ASSESSMENT OF STIFFNESS IN HUMAN BACKS

Jane Latimer, BAppSc(Phty), GradDipAppSc (ManipPhty)

A Thesis submitted in fulfilment of the requirements for the degree of
Doctor of Philosophy

School of Physiotherapy,
Faculty of Health Sciences
The University of Sydney

March 1997
Copyright in relation to this Thesis

Under the Copyright Act 1968 (several provisions of which are referred to below), this material must be used only under the normal conditions of scholarly fair dealing for the purposes of research, criticism or review. In particular no results or conclusions should be extracted from it, nor should it be copied or closely paraphrased in whole or in part without the written consent of the author. Proper written acknowledgement should be made for any assistance obtained from this material.

Under Section 35 (2) of the Copyright Act 1968 'the author of a literary, dramatic, musical or artistic work is the owner of any copyright subsisting in the work'. By virtue of Section 32 (1) copyright 'subsists in an original literary, dramatic, musical or artistic work that is unpublished' land of which the author was an Australian citizen, an Australian protected person or a person resident in Australia.

The Act, by Section 36 (1) provides: 'Subject to this Act, the copyright in a literary, dramatic, musical or artistic work is infringed by a person who, not being the owner of the copyright and without the licence of the owner of the copyright, does in Australia, or authorises the doing in Australia of, any act comprised in the copyright'.

Section 31 (1) (a) (i) provides that copyright includes the exclusive right to 'reproduce the work in a material form'. Thus, copyright is infringed by a person who, not being the owner of the copyright, reproduces or authorises the reproduction of a work, or of more than a reasonable part of the work, in a material form, unless the reproduction is a 'fair dealing' with the work 'for the purpose of research or study' as further defined in Sections 40 and 41 of the Act.

Section 51 (2) provides that "Where a manuscript, or a copy, of material of other similar literary work that has not been published is kept in a library of a university or other similar institution or in an archives, the copyright in the material or other work is not infringed by the making of a copy of the material or other work by or on behalf of the officer in charge of the library or archives if the copy is supplied to a person who satisfies an authorized officer of the library or archives that he requires the copy for the purpose of research or study'.

* Thesis' includes 'treatise', 'dissertation' and other similar productions.
I, Jane Latimer, hereby declare that this submission is my own work and that it contains no material previously published or written by another person except where acknowledged in the text. Nor does it contain material which has been accepted for the award of another degree.

I, Jane Latimer, understand that if I am awarded a higher degree for my thesis entitled ‘Assessment of stiffness in human backs’ being lodged herewith for examination, the thesis will be lodged in the University Library and be available immediately for use. I agree that the University Librarian (or in the case of a department, the Head of Department) may supply a photocopy or microform of the thesis to an individual for research or study or to a library.

Signed.................................................. Date...............
SUPERVISORS’ STATEMENT

As supervisors of Jane Latimer’s doctoral work, we certify that we consider her thesis ‘Assessment of stiffness in human backs’ to be suitable for examination.

..................................Date........

..................................Date........

Dr Roger Adams
Chief Supervisor
Senior Lecturer
School of Physiotherapy
Faculty of Health Sciences
The University of Sydney

Michael Lee
Associate Supervisor
Lecturer
School of Exercise and Sports Sc.
Faculty of Health Sciences
The University of Sydney
# TABLE OF CONTENTS

ACKNOWLEDGEMENTS ........................................................................................................ viii
PREFACE .................................................................................................................................. ix
PUBLICATIONS .................................................................................................................. xi
LIST OF FIGURES ............................................................................................................... xiii
LIST OF TABLES ................................................................................................................... xv
ABSTRACT ........................................................................................................................... xvi

CHAPTER 1: MANUAL CLINICAL ASSESSMENT OF SPINAL POSTERIOANTERIOR STIFFNESS ............................................................. 1

The problem of low back pain: ......................................................................................... 2
   Incidence .......................................................................................................................... 2
   Cost implications ........................................................................................................... 3
   The diagnostic dilemma ............................................................................................... 4
   Historical overview of the practice of spinal manipulative therapy ....................... 5
   Physiotherapists use of spinal manipulative therapy to examine patients with spinal pain .......................................................................................................................... 11
   Measurement of stiffness
      (i) peripheral joints .................................................................................................. 19
      (ii) lumbar PA stiffness ............................................................................................. 22

