 2 Distressed depressed (DD) patients constituting a protocol violation				

4.3.2 Responders versus non-responders at 12-month follow-up

There were 29 non-responders⁴⁶ in total and these were equally distributed between the two groups with 14 (28%) in the SSSE group and 15 (32%) in the CT group. Examination of the demographic and biographic data at baseline for both responders⁴⁷ and non-responders revealed that the groups were comparable (See Appendix CC).

⁴⁶ Patients who failed to respond to 12-month follow-up

⁴⁷ Patients with 12-month follow-up

The study variables at baseline for responders and non-responders are shown in Table 4-3. No significant difference between responders and non-responders were apparent for any of the study variables. Consequently it could be assumed that the non-responders did not represent a significantly different group of patients e.g. those with higher or lower baseline scores for certain variables, and therefore the planned analyses could be undertaken.

Table 4-3: Study variables at baseline for responders and non-responders

Outcome		Responders (n=68)	Non-responders (n=29)	p value [#]
		Mean (SD)	Mean (SD)	
RMDQ		10.0 (4.0)	11.2 (4.5)	0.52
ODI		24.7 (9.0)	24.4 (11.6)	0.14
SFMPQ	Total	10.6 (6.9)	11.4 (7.1)	0.80
	Sensory	8.7 (5.2)	9.2 (5.1)	0.97
	Affective	2.0 (2.5)	2.2 (2.7)	0.71
	VAS (cm)	3.9 (2.1)	4.8 (2.3)	0.48
	PPI	2.1 (0.9)	2.0 (0.9)	0.62
NRS		5.4 (2.1)	5.8 (2.1)	0.66
U.K. SF-36	PCS	32.5 (9.9)	31.2 (8.4)	0.37
	MCS	54.3 (6.8)	57.4 (7.5)	0.18
MZ		17.7 (8.3)	19.0 (6.8)	0.17
MSPQ		4.5 (3.9)	4.0 (3.1)	0.27
Number of treatments		7.0 (2.4)	5.6 (2.6)	0.34
DRAM Category		n (%)	n (%)	
Normal		31 (46)	11 (38)	N/A
At Risk		35 (52)	18 (62)	N/A
DD		2 (3)	0 (0)	N/A
[#] Independent sample t-test				
Responder = Patients who responded to 12-month follow-up				
Non-responder = Patients who failed to respond to 12-month follow-up				

4.4 Follow-up rates

The study was powered to identify a between group difference of 5 RMDQ points at 12-month follow-up allowing for an attrition rate of 10% at each follow-up, using 90% power and 5% significance level. The final follow-up rates are shown in Table 4-4. For the purposes of analyses, 6-months equates to 26 weeks and 12-months to 52 weeks. The progression of patients through the trial is shown in Figure 3-2.

Table 4-4 indicates an even dropout rate between the two groups and supports undertaking the analyses as planned. In addition, at 12-months, the attrition rate was successfully kept within that estimated during the design of the study.

Table 4-4: Summary of follow-up rates in both groups

Time point	SSSE (n=47)	CT (n=50)	Total (n=97)
	n (%)	n (%)	n (%)
Baseline	47 (100)	50 (100)	97 (100%)
Discharge	41 (87)	39 (78)	80 (83%)
6-months	36 (77)	37 (74)	73 (75%)
12-months	33 (70)	35 (70)	68 (70%)

4.5 Change over time

Although the main focus of this study was long-term outcome at 12-months following discharge, patient progress was additionally monitored at discharge and 6-months following discharge. Discharge data were collected in order to assess the immediate impact of treatment and both discharge and 6-month data were used to calculate the AUC for the primary outcome measure and for imputation in the case of missing data. Although the main focus of this Chapter is the 12-month follow-up data, discharge and 6-month data are presented where necessary for clarity or where applicable for a specific method of analysis. A summary of the analyses undertaken for the primary and secondary outcomes is presented in Table 4-5.

Table 4-5: Summary of analyses undertaken and presentation of results

Outcomes	Analysis
All	Covariate analysis for mean change adjusted for baseline score Error plots of mean ^a score at baseline and each follow-up
Primary	ITT analysis with best and worst case scenarios Area Under the Curve [#] Examination of characteristics of 'improvers' in the SSSE group Mean change based on magnitude of baseline score (0-8 and 9-16 points)
mean ^a = mean adjusted for baseline score	
[#] Provides an indication of average level of functioning in the year following discharge	

4.5.1 Summary of mean responses at each follow-up point

In order to show the profile of outcomes at individual timepoints, as opposed to the mean change score used for the analysis of covariance, the mean response for each study variable for each group was examined. These are displayed in the next section under each outcome domain i.e. function, distress, pain and quality of life. In all error plots, means and confidence intervals have been adjusted for baseline score of the appropriate outcome measure for the discharge, 6 and 12-month follow scores. Error plots only display the results for responders at 12-months. In all error plots, the box represents the mean for baseline and adjusted mean for all other follow-up points. The whiskers represent the 95% confidence interval (CI) for the primary outcome (RMDQ) and 99% CI for all other outcomes.

Table 4-6: Summary of mean change scores from baseline to 12-months for outcome measures for each group

Outcome measure or summary		SSSE (n=33)	CT (n=35)
Function		Mean Diff^a (95% CI)	
	RMDQ	-4.5 (-6.2 to -2.9)	-5.2 (-6.7 to -3.6)
		Mean Diff^u (95% CI)	
	RMDQ	-4.5 (-6.2 to -2.8)	-5.1 (-6.9 to -3.5)
		Area Under the Curve (99% CI)	
	RMDQ AUC*	-4.6 (-6.3 to -3.5)	-5.4 (-6.7 to -4.0)
		Mean Diff^a (99% CI)	
	ODI	-6.0 (-11.9 to -0.2)	-7.1 (-12.8 to -1.5)
Distress/Psychological			
	MZ	-2.4 (-6.0 to 1.7)	-2.2 (-6.2 to 1.8)
	MSPQ	0.4 (-1.4 to 2.3)	0.6 (-1.2 to 2.4)
Pain			
SFMPQ	Total	-1.1 (-4.3 to 2.2)	-2.9 (-6.0 to 0.2)
	Sensory	-0.7 (-3.3 to 1.8)	-2.5 (-5.0 to 0.01)
	Affective	-0.35 (-1.2 to 0.5)	-0.84 (-1.7 to 1.1)
	VAS (cm)	-1.2 (-2.1 to -0.4)	-1.8 (-2.6 to -0.9)
	PPI	-0.6 (-1.1 to -0.2)	-0.8 (-1.2 to -0.3)
	NRS	-2.2 (-3.3 to -1.0)	-2.0 (-3.1 to -1.0)
Quality of life			
UK SF-36			
	PCS	8.5 (4.2 to 12.7)	8.6 (4.4 to 12.8)
	MCS	-2.8 (-5.7 to 0.01)	-3.5 (-6.3 to 0.7)

Mean Diff^a = mean difference adjusted for baseline score
^u Unadjusted mean
 CI=Confidence Intervals. 95% quoted for primary outcome and 99% for secondary outcomes and AUC
 For all measures except the UK SF-36, negative scores indicate a reduction in the dimension measured e.g. reduced levels of disability or pain.
 UK SF-36 positive figures represent an improvement in reported health status
 * AUC represents summary measure of functional level of the previous 12 months

4.6 Function

This section discusses the results obtained from the two low-back pain specific questionnaires, the Roland Morris Disability Questionnaire (RMDQ) and the Oswestry Disability Index (ODI).

4.6.1 Roland Morris Disability Questionnaire (Primary Outcome)

A clinically meaningful reduction in functional disability was achieved between baseline and the 12-month follow-up for the CT group when the RMDQ scale is considered as a whole, but not for the SSSE group based on a change score of 5 being the MCID (Stratford et al., 1998) (See Table 4-6). The SSSE group failed to reach a clinically important change by 0.5 of a RMDQ point. Group mean change in RMDQ scores was 4.5 for the SSSE group compared with 5.2 for the CT group. Analysis of these findings indicated a statistically significant reduction in functional disability for both groups, but no significant difference between the groups ($F=0.28$, $df=1$, $P=0.60$) (Table 4-9). As evidence from the literature suggests that meaningful change in RMDQ score is dependent on the magnitude of baseline score, results were examined in the groups as suggested by Stratford et al. (1998) (Stratford et al., 1998; Riddle et al., 1998). Table 4-7 shows the mean change and associated 95% CI for patients with baseline scores of 0-8 and 9-16 RMDQ points. In both groups for both score bands, a clinically meaningful reduction in functional disability was achieved. This is relevant to clinical practice, where change is judged in relation to baseline scores and on individual cases. The table also presents the MCID suggested for each of these RMDQ score groups.

Table 4-7: Summary of mean RMDQ change scores by baseline grouping

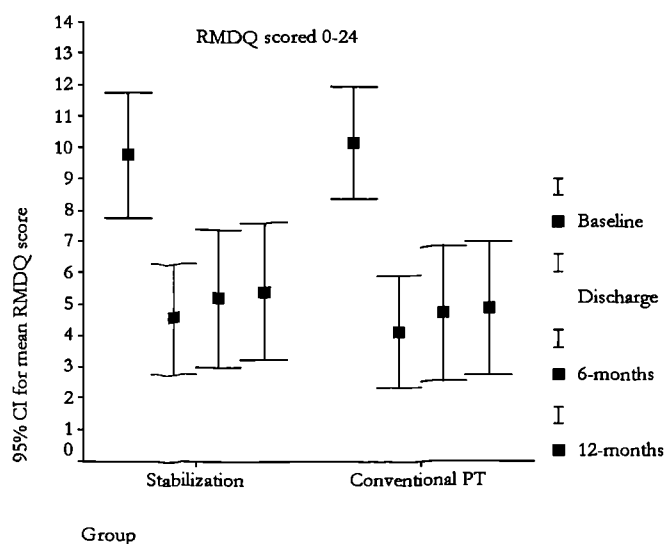
Baseline score	MCID*	Mean ^a change (95% CI)	
		SSSE n=16	CT N=12
0-8	2	-3.50 (-2.17 to -4.83)	-3.75 (-2.54 to -4.97)
		n=15	N=22
9-16	4	-4.13 (-1.26 to -7.00)	-6.05 (-3.36 to -8.73)
Mean ^a = adjusted mean change			
*MCID = Minimal clinically important difference {Based on Stratford et al. 1998}			
Baseline scores of 17-24 too few patients to analyse			

This contrasts with the mean change for the whole scale achieved by the SSSE group, which, although showing a decrease in disability, just failed to reach a clinically significant level. Additionally, the range of the confidence intervals does not include positive values (which would indicate an increase in self-reported disability). This

assists in determining the clinical usefulness of these findings and their application to the clinical setting. However, it should be noted that the sample size for the study was not calculated for subgroup analysis, although it was felt to be of clinical importance and therefore appropriate to report.

As can be seen from Figure 4-2 both groups demonstrated a marked decrease in the RMDQ scores indicating an increase in their functional ability following discharge from treatment. Both groups showed a slight worsening of functioning during the follow-up period after discharge illustrated by an increase in the mean scores, but this is minimal and does not represent a clinically meaningful deterioration.

Figure 4-2: 95% CI for RMDQ for 12-month responders



The area under the curve (AUC) was calculated for the RMDQ scores and a summary of the equation and calculations is presented in Appendix DD. The AUC gave an indication of 'average' functioning of individuals and groups over the year following discharge from treatment. It could only be calculated for those with 12-month follow-up data, as the final time-point is required for the calculation (Altman, 1992). The AUC showed similar results to the other analyses, indicating that patients in both groups reporting a significant improvement in function over the year following discharge from treatment compared to baseline levels (Table 4-6).

4.6.1.1 Results of the intention-to-treat analysis with imputation

Two separate scenarios were used for the intention to treat (ITT) analyses and the results of this and two other analyses for the primary outcome are summarised in Table 4-8. As detailed in Chapter 3, a carry forward imputation was undertaken for the 'best' case scenario, return to baseline as the 'worst' case and responders only i.e. patients with 12-month point follow-up data.

Table 4-8: Summary of results of intention-to-treat analyses for RMDQ scores

Scenario	Imputed Value (carried forward)	Mean change (95% CI)		F*	p-value*
		SSSE (n=47)	CT (n=50)		
Best case (ITT)	Last value	-4.1 (-5.5 to -2.7)	-3.7 (-5.0 to -2.3)	0.22	0.64
Worst case (ITT)	Baseline value	-3.1 (-4.5 to -1.8)	-3.6 (-4.9 to -2.4)	0.31	0.58
Responders only	None	-4.5 (-6.2 to -2.9)	-5.2 (-6.7 to -3.6)	0.28	0.60
	Area Under Curve	-4.6 (-6.3 to -3.5)	-5.4 (-6.7 to -4.0)	0.26	0.61

* Comparison of change across groups using analysis of covariance

As can be seen from Table 4-8, the three analyses produced similar findings. All indicated no statistical difference between the two treatment groups. However, the mean⁴⁸ change consistently shows an improvement in reported functional level, and 95% CI are relatively narrow and exclude the value zero, indicating a true improvement in function in both groups. Results from carry-forward imputation (Best case ITT) and the responders only are very similar for the SSSE group, but differ for the CT group. As would be expected, the 'Worst' case scenario ITT shows a smaller improvement. However, the small difference between the 'Best' and 'Worst' for the CT group may be due to the timing and pattern of loss from this group where 11/15 of the losses occurred before the first follow-up resulting in the baseline figure being inputted in both scenarios.

⁴⁸ Adjusted mean change for baseline RMDQ score

For the SSSE group, individual ‘time-series’ graphs were drawn for responders who were classified as ‘improvers’. For the purposes of this, a minimum change score of 3 RMDQ points, which was maintained at the 12-month follow-up point, was used to identify ‘improvers’ based on literature published since the design of the trial (Appendix EE). The demographic and biographic profiles of these patients at baseline were examined in order to identify any factors that might predict which patients would benefit from specific stabilisation training to inform future research directions. Preliminary examination did not reveal any obvious patterns. However, further exploration of these data may be undertaken in the future.

ITT and responders analyses for the primary outcome, the RMDQ, showed greater discrepancies across the 3 analyses for the CT group than the SSSE group, but did not show a clinically meaningful mean change (based on a reduction of 5 or more RMDQ points) for either group using ‘worst’ or ‘best’ case imputation.

4.6.2 Covariate analysis

Table 4-6 summarises the mean change in score, i.e. the post-treatment value minus the baseline value, adjusted for baseline score. In keeping with recommendations for analysis of multiple outcomes, the primary outcome (RMDQ) was analysed using a 5% alpha level and therefore 95% CI are quoted and secondary outcomes were analysed a 1% alpha and 99% CI are quoted in order to allow for the increased risk of a Type I error⁴⁹ associated with multiple testing. The unadjusted mean change is also shown for the primary outcome in order to allow comparison.

Table 4-9 summarises the adjusted mean difference in mean change scores between the two groups with 95% CIs for the primary and 99% CIs for the secondary outcome measures as described previously. The results for the analysis of covariance with baseline score as the covariate is also shown. As can be seen in Table 4-9, no significant changes were demonstrated between groups for any of the outcome variables as evidenced by p-values greater than 0.01. Additionally, all confidence

⁴⁹ Type I error = false positive. Rejecting the null hypothesis when it is in fact true and should be retained (Friedman, 1998)

intervals contain zero with no upper or lower limit being close to zero. However, the within group mean differences and CI indicate meaningful changes have been achieved in each group.

Analysis of covariance for the AUC analysis of the RMDQ scores with baseline RMDQ scores as the covariate, revealed no differences between the groups ($F=0.26$, $df=1$, $P=0.60$) (Table 4-9).

Table 4-9: Summary of adjusted mean difference in change scores for the two groups, and 95% (or 99%) CI, and results of the analysis of covariance

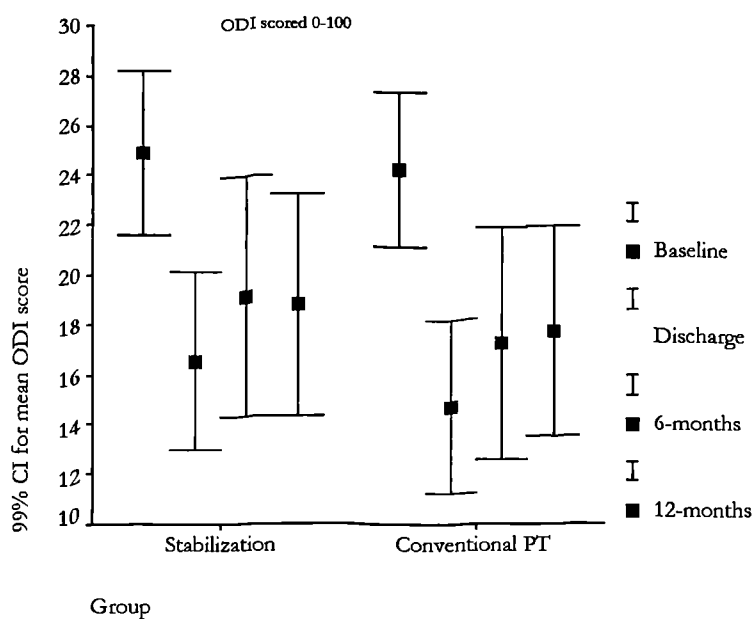
Outcome or summary score		Diff between mean diffs ^a (95% CI) ¹	F ratio	p* value
Function				
	RMDQ	0.60 (-2.87 to 1.67)	0.28	0.60
		Diff (99% CI) ²		
	AUC³	0.94 (-2.45 to 1.46)	0.26	0.61
	ODI	-1.06 (-9.17 to 7.05)	0.12	0.73
Distress/Psychological				
	MZ	0.11 (-5.37 to 5.9)	0.03	0.96
	MSPQ	-0.16 (-2.4 to 2.75)	0.27	0.87
Pain				
	SFMPQ	-1.83 (-6.35 to 2.68)	1.16	0.29
	Sensory	-1.74 (-5.29 to 1.80)	1.70	0.20
	Affective	-0.49 (-1.72 to 0.73)	1.15	0.29
	VAS	-0.52 (-1.75 to 0.72)	1.24	0.27
	NRS	0.12 (-1.45 to 1.68)	0.04	0.84
Quality of life				
SF-36	PCS	-0.11 (-8.08 to 7.87)	0.00	0.97
	MCS	0.70 (-4.66 to 6.07)	0.12	0.73

¹ Primary outcome analysed using a 5% alpha level and 95% CI are quoted
² Secondary outcomes and AUC analysed using a 1% alpha level and 99% CI are quoted
Mean diffs^a = mean adjusted for baseline score
³ AUC represents summary measure of functional level of the previous 12 months
* for comparison across groups

4.6.3 Oswestry Disability Index (ODI)

Mean change across groups in the second back-pain specific disability questionnaire, the ODI, showed a statistically significant decrease, corresponding to an increase in self-reported functioning. Mean reduction in the ODI score in each group, exceeds the minimal clinically meaningful improvement (4-6%), suggested by Beurskens et al., 1996. Table 4-6 shows the mean change and 99% CI and Figure 4-3 illustrate these for each follow-up point. As with the other functional outcome measures, no differences between groups were demonstrated ($F=0.12$, $df=1$, $P=0.73$) (Table 4-9).

Figure 4-3: 99% CI for mean ODI for 12-month responders



4.6.4 Summary of Functional Assessments

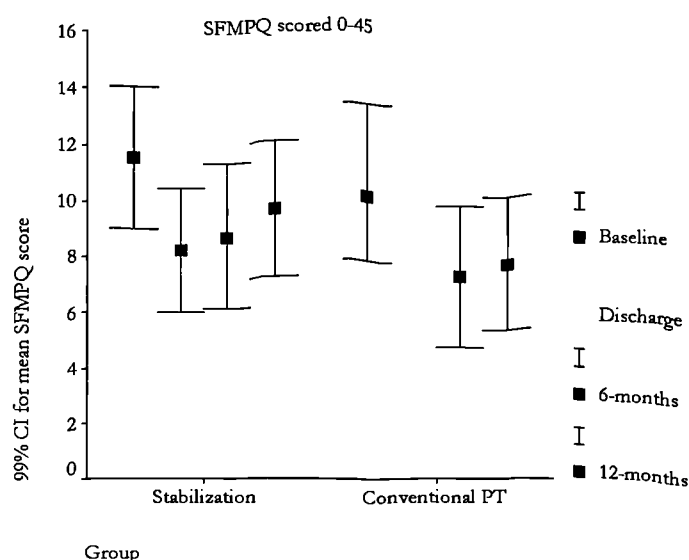
A marked reduction in disability was demonstrated by both groups following discharge from treatment and was maintained at 12-month follow-up for at least one functional outcome variable. This reached a clinically meaningful level based on the

definition of Stratford et al. (1998) for the CT group but just failed to reach this level in the SSSE group. Change scores for sub-grouping of the RMDQ, based on the magnitude of baseline scores, revealed a MCID reduction across both the RMDQ categories examined (0-8 and 9-16) for both treatment groups.

4.7 Pain outcome

Two measures were used to gain insight into the pain experienced by patients; the short form McGill pain questionnaire (SFMPQ) and the Numerical Rating Scale (NRS) for 'usual' pain (See Appendix D and Appendix E). The SFMPQ consists of 15 descriptor words, 11 (1-11) covering sensory and 4 (12-15) covering affective components of pain, a visual analogue scale and a present pain intensity index of 5 levels from "No Pain" to "Excruciating". The pain descriptors are scored from 0 (none), 1 (mild) 2 (moderate) and 3 (severe) giving a possible total of 45; 33 from sensory components and 12 from affective components. Figure 4-4 through to Figure 4-7 show the mean (adjusted for baseline score) for the total, sensory and affective components of the SFMPQ and the VAS (Melzack, 1987). The adjusted mean for the NRS at all timepoints is shown in Figure 4-8.

Figure 4-4: 99% CI for mean SFMPQ for 12-month responders



The assessment of pain is presented as the total score from the pain descriptors on the SFMPQ (as recommended (Melzack, 1987)) and the sensory and affective subscales, VAS (cm) and NRS score. The mean change in the components scales of the SFMPQ and VAS are shown in Table 4-6. No differences in change from baseline between the groups were demonstrated for the total score ($F=1.16$, $df=1$, $P=0.29$), sensory component ($F=1.70$, $df=1$, $P=0.20$) or the affective component ($F=1.15$, $df=1$, $P=0.29$) of the SFMPQ although each group demonstrated a statistically significant reduction in all three components. Both groups showed a statistically significant and clinically meaningful reduction in pain intensity measured on the VAS, of between 1.0 and 1.8 cm, as reported by Beurskens et al. (1996). The VAS scores showed a reduction of 1.2 and 1.8 for the SSSE and CT groups, respectively, but no significant difference in change between groups ($F=1.24$, $df=1$, $P=0.27$) (Table 4-9).

Figure 4-5: 99% CI for mean SFMPQ (Sensory) for 12-month responders

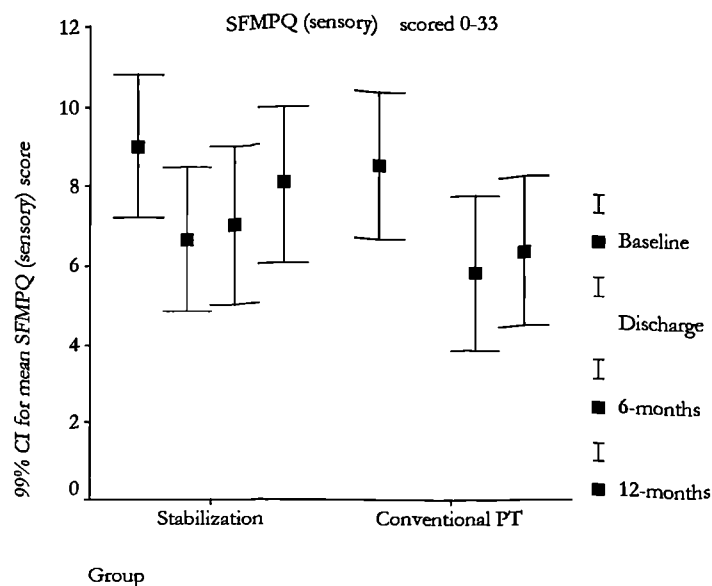
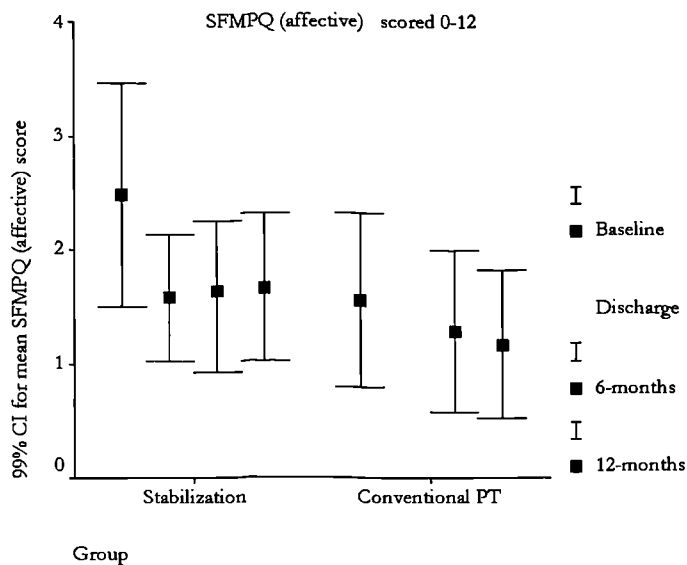
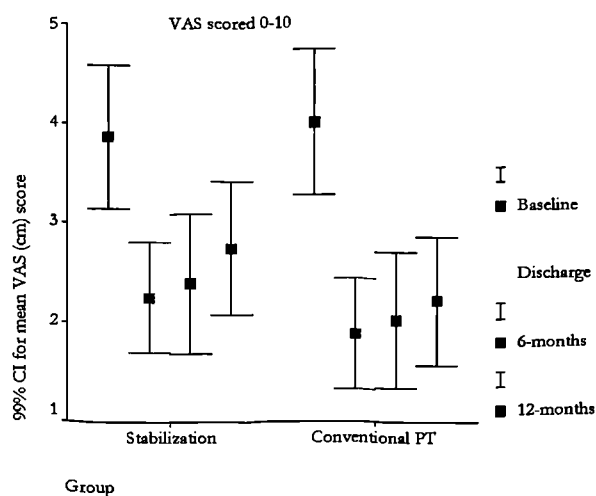


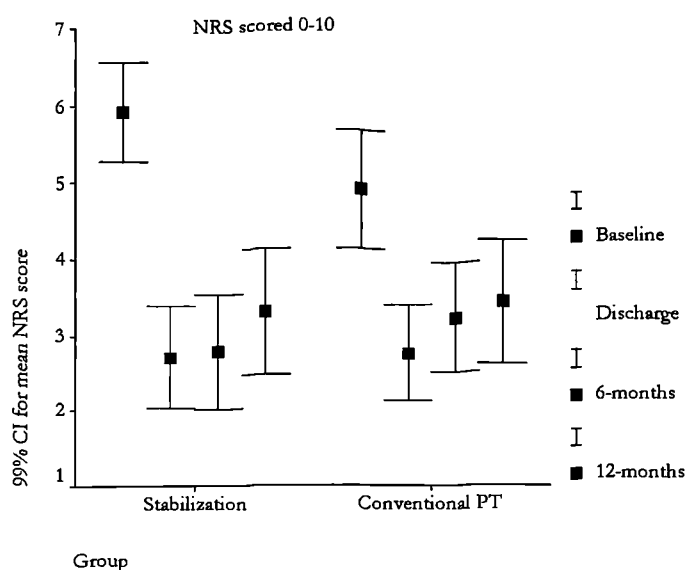
Figure 4-6: 99% CI for mean SFMPQ (Affective) for 12-month responders

The change in the VAS over the three follow-up points is shown in Figure 4-7. The reduction in the level of pain reported using the VAS was slightly greater in the CT group than the SSSE group but in contrast the NRS, again showing a statistically significant reduction in both groups, shows the SSSE group reporting a slightly greater reduction. Figure 4-8 shows the 99% CI for means for the NRS.

Figure 4-7: 99% CI for mean VAS for 12-month responders

The second pain outcome measure examined the concept of ‘usual’ pain in order to gain insight into average pain rating as opposed to a ‘snap-shot’ on the day of assessment, using an 11-point NRS. This showed a mean⁵⁰ reduction of 2.2 for the SSSE group and 2.0 for the CT group, with narrow confidence intervals (Table 4-6). Analysis of covariance showed no statistically significant difference in mean change scores between the two groups ($F=0.04$, $df=1$, $P=0.84$) (Table 4-9).

Figure 4-8: 99% CI for mean NRS for 12-month responders



4.8 Psychological variables

Results for the measure of distress are presented for all patients screened for inclusion to the trial. Figure 4-9 displays the DRAM categories for all patients screened and Figure 4-10 shows the distribution of the DRAM for each treatment group. These results are included as they illustrate the high percentage of patients, over one-third of those screened who were excluded from entry to the trial because of their levels of distress as measured by the DRAM.

⁵⁰ The mean reductions on the VAS and NRS are both adjusted for baseline score

Of the 221 patients screened, complete data were available for 212 and 97 were eligible for entry to the RCT and provided written informed consent. Of the 212 patients, 71 were classified as distressed and therefore fulfilled one of the main exclusion criteria, as outlined in Chapter 3. Figure 3-2 summarizes the reasons for the remaining exclusions.

Figure 4-9: DRAM categories of patients screened for inclusion to RCT

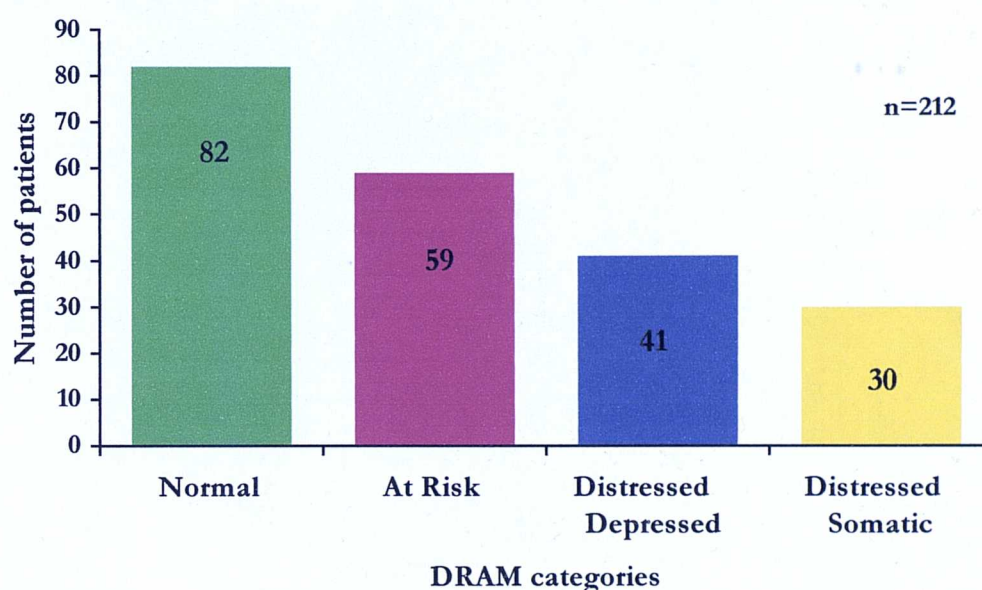


Figure 4-10 shows that both treatment groups were comparable at baseline with respect to the distribution of patients categorised as 'normal' or 'at risk' using the DRAM. The figure shows that two patients were entered into the trial who were classified as distressed depressed (DD) which was in violation of the inclusion criteria. Both patients were randomised into the CT treatment group and were followed up as normal.

Figure 4-10: Distribution of the DRAM categories at baseline per group

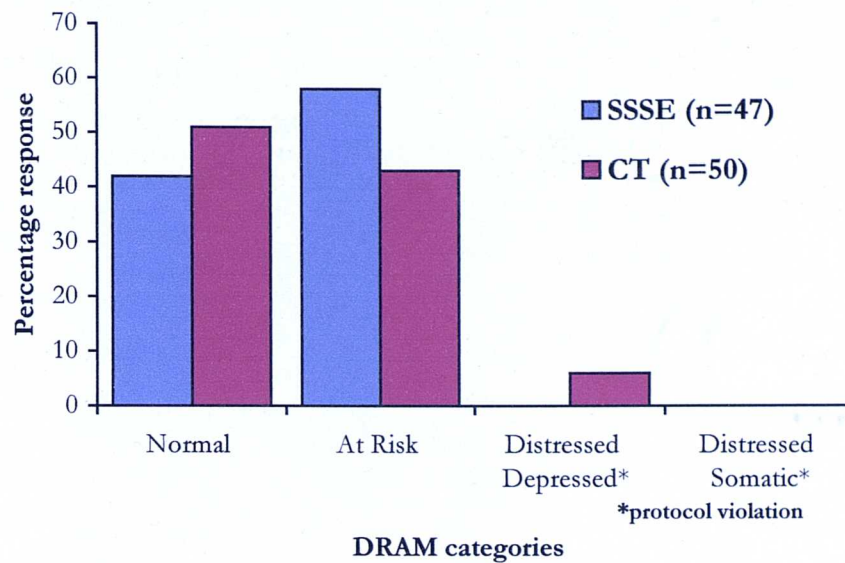


Figure 4-11 and Figure 4-12 display the 95% CI for the mean scores (adjusted for baseline score) of the MZ and MSPQ; the scores of which are used to calculate DRAM categorisation. The MZ is measured between 0 and a maximum of 69 and showed a small reduction in scores following treatment (-2.4 and -2.2 for the SSSE and CT groups respectively), which indicates a reduction in negative feelings (Figure 4-11). However, this reduction was not statistically significant (Paired t-test: $t=2.0$, $P=0.05$) and unlikely to be clinically relevant. The change in mean MSPQ score following treatment was small (0.4 and 0.6 for the SSSE and CT group respectively). This increase reporting somatic feelings was not maintained at 12 months, was also not statistically significant (Paired t-test: $t=-1.0$, $P=0.33$) and unlikely to represent a meaningful change.

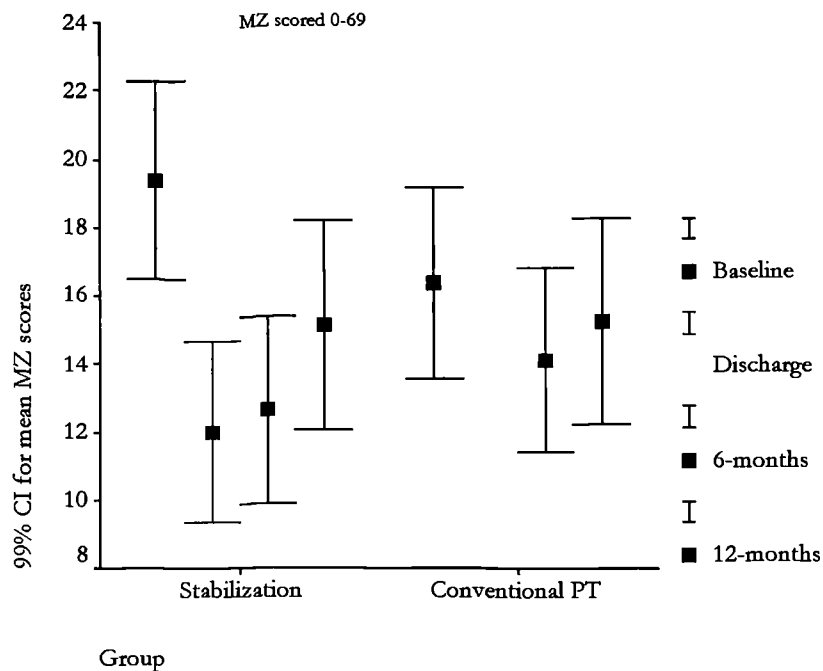
Figure 4-11: 99% CI for mean MZ score for 12-month responders

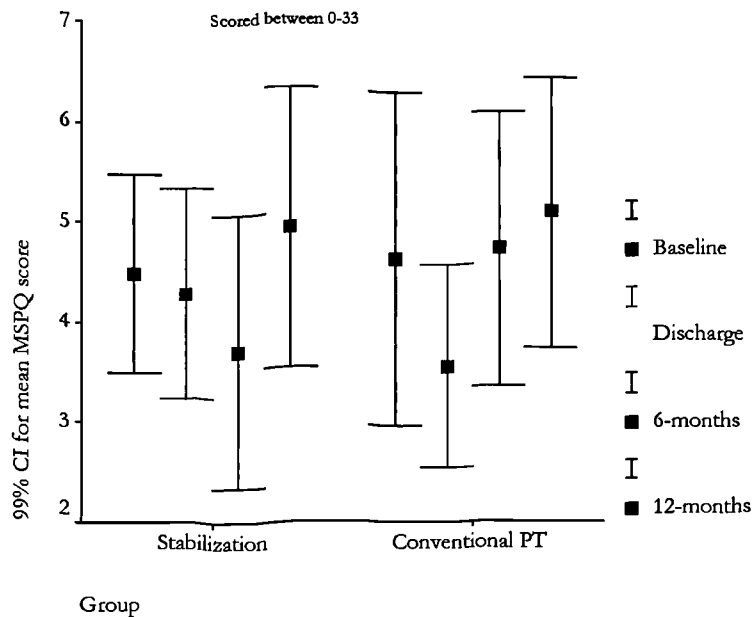
Table 4-10 demonstrates that the number of patients classified as showing no evidence of distress or 'normal' increased in both groups from baseline to 12-month follow-up, whereas the number 'at risk' of developing distress decreased. These results indicate a reduction in the number of patients at risk of developing distress but also an increase in those who are actually distressed. Results from the analysis of covariance for both components of the DRAM (MZ and MSPQ) showed no significance difference between the groups for depression or anxiety scores ($F=0.03$, $df=1$, $P=0.96$ and $F=0.27$, $df=1$, $P=0.87$) respectively across the groups.