CHAPTER 2: INSTRUMENTED MEASUREMENT OF SPINAL POSTERIOANTERIOR STIFFNESS ......................................................... 28

Abstract ............................................................................................................................. 29
Introduction ........................................................................................................................ 30
Development of the portable stiffness assessment device ............................................. 32
Calibration .......................................................................................................................... 39
Data analysis ..................................................................................................................... 40
Safety features of the device ......................................................................................... 45
Contraindications to testing ......................................................................................... 47
Discussion .......................................................................................................................... 47
Appendix A: Circuit diagram of the stiffness assessment device .................................. 50
CHAPTER 5: TRAINING WITH IMMEDIATE FEEDBACK AND PHYSIOTHERAPY STUDENTS' ABILITY TO JUDGE NON BIOLOGICAL ELASTIC STIFFNESS.................................95
  Abstract.................................................96
  Introduction............................................97
  Method..................................................99
    Research design......................................99
    Subjects..............................................99
    Stimuli...............................................100
    Procedure..........................................106
    Data analysis........................................108
  Results...............................................108
  Discussion..........................................112
  Conclusion..........................................116
  Appendix C: Raw data for elastic stiffness training study........117

CHAPTER 6: PLINTH SURFACE EFFECTS ON MEASURES OF POSTEROANTERIOR STIFFNESS.................................................................121
  Abstract.................................................122
  Introduction..........................................123
  Method..................................................125
    Research design......................................125
    Subjects..............................................125
    Measurement instruments..........................126
    Procedure..........................................127
    Data analysis........................................129
  Results...............................................129
  Discussion..........................................132
  Conclusion..........................................136
  Appendix D: Schematic diagram of modified stiffness assessment device..................................................137

CHAPTER 7: THE EFFECTS OF HIGH AND LOW LOADING FORCES ON MEASURED VALUES OF POSTEROANTERIOR STIFFNESS.............................138
  Abstract.................................................139
CHAPTER 8: AN INVESTIGATION OF THE RELATIONSHIP BETWEEN LOW BACK PAIN AND LUMBAR POSTEROANTERIOR STIFFNESS
Abstract ................................................................. 159
Introduction .............................................................. 160
Method ................................................................. 162
Research design ......................................................... 162
Subjects ............................................................... 163
Measurement instruments ............................................. 167
Measurement of stiffness ............................................ 167
Measurement of pain .................................................. 168
Procedure ............................................................. 170
Data extraction ......................................................... 172
Statistical analysis ..................................................... 173
Results ................................................................. 174
Discussion ............................................................ 182
Conclusion ............................................................. 187

CHAPTER 9: CONCLUDING REMARKS ........................................... 188
Overview of problem .................................................. 189
Current status of posteroanterior stiffness assessment in human backs ............................................ 193
Directions for future research ........................................ 194

REFERENCES .................................................................. 197
ACKNOWLEDGEMENTS

I would like to thank my supervisors, Michael Lee and Roger Adams, for the tireless supervision and expert guidance they provided during the completion of my doctoral studies. They were incredibly generous with both their time and enthusiasm, and for this I am very grateful. I would also like to thank Chris Maher for his interest and readiness at anytime to discuss issues related to the assessment of spinal stiffness. In addition, special thanks are extended to Louise Ada for her interest and comments on earlier drafts of work in preparation, and to Michaelene Goodsell, for the many friendly discussions held at her practice and for allowing me to access subjects through her practice. Similarly, I would like to thank Professor Margareta Nordin, Director of The Occupational and Industrial Orthopaedic Centre, New York, for her comments regarding my clinical study and for allowing me access to her patients. Finally, I would like to thank all the subjects (low back pain sufferers and those without symptoms) who participated in these studies, for their involvement has been critical to gaining knowledge about the assessment of stiffness in human spines.
PREFACE

This thesis contains nine chapters each of which may be read independently. The first chapter provides an overview of the problem of low back pain and an introduction to the area of spinal manipulative therapy, including a discussion of the rationale for the assessment of posteroanterior (PA) stiffness by physiotherapists and the various methods for assessing PA stiffness. The last chapter (Chapter 9) incorporates concluding remarks based on the research reported in the empirical chapters and provides some recommendations for clinical practice arising from this work.

The research work for this doctorate is comprised of two main steps. The first step began as an attempt to improve and facilitate measurement of human PA stiffness by developing and evaluating a portable, mechanical device which would be able to accurately and reliably measure PA stiffness (Chapters 2 and 3). The second step was to use this device, firstly in an attempt to train physiotherapy students to more accurately judge PA stiffness (Chapters 4 and 5), and secondly, to research aspects of stiffness testing. Chapters 6 and 7 report on studies conducted to further improve the accuracy and reliability of both manual and mechanical assessment of PA stiffness by investigating (i) the effect of plinth surface (Chapter 6), and (ii) the effect of the magnitude of applied testing force (Chapter 7) on measures of PA stiffness.

Finally, the mechanical stiffness testing device was employed in an investigation of the hypothesized relationship between back pain and measures of lumbar PA stiffness in a physiotherapy clinic with a group of subjects with low back pain (Chapter 8).
All the research studies reported here received ethical approval from the University of Sydney Human Ethics Committee. The study reported in Chapter 8 also received approval from the Institutional Review Board of the Hospital for Joint Diseases Orthopaedic Institute, New York.

A number of publications have arisen from these PhD studies and these are listed overleaf. The roles of the co-authors on these papers are as follows. B. Wilkinson, M. Goodsell and C Maher are co-authors on one of the published papers which was based on chapters 2 and 3. Barry Wilkinson was the electrical technician contracted by me to manufacture the portable stiffness testing device, and who then participated in the trial and error and discussions with me as we developed solutions to the many electrical and mechanical problems which arose. Michalene Goodsell and Chris Maher are members of the Spinal Manipulation Research Group and were involved in early discussions relating to the design of the stiffness device. Mathew Holland, a co-author on the paper based on chapter 6 was at the time an undergraduate physiotherapy student gaining experience for his honours project by acting as a research assistant.
PUBLICATIONS

The work arising from these PhD studies has resulted in the following refereed publications and manuscripts, each being based on a separate chapter from the thesis.


# LIST OF FIGURES

1.1 The PA pressure test. The rater uses either their thumbs or the heel of the hand just distal to the pisiform to perform stiffness testing.................................16

2.1 The final stiffness assessment device being used in testing at L3 by the author ..........33

2.2 Line drawing of the stiffness assessment device ..............................................34

2.3 Schematic diagram of the components of the mechanical head ........................37

2.4 Typical force/displacement loading and unloading curves obtained when testing the lumbar spine.................................................................42

2.5 Force/displacement data for loading curve for typical subject with fitted straight line above 30 N ..................................................................................43

2.6 Force/displacement data for typical subject with fitted parabola above 30 N .......44

3.1 Accuracy testing: measurement of stiffness of aluminium beams using weights and dial indicator .................................................................................59

3.2 Accuracy testing: measurement of PA stiffness using portable stiffness assessment device ..........................................................................................60

4.1 Performance of PA stiffness testing by rater prior to receipt of feedback. The card with the measured value can be seen face down at the end of the plinth.........80

4.2 The mean absolute error (N/mm) at each session for each rater (R1-R4) ............83

4.3 The mean absolute error and SE (N/mm) at each session for the group of raters. ....84

5.1 Mechanical device used to generate stiffness stimuli. The bottom photograph shows a rater judging the stimulus using a pisiform grip ..............................101

5.2 Plan views of the elastic stiffness generating device ........................................102

5.3 Force/displacement curves for low stiffness stimulus (solid triangles) and high stiffness stimulus (closed circles) obtained by testing with the stiffness assessment device over the foam covered contact pad ...........................................104

5.4 Rating spring stiffness ..................................................................................107

5.5 Mean absolute error (N/mm) at each session for each rater (R1-R4) judging elastic stiffness .........................................................................................110

5.6 Mean absolute error and SE (N/mm) at each session for the group of raters judging elastic stiffness .................................................................................111

5.7 Mean absolute error and SE (N/mm) at each session for the group of raters judging elastic stiffness (open circles) and judging lumbar stiffness (closed circles).......113
6.1 Force/displacement characteristics of the loading curve (dotted) obtained when performing PA stiffness testing of the padded plinth surface using a torso-sized, rigid board. The regression line used for stiffness analysis is shown (solid) ........................................128