Table 4-10: DRAM classifications at baseline and 12-month follow-up

	SSSE (n=33)			CT (n=35)		
	Baseline	12-months	% change	Baseline	12-months	% change
	n(%)			n(%)		
Normal	14 (42)	18 (55)	+13	18 (51)	23 (66)	+15
At Risk	19 (58)	11 (33)	-25	15 (43)	9 (26)	- 6
DD	0	2 (6)	+6	2 (6)*	2 (6)	0
DS	0	2 (6)	+6	0 (0)	1 (3)	+3

* Represents protocol violations

Figure 4-12: 99% CI for mean MSPQ score for 12-month responders



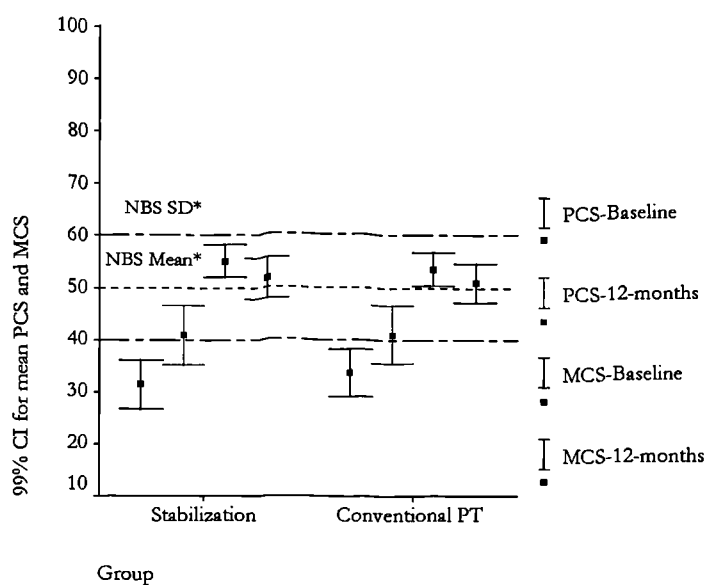
4.9 Generic Health Measure

The generic quality of life measure used was the SF-36. Only the two subscales the Physical Component Summary (PCS) and Mental Component Summary (MCS) are presented since the individual items comprising the eight scales/domains prior to statistical analysis revealed an unacceptable number of items displaying a ceiling or

floor effect (Appendix FF). The summary scales (PCS and MCS) were used as it has been suggested that these reduce the influence of such effects (Ware, 2000).

The mean score and 99% CI are shown for both the PCS and MCS in Figure 4-13. As can be seen, the PCS scores show an increase in both groups, equating to an increase in self-reported physical functioning whilst the MCS shows a decrease in score equating to deterioration in self-reported mental health functioning.

Figure 4-13: 99% CI for mean PCS and MCS at baseline and 12-months



Key to Figure 4-13

NBS*= Norm based scoring.

Linear transformation undertaken to transform scores to have a mean of 50 and standard deviation (SD) of 10, using data calculated from a UK population (Jenkinson et al., 1993; Jenkinson et al., 1996).

Figure 4-14 shows the norm based scoring for the summary scales PCS and Figure 4-15 for the MCS. This illustrates scores, which can be compared to a population norm where the mean health state is set at 50 and the SD is 10 (Jenkinson et al., 1993; Jenkinson et al., 1996). Norm-based scoring allows meaningful comparisons of research findings and calculation of these summary scales was undertaken using a UK population norm (Wright et al., 1992).

As can be seen by Figure 4-14 and Table 4-6, the PCS showed a fairly substantial mean increase (8.5 and 8.6 for the SSSE and CT groups, respectively), corresponding to improved perceived physical health. In contrast, the MCS (Figure 4-15) showed a slight decrease (-2.8 and -3.5 respectively) indicating a small deterioration in perceived mental health, but this is unlikely to represent a clinically relevant change. Analysis of covariance showed no statistical difference between the groups.

Figure 4-14: Physical Component Summary at Baseline and 12-months

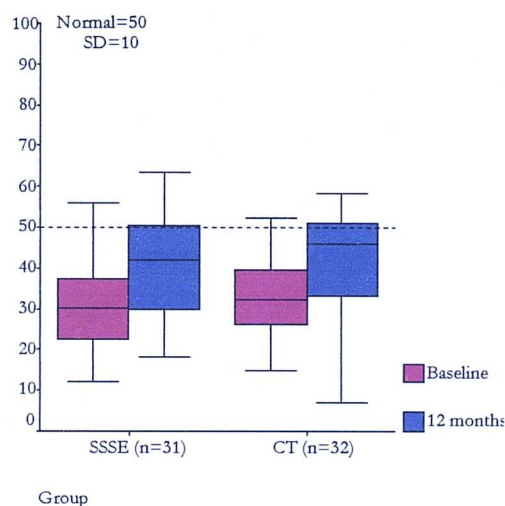
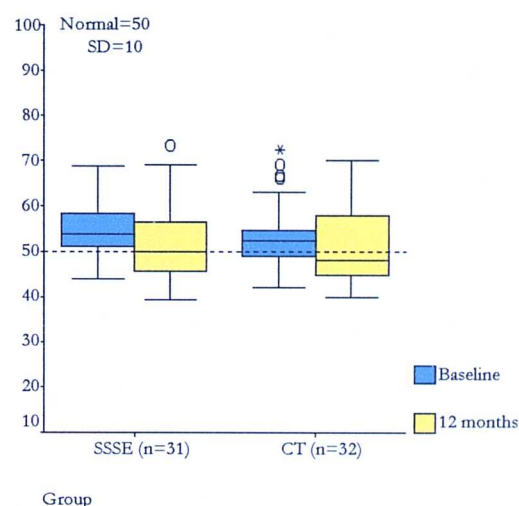


Figure 4-15: Mental Component Summary at Baseline and 12-months



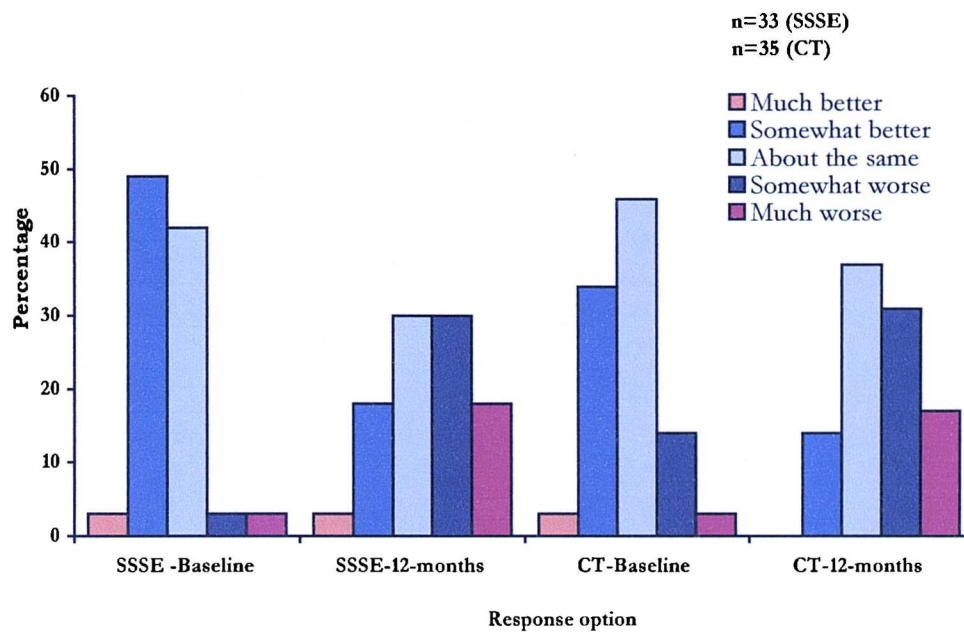
Key for Figures 4-14 and 4-15

In the above box-plots, the upper bar represents the upper quartile, the lower bar, the lower quartile and the height of the box the interquartile (indicating variability) with the horizontal line representing the median value

The second question on the SF-36 (Appendix H) reports change in health status compared to one year ago and is not reported as means and standard deviations (Jenkinson et al., 1996). It has a five item forced choice option scored from 1 (Much better now than one year ago), 2 (somewhat better now than one year ago) through to 5 (much worse now than one year ago). Figure 4-16 shows the percentage of the responses at baseline and 12-months for responders at 12-months in both groups. The response and percentages are presented in Table A 8 in Appendix FF. As can clearly be seen, neither intervention had any positive effect on self-reported change in health, with nearly 50% of patients in each group reporting deterioration in their

health compared to one year ago. This contrast to the reported improvement in the PCS, but this may be explained by examination of the items and scales, which contribute to the calculation of the PCS which ask specifically about physical functioning compared to change in perceived health (question 2) which asks about health 'in general'.

Figure 4-16: Percentage response to U.K. SF-36 change in health status



4.10 Treatment Details

In order to be able to fully describe the nature and components of the two treatment packages, a record was kept of the number of sessions, time to completion of treatment and the number of treatment modalities utilised for patients for each group. This information is summarized in Table 4-11. A summary of the number of modalities used in each group is also displayed in Figure 4-17.

Figure 4-17: Summary of total number of treatment modalities used in each group

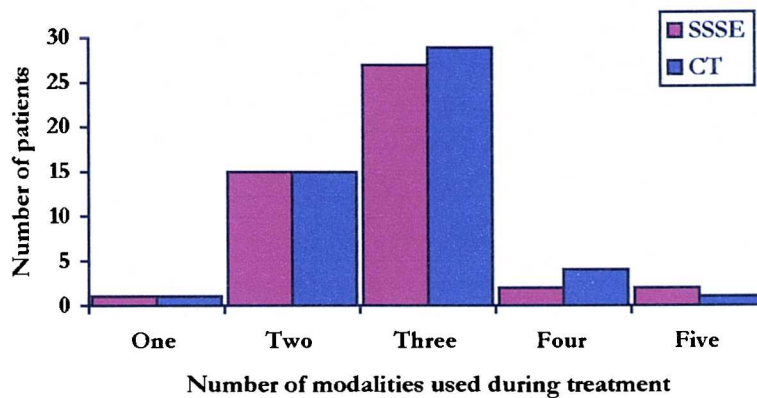


Table 4-11: Summary of treatment details for each group

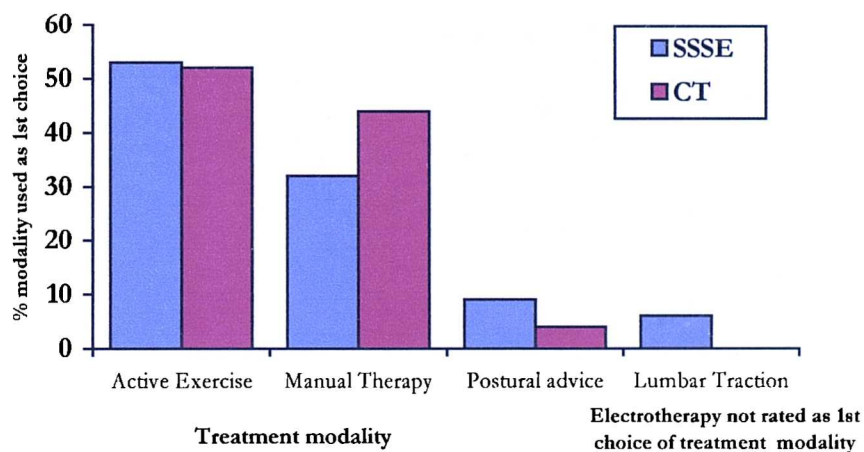
	SSSE (n=47)		CT (n=50)	
	Mean (SD)	Range	Mean (SD)	Range
Number of treatment sessions	7.5 (2.5)	1-12	5.9 (2.3)	2-11
Duration of treatment period#	11.2 (3.6)	1-20*	8.0 (3.6)	0-16*
	Mode	Range	Mode	Range
Number of modalities used	3	1-5	3	1-5
	Frequency	n (%)	n (%)	
One	1	(2)	1	(2)
Two	15	(32)	15	(30)
Three	27	(57)	29	(58)
Four	2	(4)	4	(8)
Five	2	(4)	1	(2)
TOTAL	47	(100)	50	(100)

*Protocol violations- 12 weeks maximum time period for completion of treatment
Duration in weeks

The specific spinal stabilisation exercise training taught to the SSSE group was not ranked as these formed the main base to this treatment. Excluding this, both groups were comparable in respect to the total number of modalities used. Of the five modalities available to therapists, the two least often used were electrotherapy and mechanical lumbar traction. A treatment record completed by the treating therapist

for each patient ranked the frequency of use of five treatment modalities outlined in Table 3-1. Frequency of use was ranked from 1 (most frequently used) to 5 (least frequently used). The frequency for the first choice (most frequently used) modality is shown in Figure 4-18. For both groups, the two most common treatment modalities used were active exercise and manual therapy. Active exercise was used as the first choice of modality in 50% (48/97) of cases, as chosen by the treating therapist, and manual therapy the first choice of modality used in 38% (37/97) of cases. The same pattern was seen with the choice of second modality used, with 38% (37/97) using active exercises and 31% (30/97) using manual therapy. Comparison between groups revealed a pattern of usage similar for both. The third most commonly used modality was ergonomic/postural advice, used in 46% (45/97) of cases.

Figure 4-18: Percentage usage of first choice modality for each group



The findings support the assertion that physiotherapists at all three clinical sites participating in the RCT tended to implement current national guidelines for the treatment of acute LBP, as the most frequently used treatment modalities were active exercise and manual therapy (incorporating manipulation) with very little use of passive treatment modalities such as electrotherapy and traction. The techniques incorporated in each of the five categories are presented in Table 3-1.

The mean number of treatment sessions received by the CT group (5.9, SD=2.3) was less than the SSSE group (7.5, SD=2.5). The time period over which the SSSE group received treatment was longer than that of the CT group, 11.2 (3.6) weeks compared to 8.0 (3.6) weeks, with a mean difference of 3.9 weeks and 95% CI of 2.4 to 5.4 weeks. Excluding the specific spinal stabilisation training taught to the SSSE group, the number and type of treatment modalities used was comparable across both groups, which is important as it demonstrates that the main difference between the two groups was the spinal stabilisation training.

4.11 Summary of results

In summary, when baseline and 12-month data were compared, all results indicated that there was no difference between the two treatment groups. Both groups showed an improvement in the primary outcome measure (functional disability measured using the RMDQ) between baseline and 12-months but only the CT group achieved the pre-set MCID. No statistically significant differences between the groups were demonstrated.

Secondary outcomes showed similar results. A clinically meaningful reduction in the other functional outcome measure, the ODI, was achieved and maintained at 12-months but with no statistical difference between groups. Similarly results were obtained for the VAS and NRS, which showed a clinically meaningful change for both groups but no statistical difference between the two groups. The other measure of pain, the SFMPQ showed small reductions in all components of self-reported pain (total, sensory and affective) following treatment but this did not necessarily represent a clinically meaningful change and no statistically significant difference between groups was demonstrated.

The measure of distress and anxiety, the MZ and MSPQ, showed little change, and no statistically significant difference between groups. The combined scores from these two questionnaires, the DRAM categorisations, showed an increase in the number of patients showing no evidence of distress, at 12-months compared to before treatment, and a decrease in the number of categorised as 'At risk' of

becoming distressed over the same period but an increase in the number of patients exhibiting distress.

In summary, both groups showed an improvement in clinically relevant measures, reporting an increased level of functioning, decreased pain levels and an improvement in physical functioning, measured on the SF-36. However, no statistical difference between groups was demonstrated. There was a reduction in the number of patients at risk of becoming distressed but an increase in the percentage of patients reporting of distress.

This appears to indicate that both approaches were effective in the treatment of recurrent LBP with benefits maintained at 12-month following discharge from treatment. The addition of a spinal stabilisation training regime, to conventional physiotherapy treatment package (consisting of mostly active exercise and manual therapy) in this sample of LBP patients, did not add any additional benefit.

The next chapter will discuss these results and explore competing rationales for the obtained results. It will consider the results in relation to other studies and attempt to draw some conclusions and present ideas for the future direction of research into specific spinal stabilisation exercises.

5. Chapter 5: Discussion

5.1 Summary

A pragmatic, single blind, multi-centre RCT was undertaken comparing two physiotherapy treatment packages for patients with recurrent LBP. The two packages of ‘conventional’ physiotherapy, consisting of any physiotherapeutic technique currently in use in the UK, excluding acupuncture and back school, and ‘conventional’ physiotherapy with the addition of exercises designed to increase the endurance of the deep abdominal and trunk muscles. The primary outcome for the trial was self-reported level of back-related functional disability, measured using the RMDQ, and the end-point was 12-months following discharge from treatment. Secondary outcomes included measures of pain, psychological distress, and quality of life.

Results indicated that both treatment packages produced a reduction in functional disability and pain intensity which was clinically meaningful for the CT group. Quality of life measurements showed improvements in self-reported physical functioning but little change in mental summary scales. Statistical analysis, using covariate analysis, adjusted for baseline scores, revealed no significant differences between the two groups on any outcome variable.

Although no statistical differences between the two treatment packages were demonstrated, these results appear to support the effectiveness of physiotherapy in the management of patients with recurrent LBP without evidence of psychological distress. However the results do not indicate which components of the physiotherapy treatment studied are most effective. When the physiotherapy packages were explored, most therapists used exercise (such as abdominal and trunk strengthening and stretching) and manual therapy (such as Maitland mobilisations). Very little use was made of more passive modalities such as electrotherapy or mechanical lumbar traction.

5.2 Introduction

The clinical and methodological issues of these results will be discussed, and possible rationales for the results obtained discussed. Comparisons with the results from other trials will be made and implications and directions for future research discussed.

5.3 Evaluation of results

The results clearly indicate that physiotherapy and advice, as delivered in the two treatment packages within the current trial, are effective in reducing functional disability and pain intensity in patients with recurrent LBP, without significant levels of psychological distress. Results also indicate that provision of a specific spinal stabilisation training programme in addition to conventional physiotherapy treatment does not provide any further benefit in the patient population studied. The inclusion of evidence-based, standardized written advice in the form of *The Back Book* (RCGP, 1996a), may have influenced the results obtained as recent studies have shown long-term positive effects with use of this booklet.

Alongside the improvement in functioning that was demonstrated, both groups showed an improvement in the physical summary component of the quality of life measure (SF-36), indicating that both treatment packages had a positive influence on quality of life.

Despite, neither package specifically utilizing a biopsychsocial approach, both demonstrated a positive effect on distress, with an increase in the number of patients showing no evidence of psychological distress and a reduction in the number at risk of becoming distressed. This may have resulted from a number of factors over and above any specific 'treatment' effects such as reassurance of the benign nature of the LBP following assessment or reassurance received throughout treatment. Non-specific treatment or placebo effects may have also influenced this observed change.

In summary, the findings from this study appear to support the use of physiotherapy and advice in a recurrent LBP population without significant levels of psychological distress. However, they do not support the superiority of either package, or the addition of specific spinal stabilisation exercises to a conventional physiotherapy treatment package, plus provision of an advice booklet, in a recurrent LBP population.

5.4 Implications for clinical practice

This section focuses on the clinical implication of the results and discusses possible reasons why the results were obtained and considers the results in relation to the theory behind spinal stabilisation training.

5.4.1 Pain and function

Reduction or management of pain has been reported as the goal of many LBP interventions (Foster et al., 1999; Deyo and Weinstein, 2001). However, functional disability can often decrease with very little change in associated pain levels (Moffett et al., 1999). Two measures of functional disability (or alternatively ability) were used, the RMDQ and ODI. Both are well-researched, validated back-pain specific questionnaires and are recommended for use in LBP research (Deyo et al., 1998). The findings that both groups achieved a clinically meaningful reduction in functional disability and that this improvement was maintained at long-term follow-up appears to provide support for active physiotherapy interventions in recurrent LBP patients, with no evidence of distress. Although the SSSE group failed to reach the strict MCID set at the start of the trial (of 5), other authors support lower change scores to represent important change. Change scores of 2-3 points (Roland and Morris, 1983a; Bombardier, 2000b; Roland and Fairbank, 2000), 2.5 to 5 for individual patients (Beurskens et al., 1996) and 4.4 in acute LBP patients reporting return to full activities (Deyo and Centor, 1986) have all been suggested as representing important change. Examination of the 95% confidence intervals for both mean changes support the change detected representing a clinically meaningful one. The confidence intervals are narrow, they do not include zero and the lower limit (-3) represents a change that a number of authorities regard as clinically important (Roland and Morris, 1983a; Beurskens et al., 1996; Bombardier, 2000b; Roland and Fairbank, 2000). Additionally, examination of the change scores based on magnitude of initial scores easily exceeded the MCID that have previously been identified for 0-8 and 9-16 bands (Stratford et al., 1998; Riddle et al., 1998). The MCID was also exceeded for the 17-24 initial score band but, as this only contained 5 patients, little meaningful conclusions can be drawn. The issue of important change has been the topic of

many papers (Roland and Morris, 1983b; Deyo and Centor, 1986; Beurskens et al., 1996; Stratford et al., 1996b; Stratford et al., 1996a; Stratford et al., 1998; Bombardier, 2000a; Beaton, 2000), although much controversy exists, what appears obvious from these articles is that there is not an agreed answer to the question of what constitutes an 'important change?' (Bombardier et al., 2001). Also, it is clear that what constitutes an important change varies depending on the population studied, e.g. acute versus chronic, and that important change varies dependent on what comparisons are being considered, e.g. individual, within group or between groups (Beaton, 2000). The change scores from the secondary back-pain specific functional disability index, the ODI, supports the assertions that the overall changes in functional disability represent a meaningful improvement. Both groups showed a mean reduction in the percentage disability equal to or exceeding the 4-6% that Beurskens et al. (1996) suggested discriminated between improved and non-improved patients' outcomes. Although this cut-off is perhaps more appropriately applied to individual patient change, because of the method by which it was calculated, it does provide an indication for assessing important change in clinical practice. As this trial was pragmatic with the treatment packages studied representing current clinical practice, it would seem appropriate to accept the reported change scores as clinically important in light of this.

The choice to use both rather than a single measure for function was pragmatic and two-fold. Firstly, the ODI was well established in the main research centre (ROH), both within the spinal surgical and physiotherapy departments, including the physiotherapy led Back Pain Clinic (See Figure 3-1). As the latter clinic was a regional centre, patients were regularly referred from it to the other two trial units (QEH) and (SOH), including details of ODI scores at assessment. It was considered that to continue with the use of this well-established measure would allow familiarity with at least one of the outcome measures and would be undertaken by a large proportion of the potential trial patients routinely. Secondly, RMDQ was also chosen and designated as the primary outcome measure, as it has been suggested that it may be the more appropriate measure when the ultimate levels of disability are expected to be low as a result of treatment (Roland and

Fairbank, 2000), as would be expected within the current trial. It was also apparent that more information was available regarding the clinically meaningful change for the RMDQ, which would assist with sample size and power calculations.

Comparison of the physical component summary (PCS) for SF-36 with other LBP cohort studies revealed some interesting similarities. The mean PCS score for the current cohort was 32.1 (9.5) (n=92) with 30.8 (9.95) being reported in patients with sciatica (Patrick et al., 1995) and 30.4 (9.95) in a non-specific LBP population (Fanuele et al., 2000) indicating that the current cohort do not differ widely from those cohorts in respect to broad ranging physical functioning. The finding that both groups showed a marked improvement in functional level, as measured by both the RMDQ and ODI, was mirrored by the changes in the PCS. Change scores from the latter scale (8.5 and 8.6 for the SSSE and CT groups respectively) indicated an improvement in self-reporting physical functioning including physical functioning, bodily pain and general health (See Figure 2-2), but not a return to the US reported mean of 50 (Ware, 2000). Conversely, the mental component summary (MCS), including energy and vitality, social functioning and mental health, showed a slight reduction (-2.8 and -3.5 for the SSSE and CT groups respectively) indicating deterioration in self-reported mental health. However, these reductions are small, and are unlikely to represent a clinically meaningful change. When compared to the results of the DRAM classifications at baseline and 12-month follow-up, two patterns of change can be seen. A decrease in the number of patients classified as 'At risk' of developing distress was seen and an increase in the number of patients showing no evidence of distress was also seen. However, there was an increase in the number of patients who were distressed (5/68 or 7%) compared to baseline, in addition to the distress level of the two patients entered as a result of a protocol violations remaining unchanged. The lack of change in distress levels of the latter two patients is unsurprising given the lack of any element of either treatment packages specifically designed to address distress. The development of distress by 7% of patients may be a response to the failure of the allocated treatment intervention to alter symptoms. Examination of the group mean RMDQ at 12-months for these patients showed no change from baseline and would support the prior assertion, however the high degree of

individual variation with regard to improvements and deterioration in RMDQ scores challenges this assumption. It should be remembered in comparing the results from the DRAM and MCS element of the SF-36, that both reflect different constructs that are being measured and over different time periods. The questions contributing to the MCS of the SF-36 enquire about feelings over the last four weeks (See questions 9 and 6 in Appendix H), whereas the MZ asks patients to rate how they have been feeling ‘recently’ and the MSPQ to rate feelings over the past week (See Appendix F and Appendix G respectively). The finding that more patients were classified as ‘normal’ and less patients were at risk of becoming distressed was interesting and an important finding, however it could be argued that the small number who became distressed are of greater importance as they potentially will require healthcare interventions or assistance with the management of that distress.

It has previously been reported in a systematic review, that MZ scores can discriminate between recovered and non-recovered patients (Burton et al., 1995). The baseline value of both the MZ and MSPQ are in keeping with those reported from the trial on which this assertion was based (Burton et al., 1995), however that study used an acute and sub-acute LBP population. Additionally, the current study excluded patients with high MZ scores from entrance to the RCT and therefore it was unlikely that very large changes on this measure would be demonstrated. The results of the DRAM classifications for all patients screened for entry into the trial revealed an interesting trend and this will be discussed in relation to the wider literature and implications for practice later in this chapter (See Section 5.4.2).

The results obtained from the ODI, RMDQ and PCS suggests that these tools are measuring similar dimensions of physically based functioning, but with the PCS including a greater measure of function over time e.g. question four of the SF-36 pertaining to physical health over the last four weeks. Interestingly, however, question two of the SF-36 measuring change in health perception compared to one year ago, showed contrasting results. This showed that nearly 50% of patients in both groups perceived their health to be ‘Much worse’ or ‘Somewhat worse’ than

one year ago. This suggests that physical functioning is only one element of overall perceived health, and that, despite marked improvements in physical functioning, patients still perceived that their health had deteriorated. This may be for a combination of reasons. Firstly, although marked improvements in functioning were seen in both groups, residual levels of disability, as measured by the RMDQ, at 12-month follow-up were still 5.42 and 4.83 for the SSSE and CT groups respectively. Although it is impossible to attribute group levels of disability, to individuals, as response will vary with context and situations, the residual disability levels reported may have contributed to the results obtained regarding change in health perception (Question 2, SF-36). Secondly, the SF-36 measures 'generic' health and is not back pain-specific and therefore deterioration in co-morbidity or new health problems may have influenced these scores.

Examination of the results of the ITT analysis using the scenarios outlined in Chapter 3, reveal some interesting findings. Similar results were obtained for the four analysis undertaken for the primary outcome (RMDQ); namely best-case scenario ITT (last-value carried forward), worst-case scenario ITT (return to baseline), completers only (12-months responders) and Area Under the Curve (AUC) for 12-month responders. As such, it has been suggested that increased confidence in the results can be conferred (Altman, 1992). A point of interest is the similarity between the mean change (and 95% CI) reported from the 'best' and 'worst-case' for the conventional physiotherapy treatment group (-3.7, -5.0 to -2.3 and -3.6, -4.9 to -2.4 respectively) and the discrepancy between the 'best case' and the completers only analysis (-3.7, -5.0 to -2.3 and -5.1, -6.7 to -3.6). Again, a number of factors may have influenced these results. The timing and pattern of loss from this group where 11/15 of the losses occurred before the first follow-up may explain the similarity in best and worst-case ITT analyses as it resulted in the baseline figure being inputted in both scenarios; as the last-value carried forward in the best-case scenario and as return to baseline in the worst-case scenario. It is possible that a few completers in the CT group achieved large individual improvements i.e. reduction in RMDQ score, hence the group mean difference would be effected and have contributed to the similarity in results. This is unlikely to have contributed towards the similarity in the 'best' and 'worst-case' analysis, as

its effects would have influenced both analyses. However, it will have been influenced by the level of improvement achieved by patients at their last response, which was carried forward for the 'best case' analysis.

Similarly to the functional outcome measures, two measurements of pain were undertaken with two distinct aims; one to measure the sensory qualitative dimension of the pain experience (SFMPQ) (Melzack, 1987) and the other (NRS and VAS) to measure the sensory–quantitative or intensity (Bolton and Wilkinson, 1998). The SFMPQ has been shown to correlate well with the original long version and demonstrate changes over time in a similar manner (Dudgeon et al., 1993). Results demonstrated small mean reductions, indicating reduction in self-reported pain, in both groups across the three components of the SFMPQ present rating intensity (PRI) i.e. total, sensory and affective scores. These changes however, were very small, especially for the affective components, and are unlikely to represent clinically meaningful changes. The conventional physiotherapy treatment group demonstrated a larger mean change in the total PRI and sensory component PRI than those patients receiving predominantly stabilisation training. However, the changes were very small with a reduction of -2.9 and -1.1 for the total PRI and -2.5 and -0.7 for the sensory component for the CT and SSSE groups, respectively. Although no specific MCID has been identified for the SFMPQ, the minimal metrically detectable change for each component has been reported as 5.2/45 (total), 4.5/33 (sensory) and 2.8/12 (affective) (Grafton et al., 2001) and change in both groups fell well below these levels. If the 99% CI for mean change for the current study, presented in Table 4-6, are compared against these reported minimal metrically detectable changes, it is obvious that both groups fall well below these levels. As such, it is difficult to confidently conclude that these changes are clinically meaningful. The finding that a minimal reduction in pain is achieved alongside a significant reduction in functional disability, replicates the findings from other studies. Hides et al. (1996) reported a marked reduction in functional disability, but little effect on pain of a spinal stabilisation training regime in first episode acute LBP patients. Similarly, a community based

exercise programme using a cognitive-behavioural approach and encouraging normal spinal movement, currently forming part of the UK BEAM trial⁵¹, showed little influence on intensity of pain, but a positive effect on participants' ability to cope with the pain in both the short and long term and as such must be considered as clinically relevant (Moffett et al., 1999). In contrast, however the reduction in the quantitative measure or intensity of pain as represented by the VAS scores in the current study reduced by a clinically meaningful degree in both groups (Beurskens et al., 1996). Beurskens et al. (1996) have reported the clinically meaningful reduction in VAS score (for pain) to be 1.2 and Grafton et al. (2001) reported the minimal metrically detectable change for the VAS as 1.4 cm. Both groups achieved the MCID identified, and in addition, the CT group exceeded the minimal metrically detectable change. Possibly, the measure indicating intensity of pain might have been expected to be reduced with active treatment interventions, whereas tools that measure other dimensions of pain such as the affective component may not readily be affected by physically-based therapeutic interventions.

Although the current study was specifically interested in recurrent LBP problems, after reviewing the available literature, 'recurrence' was not used as an outcome measure for a number of reasons. Firstly, as discussed in previous chapters, much controversy exists as to the definition of recurrence in terms of timescales (Waddell, 1998), natural history (Von Korff, 1994) and classification. Secondly, marked problems exist regarding the assessment of recurrence with the inevitable problem of recall bias, which will occur with long-term follow-up periods. Thirdly, it is now widely accepted that LBP runs a variable and recurrent course with episodes of acute, transient and chronic symptoms (Von Korff, 1994) and as such functional level may be a more appropriate or clinically and socially useful measure of the impact of LBP than rate of recurrence. Certainly, reports exist indicating that stabilisation training has a long term effect on the number of self-reported recurrences when compared to normal 'medical' care but only in patients with

⁵¹ UKBEAM (UK Back pain, and Exercise And Manipulation) trial funded by the Medical Research Council (MRC) and NHS Research and Development programme

acute-first episode LBP (Hides et al., 2000; Hides et al., 2001). However, this study used telephone interviews for the 3-year follow-up, which has limitations due to recall bias.

Another recently reported RCT, with ongoing 2-year follow-up has reported favourable results in a specific-sub-group of LBP patients, namely peripartum pelvic pain after pregnancy (Stuge et al., 2001). This pragmatic RCT randomised 81 women with pregnancy-related pelvic pain between 6-16 weeks after delivery, using stratified randomisation for pain location⁵². The two intervention groups consisted of a predominantly exercise-based programme, the core of which was specific stabilisation training with mobilizations used where necessary, and the second group received 'different' physiotherapy treatment modalities, including ergonomic advice, mobilization, manual therapy and electrotherapy but excluding any stabilisation training. Patients randomised to the stabilisation group had to exercise 30-60 minutes a day, three days a week for 20 weeks whereas the 'control' group received treatment 'approximately' once a week or when required. Although not yet fully reported, preliminary results may support the stabilisation training with a significant reduction in pain (measured on the VAS) and functional disability measured on the Disability Rating Index (Salen et al., 1994). With no dropouts at one year, this trial represents an important indication for the possible use of stabilisation training. However, no firm conclusions can be drawn as to the effectiveness of the stabilising exercises per se as results may largely reflect the effects of a greater level of exercise compared to the control group. In keeping with the current trial, Stuge et al. (2001) used a pragmatic design with treatment tailored on an individual basis and undertook long-term follow-up. In contrast however, the patient population was highly selected. Recent pregnancy, and the associated increased ligamentous laxity, making low-grade clinical instability likely, provided a group who should respond to any form of exercise designed to improve stability such as core stabilisation training. This may explain the differences in results, in addition to the fact that the 'control' group are reported to have received

⁵² 1) Symphysis pubis pain, 2) pain from all three pain joints 3) pain from either one or both sacroiliac joints.

exclusively 'passive' modalities. Despite using the RMDQ as an outcome, there is to date no information regarding these results so no direct comparisons can be made. Additionally it is not clear if there were any differences in the amount and duration of treatment received by either group.

Another ongoing trial is examining stabilisation training in chronic LBP of at least 12-weeks duration, in a group setting as part of a large RCT (Goldby et al., 2000). Consisting of three arms with spinal stabilization, in a structured group setting, individual manual therapy and a placebo control using an educational booklet the trial is planning 3, 6, 12 and 24 month follow-up. The stabilisation training consists of a 10-week programme aimed at rehabilitating the function of the TrA, LM, pelvic floor and diaphragm. The manual therapy group can receive a maximum of 10-treatment interventions (excluding electrotherapy) and the placebo control group are provided with a booklet ('Back in Action' (Cherkin et al., 1996)), encouraged to follow the message it contains and then discharged. Preliminary results (6-month results for n=183) report a statistically significant improvement in the stabilisation group compared to the manual therapy and placebo control group in respect to self-reported LBP and numerical rating scale for LBP. However, seven separate outcomes are being used covering pain (NRS and Pain Diagram), disability (ODI), impairment (Modified Schober's test and timed walking test), handicap (LBOS) and quality of life (NHP), and no indication has yet been reported regarding the primary outcome of interest. Additionally, all patients in all groups are entered to back school, potentially confounding the effects of the different treatment interventions. However, this trial is ongoing with no full-term results to date. When reported in full with details of the primary outcome and associated sample size calculations, more informed decisions will be possible regarding its value.

The explicit purpose of pragmatic trials is to assess the *effectiveness* of packages of care in the context of routine clinical practice (Sim and Wright, 2002). Despite the obvious problems associated with this design, results from such trials should be highly applicable to clinical practice and, therefore, are potentially of great importance to the physiotherapy profession (Wakefield, 2000). As with any

research, the applicability of the results to a wider population, must be considered in relation to the population, setting and interventions studied. As discussed earlier, exclusion criteria were kept to a minimum and inclusion criteria were broad in order to maximise the clinical applicability of the results (Bland, 2000). The trial was based in three physiotherapy centres in Birmingham, UK, the second largest city in the UK and largest city in the Midlands, with a typical metropolitan, multi-cultural population. Two units involved in the trial were University teaching hospitals (QEH and SOH) and one was a regional specialist orthopaedic centre (ROH). Although there is evidence to suggest that variation in medical consultation varies with geographical region, the Midlands has been reported as having an intermediate consultation rate (Walsh et al., 1992) and therefore can be considered fairly representative of a 'typical' clinical population. One of the inclusion criteria was the ability to read English to a level that allowed completion of the questionnaires used. This resulted in more exclusion from two of the units (QEH and SOH) than from ROH because of the patient profile of those specific units. Although resource limitations prohibited the required translations and validation of questionnaires, the subsequent exclusion of patients not fulfilling this criterion must be viewed as a limitation of the trial and be considered when judging the generalizability of the results. Despite this, however, the results should be generalisable to English-speaking communities in similar metropolitan areas.

The trial protocol allowed a maximum of 12 sessions over a 12-week period, which was set following review of the literature and, in keeping with the pragmatic design, to allow the therapist to alter treatment dependent on clinical need. Large variations exist in the number of treatment session patients receive for LBP from 6-11 (Croft, 1994) to 1-19 (Gracey et al., 2002) and 0-24 (Moore, 1997/8). Similarly, the average number of treatments varies, though less so, with reports of 4-6 (Foster et al., 1999), 5 (Gracey et al., 2002), 7 (van Baar et al., 1998) and 11 (Jette et al., 1994). Likewise, the number of treatment sessions used in other studies also varies greatly, from single treatments through to 10 (Cherkin et al., 2000) or 20 sessions (Lonn, 2000; Soukup et al., 2000). In view of the fact that the theory behind spinal stabilisation training is based on alteration of the endurance of the deep abdominal and trunk muscles in addition to improvement in co-

ordination (Richardson and Jull, 1995; Richardson et al., 1998; McGill et al., 1999; Comerford and Mottram, 2001a), an adequate timescale was needed during which training and any potential physiological changes could take place. The theory behind specific spinal stabilisation training is based on addressing certain dysfunctions that have been observed to exist in the presence of LBP symptoms such as loss of anticipatory activation of TrA, decreased muscle endurance and poor co-ordination of synergists, as discussed in Chapter 2. Much of the experimental evidence on which the concept of this form of training is based was undertaken on small numbers of chronic LBP patients, who were in remission or experiencing minimal levels of pain at the time of testing. However, as the current patient population, although suffering from recurrent LBP, fitted into the accepted definition of chronic LBP with a mean duration of symptoms in excess of 3 months (Waddell, 1998), they should theoretically display similar muscular dysfunction. Consequently, with respect to muscle physiology and training, given the correct amount and type of training, these dysfunctions should be ameliorated. However, whether this would result in a reduction in symptoms or change in functioning is the key question, and as yet unknown. A review of many RCTs incorporating exercise based interventions reveals similar timescales to the present study ranging from 10 weeks (O'Sullivan et al., 1997c; Danneels et al., 2001) to 13 weeks (Lonn, 2000; Soukup et al., 2000), suggesting that a realistic timescale was set in which to complete the treatment interventions. Therefore, it should be considered that adequate time was allowed for patients undertaking spinal stabilisation training to potentially achieve physiological muscular and co-ordination changes, if the exercises were undertaken correctly and sufficiently regularly, a factor that for the present study is unknown. The decision not include a specific measure of compliance was based on a review of the literature, especially previous studies investigating spinal stabilisation training. Certain factors are considered potential confounders in any research such as varying compliance, co-intervention (analgesics/self-medication) and spontaneous changes/natural history (van der Linden et al., 1991). As the design of the trial was pragmatic, some variance due to such factors is accepted, but must be considered when interpreting the results. Compliance is an important factor to consider with any RCT, especially one concerned with exercise based therapy that is primarily self-directed

and home based. A subjective record only of compliance was recorded, as a consensus from the literature regarding objectively measuring compliance seems lacking and previous studies have varied widely in their use of measures of compliance. However, the lack of evidence of compliance with exercise regimes, especially the spinal stabilisation training, mirrors current practice where at best highly questionable subjective measures of compliance are used, and at worst no attempt at measurement of compliance is made.

One of the common problems with research in LBP is the difficulty associated with demonstrating improvement in sub-acute conditions, as improvement will often occur naturally (Klaber Moffett et al., 1995; Foster et al., 1999). This will often be reflected in a gradual return to an average level of functioning or pain reporting, known as regression to the mean (Bland and Altman, 1994a). This is common in a relapsing and remitting condition such as LBP, however does not appear to be the case with the current study, as examination of the means and adjusted means for the primary outcome show very little difference, suggesting no clear evidence of regression to the mean (Bland and Altman, 1994b; Bland and Altman, 1994a). This can be interpreted as indicating that both groups achieved a true change and that a real treatment effect was achieved for both interventions, despite no difference being demonstrated between the groups. Of course, a third, no treatment or placebo group would have assisted in delineating any differences, however there is evidence, at least in preventative trials, that recruitment to trials with a placebo arm may reduce participation rates (Welton et al., 1999). Additionally, there may have been problems obtaining ethical approval for a trial with a no-treatment arm and recruitment of clinicians to participate would have been more difficult. The other important factor that must be considered is the natural history of LBP of recurrence and remission (Von Korff et al., 1993; Von Korff, 1994; Croft et al., 1998). It is well documented that LBP runs a recurrent course (Deyo and Weinstein, 2001) with reports of recurrences affecting 40% of patients within six months (for a US population) (Carey et al., 1995). It is accepted that patients will generally seek care when symptoms are at their worst and therefore functional disability levels at the start of the trial would have been

expected to be high, although were comparable with those from similar studies (Bombardier et al., 2001). It must also be considered that the results obtained are simply a reflection of the natural history rather than any specific effect from either treatment package. However, this argument is mitigated against, in view of the recurrent nature of the study population, as the effects of recurrent episodes of LBP would have resulted in a reduction in level of functioning (RMDQ score). In contrast, however, both groups maintained a marked reduction in RMDQ scores at long-term follow-up, suggesting that both physiotherapy treatment packages were effective in reducing functional disability. However, no specific element of the packages can be identified as responsible for this improvement.

The decision to incorporate the use of an advice booklet in both treatment groups stemmed from the evidence that suggests that recall of medical advice can be poor in chronically ill patients (Kravitz et al., 1993) and the provision of written information may facilitate retention. This therefore needs to be addressed in the design of investigations into treatment effectiveness. The provision of an educational booklet, *The Back Book*, for all patients was included as it incorporates an evidence-based message of early return to activity stressed in the most recent LBP guidelines (Burton et al., 1996; Burton et al., 1999) and a recent international comparison of clinical guidelines for the management of acute low back pain (Koes et al., 2001a). It has been suggested that a strong, consistent message, supported by the whole healthcare team, is most likely to change clinical outcomes (Cherkin et al., 1996), which was reinforced by a recent study that reported a negative interaction between a detailed educational booklet and advice (Little et al., 2001). Consequently, the information contained within the booklet chosen for the trial needed to be simple, positive and consistent with current practice in the three units studied. *The Back Book* fulfilled these requirements and had also been developed for use in conjunction with other treatments such as osteopathy and physical therapy (Burton et al., 1996; Burton et al., 1999). Previous randomised studies that have assessed the effects of *The Back Book* have reported favourable results (Burton

et al., 1999). Used alongside usual care by either a GP or osteopath⁵³, a statistically significant difference in the main outcome, which was fear avoidance beliefs about physical activity (Waddell et al., 1993a), was demonstrated between the group receiving the booklet and the group not receiving the booklet, and maintained at one-year following discharge. Therefore the incorporation of *The Back Book*, as the 'advice' element to the trial was considered appropriate. Limitations may exist however in relation to the uses made of the provided resource as no attempt was made to adherence to the message of the booklet. Similarly, the degree to which the advice was followed was not assessed formally, however the message of return to normal activities was stressed throughout treatment sessions for both groups. One unit (ROH) routinely used these booklets, and as all the units in the trial adhered to current guidelines, inclusion of the booklet was a natural progression and reinforced current practice.

The impact of the additional booklet provided to the SSSE group was not assessed separately, but was provided to facilitate the training process and provide some standardisation for the provision of exercises. It is generally accepted that re-training of motor control dysfunction is a cognitive process and, as such, visualisation (Richardson et al., 1998; Comerford and Mottram, 2001a), knowledge of the aims of treatment, feedback regarding correct activation patterns and precision are vital for successful interventions. Indeed, the exercise programme on which the stabilisation intervention was based (Richardson and Jull, 1995) highlights the importance of illustrations as an effective teaching aid and work undertaken by Miller and Mederios (1987) highlighted the importance of multisensory input on successful abdominal training. This importance placed on the individual and the cognitive nature of abdominal re-training also supports the use of individual rather than group training. The trial design examined two treatment packages delivered on an individual rather than group basis. Despite the popularity of group treatment for LBP, a recent Cochrane Review found only

⁵³ GP patients received reassurance and advice, sick certification and analgesics.

The osteopathic treatment consisted of general reassurance, advice, recommendations about non-prescription analgesics and sick leave as appropriate in addition to manipulative therapy (mean=4.3 sessions)

moderate evidence that back-schools were more effective in the short-term than other treatments (van Tulder et al., 2000a) although they did conclude that back school may be of use for patients with recurrent or chronic LBP. The specific nature (Richardson and Jull, 1995) of the exercises forming the mainstay of one of the treatment packages necessitated individual treatment rather than a group setting. Additionally, current practice at the time indicated that the majority of physiotherapeutic management of LBP was undertaken on an individual rather than group basis (Foster et al., 1999). Although it has been reported that outcome is not affected by the mode of delivery⁵⁴ i.e. group or individual (Rose et al., 1997), motivational issues associated with treatment in group settings may be important. However, the majority of studies using stabilisation training, to date, have used individual treatment. Interestingly, in an early, retrospective study, Saal and Saal (1989) identified the need for a meticulous technique for stabilizing training and despite the use of a class setting for rehabilitation, patients were initially taught the exercises on a one-to-one basis. This echoes the sentiments of many authors that these techniques require specificity and, at least initially, individual teaching is required in order to achieve this level of precision. Additionally, evidence indicating that activation of the more superficial abdominal muscles, such as rectus abdominis and external obliques, tends to exert an inhibitory effect on TrA activation (Richardson et al., 1992; Richardson and Jull, 1995) supports the need for very close supervision of training in the early stages. Saal and Saal (1989) also stress the active component of the rehabilitation, stating that decisions regarding the advancement of the programme were based on functional rather than pain levels. This is interesting in light of the current results, and those from other trials, which indicate a greater reduction in functional disability levels following treatment than necessarily pain levels.

The finding that the majority of patients received a combination of treatment modalities with the most frequent combination being manual therapy and exercise reflects previous research findings (Jette et al., 1994; Moore et al., 1996; Foster et

⁵⁴ Pain management programme for chronic LBP patients based on cognitive behavioural principles.

al., 1999). Additionally, the low frequency of electrotherapy as first choice treatment modality mirrors the results reported by Moore (1997/8). They found that patients who completed a full course of treatment received less electrotherapy utilised as a first treatment strategy, compared to those patients who ceased to attend, required further investigations or who were deemed unsuitable for physiotherapy management. Of the cohort studied, 69% received some form of active exercise including re-education of muscle imbalance or 'active' exercises compared to 98% in the current trial. Similarly, Jette et al. (1994) showed that in a representative sample of over 2000 patients who received physiotherapy for LBP in the USA, 76% of them received some form of exercise therapy, however no further details were provided. Direct comparison with a UK population of these relatively high levels of exercise use is difficult due to differences in categories used in the surveys. In a survey of Irish and British physiotherapists undertaken by Foster et al. (1999), McKenzie regimes (McKenzie, 1981) and abdominal exercises were the two forms of exercise surveyed and were classified separately. With 47% and 17.5% of physiotherapists reporting McKenzie and abdominal exercises respectively as the most frequently used treatment, it was clear that exercise was being used in the management of LBP but to a lesser extent than reported by Jette et al. (1994). Additionally although respondents acknowledged the importance of exercise in the management of LBP there was poor consensus regarding the optimal type or level. The 17.5% of physiotherapists reporting using abdominal exercises as the most frequently used treatment modality (Foster, 1998) is in contrast to more recent reports (Gracey et al., 2002). Interestingly, despite the claimed increase in the use of spinal stabilisation training, a recent survey by Gracey et al. (2002) reported muscle re-education (including abdominal training) being used in the management of only 8.1% of LBP patients (n=1062) in Northern Ireland, suggesting other techniques are perceived to be more effective as first line management. Although for ease manual therapy techniques are referred to as 'passive', comparison of the level of passivity with treatment modalities such as electrotherapy and mechanical traction, may challenge this classification. Although it is accepted that manual therapy techniques are not 'active' in the way that aerobic exercise or self-directed stretching programmes are, they do include manipulative therapy, involve regular active re-assessment of the pain provocation on active

movements and a 'hand-on' approach in a way that electrotherapy and lumbar traction do not. Possibly a compromise classification might be semi-passive in contrast to exclusively passive modalities, but either way, it may be that the use of active and passive is an over simplification of what is involved in the design and delivery of physiotherapy and may hamper attempts to identify the effect elements of physiotherapy treatment packages.

The findings reported by Moore (1997/8) that electrotherapy utilisation as the most commonly used modality increased in patients who failed to complete a course of treatment, required further investigation or were deemed unsuitable for treatment, is interesting in light of the current findings. Findings from the current trial, demonstrated an increasing use of electrotherapy as secondary modality choices, possibly when patients failed to respond to the first choice modalities of exercise and manual therapy. Further examination of the results is planned to ascertain if any relationship exists between the use of more passive treatment modalities such as electrotherapy, number of treatments and relative success of treatment e.g. is increased use of electrotherapy proportional to increasing number of treatment and inversely proportional to overall success of treatment.

5.4.2 Measure of distress

Although the number of patients not showing any evidence of distress increased following treatment and those at risk of becoming distressed decreased, the treatment packages examined did not expressly include any psychologically based interventions. Possibly the reduction of fear, reassurance about the benign nature of the condition and non-specific treatment effects may all have influenced participants reporting of distress. However a small percentage of patients (5/68; 7%) of patients moved from being 'At risk' to displaying distressed during the follow-up period, although interestingly all remained in the 'At risk' at discharge and four out of the five at 6-months.

More interestingly however, was the finding that approximately one third of patients referred to a secondary healthcare setting with recurrent LBP show clear evidence of distress, as measured by the DRAM. Although consistent with findings from similar studies (Main and Watson, 1995; Grevitt et al., 1998; Hope and Forshaw, 1999), when compared with results from studies in a primary care setting, greater differences are evident, with Waxman et al. (1998) reporting a distress level of 18% and Burton et al. (1995) reporting 19%.

Two studies, one by Grevitt and co-workers (Grevitt et al., 1998) and the other by Hope and Forshaw, reported similar levels and classifications of distress to those from the current study. Of 125 new patients attending a back pain clinic, Grevitt and co-workers reported 28% as distressed, with 25 (20%) in the distressed depressive group and 10 (8%) in the distressed-somatic group. Hope and Forshaw (1999) investigated the level of distress in 160 LBP patients referred to a physical therapy department using the DRAM and found a 32% distress rate, with 26% classified as distressed depressed and 6% as distressed somatic. They subsequently re-examined the DRAM scores of the patients following treatment and found a similar level of distress. Interestingly, Hope and Forshaw (1999) also reported no meaningful change in the functional disability level of patients in the distressed groups following physical intervention, with a pre-treatment group mean of 18.4 and post-treatment score of 16.4 RMDQ points, suggesting that physical intervention in isolation does little to influence distress or disability levels in LBP patients. This finding supports the conclusion of Main et al. (1992) that there is an increasing relative risk of poor outcome with increasing levels of distress. One limitation of the study by Hope and Forshaw is the high loss to follow-up, representing 44% attrition; despite this however these authors joined the growing consensus that patients presenting with LBP and concurrent psychological distress are unlikely to have a successful outcome with physical interventions alone.

Hope and Forshaw (1999) also reported that increasing levels of distress were highly correlated with functional disability, reporting significant correlations between disability (as measured by the RMDQ) and depression, and between disability and anxiety (somatisation). Comparison of distress levels and functional

disability scores from the present study with those from Burton et al. (1995) and Hope and Forshaw (1999) reveal interesting differences. Despite reporting differing levels of distress, with 33% in the current study and 19% from Burton and et al., both studies reported similar mean RMDQ scores, of 8.0 (SD 5.2) (n=212) and 8.8 (SD 4.9) (n=212) respectively. Hope and Forshaw reported a near identical level of distress to the current study but a much higher mean RMDQ score, 18.4 (no SD reported, n=42) compared to 14.8 (SD 4.6, n=41), respectively, for patients classified as distressed depressed. One possible explanation for these discrepancies may be differing type of populations and healthcare settings in the three studies. Despite the current study and that of Hope and Forshaw (1999) being set in secondary healthcare physiotherapy departments, the latter cohort was obtained using no exclusions, whereas the current study consisted exclusively of recurrent LBP patients with no evidence of frank nerve root compression. In contrast, the population studied by Burton et al. (1995) consisted of a high proportion of acute cases within a primary care setting. Consequently, the influence of healthcare setting and stage of LBP may have an effect on the relationship between distress levels and functional disability. Another possible explanation for the why higher functional disability levels were not found, despite the relatively high levels of distress identified may be the recurrent nature of the cohort. Possibly DRAM scores were influenced by previous experience, and were reflecting greater concern regarding the current situation without translating into higher levels of functional disability. If so, the use of psychological questionnaires as part of the clinical assessment process in physical therapy would assist in the identification of concerns manifested as distress.

The current study specifically targeted recurrent LBP patients, and the higher levels of distress identified may be a result of previous experience of pain and disability due to LBP. Clinically, this presents a number of issues. It is of importance as this group of patients may be at increased risk of failing to return to normal activity secondary to high levels of psychological distress. Additionally, it certainly indicates that the management of LBP patients by physiotherapy needs to vary according to the duration and the presenting factors.

The reasons for patients seeking medical consultation are complex and multifactorial and can often influence subsequent management. In a prospective, longitudinal study of psychosocial factors associated with the consultation for LBP in the community, increasing levels of distress were found to be associated with an increased likelihood of consultation for LBP episodes of over 3-months duration (Waxman et al., 1998). The overall distress rate from the study by Waxman et al. (1998) of 18% is very similar to the 19% reported by Burton et al. (1995) who also studied LBP in the primary healthcare setting. The different levels of distress found in the primary and secondary healthcare settings as outlined in Appendix GG may, in part, be a result of the ongoing medicalisation of back pain (Borkan and Cherkin, 1996). Previous work has suggested that the level of distress and illness behaviour may influence the amount of treatment received more than the severity of symptoms. Therefore, clinically, 'objective' evidence of distress would be constructive. This is of importance to physiotherapists involved in the management of LBP. To date, there has been much written in the literature regarding the role of psychosocial factors and distress in LBP, but aside from certain acknowledged centres of expertise, there is little evidence that these factors are being recognised or that any significant change in clinical practice has occurred. Increasingly, physiotherapists are being encouraged to undertake first-contact practitioner roles, where often little or no previous screening, either formal or informal, has been undertaken. It is therefore a challenge for the physical therapy profession to undertake these processes successfully and meaningfully in the current clinical environment.

The current results indicate that approximately one third of patients with recurrent LBP referred for physiotherapy at the centres studied exhibited levels of distress that have previously been shown to increase the relative risk of poor outcome from physical intervention alone, by 3 to 4 times. These findings are in agreement with those from previous studies using populations from secondary healthcare settings. The review and summary of published papers reporting the use of the DRAM in the clinical setting has identified a fairly consistent level of distress with some variation across healthcare settings. The consistency with which significant

levels of distress have been identified strengthens the argument for the need to utilise, judiciously, questionnaires such as the DRAM in clinical practice.

Current best practice guidelines recommend the integration of a biopsychosocial approach into the physiotherapy management of LBP but provide little indication of how this should be achieved. Physiotherapists and their medical colleagues have, for many years, been informally identifying and addressing the psychological problems associated with both acute and chronic musculoskeletal pain states. The judicious use of screening questionnaires may help to formalise this and assist in identifying when liaison with other professionals, such as clinical psychologists, or further psychological assessment is necessary alongside appropriate physical interventions.

The decision to exclude patients showing evidence of psychological distress i.e. those classified as DD or DS on the DRAM, was undertaken for a number of reasons. As spinal stabilisation training were relatively new techniques, and no large-scale trial of their effectiveness in a general LBP population had been undertaken, it was considered that the inclusion of patients known to be at an increased risk of poor outcome with physical treatment alone (Main et al., 1992) would confound the assessment of the effectiveness of these techniques. However, the design could have incorporated the entry of patients showing evidence of distress into the trial and analysed them separately. This would still have allowed the assessment of the effectiveness of spinal stabilisation training in a non-distressed population, and may have provided some useful additional information.

5.5 Methodological quality

Quality of RCT design and execution and also reporting is key to the strength of this form of methodology. Numerous frameworks exist for judging the methodological quality of RCTs (Koes et al., 1995; Altman, 1996; van Tulder et al., 2000b) and it is generally accepted that there has been a gradual improvement in the design and reporting of RCTs in recent years (Koes et al., 1995; van Tulder et al., 2000b). In their recent systematic review of exercise therapy for LBP, van

Tulder and co-workers (2000) used a modified criteria list from previous reviews (See Appendix HH) to assess the quality of published RCTs and this is summarised and completed for the current trial in Table 5-1.

Table 5-1: Methodological Quality Criteria

	Criteria	Yes (+), No (-) Don't know (?)
1	Concealment of treatment allocation	+
2	Withdrawal/dropout rate	+
3	Co intervention avoided or equal	+
4	Blinding of patients	+
5	Blinding of observer	+
6	Intention-to-treat analysis	+
7	Compliance	-
8	Similarity of baseline characteristics	+
9	Blinding of care provider	?
	TOTAL	7
High Quality $\geq 5/9$? used where criteria cannot be assessed or not applicable to specific situation Adapted from; van Tulder et al. (2000b)		

A number of points are raised from examining the work on methodological quality and these will be discussed in the next sections in relation to the current study.

5.5.1 Pragmatic designs

It is widely acknowledged that many problems are associated with RCTs and particularly RCTs within LBP research, including difficulties with LBP diagnosis and classification (Riddle, 1998), the need to make methodologies clinically relevant (Koes et al., 1995) and the influence of confounding factors (van der Linden et al., 1991). Pragmatic, or management RCTs are specifically designed to describe the consequence of a given intervention under circumstances mimicking clinical

practice as closely as possible (Jadad, 2000). Testing 'packages' of care under conditions reflecting routine clinical practice, they are concerned with the effectiveness of treatment in the routine clinical setting (Sim and Wright, 2002) and as such, pragmatic trials are aimed at decision-making (Schwartz and Lellouch, 1967). However, in order to achieve this, a degree of laxity in the inclusion criteria is allowed and often an active control, e.g. conventional therapy, is used rather than a placebo. The use of a pragmatic design, although beneficial and strongly advised for evaluating current practice, may have influenced the results obtained. The conventional treatment group within this trial received management at the discretion of the treating clinicians, which could include any physiotherapeutic technique currently in use clinically in the UK, excluding acupuncture, hydrotherapy and back school. As all three units participating in the trial had actively adopted the CSAG (1994a) and subsequent guidelines and both groups received an active rehabilitation approach to treatment with minimal use of exclusively passive modalities such as electrotherapy and lumbar traction, as evidenced by the low frequency of usage recorded (See Table 3-1). Although potentially reducing the differences between the two treatment packages and making differences in outcomes between the groups more difficult to identify, this strategy does replicate closely current clinical practice which has been shown consistently to rely on combination or treatment methods in the management of LBP and other musculoskeletal disorders (Jette et al., 1994; Evans and Richards, 1996; Moore, 1997/8; Foster, 1998; Foster et al., 1999; Li and Bombardier, 2001; Gracey et al., 2002). This design also ensured maximum clinical applicability of the results. Although the main difference between the treatment packages was the use of spinal stabilisation exercise training, the lack of physical or physiological measures of muscle function, would have precluded the inference of any causal effect had differences between the groups been identified. However, the pragmatic design aimed to assess effectiveness rather than specific causal relationships and at present no reliable, accepted clinical tests to directly monitor or measure specific muscle activation and endurance are available for use in the clinical setting.

The trial protocol was explicitly designed to replicate clinical practice, and as such allowed physiotherapeutic techniques currently in use in the UK to be utilised by both groups. Although generally, co-interventions are avoided in RCT designs, Koes et al. (1995) states that inclusion in the protocol should be determined by the specific research question (Koes et al., 1995). As the research question concerned the effectiveness of stabilisation training techniques in the management of a recurrent LBP, as presently used in the UK, co-interventions, i.e. allowing some commonality of treatment techniques, were considered necessary and appropriate. Practicalities also needed to be considered when planning the protocol. Additionally, the target population presented with varying duration of symptoms and as evidence from the literature indicates that the reason for consultation varies with duration (Waxman et al., 1998) this needed to be reflected in the study design. It was felt that clinicians needed to be able to address the main presenting problem with the appropriate techniques based on their clinical expertise, knowledge and current guidelines, in the absence of consensus or evidence for the efficacy of many physiotherapeutic techniques in the treatment of LBP. A placebo-arm was not used in the design for a number of reasons. Firstly, the trial was designed to replicate current practice, and evidence suggests that patients referred for physiotherapy will normally receive some form of active intervention for between 4-6 (Foster et al., 1999) and 11 sessions (Battie, 1994; Jette et al., 1994). Secondly, recruitment to a trial with a placebo arm has been reported to adversely affect recruitment (Welton et al., 1999). Thirdly, treatments comparators need to be equally credible and acceptable to patients to facilitate the blinding of patients to treatment allocation (van Tulder et al., 2000b).

The optimum level of skill required to use specific spinal stabilisation techniques, is as yet unknown. Many trials to date reporting positive results have used 'expert' clinicians who are specialists in the application in of these techniques (O'Sullivan et al., 1995; Hides et al., 1996; Hides et al., 2001), which does not represent the current situation regarding the provision of physiotherapy treatment in the UK (Foster, 1998; Foster et al., 1999). At the time of designing this trial, for most physiotherapists, experience of stabilisation techniques was acquired primarily

through post-graduate course attendance and in-service training with the obvious associated cost implications. For the purposes of the current trial, 'experienced' clinicians were recruited who had a basic level of training in the concept of spinal stabilisation training and further instruction was provided to ensure uniformity of treatment provision across both therapist and treatment centres. Therefore it can be assumed that this particular component of the treatment packages was appropriately applied which is interesting in light of the findings that the addition of specific spinal stabilisation training, as studied, did not result in any additional demonstrable benefit. With increasingly broad undergraduate physiotherapy degree programmes, many physiotherapists continue to pursue post-graduate courses, of varying lengths, and costs, which purport to develop their skills in these specific techniques, with, as yet no evidence to support this.

In summary, the advantages afforded from a pragmatic design to answer the current question were considered to outweigh the problems associated with a level of potential confounding extraneous variables.

5.5.2 Randomisation

In all assessment of RCT quality, randomization, its design, execution and reporting are strongly stressed and weighted (Koes et al., 1995; van Tulder et al., 2000b). Altman (1992) defines random allocation as all participants having an equal chance of being assigned to either group (Altman, 1992) and the purpose is to provide a balanced baseline and homogenous groups, and reduce the risk that important factors unevenly influence the outcome (Altman, 1992; Altman and Bland, 1999; Jadad, 2000). However, it is important that if factors are known, *a priori*, to have an effect on either prognosis or outcome of treatment, they are used as stratifying factors in order that balanced groups are achieved at baseline (Birkett, 1985; Treasure and MacRae, 1998; Treasure and MacRae, 1999). A pseudo-randomisation method was utilised incorporating minimization, as discussed in the Chapter 3, with three stratifying variables. Although it has been argued that trials with less than 50 patients in each group have an increased risk of unknown important prognostic indicators being unevenly distributed across groups, the

success of this randomisation strategy, with regard to known variables, is obvious when comparison of the study variables at baseline is considered. Selection bias due to prior knowledge of the treatment allocation was avoided as allocation concealment was strictly observed. Although the use of a telephone randomisation system would have been reduced the suspicion of bias, cost implications prevented this option. However, the use of a minimization procedure provided a practical, successful solution to randomisation for the present clinically based study, providing groups that were highly comparable.

The choice of stratification variables was based on a number of factors, as discussed in Section 3.6.1. With respect to the use of RMDQ scores as a stratifying variable, in retrospect, it may have been useful to subdivide the baseline RMDQ scores down further, into the mutually exclusive categories suggested by Stratford et al. (1998) (See Table 2-2). This would have facilitated the examination of subgroups based on the initial baseline scores (0-8, 9-16 and 17-24), but would have inevitably resulted in an increased sample size requirements. However, the use of these categories suggested by Stratford et al. (1998), based on initial RMDQ score would have allowed the MCID based on the band score to have been calculated rather than using an overall change score of 5-points. Although the RMDQ change scores were examined in relation to these bands, the trial was not powered to detect this level of change.

Extent of radiating symptoms, as defined by the QTF classification, was chosen a one of the stratifying variables, as at the time of designing the trial there was evidence that laterality of radiating symptoms was a prognostic indicator of outcome. However, recent work has challenged this assumption (Padfield et al., 2000). Categorising 276 LBP patients attending for physiotherapy, using the QTF classifications, Padfield and co-workers (2000) reported that the QTF differentiated groups of LBP patients with respect to pain at initial assessment, and change in pain over the course of treatment but showed no difference between groups in the attainment of a clinically meaningful change in function (RMDQ). However, it is considered that extent of radiation remains an important factor to

consider when planning clinical trials as success of treatment should not be based solely on one outcome variable.

5.5.3 Blinding

Although double-blinding is the gold standard of RCT design (Pocock, 1983), by necessity, single-blinding is required where clinicians direct a physically based treatment. Although identified as an important quality indicator by van Tulder et al. (2000) (See Table 5-1), it is well accepted that blinding of patients and care providers in studies of exercise therapy is difficult. All eligible patients were approached regarding entry to the trial and a standard consent procedure used to reduce the possibility of discouraging consent and thus avoiding selection bias (Jadad, 2000). However, as the consenting procedure necessarily involved a description of the two treatment packages within the trial, although not officially informed of the result of randomization, a simple process of elimination on the part of the patient, would have revealed patients grouping. This may have had an effect on outcome if patients had strong preferences to one of the treatment options and preferences were not recorded. However, the number of patients approached to participate in the trial and not consenting was low ($n=5$) and therefore randomisation preference may not have yielded much useful information, although it is accepted that preferences may influence compliance and outcome (Moffett et al., 1999). The fact that both treatment groups were credible and acceptable to patients and included an element of exercise therapy may have assisted in the blinding procedure to group allocation. One possible option would have been to question patients at discharge as to which group they thought they had been randomised into, to assess the credibility of the two groups. However, this was not undertaken but may have had an effect on the issue of blinding of outcome assessment identified in Table 5-1. As outcomes were self-administered questionnaires with no physical measurements or assessments, the issue of blinding of observations is not clear-cut. Patients were naïve to the outcome of randomization, as discussed previously, and therefore in theory were 'blind' as observers when completing follow-up questionnaire. However, in reality, patients may have identified their group allocation because of the content of treatment and

this may in term have effected their perception of outcome and the effectiveness of treatment. The written patient information indicated that spinal stabilisation training (*specific tummy exercises*) were the technique under investigation and this might have influenced perceptions regarding the effectiveness of the exercises. Additionally, necessarily during the consenting process, a brief description of both treatment packages was given with some patients' expressing preferences at that time such as affirming that muscle strengthening was what they felt would help their LBP or identifying that their abdominal muscles were weak. This again may have influenced the results obtained, as there was no facility to alter treatment allocation following randomization. This may also have had an influence on patients failing to complete a course of treatment, if randomisation had resulted in allocation to a group that was perceived to be less effective. This may indicate a lack understanding of the true meaning of randomisation by some patients, despite clear written and verbal information during the consenting procedure. This was also illustrated by reports from study clinicians that a number of patients allocated to the group not receiving stabilisation training, enquired when they would start the exercises, referring to the spinal stabilisation training.

In summary, blinding of patients to allocation of treatments including exercise is difficult and fraught with problems, but procedures from the current trial ensured that participants were presented with credible treatment groups and options to facilitate the success of the procedure as far as possible.

5.5.4 Sample size

Sample sizes and calculations of the power of trials are fundamental to the assessment of quality and applicability of results. It is well accepted that small sample sizes are associated with an increased risk of Type II error i.e. failing to detect a difference when one exists (Altman, 1992; Koes et al., 1995; Freedman et al., 2001). The sample size for the trial was calculated following a pilot study, as there was a lack of reliable evidence available regarding the functional level of the target population. This provided a standard deviation for the RMDQ of 6 and as a difference between groups of 5 RMDQ points was used as a clinically meaningful change, the desired effect size was large and subsequently a relatively small sample

size was required (total 64 without attrition and 92 with attrition). The decision to use this difference between groups was based on evidence from the literature and the need to produce clinically applicable results (Stratford et al., 1996a). However, the use of an active 'control' group reduced the potential differences between the groups and, as both groups improved with treatment, a further large improvement above this common level was required in order for any statistical significance to be reached. Although powered to find the 5-point difference, one option may have been to power the trial to find a clinically meaningful within group difference (such as 5) and a smaller between group difference. The problem associated with this scheme however, is the identification of what a clinically meaningful 'between' group difference would have been, as statistical significance does not equate to clinical significance (van der Linden et al., 1991; McClure, 1998). As discussed earlier, the use of MCID based on initial RMDQ scores may have been appropriate but as over 200 patients were screened in order to achieve the sample for the trial, the increased sample size required to analyse sub-groups may have resulted in the trial becoming unfeasible.

A similar issue arose in a recent study by Klaber Moffett and co-workers, using the RMDQ as the primary outcome in a study of patients, aged between 18 and 60 years, with mechanical LBP of at least four weeks' duration but less than six months (Moffett et al., 1999). Initially, a mean between group difference of 1.5 points was chosen as clinically meaningful, requiring a sample size of 300. However, due to slower recruitment than expected, recruitment was stopped after 187 patients. Whereas a sample of 300 would have given 90% power at the 5% significance to detect a 1.5-point difference between mean change, using a standard deviation of 4, the smaller sample reduced the power to 72%, with 90% power to detect a 2-point difference. Although conceding that the mean changes in scores were small, this study reported statistically significant changes at the 6 and 12-month follow-up points. Although these changes may be clinically meaningful, evidence from the literature suggests that considering change across the entire RMDQ scale requires larger changes to correspond with clinically meaningful changes in functioning (Stratford et al., 1996a). This example highlights the

problems associated with incorporating clinically meaningful change into research designs, which are practical and feasible to complete in a clinical environment.

The current study used a power of 90% rather than 80% in order to reduce the risk of a type II error and achieved an actual power of over 95% (Altman, 1992). This indicates that the trial was powered correctly to detect the 5-point difference between group's mean change, but may have contributed to no statistical difference between groups being identified. As both treatment packages appeared effective, a high within group change was demonstrated which had not been incorporated in the initial calculations. The sample size was calculated to allow for 10% attrition at each follow-up point i.e. a total of 30%, as this is accepted as realistic for trials with longer term follow-up with measurement at multiple points (Gibaldi and Sullivan, 1997; Staquet et al., 1998) and the attrition rate was within these limits. A number of steps were undertaken in order to minimize loss to follow-up including recording the address of a relative that could be used as a contact in the event of the patients moving house etc, and recording of both home and work telephone number where applicable (Weinstein and Deyo, 2000). A major concern however with any level of attrition, is that those patients may differ in some fundamental way to the remainder of the sample e.g. represent patients who showed marked improvement or deterioration or those with higher or lower baseline scores than average. Non-responders i.e. patients who failed to respond to a second mailing, were sent a pre-paid return post-card with tick-boxes indicating the reason for non-response. Perhaps, not surprisingly, only 2/29 of these were returned, detailing resolution of symptoms and family commitments as the reason for non-response. However, as baseline variables were so comparable, results should be generalisable to a wider population. The effect size achieved in the current study, i.e. the ratio of the difference between group means divided by the total variation of the sample, (Cohen, 1977) was very small at 0.2 (0.75/3.9). This effect size is considered small (Cohen, 1977) and would have required an unrealistic sample size of approximately 1000 patients to detect this effect size and is unlikely to represent a clinically meaningful change. Again, although with the benefit of hindsight, a within group change might have been considered within the power calculation, no evidence appears to exist that identifies what combination of

within group and between group changes are clinically meaningful and therefore remains subjective decisions at present.

5.5.5 Patient management and withdrawal

Care was taken during the planning and recruitment phases of the trial to monitor the number and outcome of all patients entering the RCT as this is stressed in all quality assessment criteria (JAMA, 1994; Koes et al., 1995; van Tulder et al., 2000b). The number, progress and outcome of all patients entering the trial is clearly shown in Figure 3-2 in Chapter 3. Six patients in the SSSE group and 11 in the CT group were not followed-up at discharge with three in the former and two in the latter ceasing to attend for treatment and therefore being classified as non-treatment completers. A further two patients in each group failed to complete the course of treatment but completed discharge questionnaires equating to the 5/47 and 4/50 patients who were non-treatment completers in the SSSE and CT groups respectively. All patients were informed of their right to withdraw without prejudicing future treatment at the time of consent. Only 3 patients were withdrawn from the trial; one prior to 6-month follow-up not stating the reason, one developed cauda equina symptoms following randomisation but prior to the start of treatment and the third deteriorated significantly during treatment and required a surgical opinion. This low level of withdrawal reflects the success of the recruiting process and is in marked contrast to the high levels recorded in the pilot as described in Section 3.9.2.

Of the 97 patients entered to the trial, 8 failed to complete a course of treatment and as such could represent either those who were better or those who deteriorated markedly. Either way it is important to compare this group with those patients completing a course of treatment to ensure that there is no systematic bias. Similarly, non-responders to 12-month follow-up may differ in important prognostic indicators from responders. In neither case did any differences exist between these groups, indicating that analysis of the treatment completers and responders as planned were appropriate.