6.2 Scatterplot showing the relationship between $K_{\text{unpadded}}$ and $K_{\text{padded}}$ and the regression line ........................................................................................................131

7.1 Force/displacement curve (open circles) for a subject exhibiting predominantly linear force/displacement behaviour. Regression lines have been fitted using different force intervals but same minimum forces (left figure), and using different fixed intervals of the force/displacement curve (right figure) .............................................153

7.2 Force/displacement curve (open circles) for a subject exhibiting less linear force/displacement behaviour. Regression lines have been fitted using different force intervals but same minimum forces (left figure), and using different fixed intervals of the force/displacement curve (right figure) .............................................154

8.1 Research design ........................................................................................................163

8.2 Absolute visual analogue scale ...............................................................................170

8.3 Lumbar PA stiffness testing .....................................................................................171

8.4 Change in K for subjects with LBP. Change scores for individual subjects ranked in ascending order from left to right according to the magnitude of change ........................................................................................................175

8.5 Change in K for painfree subjects. Change scores for individual subjects ranked in ascending order from left to right according to the magnitude of change ........................................................................................................177

8.6 Change in the displacement D30 for subjects with LBP. Change scores for individual subjects ranked in ascending order from left to right according to the magnitude of change ........................................................................................................178

8.7 Change in the displacement D30 for painfree subjects. Change scores for individual subjects ranked in ascending order from left to right according to the magnitude of change ........................................................................................................180
LIST OF TABLES

3.1 Subject characteristics ..................................................................................................................57
3.2 K values (N/mm) obtained using weights and dial indicator (Method 1) and the new
    stiffness device (Method 2), and the percentage error .................................................................61
3.3 K (N/mm) and D30 (mm) values for Test 1 and Test 2 for each subject ..............................67
3.4 K values (N/mm) for cycles 2-5 for typical subjects at each tested level .............................68
3.5 D30 values (mm) for cycles 2-5 for typical subjects at each tested level .............................68
4.1 Percentage of occasions at which the rater was in exact agreement (PEA) for the
    pre- and post-tests and the maximum absolute error for each test for each rater.........82
4.2 Mean absolute error and SE (N/mm) for each rater (R1-R4) for each session.
    There was a pre-test (Pre), five training sessions (1-5) and a post test (Post).................84
5.1 K and D30 values obtained when testing stimuli generated by the spring device
    prior to performing PA oscillations (pre-test) and after the application of one
    hundred PA oscillations (post-test) ...........................................................................................105
5.2 Percentage of occasions at which the rater was in exact agreement (PEA) for the
    pre- and post-tests and the maximum absolute error for each test for each rater....109
5.3 Mean absolute error and SE (N/mm) for each rater (R1-R4) for each session.
    There was a pre-test (Pre), five training sessions (1-5) and a post test (Post).............110
6.1 The number of subjects who demonstrated an increase in stiffness when tested on
    the rigid compared with the padded plinth surfaces and the value of that
    increase ........................................................................................................................................130
7.1 K and $r^2$ values for each subject calculated using regression forces from 30-80N,
    30-150N, 30-200N .....................................................................................................................149
7.2 K and $r^2$ values for each subject calculated using regression forces from
    30-83.3N, 83.3-136.7N, 136.7-200N .......................................................................................151
8.1 LBP subject characteristics ........................................................................................................165
8.2 Painfree subject characteristics .................................................................................................166
8.3 Product-moment correlation between change in K and subject variables
    for subjects in the pain group .....................................................................................................176
8.4 Product-moment correlation between change in displacement D30, and subject
    variables for subjects in the pain group .....................................................................................179
8.5 Product-moment correlation between change in AVAS score and subject variables for subjects in the pain group .............................................................. 181
9.1 Current status of PA stiffness assessment in human backs ........................................ 193

SMILE ! YOUR HAVING A GREAT DAY
ABSTRACT

This thesis reports a series of investigations on the assessment of lumbar posteroanterior stiffness in human backs by both instrumented and manual methods. Investigation of posteroanterior stiffness has been retarded due to difficulties in manually measuring stiffness, with many studies reporting the poor reliability of manual judgements of stiffness. The first study in this thesis, therefore, describes the design and fabrication of a portable instrument for measuring posteroanterior stiffness. The findings of the second study establish the test-retest reliability of the device in measuring posteroanterior stiffness in subjects with low back pain.