Due to the pragmatic nature of the trial, inclusion criteria were kept deliberately broad and exclusion criteria to a minimum. Although exclusion and inclusion criteria were checked prior to consent, of the 97 entered into the trial, 2 were subsequently found to fulfil exclusion criteria, in both cases exhibiting distress as measured on the DRAM. Both patients were classified as distressed depressed and both were randomised into the conventional treatment techniques group. Although potentially representing a source of bias, it was decided not to exclude these patients from analysis as response to 12-month follow-up, indicated one patient had improved and one deteriorated slightly and therefore were not a source of systematic error.

5.5.6 End-points, outcomes and statistical analyses

The timing of the end point and the use of multiple outcomes were designed to improve the clinical applicability of findings. There is weight of evidence to suggest that many conservative treatments for LBP work in the short-term but little to support treatments altering the natural course of the condition. The decision to use 12-month follow-up as the primary end-point reflects the importance placed on longer-term outcomes in LBP research (Bombardier, 2000b). The use of multiple outcomes was undertaken in order to reflect the consensus that LBP research needs to reflect various domains of health e.g. function, pain reporting, quality of life (Bouter et al., 1998), and is also recommended when pragmatic designs are used (Roland and Torgerson, 1998a). However, as recommended in such situations, an *a priori* decision was made designating the functional level, as measured by the RMDQ as the primary outcome (Pocock, 1997). Controversy exists regarding the analyses of secondary outcomes in clinical trials with persuasive support both for (Davis, 1997; Prentice, 1997) and against their use (O'Neill, 1997). It has been suggested, as trials are powered based on the main outcome and sample sizes calculated accordingly, that analyses of secondary outcomes should be undertaken at a significance levels that guards against the problems associated with multiple testing (Davis, 1997; Prentice, 1997; Sim and Reid, 1999). This strategy was used in the current trial with secondary outcomes analysed using a 1% alpha compared to the 5% level used for the primary outcome.

In the present case, this is unlikely to have affected the results greatly due to the large within group change and small between group differences.

It is accepted that there are numerous potential confounders to LBP research which should be considered when clinical trials are designed and monitored including age, gender, current employment status and smoking status (Croft et al., 1996). This demographic and biographic information was collected at baseline and examined to ensure equal distribution between groups. Return to work (RTW) is accepted as an important aim in treatment for LBP and is often used as an outcome (Bombardier, 2000a), but as the majority of the cohort (67%) was working, this was not considered an appropriate outcome for the current study. Additionally, it is well accepted that RTW is associated with multifactorial issues including job, remuneration and peer support (Gronblad et al., 1996; Dionne et al., 1999; Deyo and Weinstein, 2001).

In conclusion, results from the trial indicate that both treatment packages studied were effective in reducing functional disability and pain intensity and improving quality of life, as measured by the SF-36, but showed no clinically meaningful change in levels of psychological distress. However, the results do not indicate superiority of either package or which elements of the treatment packages effected the change.

The next section will highlight the problems and limitations of the current trial. These have been discussed in the appropriate sections throughout this chapter but will be summarized here for clarity.

5.6 Limitations

As with any research, limitations are inevitable, but must be acknowledged, reported and considered when results are interpreted. As the current trial was developed in response to an important research question posed by the Chartered Society Of Physiotherapy (CSP, 1997), was twelve months in the planning and enlisted the assistance of acknowledged experts in the area of LBP research, many potential limitations had been considered prior to the start of the trial.

Consequently, some of those encountered have already been accepted as a consequence of a pragmatic design and were entered into with prior knowledge.

As has been discussed previously, the difficulties associated with identifying universally accepted sub-groups of LBP patients and the requirement to preserve the pragmatic nature of the trial, resulted in a target population and subsequent sample that was not homogenous. This may, along with other issues, have contributed to the lack of differences between the two treatment interventions although both groups were reasonably well balanced at baseline that mitigates against lack of homogeneity being a major influence. Similarly, the use of an active 'control' in the current trial as a comparator rather than a true placebo or passive control would potentially have reduced the possibility of demonstrating a difference between the groups, although as already discussed, the use of a placebo or passive control would not have represented current practice. Similarly, it is acknowledged that a high standard was set for this trial by aiming to detect a 5-point change in the primary outcome measure (RMDQ). This was based on evidence from the literature and represented a stringent test, but one that was clinically meaningful and therefore considered appropriate. However, as discussed above, both treatments were active and resulted in improvements in both groups, and therefore a further difference of 5-RMDQ points between groups was unlikely to have been demonstrated. However, at the time of designing the trial, although anecdotally conventional physiotherapy treatment appeared to be effective in the management of LBP, there was little definitive evidence (Koes et al., 1991) and therefore the design that resulted seemed appropriate. It should be stressed that although no differences between the groups were demonstrated, the improvement achieved, was maintained at the 12-month follow-up, and thus appears to provide evidence for the effectiveness of physiotherapy treatment packages the management of recurrent LBP.

A limitation that must be considered is the lack of any economic analysis in addition to the other clinical outcomes, which, given the appropriate facilities and funding, would have been undertaken. However, as with any trial, resources were limited and precluded this type of analysis but raise an important point surrounding

the funding of trials. This trial was undertaken, in partial fulfilment for a higher degree, in an innovative collaboration between the NHS and academia. With a clinically based post funded by the NHS and an academic supervision team in the University and professional clinical support, this set-up provides the ideal basis for undertaking clinically based research. Problems occurred due to the lack of any designated funding as indicated above and would need to be addressed for future trials. It could be argued that poorly or non-funded trials should not be undertaken. However, this would severely limit the ability to undertake small-scale trials like this one, which have an important place in informing future, larger scale trials, and can play an important role in providing evidence about 'real-life' practice. Although no formal cost analysis was undertaken, a rough estimate of direct costs can be inferred purely on the basis of the number of treatment sessions for each group; 7.5 (2.5) for the SSSE group and 5.9 (2.3) for the CT group (See Table 4-11). These results would appear to indicate that patients in the SSSE group received on average more treatment than patients in the CT group and over a longer period of time (11.2 (3.6) and 8.0 (3.6) weeks respectively). Therefore it could be inferred that the treatment incorporating spinal stabilisation exercises, in addition to conventional physiotherapy and advice booklet, incurred more direct costs, and resulted in no additional or less, sustained improvement in functioning than conventional physiotherapy alone, in this recurrent LBP population. However, no formal cost analysis was undertaken, and therefore no firm conclusions can be drawn.

It was decided not to use return to work (RTW) as an outcome in the current trial, but this may be seen as a limitation. National (CSAG, 1994a; RCGP, 1996b; Waddell et al., 1999) and international guidelines (Koes et al., 2001a) stress the importance of early return to work and it is widely acknowledged that increasing time off work increases the risk of a poor outcome of LBP (CSAG, 1994b; Waddell, 1998). As previously discussed, multiple issues surround the use of RTW as an outcome and in practice, the majority of the patients entered into the trial were working despite their LBP. Consequently, the use of RTW may not have contributed much additional information in the context of the current trial.

Despite the target population being recurrent LBP patients, the decision was made not to use recurrence rates as a study variable for a number of reasons. As outlined in Chapter 2, defining recurrence is controversial and highly subjective. Additionally, level of function has been suggested as a primary outcome measure of importance in LBP research and therefore may be a more appropriate measure of the impact of LBP rather than purely recurrences. Another issue, which should be considered, is the issue of length of follow-up. Although 12-month follow-up is regarded as 'longer-term', problems may exist in respect to duration of symptoms for the current cohort. The average duration of symptoms for the current cohort was 8.7 months (SD=8.1, range 1-36, n=97) and 7.8 months (SD=6.6, range 1-24, n=68) for trial completers. Therefore a 12-month follow-up may not have provided a long enough timescale to fully assess the impact of further recurrences for the whole group. However, practically, longer-term follow-up was not feasible and the 12-months used should have provided a reasonably representative timescale.

5.7 Future research

Over the last few years there has been a rejection of sub-grouping of LBP patients and the provision of specific diagnoses and a move towards more generic terminology such as non-specific or simple LBP (CSAG, 1994a). However, there now appears to be some calls for the re-introduction of sub-grouping for LBP (Deyo and Phillips, 1996), and certainly for research purposes as the current emphasis provides heterogeneous populations and decreases discriminative ability of different treatment interventions. Although this idea of sub-groups has a clear contribution to make to future research and treatment of LBP, it is not without both its problems and critics. Certainly the results from the present trial and others reporting improvements in functioning, but demonstrating no differences between stabilisation training and other conventional treatment would indicate that at present we cannot necessarily identify which patients benefit from which type of training, when or why. This therefore may provide a useful direction for future research. The identification of which LBP patients may benefit from specific interventions such as spinal stabilisation training would be useful and may initially be explored through consensus opinion regarding indications for use. This could take the form of a Delphi Technique regarding current practice for the provision of specific interventions such as stabilisation training, in contrast to surveys to date that have used 'exercise' as a generic term. By identifying what subjective and objective signs indicate to clinicians that a specific treatment intervention should be successful, future trials could be designed utilizing these and may provide clinically relevant and practical sub-grouping. Additionally, deeper exploration of the characteristics of patients who were consistent responders to SSSE in the current trial may provide information towards this end.

Although the strength of this trial lies in the pragmatic design and its representation of current clinical practice, it also presents one of the limitations, i.e. the use of an active 'control' and therefore lack of a 'non-treatment' group. As both groups showed an improvement in the primary outcome, no differences between interventions were demonstrated. While impossible to predict whether differences would have existed had a passive, or non-active, control been used,

evidence from the literature would support this, as exercise has been shown to be more effective than 'usual care' (provided by a GP) for non-acute LBP (van Tulder et al., 2000b) and therefore, would possibly be one potential direction for further research. However, perhaps more appropriate, in light of the current findings of effectiveness of both treatment packages would be an examination of the potential cost implications of incorporating specific exercise regimes within the management of LBP.

5.8 Conclusion

Findings from the current trial indicate that the patients with recurrent LBP studied, without evidence of psychological distress, responded well to physiotherapy, achieving an improvement in their level of functioning. However, the addition of specific spinal stabilisation exercises, designed to improve the endurance and functioning of the abdominal and trunk muscles, did not demonstrate any additional benefit over and above that of conventional physiotherapy management.

The positive results of stabilisation training reported from trials using very selected subgroups of LBP patients e.g. acute, first episode LBP (Hides et al., 1996), spondylolisthesis (O'Sullivan et al., 1997c) and post-partum pelvic pain (Stuge et al., 2001), indicate that these techniques may add an additional benefit to conventional treatment at least in some patients. In contrast, current findings concur with the conclusions from a recent systematic review indicating that specific exercise training is not more effective in the treatment of LBP than conventional physiotherapy (van Tulder et al., 2000b). Consequently, the results from this trial indicate that physiotherapy packages, incorporating exercise-based interventions can be effective in the management of a recurrent LBP population without evidence of distress. This is an important finding as evidence for the effectiveness of physiotherapy is not firmly established and these findings appear to support its ongoing use, however, further studies incorporating placebo group allowing for the natural history of LBP would be required in order for any firm conclusions to be drawn. Additionally, as detailed records were kept of what specific elements of physiotherapy modalities were used in each treatment arm for each patient, a portrait can begin to be built of what elements of physiotherapy are most effective. Similarly, although examination of treatments in isolation allows effectiveness to be easily assessed, it does not represent clinical practice. It may be, given the findings from the current trial, that spinal stabilisation training in isolation, is unlikely to have shown any greater effectiveness than as used here. Certainly, the superiority of spinal stabilisation training above conventional physiotherapy treatment in this LBP population has not been demonstrated.

This is important, as justification for the use of specific stabilisation exercise training in recurrent LBP patients rather than highly selected sub-groups of LBP patients still has yet to be proven. Specifically, information is needed which will help identify those patients who are most likely to benefit from specific approaches within physiotherapy such as specific spinal stabilisation and others. However, clearly, the superiority of these techniques over conventional physiotherapy treatment in this patient population has not been demonstrated.

References

1. Abenhaim, L., Rossignol, M., Valat, P., Nordin, M., Avouac, B., Blotman, F., Charlot, J., Dreiser, R., Legrand, E., Rozenberg, S. and Vautravers, P. (2000) *The Role of Activity in the Therapeutic Management of Back Pain*. Spine 25(4), 1S-33S.
2. Abenhaim, L. and Suissa, S. (1987) *Importance and economic burden of occupational back pain: a study of 2,500 cases representative of Quebec*. J Occup Med 29(8), 670-4.
3. Allison, G., Kendle, K., Roll, S., Schupelius, J., Scott, Q. and Panizza, J. (1998) *The role of the diaphragm during abdominal hollowing exercises*. Australian Journal of Physiotherapy 44(2), 95-102.
4. Altman, D. (1985) *Comparability of randomised groups*. The Statistician 34, 125-136.
5. Altman, D. (1992) *Practical Statistics for Medical Research*. Chaphall & Hall, London.
6. Altman, D.G. (1996) *Better reporting of randomised controlled trials: the CONSORT statement*. BMJ 313(7057), 570-1.
7. Altman, D.G. and Bland, J.M. (1999) *Treatment allocation in controlled trials: why randomise?* BMJ 318(7192), 1209.
8. Altman, D.G. and Dore, C.J. (1990) *Randomisation and baseline comparisons in clinical trials*. Lancet 335(8682), 149-53.
9. Aruin, A.S. and Latash, M.L. (1995) *Directional specificity of postural muscles in feed-forward postural reactions during fast voluntary arm movements*. Exp Brain Res 103(2), 323-32.
10. Bartys, S., Burton, A.K., Watson, P.J., Wright, I., MacKay, C. and Main, C. (2001) *'Black Flags' and their role in the implementation of an occupational intervention for musculoskeletal disorders*. Conference Proceedings Society for Back Pain Research, Bristol, UK.
11. Battie, M. (1994) *Managing low back pain: attitudes and treatment preferences of physical therapists*. Phys Ther 74, 219-226.

12. Battie, M.C., Videman, T., Gill, K., Moneta, G.B., Nyman, R., Kaprio, J. and Koskenvuo, M. (1991) *1991 Volvo Award in clinical sciences. Smoking and lumbar intervertebral disc degeneration: an MRI study of identical twins*. Spine 16(9), 1015-21.
13. Beaton, D.E. (2000) *Understanding the Relevance of Measured Change Through Studies of Responsiveness*. Spine 25(24), 3192-3199.
14. Beattie, P. and Maher, C. (1997) *The role of functional status questionnaires for low back pain*. Australian Journal of Physiotherapy 43(1), 29-38.
15. Begg, C., Cho, M., Eastwood, S., Horton, R., Moher, D., Olkin, I., Pitkin, R., Rennie, D., Schulz, K.F., Simel, D. and Stroup, D.F. (1996) *Improving the quality of reporting of randomized controlled trials. The CONSORT statement*. JAMA 276(8), 637-9.
16. Beiring-Sorenson, F. (1984) *Physical measurements as an indication for low-back trouble over a one-year period*. Spine 9, 1.6-19.
17. Bergmark, A. (1989) *Stability of the lumbar spine. A study in mechanical engineering*. Acta Orthop Scand Suppl 230, 1-54.
18. Bergner, M., Bobbitt, R.A., Carter, W.B. and Gilson, B.S. (1981) *The Sickness Impact Profile: development and final revision of a health status measure*. Med Care 19(8), 787-805.
19. Bergquist-Ullman, M. and Larsson, U. (1977) *Acute low back pain in industry. A controlled prospective study with special reference to therapy and confounding factors*. Acta Orthop Scand 170, 1-117.
20. Beurskens, A.J., de Vet, H.C. and Koke, A.J. (1996) *Responsiveness of functional status in low back pain: a comparison of different instruments*. Pain 65(1), 71-6.

21. Biedermann, H.J., Shanks, G.L., Forrest, W.J. and Inglis, J. (1991) *Power spectrum analyses of electromyographic activity. Discriminators in the differential assessment of patients with chronic low-back pain.* Spine 16(10), 1179-84.
22. Biering-Sorensen, F. (1983) *A prospective study of low back pain in a general population. I. Occurrence, recurrence and aetiology.* Scand J Rehabil Med 15(2), 71-9.
23. Bigos, S.J., Bowyer, O. and Braen, G. (1994) *Acute Low Back Pain in Adults: Clinical Practice Guideline No.14. AHCPR Publication No. 95-0642.* Agency for Health Care Policy and Research, Public Health Service, US Department of Health and Human Service, Rockville, MD.
24. Birkett, N. (1985) *Adaptive Allocation in Randomized Controlled Trials*, Controlled Clinical Trials, pp. 146-155. Vol. 6. Elsevier Science Publishing Co, New York.
25. Bland, J.M. and Altman, D.G. (1994a) *Regression towards the mean.* BMJ 308(6942), 1499.
26. Bland, J.M. and Altman, D.G. (1994b) *Some examples of regression towards the mean.* BMJ 309(6957), 780.
27. Bland, M. (2000) *An introduction to medical statistics.* 3rd ed. Oxford University Press, Oxford.
28. Bogduk, N. (1997) *Clinical Anatomy of the Lumbar Spine and Sacrum.* Churchill Livingstone, New York.
29. Bolton, J.E. (1999) *Accuracy of recall of usual pain intensity in back pain patients.* Pain 83(3), 533-9.
30. Bolton, J.E. and Wilkinson, R.C. (1998) *Responsiveness of pain scales: a comparison of three pain intensity measures in chiropractic patients.* Journal of Manipulative & Physiological Therapeutics 21(1), 1-7.

31. Bombardier, C. (2000a) *Outcome Assessments in the Evaluation of Treatment of Spinal Disorders: Summary and General Recommendations*. Spine 25(24), 3100-3103.
32. Bombardier, C. (2000b) *Spine Focus Issue Introduction: Outcome Assessments in the Evaluation of Treatment of Spinal Disorders*. Spine 25(24), 3097-3099.
33. Bombardier, C., Hayden, J. and Beaton, D.E. (2001) *Minimal clinically important difference. Low back pain: outcome measures*. J Rheumatol 28(2), 431-8.
34. Borenstein, D., Wiesel, S.W. and Boden, S. (1995) *Low Back Pain. Mechanical Diagnosis and Comprehensive Management*. 2nd ed. WB Saunders Company, Philadelphia.
35. Borkan, J.M. and Cherkin, D.C. (1996) *An agenda for primary care research on low back pain*. Spine 21(24), 2880-4.
36. Borkan, J.M., Koes, B., Reis, S. and Cherkin, D.C. (1998) *A report from the Second International Forum for Primary Care Research on Low Back Pain. Reexamining priorities*. Spine 23(18), 1992-6.
37. Bouisset, S. and Zattara, M. (1987) *Biomechanical study of the programming of anticipatory postural adjustments associated with voluntary movement*. J Biomech 20(8), 735-42.
38. Bouter, L.M., van Tulder, M.W. and Koes, B.W. (1998) *Methodologic issues in low back pain research in primary care*. Spine 23(18), 2014-20.
39. Brumagne, S., Lysens, R. and Spaepen, A. (1999) *Lumbosacral repositioning accuracy in standing posture: a combined electrogoniometric and videographic evaluation*. Clin Biomech 14(5), 361-3.
40. Burton, A.K., Tillotson, K.M., Main, C.J. and Hollis, S. (1995) *Psychosocial predictors of outcome in acute and subchronic low back trouble*. Spine 20(6), 722-8.
41. Burton, A.K., Waddell, G., Burt, R. and Blair, S. (1996) *Patient educational material in the management of low back pain in primary care*. Bull Hosp Jt Dis 55(3), 138-41.

42. Burton, A.K., Waddell, G., Tillotson, K.M. and 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), 2484-91.
43. Burton, K., Polatin, P.B. and Gatchel, R.J. (1997) *Psychosocial factors and the rehabilitation of patients with chronic work-related upper extremity disorders.* Journal of Occupational Rehabilitation 7(3), 139-53.
44. Butler, D. (1991) *Mobilisation of the nervous system.* Churchill Livingstone, Edinburgh.
45. Cairns, M.C., Harrison, K. and Wright, C. (2000) *Pressure biofeedback: a useful tool in the quantification of abdominal muscular dysfunction?* Physiotherapy, 86(3), 127-38.
46. Caldwell and Galanville. (1993) *Cited in: Epidemiology Review: the Epidemiology and Cost of Back Pain. The Annex to the Clinical Standards Advisory Group's Report on Low Back Pain,* Her Majesty's Stationary Office.
47. Carey, T.S., Evans, A., Hadler, N., Kalsbeek, W., McLaughlin, C. and Fryer, J. (1995) *Care-seeking among individuals with chronic low back pain.* Spine 20(3), 312-7.
48. Chatman, A.B., Hyams, S.P., Neel, J.M., Binkley, J.M., Stratford, P.W., Schomberg, A. and Stabler, M. (1997) *The Patient-Specific Functional Scale: measurement properties in patients with knee dysfunction.* Phys Ther 77(8), 820-9.
49. Cherkin, D., Eisenberg, D., Kapchuk, T., Deyo, R.A., Street, J., Barlow, W. and Sherman, K. (2000) *A Randomized Trial Comparing Acupuncture, Therapeutic Massage and Self-Care Education for Chronic Low Back Pain.* Conference Proceedings The Fourth International Forum for Primary Care Research on Low Back Pain, Eilat, Israel, 55.
50. Cherkin, D.C., Deyo, R.A., Street, J.H., Hunt, M. and Barlow, W. (1996) *Pitfalls of patient education. Limited success of a program for back pain in primary care.* Spine 21(3), 345-55.

-
51. Cohen, J. (1977) *Statistical Power: Analysis for the Behavioural Sciences*. Academic Press, New York.
52. Comerford, M.J. and Mottram, S.L. (2001a) *Functional stability re-training: principles and strategies for managing mechanical dysfunction*. *Manual Therapy* 6(1), 3-14.
53. Comerford, M.J. and Mottram, S.L. (2001b) *Movement and stability dysfunction-contemporary developments*. *Manual Therapy* 6(1), 15-26.
54. Cooper, R.G., Stokes, M.J., Sweet, C., Taylor, R.J. and Jayson, M.I. (1993) *Increased central drive during fatiguing contractions of the paraspinal muscles in patients with chronic low back pain*. *Spine* 18(5), 610-6.
55. Coste, J., Paolaggi, J.B. and Spira, A. (1992) *Classification of nonspecific low back pain. I. Psychological involvement in low back pain. A clinical, descriptive approach*. *Spine* 17(9), 1028-37.
56. Cox, M.E., Asselin, S., Gracovetsky, S.A., Richards, M.P., Newman, N.M., Karakusevic, V., Zhong, L. and Fogel, J.N. (2000) *Relationship between functional evaluation measures and self-assessment in nonacute low back pain*. *Spine* 25(14), 1817-26.
57. Cresswell, A.G., Blake, P.L. and Thorstensson, A. (1994a) *The effect of an abdominal muscle training program on intra-abdominal pressure*. *Scan J Rehab Med* 26(2), 79-86.
58. Cresswell, A.G., Grundstrom, H. and Thorstensson, A. (1992) *Observations on intra-abdominal pressure and patterns of abdominal intra-muscular activity in man*. *Acta Physiol Scand* 144(4), 409-18.
59. Cresswell, A.G., Oddsson, L. and Thorstensson, A. (1994b) *The influence of sudden perturbations on trunk muscle activity and intra-abdominal pressure while standing*. *Exp Brain Res* 98(2), 336-41.

-
60. Cresswell, A.G. and Thorstensson, A. (1994) *Changes in intra-abdominal pressure, trunk muscle activation and force during isokinetic lifting and lowering*. Eur J Appl Physiol 68(4), 315-21.
61. Croft, P. and Raspe, H. (1995) *Back pain*, Baillieres Clin Rheumatol, pp. 565-83. Vol. 9.
62. Croft, P.R., Macfarlane, G.J., Papageorgiou, A.C., Thomas, E. and Silman, A.J. (1998) *Outcome of low back pain in general practice: a prospective study*. BMJ 316(7141), 1356-9.
63. Croft, P.R., Papageorgiou, A.C., Ferry, S., Thomas, E., Jayson, M.I. and Silman, A.J. (1996) *Psychologic distress and low back pain. Evidence from a prospective study in the general population*. Spine 20(24), 2731-7.
64. CSAG. (1994a) *Clinical Standards Advisory Group on Low Back Pain*, Her Majesty's Stationary Office.
65. CSAG. (1994b) *Epidemiology Review: the Epidemiology and Cost of Back Pain. The Annex to the Clinical Standards Advisory Group's Report on Low Back Pain*, Her Majesty's Stationary Office.
66. CSP. (1991) *Standards for the use of electrophysical modalities*. The Chartered Society of Physiotherapy.
67. CSP. (1996) *Rules of Professional Conduct*. Chartered Society of Physiotherapy.
68. CSP. (1997) *Priorities for Physiotherapy Research*. Research and Clinical Effectiveness Unit, The Chartered Society of Physiotherapy.
69. Cyriax, J. (1984) *Textbook of Orthopaedic Medicine*. 11th ed. Bailliere Tindall, London.
70. Daltroy, L.H., Cats-Baril, W.L., Katz, J.N., Fossel, A.H. and Liang, M.H. (1996) *The North American spine society lumbar spine outcome assessment instrument: reliability and validity tests*. Spine 21(6), 741-9.

71. Danneels, L.A., Vanderstraeten, G.G., Cambier, D.C., Witvrouw, E.E., Bourgois, J., Dankaerts, W. and De Cuyper, H.J. (2001) *Effects of three different training modalities on the cross sectional area of the lumbar multifidus muscle in patients with chronic low back pain*. Br J Sports Med 35(3), 186-91.
72. Davis, C.E. (1997) *Secondary endpoints can be validly analyzed, even if the primary endpoint does not provide clear statistical significance*. Control Clin Trials 18(6), 557-60.
73. De Troyer, A., Estenne, M., Ninane, V., Van Gansbeke, D. and Gorini, M. (1990) *Transversus abdominis muscle function in humans*. J Appl Physiol 68(3), 1010-6.
74. Delitto, A. (1994) *Are measures of function and disability important in low back care?* Physical Therapy 74(5), 452-62.
75. Delitto, A., Erhard, R.E. and Bowling, R.W. (1995) *A treatment-based classification approach to low back syndrome: identifying and staging patients for conservative treatment*. Physical Therapy 75(6), 470-89.
76. Deschamps, M., Band, P.R. and Coldman, A.J. (1988) *Assessment of adult cancer pain: shortcomings of current methods*. Pain 32(2), 133-9.
77. Deyo, R.A. (1982) *Compliance with therapeutic regimens in arthritis: issues, current status, and a future agenda*. Semin Arthritis Rheum 12(2), 233-44.
78. Deyo, R.A. (1988) *Measuring the functional status of patients with low back pain*. Arch Phys Med Rehabil 69(12), 1044-53.
79. Deyo, R.A., Andersson, G., Bombardier, C., Cherkin, D.C., Keller, R.B., Lee, C.K., Liang, M.H., Lipscomb, B., Shekelle, P., Spratt, K.F. and et al. (1994) *Outcome measures for studying patients with low back pain*. Spine 19(18 Suppl), 2032S-2036S.
80. Deyo, R.A., Battie, M., Beurskens, A.J., Bombardier, C., Croft, P., Koes, B., Malmivaara, A., Roland, M., Von Korff, M. and Waddell, G. (1998) *Outcome*

- measures for low back pain research. A proposal for standardized use.* Spine 23(18), 2003-13.
81. Deyo, R.A. and Centor, R.M. (1986) *Assessing the responsiveness of functional scales to clinical change: an analogy to diagnostic test performance.* J Chronic Dis 39(11), 897-906.
82. Deyo, R.A. and Phillips, W.R. (1996) *Low back pain. A primary care challenge.* Spine 21(24), 2826-32.
83. Deyo, R.A. and Tsui-Wu, Y.J. (1987) *Descriptive epidemiology of low-back pain and its related medical care in the United States.* Spine 12(3), 264-8.
84. Deyo, R.A. and Weinstein, J.N. (2001) *Low back pain.* N Engl J Med 344(5), 363-70.
85. Dionne, C.E., Von Korff, M., Koepsell, T.D., Deyo, R.A., Barlow, W.E. and Checkoway, H. (1999) *A comparison of pain, functional limitations, and work status indices as outcome measures in back pain research.* Spine 24(22), 2339-45.
86. Dolan, P., Greenfield, K., Nelson, R.J. and Nelson, I.W. (2000) *Can exercise therapy improve the outcome of microdiscectomy?* Spine 25(12), 1523-32.
87. Dolan, P., Mannion, A.F. and Adams, M.A. (1995) *Fatigue of the erector spinae muscles. A quantitative assessment using "frequency banding" of the surface electromyography signal.* Spine 20(2), 149-59.
88. Donahue, M.S., Riddle, D.L. and Sullivan, M.S. (1996) *Intertester reliability of a modified version of McKenzie's lateral shift assessments obtained on patients with low back pain.* Physical Therapy 76(7), 706-16; discussion 717-26.
89. Dudgeon, D., Raubertas, R.F. and Rosenthal, S.N. (1993) *The short-form McGill Pain Questionnaire in chronic cancer pain.* Journal of Pain & Symptom Management 8(4), 191-5.
90. Ekblom, A. and Hansson, P. (1988) *Pain intensity measurements in patients with acute pain receiving afferent stimulation.* J Neurol Neurosurg Psychiatry 51(4), 481-6.

91. Erhard, R.E., Delitto, A. and Cibulka, M.T. (1994) *Relative effectiveness of an extension program and a combined program of manipulation and flexion and extension exercises in patients with acute low back syndrome*. Physical Therapy 74(12), 1093-100.
92. EuroQol-Group. (1990) *EuroQol. A new facility for the measurement of health-related quality-of-life*, Health Policy, pp. 199-208. Vol. 16.
93. Evans, G. and Richards, S. (1996) *Low Back Pain: An Evaluation of Therapeutic Interventions*. Health Care Evaluation Unit, University of Bristol.
94. Faas, A. (1996) *Exercises: which ones are worth trying, for which patients, and when?* Spine 21(24), 2874-8; discussion 2878-9.
95. Fairbank, J. (2000) *The reliability of the shuttle walking test, Swiss spinal stenosis score, Oxford spinal stenosis score and Oswestry disability index in the assessment of patients with lumbar spine stenosis*, Society of Back Pain Research, Manchester, UK.
96. Fairbank, J.C., Couper, J., Davies, J.B. and O'Brien, J.P. (1980) *The Oswestry low back pain disability questionnaire*. Physiotherapy 66(8), 271-3.
97. Fanuele, J.C., Birkmeyer, N.J., Abdu, W.A., Tosteson, T.D. and Weinstein, J.N. (2000) *The impact of spinal problems on the health status of patients: have we underestimated the effect?* Spine 25(12), 1509-14.
98. Fisher, K. and Johnston, M. (1997) *Validation of the Oswestry Low Back Pain Disability Questionnaire, its sensitivity as a measure of change following treatment and its relationship with other aspects of the chronic pain experience*. Physiotherapy Theory & Practice 13(1), 67-80.
99. Floman, Y. (2000) *Progression of lumbosacral isthmic spondylolisthesis in adults*. Spine 25(3), 342-7.
100. Flynn, C., Lew, P. and Story, I. (1997) *A comparison of the effectiveness of two conservative treatment regimes for low back pain patients who demonstrate a pattern of instability*.

- Conference Proceedings 10th Biennial Conference of the Manipulative Physiotherapists Association of Australia, Melbourne, Australia, 53-55.
101. Foster, N.E. (1998) *Current clinical management of low back pain and hypoalgesic effects of transcutaneous electrical nerve stimulation*. DPhil Thesis, University of Ulster, Jordonstown, N. Ireland.
102. Foster, N.E., Thompson, K.A., Baxter, G.D. and Allen, J.M. (1999) *Management of nonspecific low back pain by physiotherapists in Britain and Ireland. A descriptive questionnaire of current clinical practice*. Spine 24(13), 1332-42.
103. Freedman, K.B., Back, S. and Bernstein, J. (2001) *Sample size and statistical power of randomised, controlled trials in orthopaedics*. J Bone Joint Surg Br 83(3), 397-402.
104. Friedli, W.G., Cohen, L., Hallett, M., Stanhope, S. and Simon, S.R. (1988) *Postural adjustments associated with rapid voluntary arm movements. II. Biomechanical analysis*. J Neurol Neurosurg Psychiatry 51(2), 232-43.
105. Friedman, L., Furberg, C. and DeMets, D. (1998) *Fundamentals of Clinical Trials*. Springer-Verlag, New York.
106. Friedrich, M., Gittler, G., Halberstadt, Y., Cermak, T. and Heiller, I. (1998) *Combined exercise and motivation program: effect on the compliance and level of disability of patients with chronic low back pain: a randomized controlled trial*. Archives of Physical Medicine & Rehabilitation 79(5), 475-87.
107. Frost, H., Klaber Moffett, J.A., Moser, J.S. and Fairbank, J.C. (1995) *Randomised controlled trial for evaluation of fitness programme for patients with chronic low back pain*. BMJ 310(6973), 151-4.
108. Frymoyer, J.W. and Pope, M.H. (1991) *Segmental Instability*. Semin Spine Surgery 3(109-118).

-
109. Gardner-Morse, M.G. and Stokes, I.A. (1998) *The effects of abdominal muscle coactivation on lumbar spine stability*. Spine 23(1), 86-91.
110. Garratt, A.M., Klaber Moffett, J. and Farrin, A.J. (2001) *Responsiveness of Generic and Specific Measures of Health Outcome in Low Back Pain*. Spine 26(1), 71-77.
111. Gibaldi, M. and Sullivan, S. (1997) *Intention-to-treat analysis in randomized trials: who gets counted?* Issues Clin Pharmacol 37, 667-672.
112. Gilbody, S.M., House, A.O. and Sheldon, T.A. (2001) *Routinely administered questionnaires for depression and anxiety: systematic review*. BMJ 322(7283), 406-9.
113. Gill, K.P. and Callaghan, M.J. (1998) *The measurement of lumbar proprioception in individuals with and without low back pain*. Spine 23(3), 371-7.
114. Glass, K., Allison, G. and Edmondston, S. (1997) *Physiotherapist's ability to differentiate anterolateral abdominal control strategies*, 10th Biennial Conference of the Manipulative Physiotherapists Association of Australia, pp. 58-59, Melbourne, Australia.
115. Goel, V.K., Kong, W., Han, J.S., Weinstein, J.N. and Gilbertson, L.G. (1993) *A combined finite element and optimization investigation of lumbar spine mechanics with and without muscles*. Spine 18(11), 1531-41.
116. Goldberg, M.S., Scott, S.C. and Mayo, N.E. (2000) *A review of the association between cigarette smoking and the development of nonspecific back pain and related outcomes*. Spine 25(8), 995-1014.
117. Goldby, L., Moore, A., Doust, J. and Trew, M. (2000) *An RCT investigating the efficacy of manual therapy, exercises to rehabilitate spinal stabilisation and an educational booklet in the treatment of chronic low back pain*. Conference Proceedings International Federation of Orthopaedic Manipulative Therapists, Perth, Australia, 47.

118. Gracey, J.H., McDonough, S.M. and Baxter, G.D. (2002) *Physiotherapy management of low back pain: a survey of current practice in northern ireland*. Spine 27(4), 406-11.
119. Grafton, K., Foster, N. and Wright, C. (2001) *An Evaluation of the test-retest reliability of the Short-Form McGill Pain Questionnaire*. Conference Proceedings Chartered Society of Physiotherapy Annual Conference, Birmingham, UK.
120. Gray, H. (1980) *Gray's Anatomy*. 36th ed, edited by P. Williams and R. Warwick.
121. Greenfield, K. (2001) *Pain, disability and psychological status related to lumbar disc pathology*. Conference Proceedings Chartered Society of Physiotherapy, Birmingham, UK.
122. Greenough, C.G. and Fraser, R.D. (1992) *Assessment of outcome in patients with low-back pain*. Spine 17(1), 36-41.
123. Grevitt, M., Pande, K., O'Dowd, J. and Webb, J. (1998) *Do first impressions count? A comparison of subjective and psychologic assessment of spinal patients*. Euro Spine J 7(3), 218-23.
124. Grieve, G.P. (1982) *Lumbar instability*. Physiotherapy 68, 2-9.
125. Gronblad, M., Jarvinen, E., Airaksinen, O., Ruuskanen, M., Hamalainen, H. and Kouri, J.P. (1996) *Relationship of subjective disability with pain intensity, pain duration, pain location, and work-related factors in nonoperated patients with chronic low back pain*. Clinical Journal of Pain 12(3), 194-200.
126. Grushka, M. and Sessle, B.J. (1984) *Applicability of the McGill Pain Questionnaire to the differentiation of 'toothache' pain*. Pain 19(1), 49-57.
127. Guyatt, G.H., Feeny, D.H. and Patrick, D.L. (1993) *Measuring health-related quality of life*. Ann Intern Med 118(8), 622-9.

128. Hagins, M., Adler, K., Cash, M., Daugherty, J. and Mitrani, G. (1999) *Effects of practice on the ability to perform lumbar stabilization exercises*. Journal of Orthopaedic and Sports Physical Therapy 29(9), 546-555.
129. Heft, M.W. and Parker, S.R. (1984) *An experimental basis for revising the graphic rating scale for pain*. Pain 19(2), 153-61.
130. Herr, K.A. and Mobily, P.R. (1993) *Comparison of selected pain assessment tools for use with the elderly*. Appl Nurs Res 6(1), 39-46.
131. Hides, J., Richardson, B., Jull, G. and Davies, S. (1995a) *Ultrasound imaging in rehabilitation*. Aust J Phys Ther 41(3), 187-193.
132. Hides, J., Richardson, C. and Jull, G. (2000) *Long-term effects of specific stabilizing exercises for first episode low back pain*. Conference Proceedings International Federation of Orthopaedic Manipulative Therapists, Perth, Australia, 47.
133. Hides, J.A., Jull, G.A. and Richardson, C.A. (2001) *Long-term effects of specific stabilizing exercises for first-episode low back pain*. Spine 26(11), E243-8.
134. Hides, J.A., Richardson, C.A. and Jull, G.A. (1995b) *Magnetic resonance imaging and ultrasonography of the lumbar multifidus muscle. Comparison of two different modalities*. Spine 20(1), 54-8.
135. Hides, J.A., Richardson, C.A. and Jull, G.A. (1996) *Multifidus muscle recovery is not automatic after resolution of acute, first-episode low back pain*. Spine 21(23), 2763-9.
136. Hides, J.A., Richardson, C.A. and Jull, G.A. (1998) *Masterclass. Use of real-time ultrasound imaging for feedback in rehabilitation*. Manual Therapy 3(3), 125-31.
137. Hides, J.A., Stokes, M.J., Saide, M., Jull, G.A. and Cooper, D.H. (1994) *Evidence of lumbar multifidus muscle wasting ipsilateral to symptoms in patients with acute/subacute low back pain*. Spine 19(2), 165-72.

-
138. Hobby, J.L., Lutchman, L.N., Powell, J.M. and Sharp, D.J. (2001) *The distress and risk assessment method (DRAM)*. J Bone Joint Surg Br 83(1), 19-21.
139. Hodges, P. (2001) *Experimentally induced low back pain causes changes in motor control of the trunk muscles*. Conference Proceedings 4th Interdisciplinary World Congress on Low Back and Pelvic Pain, Montreal, Canada, 184-185.
140. Hodges, P., Cresswell, A. and Thorstensson, A. (1999) *Preparatory trunk motion accompanies rapid upper limb movement*. Exp Brain Res 124(1), 69-79.
141. Hodges, P., Richardson, C. and Jull, G. (1996) *Evaluation of the relationship between laboratory and clinical tests of transversus abdominis function*. Physiother Res Int 1(1), 30-40.
142. Hodges, P.W. (1999) *Is there a role for transversus abdominis in lumbo-pelvic stability?* Manual Therapy 4(2), 74-86.
143. Hodges, P.W., Butler, J.E., McKenzie, D.K. and Gandevia, S.C. (1997) *Contraction of the human diaphragm during rapid postural adjustments*. J Physiol (Lond) 505(Pt 2), 539-48.
144. Hodges, P.W., Cresswell, A.G., Daggfeldt, K. and Thorstensson, A. (2000) *Three dimensional preparatory trunk motion precedes asymmetrical upper limb movement*. Gait & Posture 11(2), 92-101.
145. Hodges, P.W. and 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), 2640-50.
146. Hodges, P.W. and Richardson, C.A. (1997a) *Contraction of the abdominal muscles associated with movement of the lower limb*. Physical Therapy 77(2), 132-44.

147. Hodges, P.W. and Richardson, C.A. (1997b) *Feedforward contraction of transversus abdominis is not influenced by the direction of arm movement*. Exp Brain Res 114(2), 362-70.
148. Hodges, P.W. and Richardson, C.A. (1997c) *Relationship between limb movement speed and associated contraction of the trunk muscles*. Ergonomics 40(11), 1220-30.
149. Hodges, P.W. and Richardson, C.A. (1998) *Delayed postural contraction of transversus abdominis in low back pain associated with movement of the lower limb*. J Spinal Disord 11(1), 46-56.
150. Hodges, P.W. and Richardson, C.A. (1999) *Transversus abdominis and the superficial abdominal muscles are controlled independently in a postural task*. Neurosci Lett 265(2), 91-4.
151. Hollis, S. and Campbell, F. (1999) *What is meant by intention to treat analysis? Survey of published randomised controlled trials*. BMJ 319(7211), 670-4.
152. Hoogendoorn, W.E., van Poppel, M.N., Bongers, P.M., Koes, B.W. and Bouter, L.M. (2000) *Systematic review of psychosocial factors at work and private life as risk factors for back pain*. Spine 25(16), 2114-25.
153. Hope, P. and Forshaw, M. (1999) *Assessment of psychological distress is important in patients presenting with low back pain*. Physiotherapy 85(12), 563-570.
154. Hothersall, A., Knight, M. and Goswami, A. (2001) *The Use of the Distress Risk Assessment Method in Patients with Chronic Back Pain. Its failure to predict the response to endoscopic spinal surgery (ELF)*. Conference Proceedings Society for Back Pain Research, Bristol, UK.
155. Hunt, S.M., McEwen, J. and McKenna, S.P. (1985) *Measuring health status: a new tool for clinicians and epidemiologists*. J R Coll Gen Pract 35(273), 185-8.

-
156. Jadad, A. (2000) *Randomised Controlled Trials. A User's Guide*. BMJ Publications, London.
157. Jaeschke, R., Singer, J. and Guyatt, G.H. (1989) *Measurement of health status. Ascertaining the minimal clinically important difference*. *Control Clin Trials* 10(4), 407-15.
158. Jakobs, T., Miller, J.A. and Schultz, A.B. (1985) *Trunk position sense in the frontal plane*. *Exp Neurol* 90(1), 129-38.
159. JAMA. (1994) *A proposal for structured reporting of randomized controlled trials. The Standards of Reporting Trials Group*. *JAMA* 272(24), 1926-31.
160. Janda, V. (1977) *Muscles, central nervous motor regulation and back problems*. In: Korr, IV (Ed) *The Neurobiological Mechanisms in Manipulative Therapy*, Plenum Press, New York.
161. Jenkinson, C., Carroll, D., Egerton, M., Frankland, T., McQuay, H. and Nagle, C. (1995) *Comparison of the sensitivity to change of long and short form pain measures*. *Qual Life Res* 4(4), 353-7.
162. Jenkinson, C., Coulter, A. and Wright, L. (1993) *Short form 36 (SF 36) health survey questionnaire: normative data for adults of working age*. *BMJ* 306, 1437-40.
163. Jenkinson, C., Layte, R., Wright, L. and Coulter, A. (1996) *The UK SF-36: An analysis and interpretation manual*, Oxford.
164. Jensen, M.P., Karoly, P. and Braver, S. (1986) *The measurement of clinical pain intensity: a comparison of six methods*. *Pain* 27(1), 117-26.
165. Jensen, M.P., Turner, J.A. and Romano, J.M. (1994) *What is the maximum number of levels needed in pain intensity measurement?* *Pain* 58(3), 387-92.
166. Jette, A.M., Smith, K., Haley, S.M. and Davis, K.D. (1994) *Physical therapy episodes of care for patients with low back pain*. *Physical Therapy* 74(2), 101-10;.

167. Jull, G. and Richardson, C. (1994) *Rehabilitation of Active Stabilization of the Lumbar Spine. In: Physical Therapy of the Lumbar Spine.*, edited by L.T. Twomey, JR. Churchill Livingstone, New York.
168. Jull, G., Richardson, C., Toppenberg, R., Comerford, M. and Bui, B. (1993) *Towards a measurement of active muscle control for lumbar stabilisation.* Australian Journal of Physiotherapy 39(3), 187-93.
169. Jull, G.A. and Richardson, C.A. (2000) *Motor control problems in patients with spinal pain: a new direction for therapeutic exercise.* Journal of Manipulative & Physiological Therapeutics 23(2), 115-7.
170. Kaltenborn, F. (1970) *Mobilisation of the spinal column.* University Press, New Zealand.
171. Kaser, L., Mannion, A.F., Rhyner, A., Weber, E., Dvorak, J. and Muntener, M. (2001) *Active therapy for chronic low back pain: part 2. Effects on paraspinal muscle cross-sectional area, fiber type size, and distribution.* Spine 26(8), 909-19.
172. Kendall, N., Linton, S. and Main, C.J. (1997) *Guide to assessing psychosocial yellow flags in acute low back pain: risk factors for long-term disability and work loss.*, Accident Rehabilitation and Compensation Insurance Corporation of New Zealand and the National Health Committee.
173. Kendall, N.A. (1999) *Psychosocial approaches to the prevention of chronic pain: the low back paradigm.* Baillieres Best Pract Res Clin Rheumatol 13(3), 545-54.
174. Klaber Moffett, J., Richardson, G., Sheldon, T. and Maynard, A. (1995) *Back Pain. Its Management and Cost to Society.* Discussion Paper.
175. Koes, B.W., Bouter, L.M., Beckerman, H., van der Heijden, G.J. and Knipschild, P.G. (1991) *Physiotherapy exercises and back pain: a blinded review.* BMJ 302(6792), 1572-6.

-
176. Koes, B.W., Bouter, L.M. and van der Heijden, G.J. (1995) *Methodological quality of randomized clinical trials on treatment efficacy in low back pain*. Spine 20(2), 228-35.
177. Koes, B.W., van Tulder, M.W., Ostelo, R., Kim Burton, A. and Waddell, G. (2001a) *Clinical guidelines for the management of low back pain in primary care: an international comparison*. Spine 26(22), 2504-13.
178. Koes, B.W., van Tulder, M.W., Ostelo, R., Kim Burton, A. and Waddell, G. (2001b) *Clinical guidelines for the management of low back pain in primary care: an international comparison*. Spine 26(22), 2504-13; discussion 2513-4.
179. Kopec, J.A. and Esdaile, J.M. (1995) *Functional disability scales for back pain*. Spine 20(17), 1943-9.
180. Kravitz, R.L., Hays, R.D., Sherbourne, C.D., DiMatteo, M.R., Rogers, W.H., Ordway, L. and Greenfield, S. (1993) *Recall of recommendations and adherence to advice among patients with chronic medical conditions*. Arch Intern Med 153(16), 1869-78.
181. Leboeuf-Yde, C. (1999) *Smoking and low back pain. A systematic literature review of 41 journal articles reporting 47 epidemiologic studies*. Spine 24(14), 1463-70.
182. Leboeuf-Yde, C. and Kyvik, K.O. (1998) *At what age does low back pain become a common problem? A study of 29,424 individuals aged 12-41 years*. Spine 23(2), 228-34.
183. Lewis, J.A. and Machin, D. (1993) *Intention to treat—who should use ITT?* Br J Cancer 68(4), 647-50.
184. Lexell, J. and Downham, D. (1992) *What is the effect of ageing on type 2 muscle fibres?* J Neurol Sci 107(2), 250-1.
185. Li, L.C. and Bombardier, C. (2001) *Physical therapy management of low back pain: an exploratory survey of therapist approaches*. Phys Ther 81(4), 1018-28.

186. Linton, S.J. (2000) *A review of psychological risk factors in back and neck pain*. Spine 25(9), 1148-56.
187. Linton, S.J. and Hallden, K. (1998) *Can we screen for problematic back pain? A screening questionnaire for predicting outcome in acute and subacute back pain*. Clin J Pain 14(3), 209-15.
188. Linton, S.J. and van Tulder, M.W. (2001) *Preventive interventions for back and neck pain problems: what is the evidence?* Spine 26(7), 778-87.
189. Little, P., Somerville, J., Williamson, I., Warner, G., Moore, M., Wiles, R., George, S., Smith, A. and Peveler, R. (2001) *Randomised controlled trial of self management leaflets and booklets for minor illness provided by post*. BMJ 322(7296), 1214-6, 1217.
190. Loney, P.L. and Stratford, P.W. (1999) *The prevalence of low back pain in adults: a methodological review of the literature*. Phys Ther 79(4), 384-96.
191. Lonn, J. (2000) *Active Back School, Prophylactic Management for Low Back Pain. Three year follow-up of a randomized controlled trial*. Conference Proceedings The Fourth International Forum for Primary Care Research on Low Back Pain, Eilat, Israel, page 52.
192. Maffey-Ward, L., Jull, G. and Wellington, L. (1996) *Toward a clinical test of lumbar spine kinesthesia*. J Orthop Sports Phys Ther 24(6), 354-8.
193. Maher, C., Latimer, J. and Refshauge, K. (1999) *Prescription of activity for low back pain: what works?* Australian Journal of Physiotherapy 45(2), 121-32.
194. Main, C. and Burton, A.K. (1998) *Pain Mechanisms*. In: R. McGraig and M. Harrinton (Eds), *The changing face of occupational health*. Health & Safety Executive, London.
195. Main, C.J. (1983) *The Modified Somatic Perception Questionnaire (MSPQ)*. J Psychosom Res 27(6), 503-14.

196. Main, C.J. and Watson, P.J. (1995) *Screening for patients at risk of developing chronic incapacity*. Journal of Occupational Rehabilitation 5(4), 207-17.
197. Main, C.J., Wood, P.L., Hollis, S., Spanswick, C.C. and Waddell, G. (1992) *The Distress and Risk Assessment Method. A simple patient classification to identify distress and evaluate the risk of poor outcome*. Spine 17(1), 42-52.
198. Maitland, G. (1986) *Vertebral Manipulation*. 5th ed. Butterworth Heinemann, Oxford.
199. Maniadakis, N. and Gray, A. (2000) *The economic burden of back pain in the UK*. Pain 84(1), 95-103.
200. Mannion, A.F. and Dolan, P. (1994) *Electromyographic median frequency changes during isometric contraction of the back extensors to fatigue*. Spine 19(11), 1223-9.
201. Mannion, A.F., Dolan, P. and Adams, M.A. (1996) *Psychological questionnaires: do "abnormal" scores precede or follow first-time low back pain?* Spine 21(22), 2603-11.
202. Mannion, A.F., Dumas, G.A., Cooper, R.G., Espinosa, F.J., Faris, M.W. and Stevenson, J.M. (1997) *Muscle fibre size and type distribution in thoracic and lumbar regions of erector spinae in healthy subjects without low back pain: normal values and sex differences*. J Anat 190(Pt 4), 505-13.
203. Mannion, A.F., Junge, A., Taimela, S., Muntener, M., Lorenzo, K. and Dvorak, J. (2001a) *Active therapy for chronic low back pain: part 3. Factors influencing self-rated disability and its change following therapy*. Spine 26(8), 920-9.
204. Mannion, A.F., Kaser, L., Weber, E., Rhyner, A., Dvorak, J. and Muntener, M. (2000) *Influence of age and duration of symptoms on fibre type distribution and size of the back muscles in chronic low back pain patients*. Eur Spine J 9(4), 273-81.
205. Mannion, A.F., Muntener, M., Taimela, S. and Dvorak, J. (1999) *A randomized clinical trial of three active therapies for chronic low back pain*. Spine 24(23), 2435-48.

206. Mannion, A.F., Taimela, S., Muntener, M. and Dvorak, J. (2001b) *Active therapy for chronic low back pain part 1. Effects on back muscle activation, fatigability, and strength.* Spine 26(8), 897-908.
207. Matthews, J.N., Altman, D.G., Campbell, M.J. and Royston, P. (1990) *Analysis of serial measurements in medical research.* BMJ 300(6719), 230-5.
208. Mawson, S. (1995) *What is the SF-36 and Can It Measure the Outcome of Physiotherapy?* Physiotherapy 81(4), 208-212.
209. McClure, P.W. (1998) *Research for the clinician. Evaluating research when "no significant differences were found": the issue of statistical power.* Journal of Hand Therapy 11(3), 212-3.
210. McComas, A. (1996) *Skeletal Muscle. Form & Function.* Human Kinetics, USA.
211. McGill, S.M. (1991) *Kinetic potential of the lumbar trunk musculature about three orthogonal orthopaedic axes in extreme postures.* Spine 16(7), 809-15.
212. McGill, S.M., Childs, A. and Liebenson, C. (1999) *Endurance times for low back stabilization exercises: clinical targets for testing and training from a normal database.* Archives of Physical Medicine & Rehabilitation 80(8), 914-4.
213. McGlashen, K., Ashton-Miller, J.A., Green, M. and Schultz, A.B. (1991) *Trunk positioning accuracy in the frontal and sagittal planes.* J Orthop Res 9(4), 576-83.
214. McKenzie, R. (1981) *The lumbar spine. Mechanical diagnosis and therapy.* Spinal Publications, Waikanae, New Zealand.
215. Melzack, R. (1987) *The short-form McGill Pain Questionnaire.* Pain 30(2), 191-7.
216. Melzack, R. and Katz, J. (1994) *Pain measurement in persons in pain.* In: P. Wall and R. Melzack (Eds), Textbook of Pain, 3rd ed.

217. Miller, M.I. and Medeiros, J.M. (1987) *Recruitment of internal oblique and transversus abdominis muscles during the eccentric phase of the curl-up exercise*. Phys Ther 67(8), 1213-7.
218. Moffett, J.K., Torgerson, D., Bell-Syer, S., Jackson, D., Llewlyn-Phillips, H., Farrin, A. and Barber, J. (1999) *Randomised controlled trial of exercise for low back pain: clinical outcomes, costs, and preferences*. BMJ 319(7205), 279-83.
219. Moffroid, M.T., Haugh, L.D., Henry, S.M. and Short, B. (1994) *Distinguishable groups of musculoskeletal low back pain patients and asymptomatic control subjects based on physical measures of the NIOSH Low Back Atlas [published erratum appears in Spine 1994 Sep 15;19(18):2137]*. Spine 19(12), 1350-8.
220. US Dept of Health and Human Services, Public Health Services, Centre for Disease Control, National Institute for Occupational Safety and Health. NIOSH Final Report RFP 200-89-2917. (1992) *Sensitivity and Specificity of the NIOSH Low Back Pain Atlas*. Moffroid, M.T., Haugh, L.D. and Hodous, T. May.
221. Moore, A. (1997/8) *An audit of the outcome of physiotherapy intervention for outpatients with back pain against set clinical standards*, South Thames Clinical Audit Programme.
222. Moore, R.J., Vernon-Roberts, B., Fraser, R.D., Osti, O.L. and Schembri, M. (1996) *The origin and fate of herniated lumbar intervertebral disc tissue*. Spine 21(18), 2149-55.
223. Mottram, S. and Comerford, M. (1998) *Stability Dysfunction and Low Back Pain*. J Orthopaedic Medicine 20(2), 13-18.
224. Mulligan, B. (1989) *Manual Therapy 'NAGS', 'SNAGS', 'PRPS' etc*. Plan View Services, Wellington.
225. Ng, G.Y.F. and Richardson, C.A. (1990) *The effects of training triceps surae using progressive speed loading*. Physiotherapy Theory & Practice 6(2), 77-84.

226. Nicolaisen, T. and Jorgensen, K. (1985) *Trunk strength, back muscle endurance and low-back trouble*. Scand J Rehabil Med 17(3), 121-7.
227. Norman, J. and Streiner, D. (1998) *Biostatistics. The Bare Essentials*. Mosby.
228. Norris, C.M. (1995) *Spinal stabilisation: 5. An exercise programme to enhance lumbar stabilisation*. Physiotherapy 81(3), 138-46.
229. O'Sullivan, P., Twomey, L. and Allison, G. (1995) *Evaluation of Specific Stabilising Exercises in the Treatment of Chronic Spondylosis or Spondylolithesis*. Conference Proceedings Manipulative Physiotherapy Association of Australia 9th Biennial Conference, Gold Coast, Australia, 113-114.
230. OCPS. (1997) *Census and Labour Force Survey*, Office of Population and Census and Surveys. HMSO, London.
231. Ohnmeiss, D.D., Vanharanta, H., Estlander, A.M. and Jamsen, A. (2000) *The relationship of disability (Oswestry) and pain drawings to functional testing*. Eur Spine J 9(3), 208-12.
232. Oldridge, N.B. (1982) *Compliance and exercise in primary and secondary prevention of coronary heart disease: a review*. Preventive Medicine 11(1), 56-70.
233. O'Neill, R.T. (1997) *Secondary endpoints cannot be validly analyzed if the primary endpoint does not demonstrate clear statistical significance*. Control Clin Trials 18(6), 550-6.
234. O'Sullivan, P., Twomey, L. and Allison, G. (1997a) *Dysfunction of the Neuro-Muscular System in the presence of Low Back Pain-Implications for Physical Therapy Management*. J Man & Manip Ther 5(1), 20-26.
235. O'Sullivan, P., Twomey, L., Allison, G., Sinclair, J., Miller, K. and Knox, J. (1997b) *Altered patterns of abdominal muscle activation in patients with chronic low back pain*. Australian Journal of Physiotherapy 43(2), 91-8.

236. O'Sullivan, P.B. (2000) *Lumbar segmental 'instability': clinical presentation and specific stabilizing exercise management*. *Manual Therapy* 5(1), 2-12.
237. O'Sullivan, P.B., Twomey, I. and Allison, G.T. (1998) *Altered abdominal muscle recruitment in patients with chronic back pain following a specific exercise intervention*. *J Orthop Sports Phys Ther* 27(2), 114-24.
238. O'Sullivan, P.B., Twomey, L.T. and Allison, G.T. (1997c) *Evaluation of specific stabilizing exercise in the treatment of chronic low back pain with radiologic diagnosis of spondylolysis or spondylolisthesis*. *Spine* 22(24), 2959-67.
239. Padfield, B., Butler, R. and Chesworth, B. (2000) *Validity of the Quebec Task Force classification for spinal disorders*. In: K. Singer (Ed), *International Federation of Orthopaedic Manipulative Therapists*, Perth, Australia, pp. 63.
240. Panjabi, M. (1995) *Clinical Spinal Instability: Past Achievements and Recent Advances*. Conference Proceedings Manipulative Physiotherapy Association of Australia 9th Biennial Conference, Gold Coast, Queensland, Australia, 119-122.
241. Panjabi, M., Abumi, K., Duranceau, J. and Oxland, T. (1989) *Spinal stability and intersegmental muscle forces. A biomechanical model*. *Spine* 14(2), 194-200.
242. Panjabi, M.M. (1992a) *The stabilizing system of the spine. Part I. Function, dysfunction, adaptation, and enhancement*. *J Spinal Disord* 5(4), 383-9.
243. Panjabi, M.M. (1992b) *The stabilizing system of the spine. Part II. Neutral zone and instability hypothesis*. *J Spinal Disord* 5(4), 390-6.
244. Papageorgiou, A.C., Croft, P.R., Thomas, E., Silman, A.J. and Macfarlane, G.J. (1998) *Psychosocial risks for low back pain: are these related to work?* *Ann Rheum Dis* 57(8), 500-2.
245. Parnianpour, M., Nordin, M., Kahanovitz, N. and Frankel, V. (1988) *1988 Volvo award in biomechanics. The triaxial coupling of torque generation of trunk muscles during*

- isometric exertions and the effect of fatiguing isoinertial movements on the motor output and movement patterns.* Spine 13(9), 982-92.
246. Patrick, D.L., Deyo, R.A., Atlas, S.J., Singer, D.E., Chapin, A. and Keller, R.B. (1995) *Assessing health-related quality of life in patients with sciatica.* Spine 20(17), 1899-908.
247. Petersen, T., Thorsen, H., Manniche, C. and Ekdahl. (1999) *Classification of non-specific low back pain: a review of literature on classification systems relevant to physiotherapy.* Phys Ther Rev 4, 265-281.
248. Pincus, T., Burton, A.K., Vogel, S. and Field, A.P. (2002) *A Systematic Review of Psychological Factors as Predictors of Chronicity/ Disability in Prospective Cohorts of Low Back Pain.* Spine 27(5), E109-E120.
249. Pocock, S. (1983) *Clinical Trials.* John Wiley & Sons, Chichester.
250. Pocock, S.J. (1997) *Clinical trials with multiple outcomes: a statistical perspective on their design, analysis, and interpretation.* Control Clin Trials 18(6), 530-45.
251. Prentice, R. (1997) *On the Role and Analysis of Secondary Outcomes in Clinical Trials.* Control Clin Trials 18, 561-567.
252. Price, D.D., McGrath, P.A., Rafii, A. and Buckingham, B. (1983) *The validation of visual analogue scales as ratio scale measures for chronic and experimental pain.* Pain 17(1), 45-56.
253. RCGP. (1996a) *The Back Book.* Her Majesty's Stationery Office. ISBN 0-11-702078-8.
254. Royal College of General Practitioners. (1996b) *Clinical guidelines for the management of acute low back pain.* RCGP.

255. Ren, X.S., Selim, A.J., Fincke, G., Deyo, R.A., Linzer, M., Lee, A. and Kazis, L. (1999) *Assessment of functional status, low back disability, and use of diagnostic imaging in patients with low back pain and radiating leg pain*. J Clin Epidemiol 52(11), 1063-71.
256. Richardson, B. and Jull, G. (1994) *Concepts of assessment and rehabilitation for active lumbar stability*. In: P. Boyling and N. Palastagna (Eds), *Grieve's Modern Manual Therapy*, 2nd ed., pp. 715-720. Churchill Livingstone, Edinburgh.
257. Richardson, C. (1995) *Muscle Control on Spinal Instability: Advances Through Interaction of Clinical Skills and Scientific Research*. Conference Proceedings Manipulative Physiotherapy Association of Australia 9th Biennial Conference, Gold Coast, Queensland, Australia, 127-131.
258. Richardson, C. and Bullock, M.I. (1986) *Changes in muscle activity during fast, alternating flexion-extension movements of the knee*. Scan J Rehab Med 18(2), 51-8.
259. Richardson, C., Jull, G., Hodges, P. and Hides, J. (1998) *Therapeutic Exercise for Spinal Stabilization in Low Back Pain*. Churchill Livingstone, Edinburgh.
260. Richardson, C., Jull, G. and Richardson, B. (1995) *A Dysfunction of the Deep Muscles Exists in Low Back Pain Patients*. Conference Proceedings World Congress of Physical Therapy, Washington, 932.
261. Richardson, C., Jull, G., Toppenberg, R. and Comerford, M. (1992) *Technique for active stabilisation for spinal protection: A Pilot Study*. Aust J Phys Ther 38(2), 105-112.
262. Richardson, C.A. and Jull, G.A. (1995) *Muscle control - pain control. What exercises would you prescribe?* Manual Therapy 1(1), 2-10.
263. Riddle, D.L. (1998) *Classification and low back pain: a review of the literature and critical analysis of selected systems*. Phys Ther 78(7), 708-37.

264. Riddle, D.L., Stratford, P.W. and Binkley, J.M. (1998) *Sensitivity to change of the Roland-Morris Back Pain Questionnaire: part 2*. Physical Therapy 78(11), 1197-207.
265. Roberts, C. and Torgerson, D.J. (1999) *Understanding controlled trials: baseline imbalance in randomised controlled trials*. BMJ 319(7203), 185.
266. Roland, M. and Fairbank, J. (2000) *The Roland-Morris Disability Questionnaire and the Oswestry Disability Questionnaire*. Spine 25(24), 3115-3124.
267. Roland, M. and Morris, R. (1983a) *A study of the natural history of back pain. Part I: development of a reliable and sensitive measure of disability in low-back pain*. Spine 8(2), 141-4.
268. Roland, M. and Morris, R. (1983b) *A study of the natural history of low-back pain. Part II: development of guidelines for trials of treatment in primary care*. Spine 8(2), 145-50.
269. Roland, M. and Torgerson, D. (1998a) *Understanding controlled trials: what outcomes should be measured?* BMJ 317(7165), 1075.
270. Roland, M. and Torgerson, D.J. (1998b) *What are pragmatic trials?* BMJ 316(7127), 285.
271. Roland, M., Waddell, G., Klaber Moffett, J., Burton, K., Main, C. and Cantrell, T. (1996) *The Back Book*. The Stationery Officer Ltd.
272. Rose, M.J., Reilly, J.P., Pennie, B., Bowen-Jones, K., Stanley, I.M. and Slade, P.D. (1997) *Chronic low back pain rehabilitation programs: a study of the optimum duration of treatment and a comparison of group and individual therapy*. Spine 22(19), 2246-51.
273. Rosenstiel, A.K. and Keefe, F.J. (1983) *The use of coping strategies in chronic low back pain patients: relationship to patient characteristics and current adjustment*. Pain 17(1), 33-44.
274. Rothstein, J.M. (1993) *Patient classification*. Phys Ther 73(4), 214-5.

275. Roy, S.H., De Luca, C.J. and Casavant, D.A. (1989) *Lumbar muscle fatigue and chronic lower back pain*. Spine 14(9), 992-1001.
276. Roy, S.H., De Luca, C.J., Snyder-Mackler, L., Emley, M.S., Crenshaw, R.L. and Lyons, J.P. (1990) *Fatigue, recovery, and low back pain in varsity rowers*. Med Sci Sports Exerc 22(4), 463-9.
277. Ruta, D.A., Garratt, A.M., Wardlaw, D. and Russell, I.T. (1994) *Developing a valid and reliable measure of health outcome for patients with low back pain*. Spine 19(17), 1887-96.
278. Saal, J.A. and Saal, J.S. (1989) *Nonoperative treatment of herniated lumbar intervertebral disc with radiculopathy. An outcome study*. Spine 14(4), 431-7.
279. Sahrman, S. (1993) *Movement as a cause of musculoskeletal pain*. Conference Proceedings 8th Biennial Conference of the Manipulative Physiotherapists Association of Australia, Perth, Western Australia, 69-74.
280. Salen, B.A., Spangfort, E.V., Nygren, A.L. and Nordemar, R. (1994) *The Disability Rating Index: an instrument for the assessment of disability in clinical settings*. J Clin Epidemiol 47(12), 1423-35.
281. Sapsford, R. (2001) *The Pelvic Floor. A clinical model for function and rehabilitation*. Physiotherapy 87(12), 620-630.
282. Sapsford, R., Hodges, P., Richardson, C., Cooper, D., Jull, G. and Markwell, S. (1997) *Activation of Pubococcygeus During a Variety of Isometric Abdominal Exercises*. Conference Proceedings Int Continence Soc, Japan.
283. Sapsford, R.R. and Hodges, P.W. (2001) *Contraction of the pelvic floor muscles during abdominal maneuvers*. Arch Phys Med Rehabil 82(8), 1081-8.
284. Schwartz, D. and Lellouch, J. (1967) *Explanatory and pragmatic attitudes in therapeutic trials*. J Chron Dis 20, 637-648.

285. Schwartz, R., Rohan, J. and Hayden, F. (1999) *Diagnostic paraspinal musculoskeletal ultrasonography*. J Back and Musculoskeletal Rehabilitation 12, 25-33.
286. Scott, J. and Huskisson, E.C. (1976) *Graphic representation of pain*. Pain 2(2), 175-84.
287. Scott, S.C., Goldberg, M.S., Mayo, N.E., Stock, S.R. and Poitras, B. (1999) *The association between cigarette smoking and back pain in adults*. Spine 24(11), 1090-8.
288. Selim, A.J., Ren, X.S., Fincke, G., Deyo, R.A., Rogers, W., Miller, D., Linzer, M. and Kazis, L. (1998) *The importance of radiating leg pain in assessing health outcomes among patients with low back pain. Results from the Veterans Health Study*. Spine 23(4), 470-4.
289. Sell, P. (1999a) *An independence evaluation of the Distress Risk Assessment using outcome data*. Conference Proceedings Brit Spine, Manchester, UK.
290. Sell, P. (1999b) *A prospective evaluation of bio social factors and Distress Risk Assessment in the outcome of spinal surgery*. Conference Proceedings Brit Spine, Manchester, UK.
291. Sim, J. and Reid, N. (1999) *Statistical Inference by Confidence Intervals: Issues of Interpretation and Utilization*. Phys Ther 79(2), 186-195.
292. Sim, J. and Waterfield, J. (1997) *Validity, reliability and responsiveness in the assessment of pain*. Physiotherapy Theory and Practice 13, 23-37.
293. Sim, J. and Wright, C. (2000) *Research in Health Care*. Stanley Thornes, Cheltenham.
294. Sim, J. and Wright, C. (2002) *Intention to Treat Approach to Data from Randomized Controlled Trials; a Sensitivity Analysis*. In preparation.
295. Singh, S.J., Morgan, M.D., Scott, S., Walters, D. and Hardman, A.E. (1992) *Development of a shuttle walking test of disability in patients with chronic airways obstruction*. Thorax 47(12), 1019-24.

296. Skinner, A. and Thomson, A. (1994) *Hydrotherapy for spinal problems*. In: J. Boyling and N. Palastagna (Eds), *Grieve's Modern Manual Therapy*, 2nd ed., pp. 823-835. Churchill Livingstone, Edinburgh.
297. Sluijs, E.M., Kok, G.J. and van der Zee, J. (1993) *Correlates of exercise compliance in physical therapy*. *Physical Therapy* 73(11), 771-86.
298. Sokal, R. and Rolf, F. (1969) *Biometry: the principles and practice of statistics in biological research*. WH Freeman and Company, USA.
299. Soukup, M., Lonn, J., Glomsrod, B., Bo, K. and Larsen, S. (2000) *Long-Term Effect of Exercise and Education on Recurrent Low Back Pain*. Conference Proceedings The Fourth International Forum for Primary Care Research on Low Back Pain, Eilat, Israel, 51.
300. Spitzer, W., LeBlanc, F. and Dupis, M. (1987) *Scientific approach to the assessment and management of activity-related spinal disorders. A monograph for clinicians. Report of the Quebec Task Force on Spinal Disorders*. *Spine* 12(7 Suppl), S1-59.
301. Staquet, M., Hays, R. and Fayers, P. (1998) *Quality of life assessment in clinical trials*. Oxford University Press, Oxford.
302. Stevans, J. and Hall, K.G. (1998) *Motor skill acquisition strategies for rehabilitation of low back pain*. *J Orthop Sports Phys Ther* 28(3), 165-7.
303. Stratford, P.W., Binkley, J., Solomon, P., Finch, E., Gill, C. and Moreland, J. (1996a) *Defining the minimum level of detectable change for the Roland-Morris Questionnaire*. *Physical Therapy* 76(4), 359-68.
304. Stratford, P.W. and Binkley, J.M. (1999) *Applying the results of self-report measures to individual patients: an example using the Roland-Morris Questionnaire*. *J Orthop Sports Phys Ther* 29(4), 232-9.

305. Stratford, P.W., Binkley, J.M., Riddle, D.L. and Guyatt, G.H. (1998) *Sensitivity to change of the Roland-Morris Back Pain Questionnaire: part 1*. Physical Therapy 78(11), 1186-96.
306. Stratford, P.W., Finch, E., Solomon, P., Binkley, J., Gill, C. and Moreland, J. (1996b) *Using the Roland-Morris Questionnaire to make decisions about individual patients*. Physiotherapy Canada 48(2), 107-10.
307. Stuge, B., Vollestad, N., Laerum, E. and Kirkesola, G. (2001) *The efficacy of a specific spinal stabilising exercise program in the treatment of patients with peripartum pelvic pain after pregnancy. A randomised controlled trial*. Conference Proceedings 4th Interdisciplinary World Congress on Low Back and Pelvic Pain, Montreal, Canada, 334-338.
308. Sullivan, M.J., Reesor, K., Mikail, S. and Fisher, R. (1992) *The treatment of depression in chronic low back pain: review and recommendations*. Pain 50(1), 5-13.
309. Swinkels, A. and Dolan, P. (1998) *Regional assessment of joint position sense in the spine*. Spine 23(5), 590-7.
310. Swinkels, A. and Dolan, P. (2000) *Spinal position sense is independent of the magnitude of movement*. Spine 25(1), 98-104.
311. Symonds, T.L., Burton, A.K., Tillotson, K.M. and Main, C.J. (1996) *Do attitudes and beliefs influence work loss due to low back trouble?* Occup Med (Lond) 46(1), 25-32.
312. Taimela, S., Kankaanpää, M. and Luoto, S. (1999) *The effect of lumbar fatigue on the ability to sense a change in lumbar position. A controlled study*. Spine 24(13), 1322-7.
313. Taylor, J.L. and McCloskey, D.I. (1990) *Proprioceptive sensation in rotation of the trunk*. Exp Brain Res 81(2), 413-6.
314. Taylor, S.J., Taylor, A.E., Foy, M.A. and Fogg, A.J. (1999) *Responsiveness of common outcome measures for patients with low back pain*. Spine 24(17), 1805-12.

315. Tesh, K.M., Dunn, J.S. and Evans, J.H. (1987) *The abdominal muscles and vertebral stability*. Spine 12(5), 501-8.
316. Thomas, E., Silman, A.J., Papageorgiou, A.C., Macfarlane, G.J. and Croft, P.R. (1998) *Association between measures of spinal mobility and low back pain. An analysis of new attenders in primary care*. Spine 23(3), 343-7.
317. Thorstensson, A., Oddsson, L. and Carlson, H. (1985) *Motor control of voluntary trunk movements in standing*. Acta Physiol Scand 125(2), 309-21.
318. Tillotson, K.M. and Burton, A.K. (1991) *Noninvasive measurement of lumbar sagittal mobility. An assessment of the flexicurve technique*. Spine 16(1), 29-33.
319. Treasure, T. and MacRae, K.D. (1998) *Minimisation: the platinum standard for trials? Randomisation doesn't guarantee similarity of groups; minimisation does*. BMJ 317(7155), 362-3.
320. Treasure, T. and MacRae, K.D. (1999) *Minimisation is much better than the randomised block design in certain cases*. BMJ 318(7195), 1420.
321. Triano, J., McGregor, M., Cramer, G. and Emde, D. (1993) *A Comparison of Outcome Measures for Use with Back Pain Patients: Results of a Feasibility Study*. Journal of Manipulative & Physiological Therapeutics 16(2).
322. Troup, J.D., Martin, J.W. and Lloyd, D.C. (1981) *Back pain in industry. A prospective survey*. Spine 6(1), 61-9.
323. Underwood, M.R., Barnett, A.G. and Vickers, M.R. (1999) *Evaluation of two time-specific back pain outcome measures*. Spine 24(11), 1104-12.
324. Vaccaro, A.R., Ring, D., Scuderi, G., Cohen, D.S. and Garfin, S.R. (1997) *Predictors of outcome in patients with chronic back pain and low-grade spondylolisthesis*. Spine 22(17), 2030-4.