It was hypothesized that one reason for the reported poor reliability of posteroanterior stiffness assessment might be inaccurate feedback during training of stiffness judgement skills in physiotherapy students. The third study therefore investigated whether it was possible to train students to judge lumbar stiffness by providing immediate and accurate feedback about stiffness judgements. Feedback was provided after measuring the stiffness stimuli (human lumbar spines) using the mechanical device developed earlier. The results of this study showed no significant difference in the mean absolute judgement error between the pre- and post-tests, suggesting that provision of information regarding the true posteroanterior stiffness of each lumbar spine judged did not improve the accuracy of students' judgements of lumbar stiffness. It was possible that due to the complex nature of the stiffness judgement task, which involves concurrent perceptions of both the non-linear and linear components of the lumbar force/displacement curve, attending only to the linear component of the curve was too difficult. Therefore, the fourth study investigated whether the same physiotherapy students could be trained to assess the linear elastic stiffness arising from pressing on a series of springs. The results of this fourth study found that immediate quantitative feedback training did improve physiotherapy students' ability to judge linear elastic posteroanterior stiffness, implying that the error made when judging posteroanterior
lumbar stiffness is not due to an incapacity to judge linear elastic stiffness. A further factor that may have contributed to the raters’ failure to learn in the third study related to error that may have been introduced in the feedback given to the raters. If measures of posteroanterior stiffness performed on a rigid table during instrumented testing of stiffness were not highly correlated with stiffness measures obtained on minimally padded, height-adjustable plinths where the physiotherapy raters performed their assessment, then some error may have been introduced. The results of the fifth study demonstrated that mean lumbar posteroanterior stiffness values obtained when testing using the portable stiffness assessment device on a padded plinth were significantly lower than when testing on a rigid surface. There was, however, a high correlation between the two measures suggesting that this did not explain the raters’ failure to learn in Study 3. One further hypothesis about why student raters were unable to learn to manually judge posteroanterior stiffness relates to the fact that if sufficient non-linearity existed in the force/displacement data obtained from the lumbar spine stimuli, then if the physiotherapy student raters were palpating with different forces they may have experienced different stiffness values to those gained using the instrumented device. The sixth study therefore investigated the reliability of stiffness measures obtained using different applied testing forces. The results of this study suggested that if physiotherapy raters were sampling posteroanterior stiffness using different force limits from those used during instrumented testing then, because there is some non-linearity in the force/displacement relation, a different stiffness value may have been obtained.

The final study in this thesis was a clinical study investigating whether a relationship does, as hypothesized, exist between low back pain and lumbar posteroanterior stiffness. Subjects with low back pain showed increased posteroanterior stiffness compared to when they had little or no pain at the same lumbar level, while painfree subjects showed unchanged posteroanterior stiffness over time. These data thus provide the first empirical support for the pain-stiffness link.
CHAPTER 1

MANUAL CLINICAL ASSESSMENT OF SPINAL POSTEROANTERIOR STIFFNESS
THE PROBLEM OF LOW BACK PAIN

Incidence

Low back pain poses a significant personal and economic burden on much of the adult population of industrialised countries. By middle life it is usual for adults to have experienced an episode of low back pain. Estimates from cross-sectional studies suggest that 50-70% of adults report having experienced low back pain (LBP) at some time in the past (Biering-Sorensen 1983, Deyo and Tsui-Wu 1987, Waddell 1987), with the yearly prevalence of LBP in the United States population being as high as 15-20 percent (Andersson 1991). Back pain remains one of the most common reasons for people seeking medical care. In 1990 there were almost 15 million office visits for mechanical LBP in the United States, this condition ranking fifth as a reason for all physician visits (Hart et al 1995).

It appears that physiotherapists play an important role in the management of LBP, with physiotherapy being ordered at 21% of the 15 million visits for LBP. Further indication of the role of physiotherapy in the treatment of LBP has been provided in a recent national survey conducted in the United States, which concluded that 80 percent of 1200 physicians surveyed believed physical therapy to be an effective treatment for LBP (Cherkin et al 1995).