325. van Baar, M.E., Dekker, J. and Bosveld, W. (1998) *A survey of physical therapy goals and interventions for patients with back and knee pain*. Phys Ther 78(1), 33-42.
326. van der Linden, S., Bouter, L.M. and Tugwell, P. (1991) *What are the minimal methodological and statistical requirements for a good trial? I. The clinician's view*. In: P. Schlapbach and N. Gerber (Eds), *Physiotherapy: Controlled Trials and Facts*., pp. 1-8. Vol. 14. Basel, Karer.
327. van Tulder, M., Esmil, R., Bombardier, C. and Koes, B. (2000a) *Back schools for non-specific low back pain (Cochrane Review)*. In: The Cochrane Library, Issue 4, Oxford.
328. van Tulder, M. and Koes, B. (2001) *Clinical Evidence. A compendium of the best available evidence effective healthcare*, 772-789 pp, Issue 5. BMJ Publishing group.
329. van Tulder, M., Malmivaara, A., Esmail, R. and Koes, B. (2000b) *Exercise therapy for low back pain: A systematic review within the framework of the Cochrane collaboration back review group*. Spine 25(21), 2784-96.
330. van Tulder, M.W., Esmail, R., Bombardier, C. and Koes, B.W. (2000c) *Back schools for non-specific low back pain*. Cochrane Database Syst Rev 2.
331. van Tulder, M.W., Koes, B.W. and Bouter, L.M. (1997) *Conservative treatment of acute and chronic nonspecific low back pain. A systematic review of randomized controlled trials of the most common interventions*. Spine 22(18), 2128-56.
332. Vickers, A.J. (2001) *The use of percentage change from baseline as an outcome in a controlled trial is statistically inefficient: a simulation study*. BMC Med Res Methodol 1(1), 6.
333. Vickers, A.J. and Altman, D.G. (2001) *Analysing controlled trials with baseline and follow up measurements*. BMJ 323(7321), 1123-4.
334. Von Korff, M. (1994) *Studying the natural history of back pain*. Spine 19(18 Suppl), 2041S-2046S.

335. Von Korff, M., Deyo, R.A., Cherklin, D. and Barlow, W. (1993) *Back pain in primary care. Outcomes at 1 year*. Spine 18(7), 855-62.
336. Waddell, G. (1987) *1987 Volvo award in clinical sciences. A new clinical model for the treatment of low-back pain*. Spine 12(7), 632-44.
337. Waddell, G. (1998) *The Back Pain Revolution*. Churchill Livingstone, Edinburgh.
338. Waddell, G., McCulloch, J.A., Kummel, E. and Venner, R.M. (1980) *Nonorganic physical signs in low-back pain*. Spine 5(2), 117-25.
339. Royal College of General Practitioners. (1999) *Low back pain evidence review*. Waddell, G., McIntosh, A. and Hutchinson, A.
340. Waddell, G., Newton, M., Henderson, I., Somerville, D. and Main, C.J. (1993a) *A Fear-Avoidance Beliefs Questionnaire (FABQ) and the role of fear-avoidance beliefs in chronic low back pain and disability*. Pain 52(2), 157-68.
341. Waddell, G., Newton, M., Henderson, I., Somerville, D. and Main, C.J. (1993b) *A Fear-Avoidance Beliefs Questionnaire (FABQ) and the role of fear-avoidance beliefs in chronic low back pain and disability*. Pain 52(2), 157-68.
342. Wakefield, A. (2000) *Evidence-based Physiotherapy: The case of pragmatic randomised controlled trials*. Physiotherapy 86(8), 394-396.
343. Walsh, K., Cruddas, M. and Coggon, D. (1992) *Low back pain in eight areas of Britain*. J Epidemiol Community Health 46(3), 227-30.
344. Ware, J.E., Jr. (2000) *SF-36 Health Survey Update*. Spine 25(24), 3130-3139.
345. Ware, J.E., Jr. and Sherbourne, C.D. (1992) *The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection*. Med Care 30(6), 473-83.
346. Waxman, R., Tennant, A. and Helliwell, P. (1998) *Community survey of factors associated with consultation for low back pain*. BMJ 317(7172), 1564-7.

347. Weinstein, J.N. and Deyo, R.A. (2000) *Clinical Research: Issues in Data Collection*. Spine 25(24), 3104-3109.
348. Welton, A.J., Vickers, M.R., Cooper, J.A., Meade, T.W. and Marteau, T.M. (1999) *Is recruitment more difficult with a placebo arm in randomised controlled trials? A quasirandomised, interview based study*. BMJ 318(7191), 1114-7.
349. Wilder, D.G., Aleksiev, A.R., Magnusson, M.L., Pope, M.H., Spratt, K.F. and Goel, V.K. (1996) *Muscular response to sudden load. A tool to evaluate fatigue and rehabilitation*. Spine 21(22), 2628-39.
350. Wilke, H.J., Wolf, S., Claes, L.E., Arand, M. and Wiesend, A. (1995) *Stability increase of the lumbar spine with different muscle groups. A biomechanical in vitro study*. Spine 20(2), 192-8.
351. Williams, M., Grant, R. and Main, C. (1995) *The Distress Risk Assessment Method (DRAM) as a predictor of outcome in chronically disabled workers attending a physical rehabilitation programme*. Conference Proceedings International Society for the Study of the Lumbar Spine, Helsinki, Finland, 46.
352. Wright, A.F. (1994) *Should general practitioners be testing for depression?* Br J Gen Pract 44(380), 132-5.
353. Wright, B.D. and Linacre, J.M. (1989) *Observations are always ordinal; measurements, however, must be interval*. Arch Phys Med Rehabil 70(12), 857-60.
354. Wright, L., Harwood, D. and Coulter, A. (1992) *Health and Lifestyles in the Oxford Region*. Health Services Research Unit, Oxford.
355. Zung, W.W., Richards, C.B. and Short, M.J. (1965) *Self-rating depression scale in an outpatient clinic. Further validation of the SDS*. Arch Gen Psychiatry 13(6), 508-15.

Appendices

Appendix	Page number
Appendix A: Description of spinal stabilisation training.....	234
Appendix B: Roland Morris Disability Questionnaire	236
Appendix C: Oswestry Disability Index.....	237
Appendix D: Numerical Rating Scale	239
Appendix E: Short Form McGill Pain Questionnaire	240
Appendix F: Modified Zung.....	241
Appendix G: Modified Somatic Perception Questionnaire	242
Appendix H: U.K. SF-36 Questionnaire.....	243
Appendix I: Detailed anatomical attachments of Transversus Abdominis	248
Appendix J: Notes on clinical instability	249
Appendix K: Outline for Specific Spinal Stabilisation Regime.....	251
Appendix L: Spinal Stabilisation booklet.....	252
Appendix M: Indications for referral for diagnostic ultrasound	265
Appendix N: Consent Form	266
Appendix O: Patient information sheet.....	267
Appendix P: Data collection sheet – Pilot Study	270
Appendix Q: Data collection sheet - Main Trial.....	271
Appendix R: Non-responders follow-up post-cards.....	272
Appendix S: Outcome Sheet for main trial	273
Appendix T: Treatment Record	274
Appendix U: RCT Discharge Sheet.....	275
Appendix V: Clinicians' Questionnaires	276
Appendix W: Power calculations	277
Appendix X: Summary of data check.....	278
Appendix Y: Distributions of demographic and biographic variables at baseline.....	281
Appendix Z: Socio-demographics baseline information.....	283
Appendix AA: Baseline Distributions - Study variables	286
Appendix BB: Demographics of treatment completers versus treatment non-completers	289
Appendix CC: Baseline demographics and biographics of responders and non-responders at 12-month follow-up	291
Appendix DD: Calculation of Area Under the Curve (AUC)	292
Appendix EE: Individual time-series curves for each group	294

Appendix FF: Supplementary data for U.K. SF-36.....	296
Appendix GG: Summary of studies using the DRAM to assess distress in various LBP populations.....	300
Appendix HH: Methodological Quality Criteria.....	302

Appendices Figures & Tables	Page number
Figure A 1: Diagrammatic representation of abdominal drawing-in procedure.....	235
Figure A 2: Boxplot of age by group	281
Figure A 3: Boxplot of height by group.....	281
Figure A 4: Boxplot of weight by group	282
Figure A 5: Boxplot of total duration of LBP symptoms.....	282
Figure A 6: Boxplot of current duration of LBP symptoms.....	282
Figure A 7: Summary of employment status for each group	284
Figure A 8: Summary of number of previous pregnancies	285
Figure A 9: Boxplot of RMDQ scores at baseline.....	286
Figure A 10: Boxplot of ODI scores at baseline	286
Figure A 11: Boxplot of SFMPQ at baseline.....	287
Figure A 12: Boxplot of SFMPQ (Sensory component) at baseline.....	287
Figure A 13: Boxplot of SFMPQ (Affective component) at baseline	287
Figure A 14: Boxplot of SFMPQ (PPI) at baseline	287
Figure A 15: Boxplot of NRS scores at baseline.....	288
Figure A 16: Boxplot of VAS scores at baseline.....	288
Figure A 17: Boxplot of MZ scores at baseline	288
Figure A 18: Boxplot of MSPQ scores at baseline	288
Figure A 19: QTF classification distributions for completers and non-completers.....	290
Figure A 20: QTF classifications for responders and non-responders at 12-month follow-up	291
Figure A 21: Diagrammatic representation of AUC with all known values	293
Figure A 22: Diagrammatic representation of AUC with interpolated values	293
Figure A 23: Distribution of individual items of the Physical Functioning scale	297
Figure A 24: Distribution of individual items Social Functioning scale.....	297
Figure A 25: Distribution of individual items Mental Health scale.....	297
Figure A 26: Distribution of individual items Energy/Vitality scale	297
Figure A 27: Distribution of individual items Pain scale	297
Figure A 28: Distribution of individual items General Health scale.....	297

Figure A 29: Scales comprising the PCS at baseline and 12 months follow-up	298
Figure A 30: Scales comprising the MCS at baseline and 12 months follow-up.....	298

Appendix A: Description of spinal stabilisation training

This appendix presents a description of specific stabilisation training, the type and level of contractions, positioning and progression. It is not intended to be an exhaustive review, but provide an overview for readers not familiar with these training techniques. Readers requiring more in depth descriptions are directed to the following sources; (Richardson and Jull, 1995; Norris, 1995; Comerford and Mottram, 2001a).

1. Type of contraction

As the postural, stability muscles are the target of retraining, prolonged tonic holding contractions at a low maximum voluntary contraction (MVC), usually between 25-40% are used. Contractions are undertaken in an isolated fashion and incorporated into co-contraction e.g. transversus abdominis (TrA) and lumbar multifidus (LM).

2. Body position and resistance

In contrast to strength training where external force is often purposely applied for therapeutic purposes, minimal external loading is used in stabilisation retraining, as only low levels, 30-40% MVC, are required to activate tonic muscle fibres. Antigravity body positions such as four-point kneeling or prone lying are also used in order to reduce external loads.

3. Number of repetitions

As stabilization training aims to improve endurance, it is suggested that for maximum benefit, exercises need to be repeated as often as possible throughout the day. To date there is no documented evidence of the optimum number of repetitions needed to facilitate a change in muscle function.

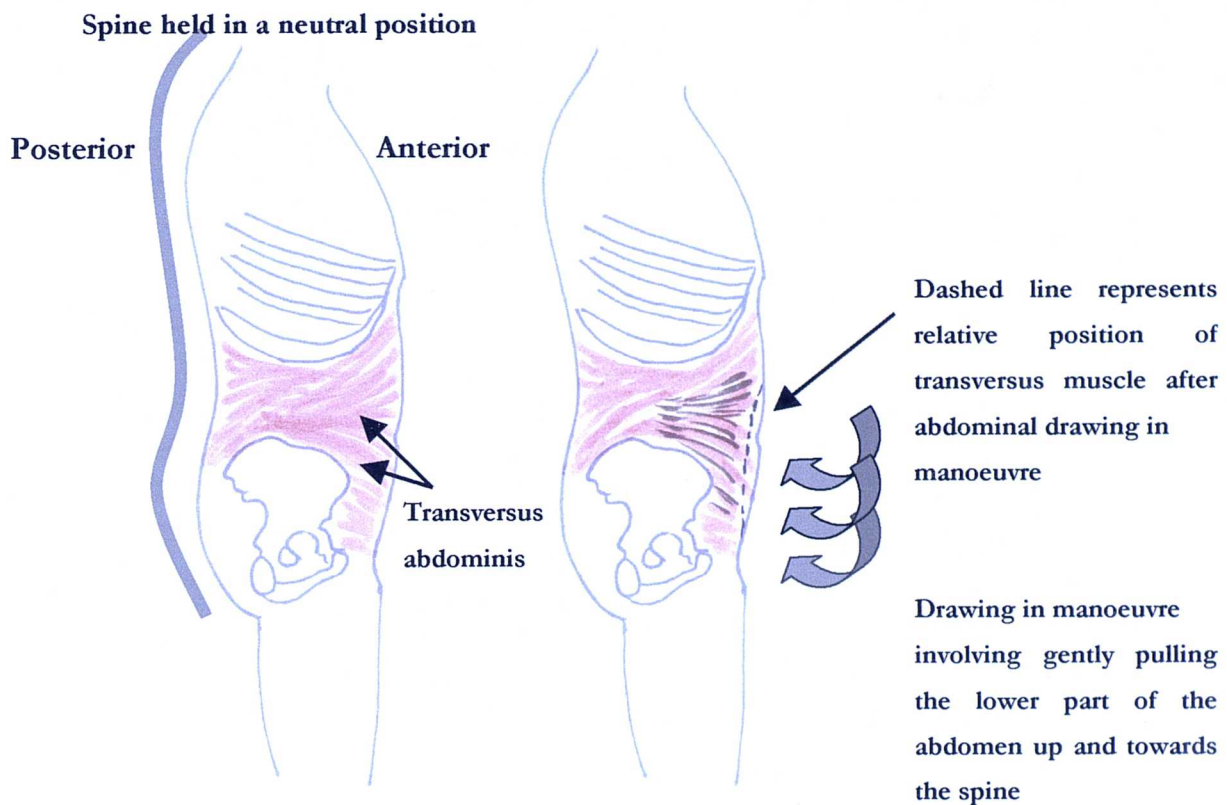
4. Method of progression

Initially, because increased endurance is a major aim of the retraining programme, the length of hold and number of repetitions are increased. The setting procedure is then progressed from antigravity to more functional positions with the gradual incorporation of graded external loading and finally to holding co-contractions during dynamic functional movements. Static training with leg loading and with direct resistance to the trunk can be used.

5. Teaching methods

It is acknowledged that there are only a few methods of achieving a relatively isolated co-contraction of the local stabilising muscles independently of the global muscles (See 2.14.5). Figure A 1 below is a diagrammatic representation of the abdominal drawing-in procedure, which has been shown to facilitate a relatively isolated TrA contraction.

Figure A 1: Diagrammatic representation of abdominal drawing-in procedure



A number of techniques can be used to assist patients in achieving the specific contraction required for specific spinal stabilization training including;

- Visualisation
- Cueing correct activation
- Focussing on precision
- Facilitation techniques

Appendix B: Roland Morris Disability Questionnaire

When your back hurts, you may find it difficult to do some things you normally do. This list contains 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, put a tick against it. If the sentence does not describe you, then leave the space blank and go on to the next one. Remember, only tick the sentence if you are sure that it describes you today.

Because of my back or leg pain (sciatica) today:

YES/NO

1. I stay at home most of the time because of my back.
2. I change position frequently to try and get my back comfortable.
3. I walk more slowly than usual because of my back.
4. Because of my back, I am not doing any of the jobs that I usually do around the house.
5. Because of my back, I use a handrail to get upstairs.
6. Because of my back, I lie down to rest more often.
7. Because of my back, I have to hold on to something to get out of an easy chair.
8. Because of my back, I try to get other people to do things for me.
9. I get dressed more slowly than I used to because of my back.
10. I only stand up for short periods because of my back.
11. Because of my back, I try not to bend or kneel down.
12. I find it difficult to get out of my chair, because of my back.
13. My back is painful all of the time.
14. I find it difficult to turn over in bed because of my back.
15. My appetite is not very good because of my back.
16. I have trouble putting on my socks (or stockings) because of the pain in my back.
17. I only walk short distances because of my back.
18. I sleep less well because of my back.
19. Because of my back pain, I get dressed with help from someone.
20. I sit down for most of the day because of my back.
21. I avoid heavy jobs around the house because of my back.
22. Because of my back pain, I am more irritable and bad tempered with people than usual.
23. Because of my back, I go upstairs more slowly than usual.
24. I stay in bed most of the time because of my back.

Reference: Roland, M. and Morris, R. (1983) A study of the natural history of back pain. Part I: development of a reliable and sensitive measure of disability in low-back pain. *Spine* 8 (2), 141-4.

Appendix C: Oswestry Disability Index

Could you please complete this questionnaire. The questionnaire is designed to give us information as to how your back (or leg) trouble has affected your ability to manage in everyday life. Please answer every section. Mark one box only in each section that most closely describes you today.

Section 1 – Pain Intensity

- ☐ I have no pain at the moment
- ☐ The pain is mild at the moment
- ☐ The pain is moderate at the moment
- ☐ The pain is fairly severe at the moment
- ☐ The pain is very severe at the moment
- ☐ The pain is the worst imaginable at the moment

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 is very painful
- ☐ 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 weights but it gives extra pain
- ☐ Pain prevents me from lifting heavy weights off the floor but I can manage if they are conveniently positioned, e.g. 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 a mile
- ☐ Pain prevents me walking more than ¼ of a mile
- ☐ Pain prevents me walking more than 100 yards
- ☐ 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 any chair as long as I like
- ☐ I can sit in my favourite chair as long as I like
- ☐ Pain prevents me from sitting for more than 1 hour
- ☐ Pain prevents me from sitting for more than half an hour
- ☐ Pain prevents me from sitting for more than 10 minutes
- ☐ 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 half an hour
- ☐ Pain prevents me from standing for more than 10 minutes
- ☐ Pain prevents me from standing at all

Section 7 – Sleeping

- ☐ My sleep is never disturbed by pain
- ☐ My sleep is occasionally disturbed by pain
- ☐ Because of pain I have less than 6 hours sleep
- ☐ Because of pain I have less than 4 hours sleep
- ☐ Because of pain I have less than 2 hours sleep
- ☐ Pain prevents me from sleeping at all

Section 8 – Sex Life (if applicable)

- ☐ My sex life is normal and causes no extra pain
- ☐ My sex life is normal and but causes some extra pain
- ☐ My sex life is nearly normal but is very painful
- ☐ My sex life is severely restricted
- ☐ 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 causes 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 e.g. sport, etc.
- ☐ 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 pain
- ☐ I can travel anywhere but it gives extra pain
- ☐ Pain is bad but I manage journeys over two hours
- ☐ Pain restricts me to journeys of less than one hour
- ☐ Pain restricts me to short necessary journeys under 30 minutes
- ☐ Pain prevents me from travelling except to receive treatment

Section 11 – Previous treatment

Over the past three months have you received treatment, tablets or medicines of any kind for your back or leg pain?

Please tick the appropriate box. ☐ No ☐ Yes (If yes, please state the type of treatment you have received)

Reference: Fairbank, J.C., Couper, J., Davies, J.B. and O'Brien, J.P. (1980) The Oswestry low back pain disability questionnaire. *Physiotherapy* 66 (8), 271-3.

Appendix D: Numerical Rating Scale

Please circle one number only on the following scale for your *usual* pain level

No pain	0	1	2	3	4	5	6	7	8	9	10	Worst Pain Possible
---------	---	---	---	---	---	---	---	---	---	---	----	------------------------

Reference: Ekblom, A. and Hansson, P. (1988) Pain intensity measurements in patients with acute pain receiving afferent stimulation. *J Neurol Neurosurg Psychiatry* 51(4), 481-6.

Appendix E: Short Form McGill Pain Questionnaire

Short Form McGill Pain Questionnaire
Roland Melzack

	<u>None</u>	<u>Mild</u>	<u>Moderate</u>	<u>Severe</u>	
i. Throbbing	0)	1)	2)	3)	
ii. Shooting	0)	1)	2)	3)	
iii. Stabbing	0)	1)	2)	3)	
iv. Sharp	0)	1)	2)	3)	
v. Cramping	0)	1)	2)	3)	
vi. Gnawing	0)	1)	2)	3)	
vii. Hot-Burning	0)	1)	2)	3)	
viii. Aching	0)	1)	2)	3)	
ix. Heavy	0)	1)	2)	3)	
x. Tender	0)	1)	2)	3)	
xi. Splitting	0)	1)	2)	3)	
xii. Tiring- Exhausting	0)	1)	2)	3)	
xiii. Sickening	0)	1)	2)	3)	
xiv. Fearful	0)	1)	2)	3)	
xv. Punishing-Cruel	0)	1)	2)	3)	

The words on
the left can all
describe pain

Please tick
from "None"
to "Severe"
for all words

Please mark along the line how you would rate your pain
Please circle the option below that describes your pain

NO		WORST
PAIN		POSSIBLE
		PAIN

PPI

- 0 No pain
- 1 Mild
- 2 Discomforting
- 3 Distressing
- 4 Horrible
- 5 Excruciating

Reference: Melzack, R. (1987) The short-form McGill Pain Questionnaire. Pain 30 (2), 191-7.

Appendix F: Modified Zung

Please indicate for each of these questions which answer best describes how you have been feeling recently				
	Rarely or none of the time (less than 1 day per week)	Some or little of the time (1-2 days per week)	A moderate amount of the time (3-4 days per week)	Most of the time (5-7 days per week)
I feel down hearted and sad	0	1	2	3
Morning is when I feel best	3	2	1	0
I have crying spells or feel like it	0	1	2	3
I have trouble getting to sleep at night	0	1	2	3
I feel that nobody cares	0	1	2	3
I eat as much as I used to	3	2	1	0
I still enjoy sex	3	2	1	0
I notice I am losing weight	0	1	2	3
I have trouble with constipation	0	1	2	3
My heart beats faster than usual	0	1	2	3
I get tired for no reason	0	1	2	3
My mind is as clear as it used to be	3	2	1	0
I tend to wake up too early	0	1	2	3
I find it easy to do the things I used to	3	1	2	0
I am restless and can't keep still	0	1	2	3
I feel hopeful for the future	3	2	1	0
I am more irritable than usual	0	1	2	3
I find it easy to make a decision	3	2	1	0
I feel quite guilty	0	1	2	3
I feel that I am useful and needed	3	2	1	0
My life is pretty full	3	2	1	0
I feel that others would be better off if I were dead	0	1	2	3
I am still able to enjoy the things I used to	3	2	1	0

Questionnaire is given to patient without the scoring on.

Reference: Zung, W.W., Richards, C.B. and Short, M.J. (1965) Self-rating depression scale in an outpatient clinic. Further validation of the SDS. Arch Gen Psychiatry 13(6), 508-15

Appendix G: Modified Somatic Perception Questionnaire

Please describe how you felt during the PAST WEEK by making a check mark (x) in the appropriate box. Please answer all the questions. Do not think too long before answering				
	Not at all	A little, slightly	A great deal, quite a bit	Extremely, could not have been worse
Heart rate increased				
Feeling hot all over	0	1	2	3
Sweating all over	0	1	2	3
Sweating in a particular part of the body				
Pulse in the neck				
Pounding in head				
Dizziness	0	1	2	3
Blurring of vision	0	1	2	3
Feeling faint	0	1	2	3
Everything appearing unreal				
Nausea	0	1	2	3
Butterflies in stomach				
Pain or ache in stomach	0	1	2	3
Stomach churning	0	1	2	3
Desire to pass water				
Mouth becoming dry	0	1	2	3
Difficulty swallowing				
Muscles in neck aching	0	1	2	3
Legs feeling weak	0	1	2	3
Muscles twitching or jumping	0	1	2	3
Tense feeling across forehead	0	1	2	3
Tense feeling in jaw muscles				

Questionnaire is given to patient without the scoring or shading on. Only non-shaded items are used in the scoring

Reference: Main, C.J. (1983) The Modified Somatic Perception Questionnaire (MSPQ). J Psychosom Res 27(6), 503-14.

Appendix H: U.K. SF-36 Questionnaire

SF-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:

(circle one)

Excellent 1
Very good..... 2
Good..... 3
Fair 4
Poor..... 5

2. Compared to one year ago, how would you rate your health in general now?

(circle one)

Much better now than one year ago..... 1
Somewhat better now than one year ago..... 2
About the same as one year ago..... 3
Somewhat worse now than one year ago 4
Much worse now than one year ago 5

3. The following questions 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)

<u>ACTIVITIES</u>	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 a mile	1	2	3
h. Walking half a mile	1	2	3
i. Walking one hundred yards	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 (for example, 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 daily 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 on 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 health 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

A little bit 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 one 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 life?	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 downhearted and low?	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 ill more easily than other 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

5/93 IQOLA SF-36 Standard U.K.
 Version 1.0

Reference: Ware, J.E., Jr. and Sherbourne, C.D. (1992) The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care* 30(6), 473-83.

Appendix I: Detailed anatomical attachments of Transversus Abdominis

The deepest of the abdominal muscles, the transversus abdominis has extensive attachments with fibres running transversely from the lateral 1/3rd of the inguinal ligament, anterior 2/3rds of the iliac crest, thoraco-lumbar fascia between the iliac crest and the 12th rib, and anterior aspects of the lower costal cartilages. Blending with the fibres of the internal oblique to form an aponeurotic conjunct tendon it attaches to the crest and the pecten of the pubis forming the flax inguinalis. The remainder of the aponeurosis passes horizontally to the medial plane, blends with the linea alba with the upper 3/4 lying behind the RA and the lower 1/4 lying in front.

The upper 3/4 of TrA blends with the posterior lamina of the aponeurosis of IO with the posterior fibres of IO attaching to the posterior thoraco-lumbar fascia. Fusiform deficits can occur in the lower muscular and aponeurotic parts of both TrA and IO and are filled with fascia. The two muscles can be fused or the TrA may be absent. Both are innervated by the anterior primary rami of T7-T12 and L1 (Gray, 1980).

Appendix J: Notes on clinical instability

This appendix presents details of the concept of clinical instability and details suggested criteria for diagnosing/identifying clinical instability.

Two basic forms of instability can be identified; *terminal instability* considered to be decreased resistance to movement, and *instability within a normal range of movement* such as altered ratios between translation and rotation and an increased neutral zone (See page 50). Early concepts of instability were based primarily on the loss of 'stiffness' but later concepts introduced the notion of the loss of segmental leading to the application of greater forces and excessive displacement compared to normal and resulting in a painful condition (Frymoyer and Pope, 1991). This concept has now been somewhat superseded by the concept of the neutral zone, which was discussed in Section 2-13 (Panjabi et al., 1989; Panjabi, 1992b).

A number of sets of clinical signs and symptoms have been suggested to indicate the presence of clinical instability with varying levels of acceptance from the medical profession. Diagnosis of terminal instabilities such as those associated with spondylolisthesis are often undertaken using radiological measurements but 'diagnosis' of clinical instabilities is more controversial. Combinations of clinical presentation and symptoms have been suggested, such as those listed (1 to 10) below, but a more 'objective' measure of clinical instability is not currently available. The reader is directed to specific articles of the subject, and an example of is presented below (Flynn et al., 1997; O'Sullivan, 2000).

- 1 A history of at least 3 episodes of acute LBP
- 2 The pain may change from one side to the other with different episodes
- 3 A painful episode triggered by a trivial incident or a sudden onset due to unknown cause
- 4 A severe catch pain, locking or giving way sensation in the lower lumbar spine
- 5 A reversible scoliosis/deviation in the lumbar spine
- 6 The presence of an arc or deviation on forward flexion

- 7 The patient having to use his/her hands on the thighs to return from forward flexion
- 8 Improved range of movement and less pain with co-contraction of abdominal and para-spinal muscles
- 9 Marked localised pain and muscle spasm on palpation of the affected segment
- 10 Increased pain and further decreased active movement after a sustained (20 seconds or longer) posterior-anterior mobilisation of the affected segment
- 11 Traction spur on X-ray at the corresponding level
- 12 The presence of spondylolysis or spondylolisthesis of the corresponding level of the suspected instability

Taken from: (Flynn et al., 1997)

Appendix K: Outline for Specific Spinal Stabilisation Regime

First session - outline of plan of treatment

Anatomy and function of transversus and multifidus explained with appropriate visual aids.

Stress the need for active participation and cognitive motor relearning skill

Position

Choice of starting position is dependent on individual patients

- Four Point kneeling

- Crook Side Lying

- Crook supine

Identify suitable facilitation techniques

- Pelvic floor contraction

- Visualisation

- Palpation

Identify any substitution techniques i.e. pelvic tilt

Progression

Position in which TrA is activated should be progressed from a supine through sitting and standing

Incorporation of functional and provocative positions should be started once good activation and endurance is achieved

An arbitrary 10x10 goal can be used i.e. 10 repetition of 10-second holds

Home exercises

Patient must be 100% sure of what they are doing

In patient booklet write:

- position

- holds

- repetitions

- number per day

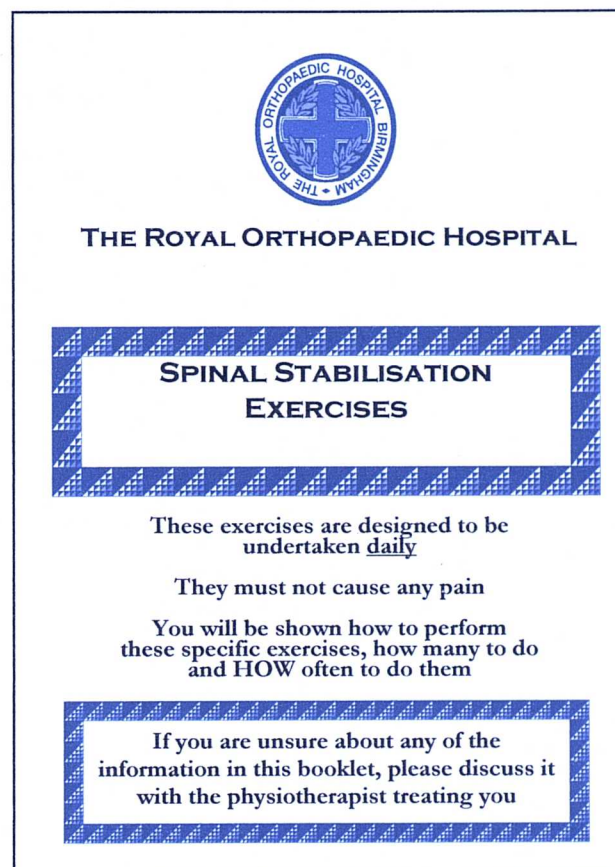
- order of contraction

Based on: Richardson, C.A. and Jull, G.A. (1995) Muscle control - pain control. What exercises would you prescribe? *Manual Therapy* 1(1), 2-10.

Appendix L: Spinal stabilisation booklet

Booklet given to all patients randomised into the SSSE group for treatment.

Patients were requested to bring the booklet to each appointment and the treating therapist noted any specific instructions in the spaces available in the booklet.



Introduction

This booklet has been designed to help explain the concepts behind deep muscle training that will form *part* of your treatment. This form of treatment has been extensively researched over the last 10 years and has been shown to be beneficial in reducing the recurrence of back pain when the exercises are performed *correctly and regularly*.

Alongside these exercises your physiotherapist will use other techniques designed to help your joints and muscles move more freely.

It is intended that this booklet provides a brief background to the exercises and training

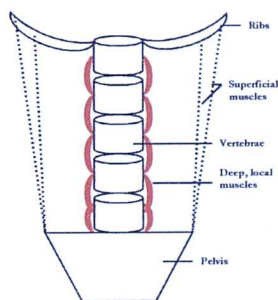
Local and superficial muscles

There are two main sets of muscles that control the spine. The *superficial muscles* attach from the ribs to the pelvis and predominantly provide control and support the whole trunk and the spine during heavy tasks such as lifting.

The other group are the *deep, local muscles*. These attach directly onto the spine and are responsible for providing support and protection for the individual parts of the spine. Normally these muscles work all day to support the spine and ironically it is these muscles that stop working properly when you get back pain. This leaves your back vulnerable to further strain. When this occurs the more superficial muscles begin to try to compensate, and as they are not designed for this, they often get sore and cause pain.

Page 1

These exercises are therefore specifically designed to improve the functioning of the deep, local muscles



Representation of the deep, local and superficial muscles of the trunk

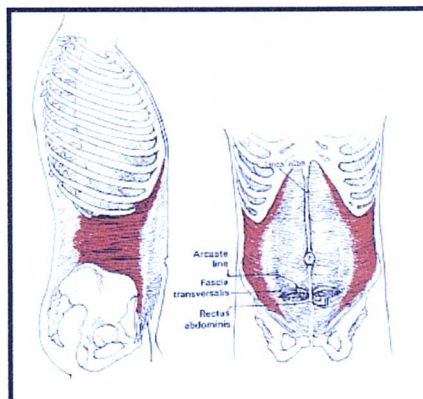
Page 2

The local deep muscles - *Transversus Abdominis*

The transversus abdominis is the deepest abdominal muscles. It attaches to the vertebrae and wraps around the stomach. It can be likened to a natural corset both from its appearance and function in supporting the back.

Research has shown that normally the brain unconsciously activates this muscle before any limb movement in order to provide the spine with support. It has also been shown that in people with back pain, this muscle does not function properly, therefore putting the spine at risk of repetitive damage.

Position of transversus muscle



Source: Kendall, FP, McCreary, E, Provance, PG (1993)
Muscles Testing and Function. Williams & Wilkins, Baltimore

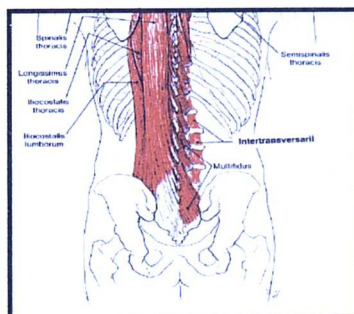
Page 3

The local deep muscles—*Lumbar Multifidus*

This is a small muscle in your back which runs between adjacent spinal bones (vertebrae). It is positioned very close to the spine. It works with the transversus abdominis muscle described earlier to provide support and protection to the spine. It gives your spine feedback about its movement and also helps prevent repetitive damage.

There are deep and superficial parts of this muscle and it is specifically the deep parts that the exercise is designed to improve as these are the ones that protect the spine.

Position of multifidus muscles



Source: Kendall, FP, McCreary, E, Provance, PG (1993)
Muscles Testing and Function. Williams & Wilkins, Baltimore

Page 4

The Exercises

All these exercises are designed to be undertaken at a low level of exertion but repeated often. If the maximum effort is 100%, then these exercises should be performed using only 25% effort.

**All exercises need to be
repeated regularly,
but very slowly and gently**

These exercises are designed to improve the endurance and co-ordination of the local, deep muscles that help support the spine.

Initially this will involve consciously thinking about working them. In order to achieve this you will need feedback and correction of your technique which your physiotherapist will provide. They will also be able to tell you when to progress to the next level. **BUT.....**

**You are the only one who can perform the exercises and make a difference
to your back problem in the longer term.**

Whilst undertaking these exercises it is very important that your spine is kept in a good position. This minimises the stresses and strains on your spine and gives it time to heal.

Page 5

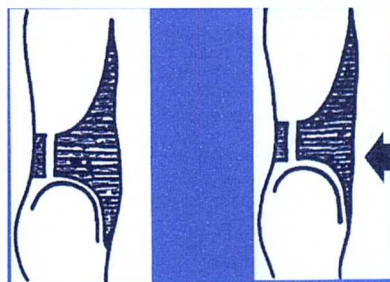
The exercise for transversus abdominis - *abdominal drawing-in*

- ◆ Relax your stomach and slowly and gently breathe
- ◆ Gently sink your fingers into your stomach about 1" from your hip bones to 'feel' the deep muscle contraction
- ◆ Breathe in then gently out then cease breathing to concentrate on achieving the muscle contraction
- ◆ *SLOWLY & GENTLY* draw in the lower part of the tummy (i.e. *the region of your tummy below the navel*) up and in towards your back bone
- ◆ Hold the contraction and resume normal breathing, aiming for a 10-second hold
- ◆ If you are unable to feel the deep muscle working, your physiotherapist may teach you some other alternative methods of activating this muscle.

Abdominal drawing-in exercise

Relaxed abdomen

Correct 'drawing-in'



Source: Richardson, Jull & Hides (1998) Educational Material

Page 6

Abdominal drawing-in - continued**NOTE:**

You have not contracted your corset muscle if:

- Your stomach above the navel is the part that draws in
- You feel a hard, fast contraction pushing against your fingers
- Your ribs flare out or up

TRY AGAIN

Individual information:

Position:

Repetitions:

Holds:

x/day:

Page 7

Individual information:

Position:

Repetitions:

Holds:

x/day:

Individual information:

Position:

Repetitions:

Holds:

x/day:

Page 8

<p>Individual information:</p> <p>Position:</p> <p>Repetitions:</p> <p>Holds:</p> <p>x/day:</p>
<p>Individual information</p> <p>Position:</p> <p>Repetitions:</p> <p>Holds:</p> <p>x/day:</p>

Notes on transversus exercise

Page 10

The exercise for Lumbar Multifidus

- ◆ Relax your stomach and breathing
- ◆ Sink fingers in the muscle adjacent to the back bone to 'feel' the contraction
- ◆ Breath in then gently out
- ◆ *SLOWLY & GENTLY* swell out the lower back muscle against your fingers and hold the contraction

Individual information:**Position:****Repetitions:****Holds:****x/day:****Page 11****Individual information:****Position:****Repetitions:****Holds:****x/day:****Individual information:****Position:****Repetitions:****Holds:****x/day:****Page 12**

Notes on activating multifidus

Page 13

Lumbar multifidus - continued

NOTE:

You have not performed the contraction correctly if:

- You feel a quick, sudden contraction
- You move your pelvis or just push your back outwards

TRY AGAIN

Additional Points

Breathing

From the start of your exercises it is essential that you do not use breathing-in as a way of cheating because this encourages contraction of the wrong muscles. *It must therefore be avoided.*

In order to achieve this, gently breathe out before you perform any of the exercises and once you have slowly contracted the muscle, gently resume breathing whilst holding the muscle contraction.

Progression

Once you have mastered the basic muscle contractions, it is important that you progress your exercises. This will include 'setting' your deep muscles during activities and positions that previously aggravated your back problem. This must still be painfree as pain will cause the muscles to be inhibited.

The rate at which people are able to train these specific muscles is very variable.

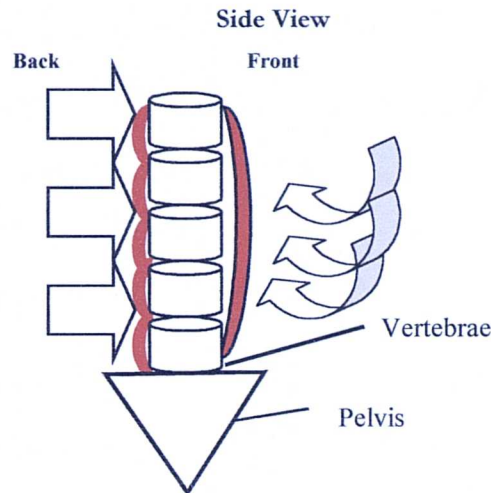
Page 14

Co-Constrictions - the corset effect

The two muscles described earlier work together as a pair in people with no back pain to protect the spine from damage.

Once you are able to contract both muscles separately it is important to make them work together. This can be done in different positions and your physiotherapist will tell you which is best for you.

Corset Effect of the deep abdominal and back muscles



Page 15

How Long will it take ????

As with learning any new skill, *practice makes perfect* and in the initial stages it may be very hard to isolate the specific muscles. This is because there are a lot of factors involved in learning a new skill so don't get disheartened. The length of time you have had pain, the type of problem and the level of dysfunction of the muscles will all influence how long it takes to get these exercises right. It may take between **10** and **12** weeks to see improvement.

Although it may seem a lot of effort with no immediate effect, research has shown that improving the function and endurance of these muscles can decrease the pain levels of patients with back pain and improve their overall functioning in daily activities. Importantly it has been shown that using these exercises can reduce the number of episodes of back pain that people suffer.

Facilitation

This is simply the term for techniques which your physiotherapist might use to help you contract the correct muscle. Everyone is different and any specific instructions will be written down to help you remember them.

Page 16

Prevention

Once you have mastered the exercises they need to be considered as preventative, like cleaning your teeth to prevent decay. Check every week that you can still contract the muscles correctly. If they are not working too well then resume the exercises for another week or until they improve.

Posture

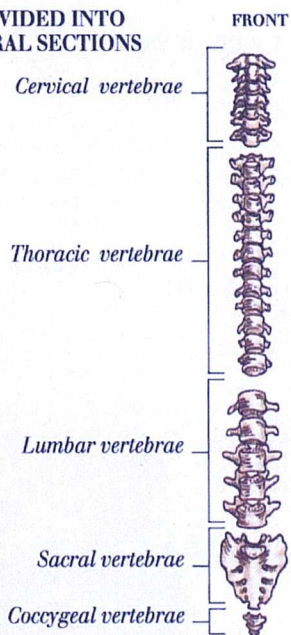
If you habitually have poor posture, you are much more likely to have back pain.

A lumbar roll or rolled towel can be used to assist good posture. Throughout the day it should become habit to be *aware* of the position of your spine and consciously correct it by rolling your pelvis forward to produce a curve in your low back. This will allow your muscles to work more efficiently.

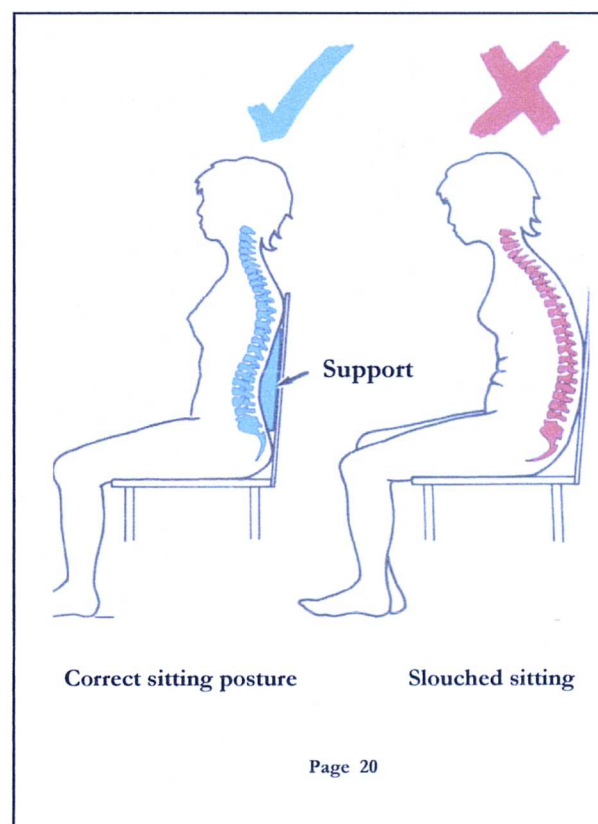
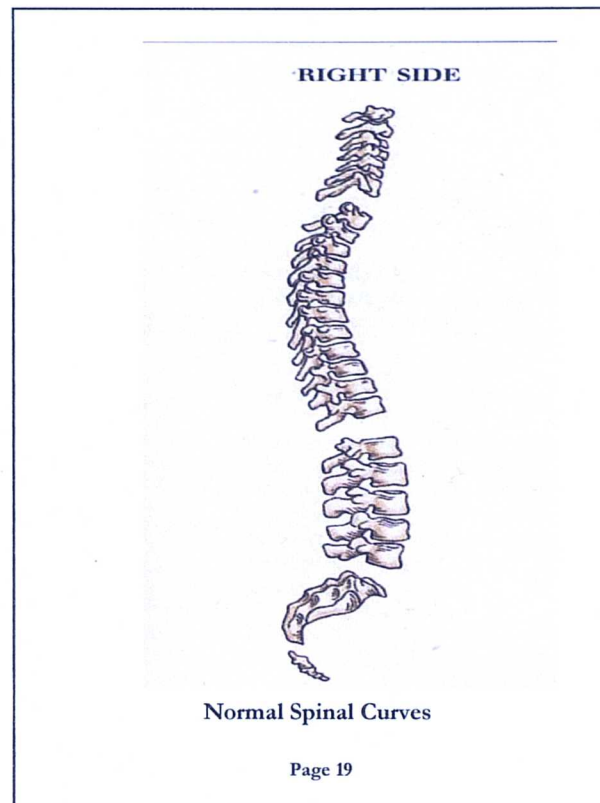
The exercise method that you will be taught to improve your back problem is based on good postural awareness and the continual correction of poor postures.

Page 17

SPINE DIVIDED INTO VERTEBRAL SECTIONS



Page 18



General Advice

Although research has shown that undertaking these specific exercises regularly can benefit people with low back pain, many other factors influence back pain. A number of activities/postures are known to aggravate back pain such as:

◆ **sustained flexion/bending** e.g. bending gardening for long periods



◆ **poor lifting and manual handling** e.g. not bending your knees when lifting/lifting above shoulder height

◆ **poor sitting and standing posture**
e.g. sitting in one position for long periods - whilst studying or travelling sitting without maintaining the curve in your low back



◆ **doing these exercises incorrectly!**
Although we know that doing these exercise correctly can help your back pain, performing them too hard or incorrectly can also make you worse. It is therefore essential that you perform them exactly as you have been advised, and if in doubt - ASK.

Page 21

Notes

Page 22

Additional Information

Prepared by Mandy Cairns in collaboration with:
Carolyn Richardson, Gwen Jull, Julie Hides & Suzanne Roll at;
The Department of Physiotherapy
University of Queensland
Brisbane
Australia

No Part of this booklet may be reproduced without
permission from MINDY CAIRNS at The Royal Orthopaedic Hospital

Appendix M: Indications for referral for diagnostic ultrasound

The purpose of using real-time ultrasound during the trial is to aid the patient in the visualisation and conscious activation of both transversus abdominis and multifidus muscles. However, additionally, during the pilot it will also serve as a teaching/learning tool for therapists prior to the main study.

When to refer?

- By the second session of using/teaching spinal stabilization concepts using your clinical judgement, the patients has no firm grasp of the concept of transversus abdominis, its position, anatomy or function

- Able to activate transversus abdominis in crook, prone or side lying but unable to progress position

- Good transversus abdominis activation but unable to progress to lumbar multifidus activation

Appendix N: Consent Form

The Royal Orthopaedic Hospital

NHS Trust

**A prospective randomized controlled trial into the effectiveness of spinal
stabilization exercises on recurrent low back pain**

Researcher: Mindy C Cairns MCSP

Consent Form

I(PRINT NAME)

Willingly consent to participate in the above study.

I understand that;

All data collected during this study will be anonymous and will remain confidential.

I am free at anytime to withdraw from this study.

Withdrawal from the study will in no way adversely affect any future treatment I may receive.

Signed Date

Witness Date.....

(Print Name)

(Consent form to be kept in records)

The Royal Orthopaedic Hospital

NHS Trust

Appendix O: Patient information sheet

In the physiotherapy department we are currently undertaking a research project. This information sheet is designed to explain the study and to help you decide if you want to participate.

(I) What is the study about?

The study is examining the effect of specific tummy exercises on low back pain in people who have had at least one previous episode of back pain.

(II) What will I have to do?

If you participate in the study you will be asked to complete some simple questionnaires prior to any treatment. You will then be assessed by a physiotherapist and receive one of two different treatments. As with any physiotherapy, you will be required to actively participate and undertake gentle exercises to help improve your back problem. A few patients, if necessary as part of the treatment, will be asked to attend The Royal Orthopaedic Hospital for a session using ultrasound to view the deep tummy muscles. This can help in understanding where the deep muscle layers are and is painless and has no adverse side-effects.

(III) What are the benefits?

Either treatment should help to speed your recovery and enable you to understand more about your back problem. All we are looking at, is which one is most effective so that we can continue to improve patient care.

(IV) What are the risks?

There are no specific known risks associated with the type of therapy we are investigating and the exercises should be completely pain-free. However, if you have not been moving very much because of your back problem you may experience short-term 'treatment soreness', similar to if you walk further than normal or undertake a new form of exercise.

(V) What are the alternatives?

Along with the specific exercises we are investigating, your therapist can use all their other skills to help improve your back problems so all options are available.

(VI) What if I do not want to take part?

You are completely free to choose not to take part and this decision will not influence any future treatment or care.

(VII) What happens to the information?

The information collected is all anonymous so you will not be able to be identified from it. The information will be examined and hopefully it will help us decide which treatment is most effective.

(VIII) Who else is taking part?

All patients who fit certain criteria are being invited to participate in this study and those consenting are being included.

(IX) What if something goes wrong?

It is very unlikely that anything will go wrong, however if you are unhappy about anything to do with the study you are at liberty to withdraw.

(X) What happens at the end of the study?

At the end of the treatment period you will be asked to complete some more questionnaires and then discharged. You will be contacted by post, 6 and 12 months after discharge to find out how your back problem is. Even if you do not complete the full course of treatment you will still be sent follow-up questionnaires. All responses to these questionnaires will be confidential and a pre-paid return envelope will be included so it will not cost you any money.

(XI) What if I have more questions or do not understand something?

It is **very** important that you fully understand everything to do with the study and what it will involve. If anything is unclear please do not hesitate to contact me on the number below.

(XII) What happens now if I decide to take part?

If you decide to take part you will need to complete a consent form stating that you are willing to participate. Certain conditions/illnesses mean that you can not join the study therefore you will need to read the list on the accompanying page to ensure you can be included. If this is all fine then you will be randomly allocated to one treatment group i.e. you have equal chance of receiving either treatment. You will not be advised which group you are in as this may influence the results of the study. Treatment will consist of between 6-12 sessions over a maximum of twelve weeks. Initial assessment will take approximately 1 hour with follow-up sessions of 30 minutes.

(XIII) What happens if I change my mind during the study?

Although we are keen that as many people as possible complete the study, if you decide you do not want to continue with the study you are free to withdraw at any time. Withdrawal from the study will not affect any future treatment. However, as already stated, even if you withdraw from the study you will still be sent follow-up questionnaires at 6 and 12-months.

Contact name and number

Mindy Cairns (Research Physiotherapist)

The Royal Orthopaedic Hospital

Bristol Road South

Birmingham, B31 2AP

0121 685 4120

Appendix P: Data collection sheet – Pilot Study

ID

Unit

DRAM Status:

Normal

At Risk

MZ

MSPQ

MZ

MSPQ

Gender:

Male/Female

Age:

Height (cm):

Weight (kg):

Low back pain classification

1

2

3

4

Duration of current episode..... Total Duration

No of pregnancies

No of caesareans

Smoker:

Yes

No

Employment status:

Employed: <i>At work</i>		Full-time carer	
Employed: <i>Off work</i>		Retired	
Unemployed: Because of LBP		Student	
Unemployed: Unrelated to LBP			
	Assessment- Date		D/C – Date
ODI			
RMDQ			
NRS			
U.K. SF-36			
SFMPQ	1-11	12-15	VAS PPI
	1-11	12-15	VAS PPI

ID: UNIT	CT/ SSSE
-------------	----------

Appendix Q: Data collection sheet - Main Trial

DRAM Status:

Normal

MZ =	MSPQ =
------	--------

At Risk

MZ =	MSPQ =
------	--------

Gender:

Male/Female

Age (years):

Height (cm):

Weight (kg):

QTF classification

1	2	3	4
---	---	---	---

Provisional Diagnosis

Duration of current episode..... Total Duration

No of pregnancies

No of caesareans

Smoker:

Yes	No
-----	----

Employment status:

Employed:	At work	Full-time carer	
	Off work		
Unemployed:	Because of LBP	Student	
	Unrelated to LBP		

Appendix R: Non-responders follow-up post-cards

Recurrent Low Back Pain Study—non-response

Between April 1999 and October 2000 you were treated for low back pain at The Royal Orthopaedic, Queen Elizabeth or Selly Oak Hospital Physiotherapy Departments. At that time, you agreed to participate in a research project about low back pain and as part of this I am looking at why you have not responded to postal follow-up.

Please take a few moments to tick one box below that best represents why you did not respond to postal follow-up. Any response will be treated in the strictest confidence and will help us when planning future research projects.

- ☐ Back pain resolved therefore questionnaires not relevant
- ☐ Back pain worsened therefore questionnaires not relevant
- ☐ Work or family commitments
- ☐ Amount of time questionnaires take to complete is too long
- ☐ Other (Please state)

Many thanks for your time

Mindy C Cairns (Research Physiotherapist)

Recurrent Low Back Pain Study—incomplete treatment

Between April 1999 and October 2000 you were treated for low back pain at The Royal Orthopaedic, Queen Elizabeth or Selly Oak Hospital Physiotherapy Departments, but did not complete a full course of treatment. At that time, you agreed to participate in a research project about low back pain and as part of this I am following-up why people didn't complete treatment.

Please take a few moments to tick one box below that best represents why you stopped attending for treatment. Any response will be treated in the strictest confidence.

- ☐ Back pain resolved
- ☐ Back pain worsened (unrelated to treatment)
- ☐ Treatment made back pain worse
- ☐ Work or family commitments
- ☐ Unsuitable appointment times
- ☐ Sought alternative treatment (Please state)
- ☐ Other (Please state)

Many thanks for your time **Mindy C Cairns (Research Physiotherapist)**

Appendix S: Outcome Sheet for main trial

	Base	D/C	6/12	1 Year
DRAM				
- MZ				
- MSPQ				
RMDQ				
ODI				
SFMPQ				
- 1-11 Sen				
- 12-15 Affect				
- VAS				
- PPI				
U.K. SF-36				
- PFC				
- MFC				
NRS (usual pain)				

Appendix T: Treatment Record

Rx No	Date	Rx administered	Additional Info	D/C
1				
2				
3				
4				
5				
6				
7				
8				
9				
10				
11				
12				
Total Rx =		6/12 Follow up date	1 year F/U date	

Appendix U: RCT Discharge Sheet

ID:	CT / SSSE
ROH SOH QEH	PT:

On discharge please complete this form regarding the patient's treatment and return to Mindy Cairns, Research Physiotherapist at The Royal Orthopaedic Hospital.

Modalities Used

(Number according to usage 1= most frequently used, 5 = least frequently used)

Manual Therapy (Including - HVT/Low velocity mobs and management of neural tissues)				
Active exercise				
Traction				
Ergonomic/postural advice				
Electrotherapy {please state}				
Total Number of Rxs				
Outcome of Treatment – Please circle one				
D/C	CTA	UTA	Withdrew	Other (Please state)

Appendix V: Clinicians' Questionnaires

You and your experience

How many years have you been qualified?

(please tick the appropriate box)

Less than 3 years	
4-6 years	
7-10 years	
11-14 years	
More than 14 years	

Name:

Site: ROH / SOH/ QEH

What is your current grade

How many years (post-graduation) have you worked in the musculoskeletal field ?

Less than 3 years	
4-6 years	
7-10 years	
11-14 years	
More than 14 years	

Please give brief details of any accredited post-graduate training you have completed.

.....

What training have you undertaken concerning muscle imbalance and dynamic stabilization of the lumbar spine? *(Please tick appropriate boxes)*

Under-graduate training	
'In-house' training (In-service etc)	
Week-end course	
Within external courses (MSc, MACP weekends)	
Long course (three weekends)	
Self taught	
Other (please specify)	

Thank you for completing this questionnaire.

Appendix W: Power calculations

Comparing two groups - factors for power calculation for main trial

$$n \geq 2 (SD/\delta)^2 (t_{\alpha} + t_{\beta})^2$$

n = number of replications

SD = estimated standard deviation

δ = the smallest difference that it is desired to detect (MCID)

α = significance level

$1-\beta$ = power

df = degrees of freedom associated with SD

t_{α} and t_{β} = values from a two tailed t-table with df degrees of freedom

Reference: Sokal, R. and Rolf, F. (1969) Biometry: the principles and practice of statistics in biological research. WH Freeman and Company, USA.

Assume:

SD= 6 (from Pilot trial)

δ = 5 (MCID for RMDQ)

α = 0.05 (5%)

Power = $1-\beta$ = 0.90 (= 90%)

β = $1- 0.90$ = 0.10

$n=32$, $df=62$, $t_{0.05}=2.00$, $t_{0.10}=1.30$

$$n \geq 2 (6/5)^2 \times (2+1.3)^2$$

$$=2(1.44 \times 10.89)=31.36$$

$N \geq 32$ in each group (no attrition)

Appendix X: Summary of data check

To ensure accuracy of raw data sheet scoring and inputting, a 20% random sample data check was undertaken by a second researcher who had not been involved in data collection. This consisted of 52 questionnaires, covering all follow-up points. SPSS syntax files had been set up to ensure that no values exceeding the score range could be inputted for any outcomes.

Checks of the RMDQ scoring (main outcome) highlighted two inputting errors. No discrepancies were found for the SFMPQ (PPI) or NRS. The U.K. SF36 was generally accurately scored and inputted. A problem was highlighted, that some of the SFMPQ used for patients recruited in the early phase of the trial did not contain a 10-cm VAS (only 8.2 cm). This affected 29 responses and these were scaled up to make them comparable to the other results i.e. $VAS \times 10/8.2$. The results are summarised in Table A 1. Table A 1 also displays the original means with the means recalculated after corrections made during the data check for baseline and 12-months follow-up. These time-points were chosen as they formed the basis of the main analyses.

Table A 1: Summary of random data check

Variable	Time point	Original data	Following data check And correction
Mean (SD)			
RMDQ	Base	8.40 (3.54)	8.53 (3.76)
	12	3.44 (3.78)	3.33 (3.81)
ODI	Base	26.27 (8.24)	26.13 (8.23)
	12	19.33 (12.61)	19.33 (12.61)*
MZ	Base	18.27 (6.72)	18.40 (6.82)
	12	19.11 (10.89)	19.33 (10.39)
MSPQ	Base	4.53 (3.16)	4.47 (3.25)
	12	4.89 (4.43)	4.67 (4.33)
SFMPQ	Base	10.31 (5.07)	10.33 (5.12)
	12	8.89 (5.33)	8.89 (5.33)*
Sensory	Base	8.13 (3.31)	8.53 (3.66)
	12	7.56 (4.42)	7.22 (4.27)
Affective	Base	2.00 (2.54)	1.80 (2.46)
	12	1.56 (1.74)	1.67 (1.66)
VAS	Base	3.65 (2.24)	3.71 (2.16)
	12	3.13 (1.75)	3.13 (1.75)*
* indicates no change in value with corrected input			
SFMPQ (PPI) and NRS not shown as no discrepancies reported			

Following the initial 20% check, a further check of the RMDQ scoring and inputting for all time points was undertaken. This revealed very few problems with either scoring or inputting and the results are summarised in Table A 2. Any discrepancies found were corrected in the main database.

Table A 2: Summary of full data check for main outcome

Patient ID	Group	Time point	Pre-check	Corrected following data check	Difference
19	SSSE	6	1	2	+1
24	CT	Base*	21	11	-10
26	SSSE	Base*	13	15	+2
36	CT	6	2	3	+1
55	SSSE	Base*	16	17	+1
78	CT	DC	12	10	-2
80	SSSE	6	2	3	+1
82	CT	12*	2	3	+1
95	SSSE	Base*	9	7	-2
97	CT	Base*	9	7	-2
Summary of changes					
			Pre check	Corrected	
	SSSE (n=47)	Baseline	10.34 (4.19)	10.36 (4.29)	0.02 (0.10)
	CT (n=50)		10.50 (4.39)	10.32 (4.10)	-0.18 (0.19)
	SSSE (n=33)	12-months	5.31 (5.10)	5.30 (5.02)	-0.01 (0.08)
	CT (n=35)		4.91 (5.47)	4.94 (5.46)	0.03 (-0.01)
* indicates error impacting on main analyses					
SSSE= Specific spinal stabilization exercise group					
CT=Conventional physiotherapy treatment group					

As can clearly be seen from Table A 2, the outcome of the full data check of the RMDQ scoring revealed few problems. These were not systematic, were comparable across both groups and resulted in tiny changes to the group means.

Appendix Y: Distributions of demographic and biographic variables at baseline

This appendix presents the distribution at baseline for demographic and biographic factors. Boxplots were constructed for all baseline variables and are shown in Figure A 2 through to Figure A 6. The box represents the interquartile range, which contains the middle 50% of values. The whiskers are lines that extend from the box to the highest and lowest values, excluding outliers, with the line across the box indicating the median. For all box plots, outliers are represented by the symbol O and extreme values by the symbol *. Outliers are defined as cases with values between 1.5 and 3 box lengths from the upper or lower edge of the box where the box length is the interquartile range. Extremes values are those more than 3 box lengths from the upper or lower edge of the box.

Figure A 2: Boxplot of age by group

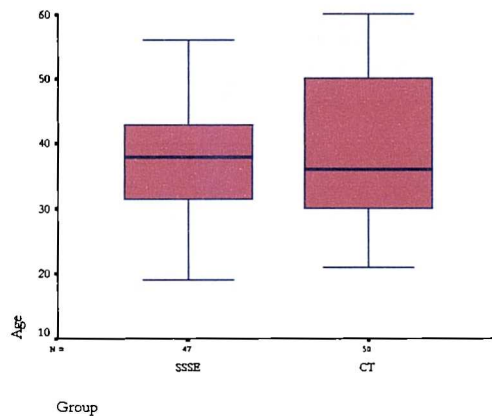
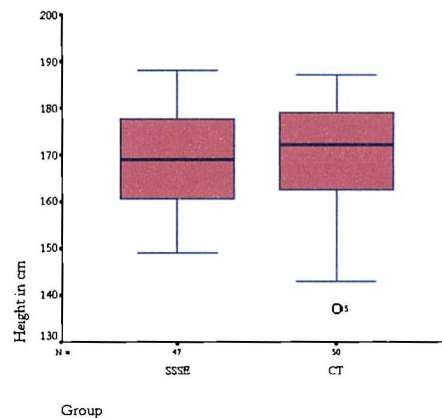
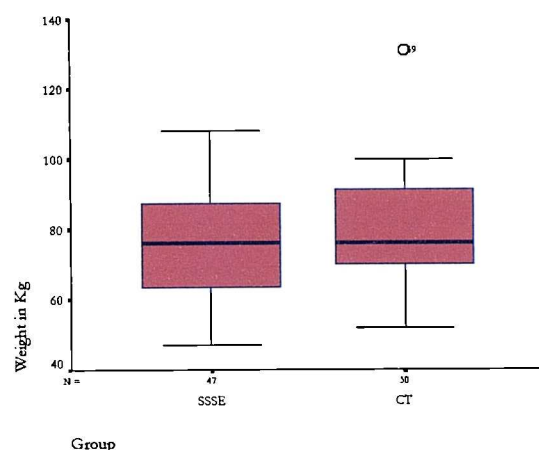
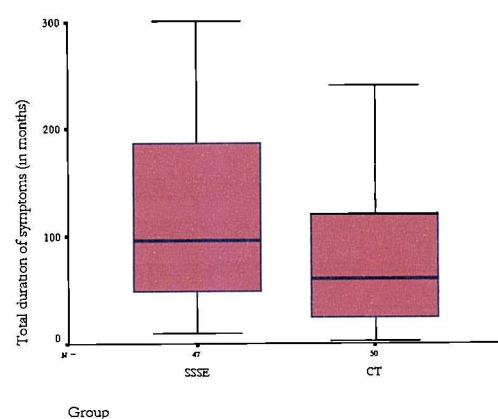
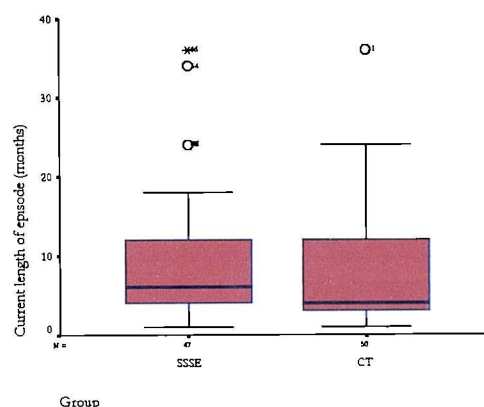


Figure A 3: Boxplot of height by group



As can be seen from Figure A 2 to Figure A 4 both groups displayed a symmetrical distribution for age, height and weight with few outliers or extreme values. The graphs suggest that the groups were comparable with respect to these demographic factors. Figure A 5 and Figure A 6 show the boxplots for distribution of total duration of symptoms and duration of current episode.

Figure A 4: Boxplot of weight by group**Figure A 5: Boxplot of total duration of LBP symptoms****Figure A 6: Boxplot of current duration of LBP symptoms**

As can be seen Figure A 5 and Figure A 6 revealed minor differences between the groups. Outliers are present in both groups, 2 in the SSSE group and 2 in the CT group, indicating that these patients reported experiencing LBP up to 3 times longer than the middle 50% of patients. Additionally, the SSSE group showed one extreme value. These figures clearly show wide variance in the duration of LBP reported. Additionally, recall of time periods and individual variations of definitions such as 'current' expose data of this nature necessarily to subjective interpretation.

Appendix Z: Socio-demographics baseline information

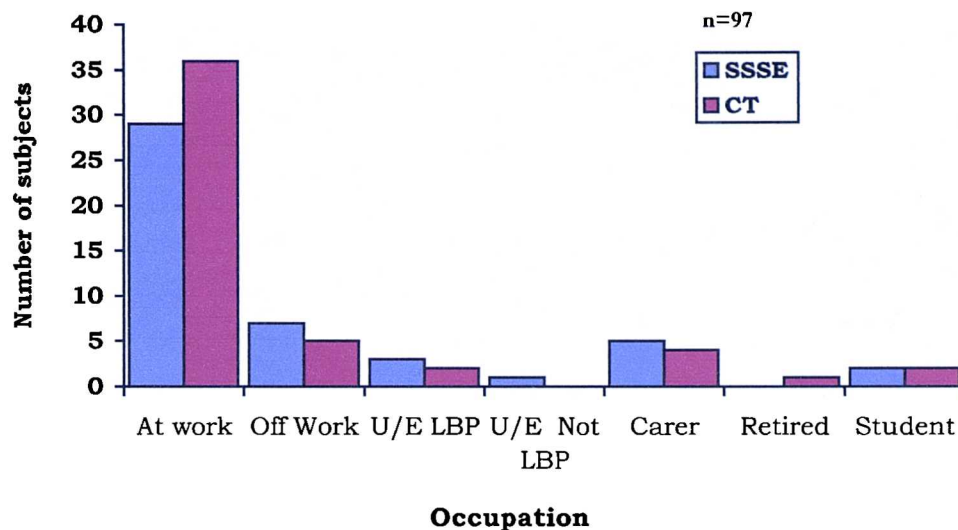
The employment status of all patients was recorded using the categories outlined on the data collection sheet (Appendix Q). Of the total 97 patients recruited, the majority 67% (65/97) continued working (either full or part time) despite their back pain, with only 12% (12/97) off work because of their back pain. A full summary is shown in Table A 3. Of the 6 patients unemployed at the start of the trial, 5 cited their back pain as the cause. A summary of the distribution of employment is displayed in Figure A 7.

Table A 3: Summary of employment status in each group at baseline

		SSSE (n=47)	CT (n=50)	Total (n=97)
		n(%)		
Employed	<i>At work</i>	29 (62)	36 (72)	65 (67)
	<i>Off work</i>	7 (15)	5 (10)	12 (12)
Unemployed	<i>Back</i>	3 (6)	2 (4)	5 (5)
	<i>Unrelated</i>	1 (2)	0 (0)	1 (2)
Full time carer		5 (11)	4 (8)	9 (9)
Retired		0 (0)	1 (2)	1 (2)
Student		2 (4)	2 (4)	4 (4)

As can be seen clearly from Table A 3 and Figure A 7 both the employment status and number of patients not working because of their back pain is comparable across the two groups.

Figure A 7: Summary of employment status for each group

**Key;**

Off work=In paid employment but off work on entrance to the trial because of current episode of LBP

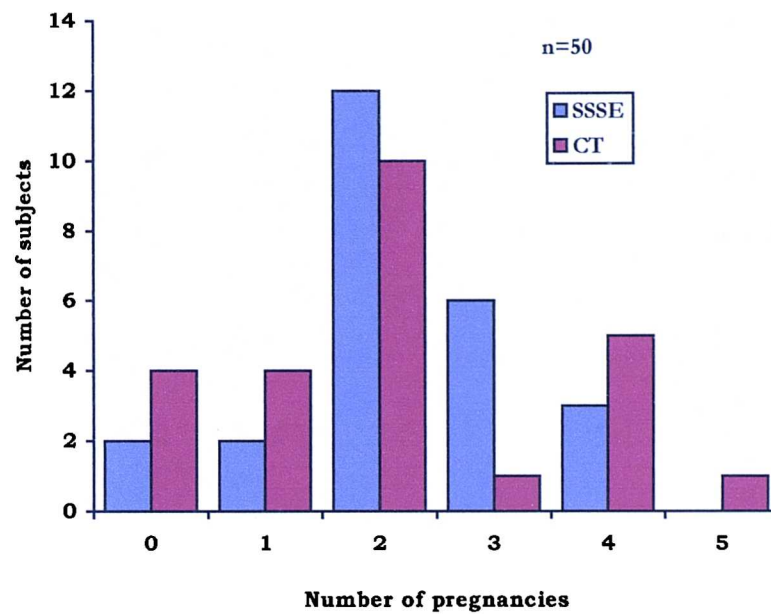
U/E LBP=Unemployed because of LBP

U/E Not LBP=Unemployed unrelated to LBP

Smoking has been linked to the onset and severity of LBP; therefore information was collected regarding current smoking status. Of the total, 33 patients currently smoked, with 16 in the SSSE group and 17 in the CT group, showing comparability.

As evidence exists to suggest that pregnancy and caesarean sections (CS) can have a profound impact of the deep abdominal muscles, information was collected (from females) regarding the number of previous pregnancies.

Figure A 8 shows the spread and number of previous pregnancies across both groups. Of the 50 female patients in the trial 44, (88%) had at least 1 previous pregnancy. Of those 44, 7 had undergone at least one CS. In the two treatment groups, 4 (17.4%) of the SSSE had undergone one CS and 2 (8.7%) had undergoing two CS. In the CT group, 1 (4.8%) patient had undergone two CS.

Figure A 8: Summary of number of previous pregnancies

In summary, the two groups were comparable with respect to the socio-demographic information collected at baseline which might be considered to have a potential influence on the outcome of treatment.

Appendix AA: Baseline Distributions - Study variables

This appendix presents the distribution at baseline for the study variables. The box represents the interquartile range, which contains 50% of values. The whiskers are lines that extend from the box to the highest and lowest values, excluding outliers with the line across the box indicating the median. For all box plots, outliers are represented by the symbol O and extreme values by the symbol *. Outliers are defined as cases with values between 1.5 and 3 box lengths from the upper or lower edge of the box where the box length is the interquartile range. Extremes values are those more than 3 box lengths from the upper or lower edge of the box.