While many investigators have examined the prevalence of LBP in Europe and North America, there is less information regarding prevalence rates of non-compensable LBP in the Australian population. One study (Eberall 1994) which examined an
Australian metropolitan sample concluded, however, that the high prevalence of adolescent LBP in studies of overseas populations was similar in Australia.

Cost Implications

The cost implications from LBP are immense. In the United States over 25 billion dollars are spent each year just to treat the condition (Frymoyer 1991). However, given the addition of disability costs, the total cost to American society is likely to be far greater, with the United States of America (USA) Agency for Health Care Policy and Research estimating it to be as much as 50 billion dollars annually (Bigos et al 1994).

When estimating the cost of LBP it is useful to distinguish between acute and chronic pain. The majority of episodes of LBP are self-limited regardless of treatment, with 75% resolving within 4 weeks and 90% within 3 months (Spitzer 1987). In approximately 5% of cases, however, LBP may persist for 6 months or more and is then defined as chronic LBP. Chronic LBP is associated with prolonged absences from work and high health care costs arising from the consequent outpatient treatment, surgery, and indirect costs related to the disability. It has been estimated that chronic LBP accounts for between 60% (Bigos et al 1994) and 95% (Webster and Snook 1990) of the total cost of LBP.
The diagnostic dilemma

Despite the prevalence of LBP there is often great difficulty in establishing its underlying cause. It appears that even after an extensive medical workup, only about 15% of patients can be given a definitive diagnosis (Bigos et al 1994, Kelsey 1982). In the remaining 85% of patients the diagnosis is simply stated as non-specific LBP (NSLBP).

One of the reasons for the difficulty in establishing an exact diagnosis is the poorly understood relationship between spinal disease and symptoms. Many studies have demonstrated that there is little relationship between the presence of any spinal disease or anatomic abnormality, as identified using imaging studies, and the presence of LBP (Biering-Sorensen et al 1985, Bigos et al 1992, Fullenlove and Williams 1957). One of the main problems with findings from imaging studies is the lack of empirically-established rules for determining if an anatomic defect seen on imaging tests is actually the cause of symptoms. Several studies have demonstrated the presence of degenerative changes and bulging or herniated discs in subjects who do not report LBP (Boden et al 1990, Hitsselberger and Witten 1968, Wiesel et al 1984). Wiesel et al (1984) examined computerized tomography (CT) scans in 52 subjects without LBP and noted herniated discs in 19.5% of subjects under the age of 40 years, and in 50% of subjects over 40. It appears that degenerative changes on X-ray are more likely to be found with increasing age rather than associated with a history of LBP. Powell et al (1986) evaluated MRI scans in 302 spinally-asymptomatic females presenting for obstetric/gynaecological investigation. Disc degeneration was identified in 34% of women aged 21-30 years, in 60% of women aged 31 to 40 years,
and in 95% of women aged 41-70 years. This study further supports the relationship being between disc degeneration and age rather than between disc degeneration and LBP symptoms.

Because of the difficulty in establishing an exact structural or pathological diagnosis, clinicians treating patients with non specific LBP (NSLBP) often establish a clinical diagnosis for the patient based on the patient’s presenting signs and symptoms (Spitzer 1987). These clinical diagnoses are then used to make treatment decisions and to measure patient progress.

**Historical overview of the practice of spinal manipulative therapy**

The laying on of hands to treat patients with spinal pain has been used since the time of Hippocrates (460-370 BC). This practice of manually applying forces to the spine to assess and treat patients suffering from LBP has more recently been called spinal manipulative therapy, but it has been used by lay and medical practitioners for some hundreds of years. Hippocrates’ treatment for spinal pain included the use of a steam bath followed by the patient lying prone on a wooden board, tractioned by two assistants at both the head and feet, while the physician provided a sharp downward thrust to the symptomatic region of the spine. The use of the physician’s foot to provide the downward thrust later became popular (Schiotz & Cyriax 1975). Five hundred years later, Galen (131-202 AD) advocated a similar treatment to that of Hippocrates for spinal pain, and became the first person to document relief of a C8
nerve root compromise, in a man who had fallen from a cart, by treating the cervical spine (Schiotz & Cyriax 1975).