Figure A 9: Boxplot of RMDQ scores at baseline

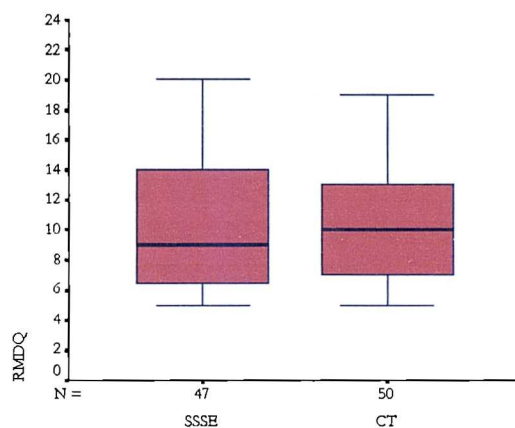


Figure A 10: Boxplot of ODI scores at baseline

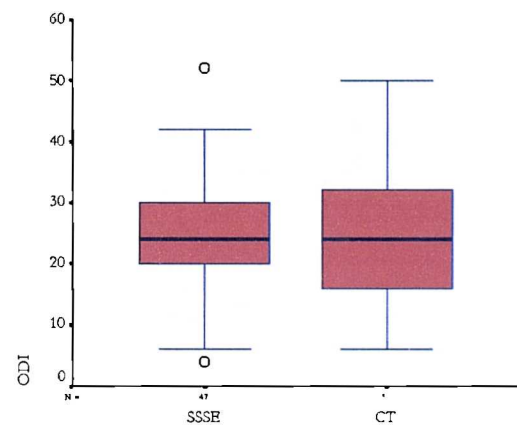


Figure A 11: Boxplot of SFMPQ at baseline

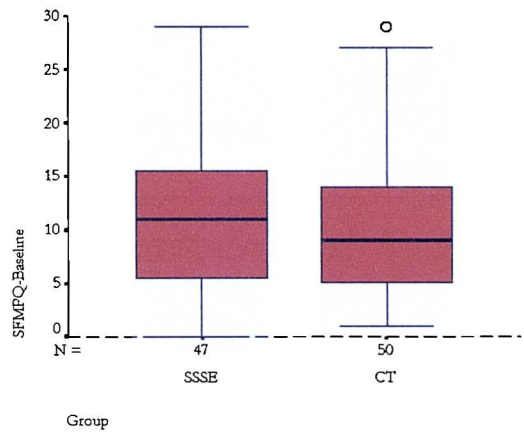


Figure A 12: Boxplot of SFMPQ (Sensory component) at baseline

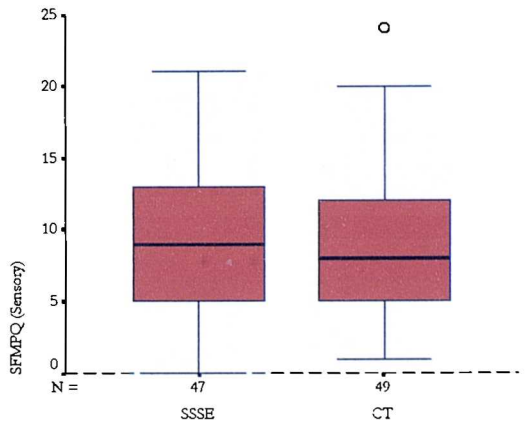


Figure A 13: Boxplot of SFMPQ (Affective component) at baseline

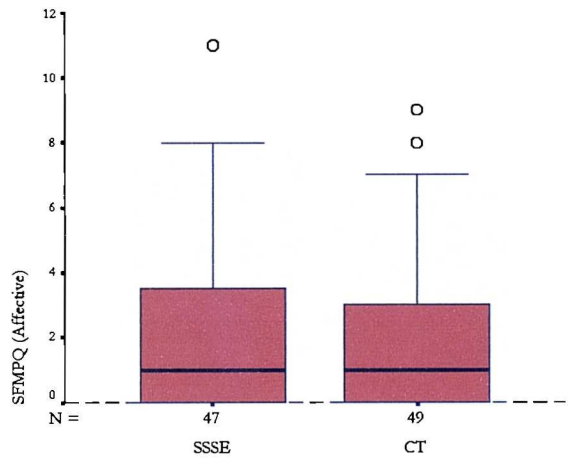


Figure A 14: Boxplot of SFMPQ (PPI) at baseline

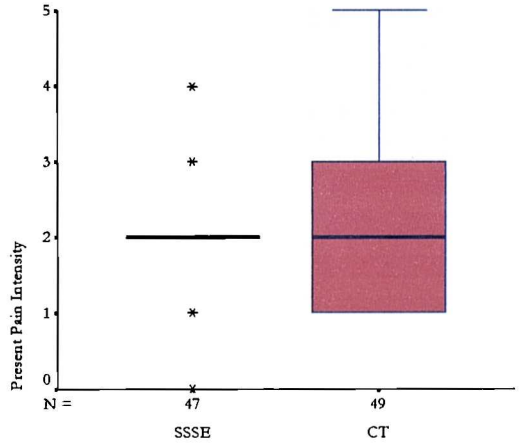
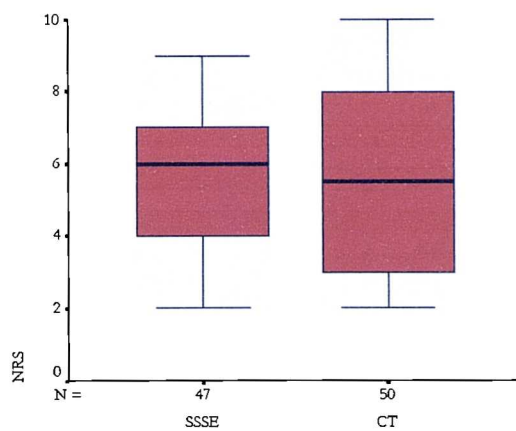
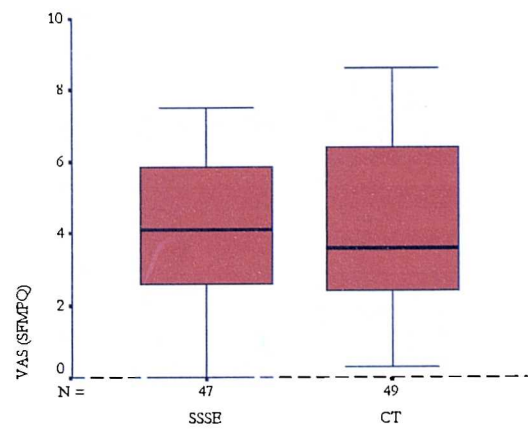
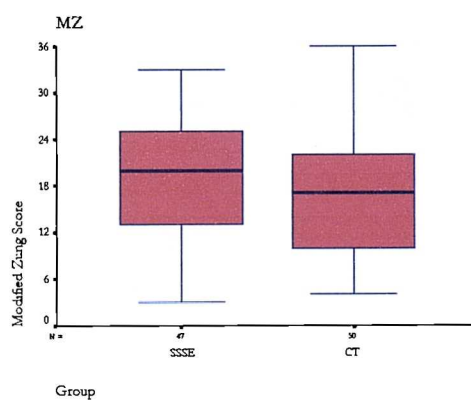
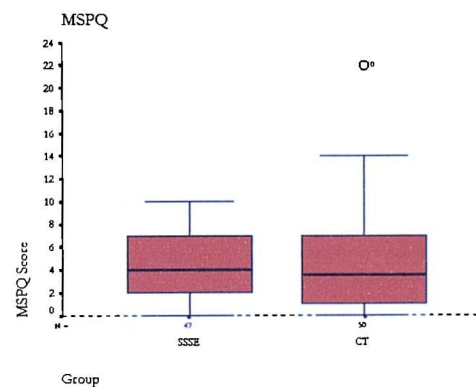


Figure A 15: Boxplot of NRS scores at baseline**Figure A 16: Boxplot of VAS scores at baseline**

The distribution for the scores for the MSPQ shown in Figure A 18 indicate some asymmetry because of the presence of 2 patients with high MSPQ scores. Due to a protocol violation, 2 patients classified as Distressed Depressed (DD) were entered into the trial, both randomised into the CT group. When the scores for these 2 patients are excluded, the distribution is symmetrical.

Figure A 17: Boxplot of MZ scores at baseline**Figure A 18: Boxplot of MSPQ scores at baseline**

Appendix BB: Demographics of treatment completers versus treatment non-completers

Demographic and biographic information were compared across patients who were discharged from treatment by joint agreement between themselves and the treating clinicians (treatment completers) and those who ceased to attend (treatment non-completers). This information is presented in Table A 4. As can be seen from Table A 4, in general, both the treatment completers and treatment non-completer groups were comparable for the demographic and biographic variables measured. It should, however, be remembered that the large differences in the sample in each group means that direct comparison may be inappropriate.

Table A 4: Demographics of treatment completers versus treatment non-completers

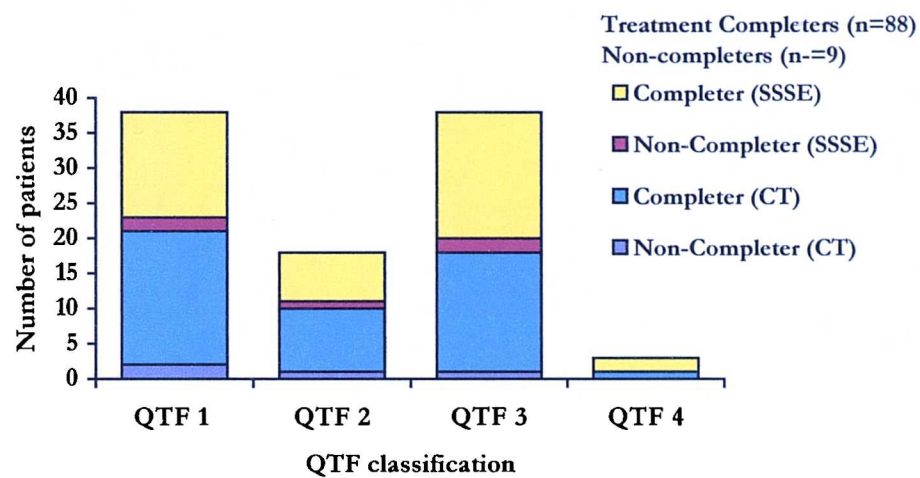
Values are mean (SD) unless otherwise stated

	Treatment completers (n=88)		Treatment non-completers (n=9)	
	SSSE (n=42)	CT (n=46)	SSSE (n=5)	CT (n=4)
Age (years)	38.0 (9.7)	40.6 (11.5)	33.2 (5.6)	32.5 (4.8)
Height (cm)	168.7 (10.1)	170.1 (10.8)	171.6 (14.7)	173.8 (9.8)
Weight (Kg)	75.9 (14.6)	78.0 (15.5)	71.3 (21.8)	81.1 (13.1)
Current duration*	9.1 (8.0)	7.6 (7.5)	14.6 (12.3)	10.5 (10.0)
Total duration*	121.0 (85.4)	81.5 (68.2)	94.8 (101.3)	92.0 (87.8)
% male/%female	45/55	48/52	60/40	75/25
*duration in months				
Shading indicates variable where differences across group are apparent				

Table A 4 shows that as with the baseline comparison across groups, the total duration of symptoms reported by the SSSE treatment completers was higher than the CT group treatment completers and the current duration for the SSSE treatment non-completers higher than the CT treatment non-completers.

As evidence exists that laterality of symptoms can influence the prognosis and outcome of treatment, the QTF classification for each patient was collected at baseline and a summary of completers and non-completers by groups is shown in Figure A 19. As can be seen, allowing for the differences in sample sizes, the four groups appear comparable.

Figure A 19: QTF classification distributions for completers and non-completers



Appendix CC: Baseline demographics and biographics of responders and non-responders at 12-month follow-up

In order to assess whether there were any differences between patients who responded and those not responding at 12-month follow-up, both demographic and study variables at baseline were compared. Table A 5 summarises the findings and Figure A 20 shows the QTF distributions between responders and non-responders.

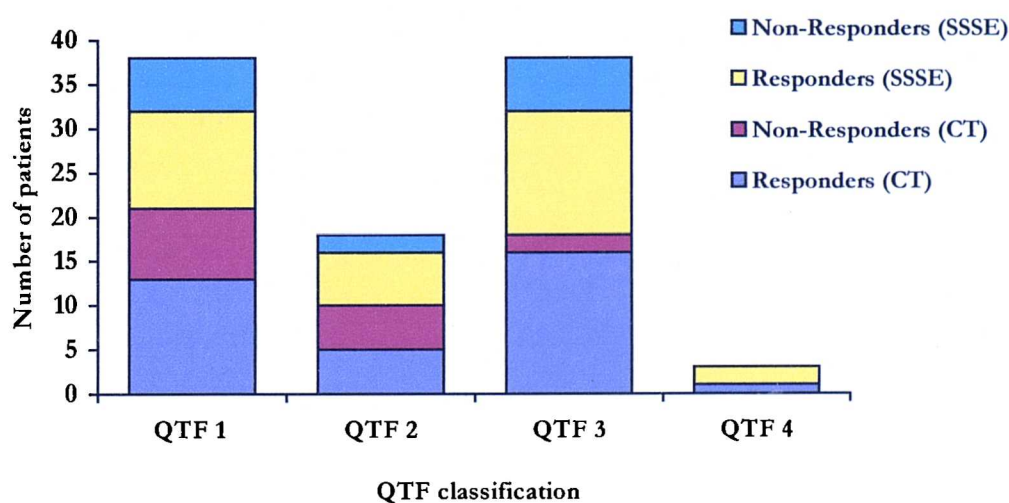
Table A 5: Demographics for responders and non-responders at 12-month follow-up

	Responders (n=68)		Non-responders (n=29)	
	SSSE (n=33)	CT (n=35)	SSSE (n=14)	CT (n=15)
Age (years)	36.82 (10.3)	42.2 (11.9)	39.1 (7.8)	34.5 (7.6)
Height (cm)	167.24 (10.4)	168.0 (11.0)	173.3 (10.5)	176.0 (7.4)
Weight (Kg)	74.39 (14.4)	77.1 (16.9)	77.9 (17.39)	81.1 (10.2)
Current duration*	8.6 (6.7)	7.1 (6.4)	12.1 (11.7)	9.7 (9.9)
Total duration*	126.7 (87.3)	88.5 (69.1)	98.2 (83.6)	67.9 (69.0)
n (%)				
male	11(33)	12(34)	11 (79)	13 (87)
female	22 (67)	23 (66)	3 (21)	2(13)

* both durations in months

Figures are mean (SD) unless otherwise stated

Figure A 20: QTF classifications for responders and non-responders at 12-month follow-up

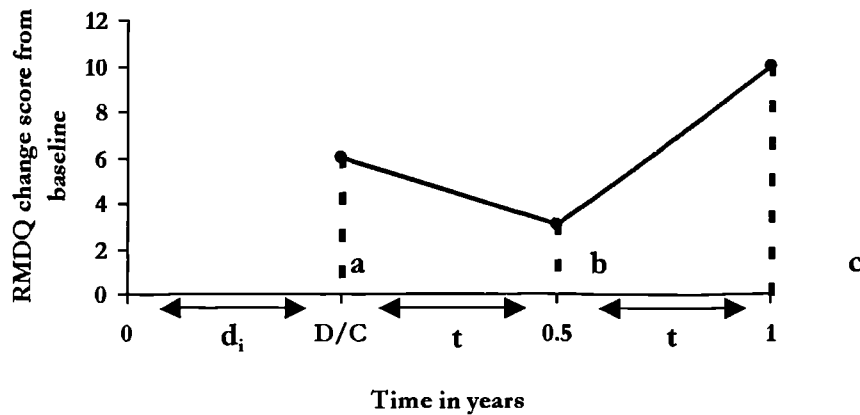


Appendix DD: Calculation of Area Under the Curve (AUC)

Differences from baseline were used to reduce the effects of varying treatment times and different baseline RMDQ scores on the AUC calculation. This allowed a curve to be constructed to illustrate change following treatment as shown in Figure A 21.

AUC was calculated for patients with 12-month follow-up. Where interim follow-up data were missing, values were estimated using linear interpolation between the two known values bordering the missing value. This is illustrated in Figure A 22. Four separate calculations were needed for the AUC i.e. for patients with no missing data, those with discharge only missing, those with 6-months only missing and those with both discharge and 6-months missing.

Figure A 21: Diagrammatic representation of AUC with all known values

**Key to Figure A 21**

Where d_i is the treatment time which varies for individuals

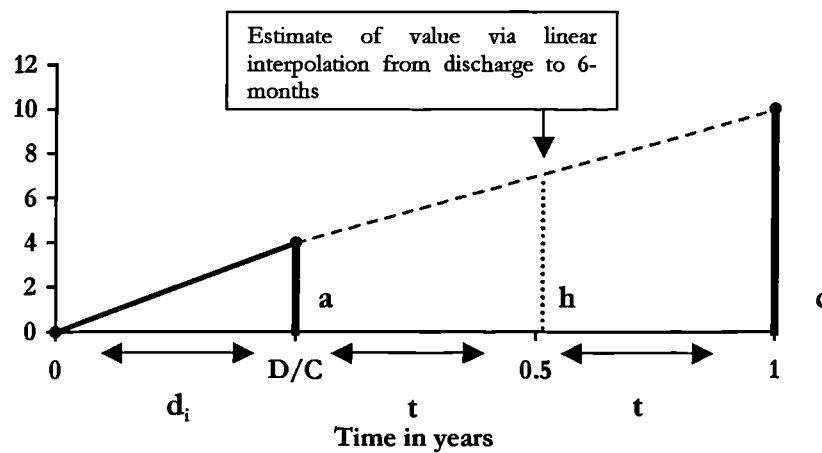
t is time between follow-up points and equals 6 months (0.5 years)

Using the formula for area of a trapezium:

$$AUC = 0.5(a+b)t + 0.5(b+c)t$$

$$AUC = 0.25(a+2b+c)$$

Figure A 22: Diagrammatic representation of AUC with interpolated values

**Key to Figure A 22**

Where d_i is the treatment time which varies for individuals

t is time between follow-up points and equals 6 months (0.5 years)

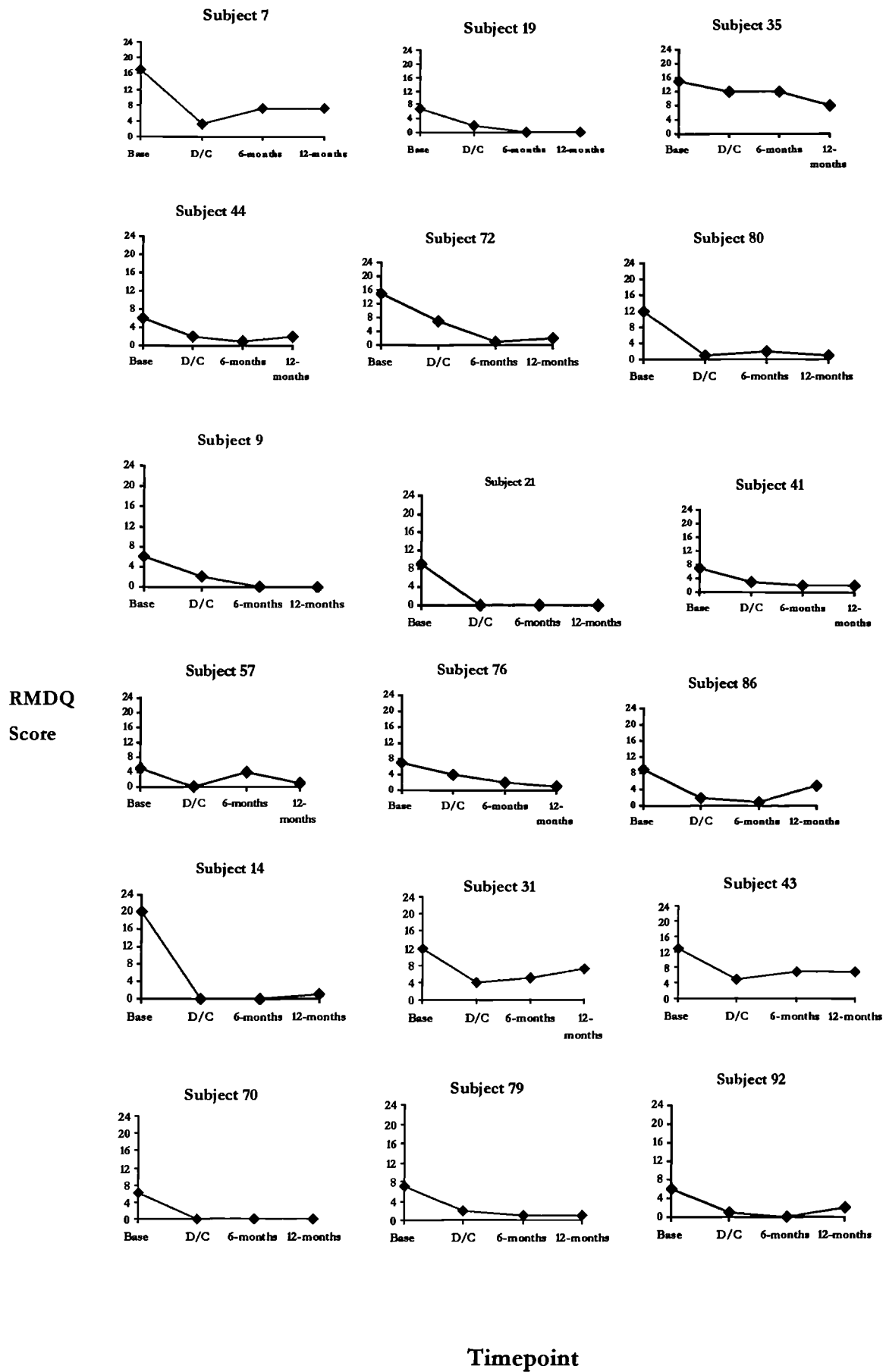
Appendix EE: Individual time-series curves for each group

This appendix presents the individual time-series curves for all patients in the SSSE group classified as ‘sustained improvers’. This was defined as patients who fulfilled all the criteria outlined in Table A 6.

Table A 6: Criteria for SSSE group ‘improvers’

Data for all follow-up points present
RMDQ change score ≥ 3 points
RMDQ score at discharge \leq RMDQ score at baseline
RMDQ score at 6-months \leq RMDQ score at baseline

In total, 18 patients fulfilled the above criteria and were identified as ‘sustained improvers’ in the SSSE group. In all graphs, the x-axis represents the follow-up time-point and the y-axis the RMDQ score (0-minimum and 24 maximum).



Appendix FF: Supplementary data for U.K. SF-36

The distribution for the individual item components for each scale is shown in Figure A 23 through to Figure A 28. In all box-plots, the upper bar represents the upper quartile, the lower bar, the lower quartile and the height of the box the interquartile (indicating variability) with the horizontal line representing the median value (Sim and Wright, 2000). The means for the baseline and 12-month follow-up are shown in Table A 7 for each group with the normal U.K values in the final column. However, looking at the boxplots, means and standard deviations, the items were unsuitable summary statistics for many of the scales.

Table A 7: Means (SD) for U.K. SF-36 at Baseline and 12-month

Scale	Baseline		12-months		OHLS*
	SSSE	CT	SSSE	CT	
	(n=33)	(n=35)	(n=33)	(n=35)	norm
Physical functioning	60.0 (18.4)	60.59 (23.0)	74.4 (21.7)	71.2 (24.1)	88.4 (18.0)
Role Physical	36.4 (41.0)	47.06 (42.5)	64.7 (40.9)	71.3 (40.4)	85.8 (29.9)
Pain	43.1 (21.1)	43.77 (19.0)	62.1 (23.3)	64.1 (23.6)	81.5 (21.7)
General Health perception	61.8 (24.2)	69.38 (16.9)	64.1 (22.8)	62.8 (18.9)	73.5 (19.9)
Energy	44.8 (22.1)	55.0 (20.6)	51.4 (25.7)	55.9 (21.7)	61.1 (19.7)
Social Functioning	69.5 (20.3)	69.7 (21.4)	70.5 (22.7)	69.4 (21.2)	88.0 (19.6)
Role Emotional	72.9 (34.33)	72.7 (40.5)	73.6 (39.2)	72.6 (40.6)	82.9 (31.8)
Mental Health	68.4 (15.3)	74.8 (18.2)	70.6 (17.6)	72.8 (22.7)	73.8 (17.2)

* Oxford Healthy Life Survey (1991/2)

Grey shading indicates scales contributing towards calculation of PCS

Red shading indicates scales contributing towards calculation of MCS

Figure A 29 and Figure A 30 show the norm based scoring for baseline and 12 months for each of the 8 U.K. SF-36 scales. This illustrates scores, which can be compared to a population norm where average is 50 and SD is 10.

Figure A 23: Distribution of individual items of the Physical Functioning scale

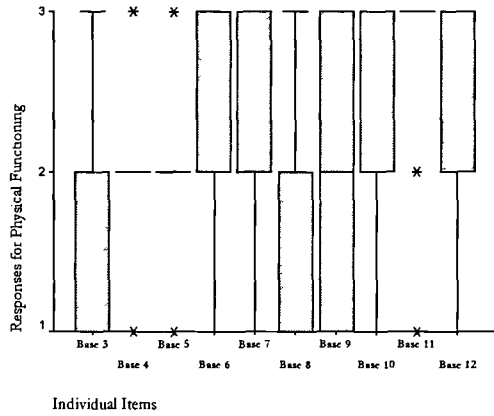


Figure A 24: Distribution of individual items Social Functioning scale

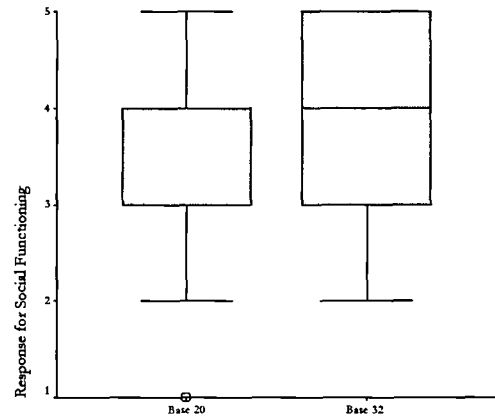


Figure A 25: Distribution of individual items Mental Health scale

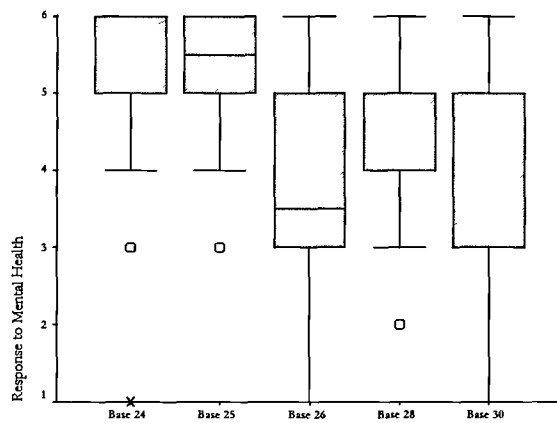


Figure A 26: Distribution of individual items Energy/Vitality scale

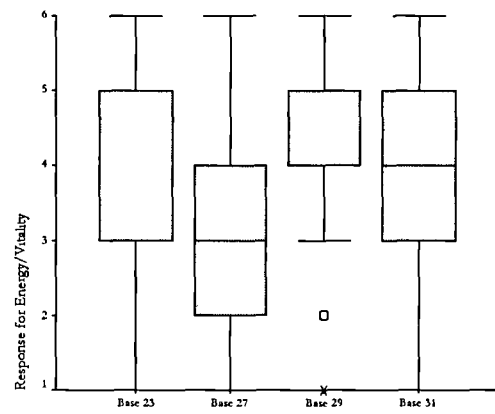


Figure A 27: Distribution of individual items Pain scale

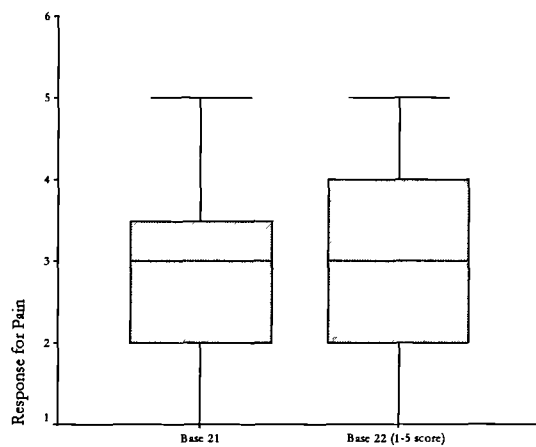


Figure A 28: Distribution of individual items General Health scale

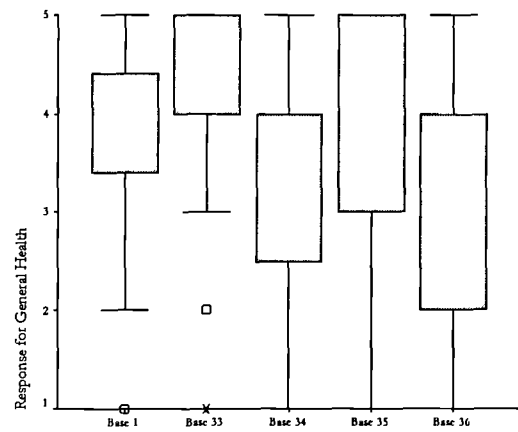


Figure A 29: Scales comprising the PCS at baseline and 12 months follow-up

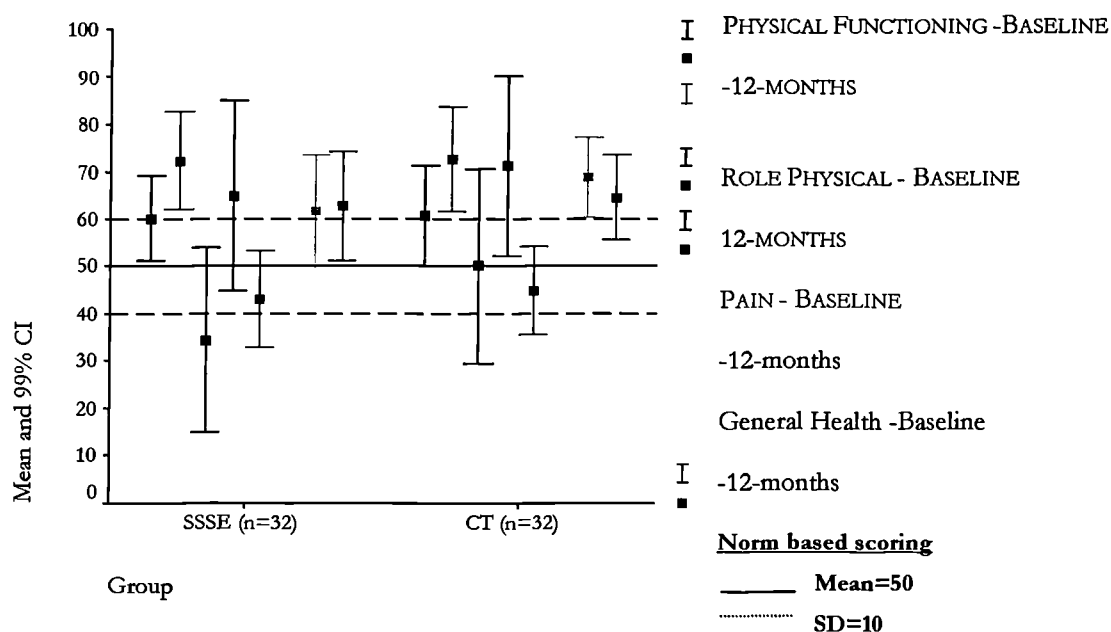
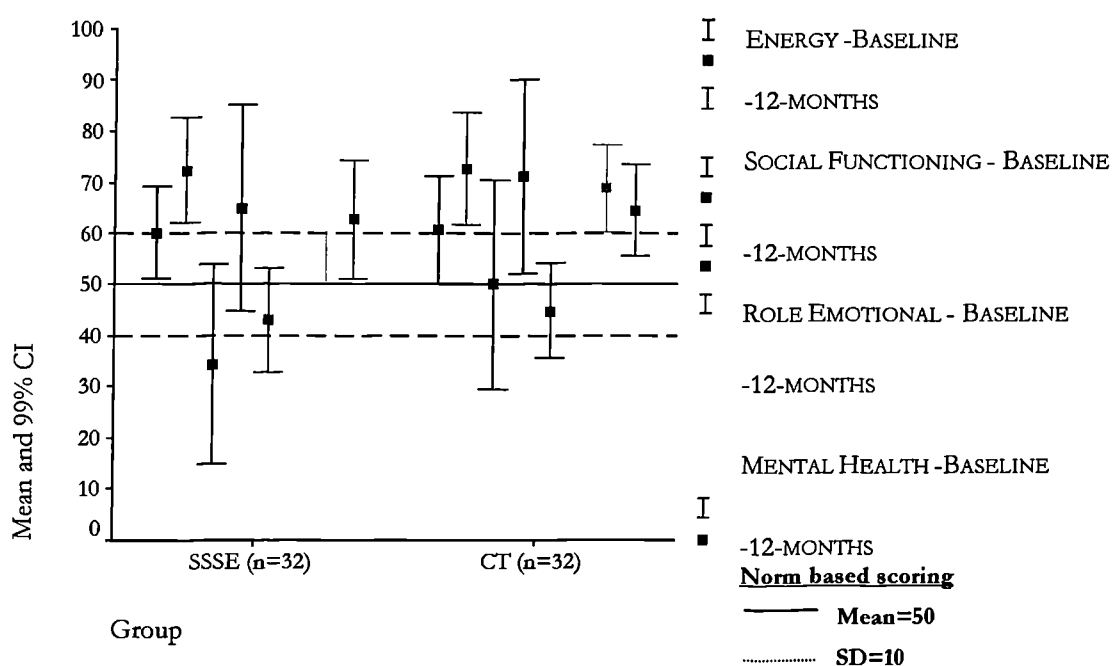


Figure A 30: Scales comprising the MCS at baseline and 12 months follow-up



Question 2 in the U.K. SF-36 inquires about change in health status compared to one year ago. Table A 8 presents the response and percentage responses for Question 2 at baseline and 12-months for responders. As can clearly be seen, neither intervention had a positive effect on self-reported change in health, with nearly 50% of patients in both group reporting a deterioration their health compared to one year ago.

Table A 8: Reported change in Health Status (U.K. SF-36)

RESPONSE	SSSE N (%)		CT N (%)	
	BASELINE	12-MONTH	BASELINE	12-MONTH
MUCH BETTER	1 (3)	1 (3)	1 (3)	0 (0)
SOMEWHAT BETTER	16 (49)	6 (18)	12 (34)	5 (14)
ABOUT THE SAME	14 (42)	10 (30)	16 (46)	13 (37)
SOMEWHAT WORSE	1 (3)	10 (30)	5 (14)	11 (31)
MUCH WORSE	1 (3)	6 (18)	1 (3)	6 (17)
TOTAL	33	33	35	35

Appendix GG: Summary of studies using the DRAM to assess distress in various LBP populations

Study	Setting	Patient Type	N=	Normal	At risk	DD	DS*	Total Distressed N %
All figures n (%) where available								
Primary Care								
(Burton et al., 1995)*	Osteopathy Clinic	New episode of LBP	212	100 (47%)	70 (33%)	22 (10%)	20 (9%)	42 19%
(Waxman et al., 1998)	Patient's Home	Episode of LBP in the previous year	389	169 (43%)	150 (39%)	41 (11%)	29 (7%)	70 18%
Secondary Care								
Main et al. (1992)	Orthopaedic Clinic	Mixed primary & tertiary referrals	449	30-39%	43-50%	7-12%	7-17%	NR 14*-29%#
	Pain Clinic	Patients selected for	118	10-22%	38-54%	20-38%	7-15%	NR 27*-53%#
	Surgical	surgery	NR	23%	65%	11%	2%	NR 13%
	Acute & Chronic Spinal Clinic	Acute & chronic LBP	125	36 (29%)	54 (43%)	25 (20%)	10 (8%)	35 28%
	Rehabilitation Programme	Chronic LBP (> 3 months)	261	37 (14%)	117 (45%)	****107 (41%)****		107 41%
Williams et al. (1995)	Physical Department	Therapy Patients referred with LBP +/- Leg pain for PT by their GP (No exclusions)	160	38 (24%)	71 (44%)	42 (26%)	9 (6%)	51 32%
Hope & Forshaw (1999)	Spinal Department	Outpatient LBP	521	99 (19%)	224 (43%)	155 (30%)	43 (8%)	198 38%
(Sell, 1999a)	Spinal Department	Outpatient LBP	521	99 (19%)	224 (43%)	155 (30%)	43 (8%)	198 38%
(Sell, 1999b)	Spinal Department	Outpatient Spinal surgical candidates	96	15 (16%)	53 (55%)	20 (21%)	8 (8%)	28 29%

(Hobby et al., 2001)	Spinal Clinic	Pre-operative surgical patients with clinical and MRI evidence of P1D	66	11 (17%)	24 (36%)	***31 (47%)***	31	47%
(Greenfield, 2001)	Spinal Clinic	Patients with leg pain (nerve root distribution) with failed conservative treatment	175	30 (17%)	90 (51%)	***55 (31%)***		31%
(Hothersall et al., 2001)	Spinal Clinic	Chronic LBP (pre endoscopic spinal surgery)	185	38 (20%)	79 (43%)	***68 (37%)***		37%
Laboratory								
(Mannion et al., 1996)	Non-Clinical Setting	Healthy patients	382	307 (80%)	71 (19%)	N/A	4 (1%)	4
								1%

DD=Distressed depressed, DS=Distressed somatic

* Baseline data (personal communication)

Minimum % calculated from published data

Maximum % calculated from published data

NR= not reported in published sources

GP = General Practitioner

PT = Physical Therapy

P1D=Prolapsed Intervertebral Disc

MRI=Magnetic Resonance Imaging

Appendix HH: Methodological Quality Criteria

This appendix summarises the criteria for judging methodological quality used by Koes et al. (1995) and on which the criteria discussed in this chapter are based.

Study Population (35 points)	
A	Description of inclusion/exclusion criteria (1 point). Restriction to a homogenous group (1 point)
B	Comparability of relevant baseline characteristics: duration of complaint, value of outcome measure, age, recurrences and radiating complaints (1 point each)
C	Randomisation procedure described (2 points) and excludes bias (2 points)
D	Drop outs described for each study group separately, including reasons for withdrawal (3 points)
E	Loss to follow-up = <20% (2 points), <10% (3 points)
F	Number of patients in smallest group, immediately after randomization: >50 (8 points), >100 (9 points)
Intervention (25 points)	
G	Physiotherapy treatment protocol established and described (5 points). All reference treatments put in a protocol and described (5 points)
H	PRAGMATIC STUDY: Comparison with other treatment modalities (5 points)
I	Co-intervention avoided in the design of the study (5 points)
J	Placebo controlled: Comparison with placebo therapy (5 points)
Measurement of effect (30 points)	
K	Patients blinded, placebo-controlled: attempt at blinding (3 points). Blinding evaluated and fully successful (2 points) Pragmatic trials: Patients fully naïve (3 points), time restriction (2 points). Naiveness evaluated and fully successful (2 points)
L	Use of pain, global measure of improvement, functional status, spinal mobility and medical consumption (2 points each)
M	Each blinded measurement under (L) – 2 points each
N	Measurement during or just after treatment (3 points) and after 6 months or longer (2 points)
Data presentation and analysis (10 points)	
O	Loss to follow-up < 10% -ITT analysis, >10% - ITT and worst case for missing data (5 points)
P	Frequencies of most important outcome presented for each treatment group. (Semi) continuous variables- mean (or medians) and SE of percentiles (5 points)
ITT=Intention-to-treat analysis, SE=standard error	