Over eight hundred years later Abu’Ali ibn Sina, an Arabian doctor who was also known as Avicenna (980-1037), compiled a document which set out the total amount of medical knowledge of his time. This document was reprinted in Latin many times between 1473 and 1608 and remained an authoritative textbook until the 17th century (Schiotz and Cyriax 1975). Part of the document contains illustrations regarding suitable treatment for spinal pain, advocating use of the same methods as those originally described by Hippocrates over one thousand years previously.

For many centuries in England lay people used spinal manipulative therapy to treat back pain. These practitioners were known as ‘bonesetters’ due to their belief that a ‘little bone’ lay out of place, and that relief with manipulation was due to relocation of this ‘little bone’ (Schiotz & Cyriax 1975). Bonesetters seem to have existed for as long as there have been records, and in England they still exist today (Bourdillion & Day 1987). Bonesetters were invariably countrymen who lacked formal training, the art being passed from parents to children as family secrets. In learning and performing bonesetting techniques it was even suggested that an hereditary trait may be of some value (Bourdillion and Day 1987) an observation perhaps prompted by the fact that some people appear to learn the practice of manipulation much more easily than others.
Few descriptive accounts of the bonesetters exist, though incidental mentions can be found. The French author, Collette (1873-1954), in a short story ‘The Rainy Moon’, discussing events occurring in 1900, speaks of ....

“.....that female bonesetter, Mademoiselle Levy, who undertook the care of bodies and souls and demanded so little money in exchange.....She practised massage and the laying on of hands in the darkest depths of pallid concierges' lodges, in variety artistes' digs in the rue Biot, and in dressing rooms in the music halls of La Fauvette.” (Collette 1900 p377).

Evidently, variety artistes and music hall performers had pain that was relieved by a bonesetters' skills.

By the 19th century, although spinal manipulative therapy was being widely practiced by the bonesetters, physicians and surgeons of the time were largely abandoning the use of manipulation for the treatment of spinal pain. Doctors became increasingly wary of applying pressure to joints. This was partly due to the fact that tuberculosis had reached extreme proportions at this time, and it was known that applying a manipulative thrust to a joint, invaded and weakened by tuberculosis, could result in damaging consequences (Haldeman 1992). Also, syphilis was prevalent and known to weaken bone, and many doctors felt it better to medicate the patient rather than to manipulate (Haldeman 1992).

One doctor who challenged his colleagues with respect to their declining use of manipulation was Wharton Hood, who had learnt the art of bonesetting from Richard Hutton, a well-known bonesetter of the early 1800’s. Following the death of Hutton in 1871, Hood published a series of articles in the Lancet describing Hutton’s
techniques. These articles were later compiled into a treatise entitled, *On Bonesetting* (Nwuga 1976).

In 1867, Sir James Paget also tried to influence the medical profession to study the ideas and techniques of the bonesetters for treating spinal pain. His famous lecture entitled “Cases that bonesetters cure” was reported in the British Medical Journal. In this lecture Paget identified the types of cases that were responsive to manipulative therapy, however his words were to no avail, the medical profession of the day believing the rationale of the bonesetters to be unsound (Nwuga 1976).

Although the medical profession in the nineteenth century moved away from spinal manipulation, patients did not. In 1895 Daniel David Palmer, the founder of chiropractic, reported that he cured the deafness of a janitor after adjusting the spine (Nwuga 1976). Despite the unusual nature of this claim there were many students keen to enrol in the first chiropractic school founded in Davenport, USA in 1897 by DD Palmer. One hundred years later, chiropractic schools now operate around the world, with chiropractors in Australia being among the most frequently consulted health care providers for patients with spinal pain (James 1994). The original philosophy of chiropractic suggests that subluxation in the spinal column interferes with nerve function and that this is a significant factor in disease causation (Waagen and Strang 1992). Because of this, the chiropractic profession believe that manipulation can be used to treat organic disease of viscera such as diabetes, cardiac abnormalities, hypertension etc (Jamison et al 1992).
The osteopathic profession also uses spinal manipulative therapy to treat patients with spinal pain. Osteopathy was founded by Andrew Still (1828-1917) who was also a medical doctor. He described the ‘osteopathic lesion’, a region of the spine exhibiting increased tenderness or reduced motion. He believed that spinal manipulation directed to this lesion would ameliorate symptoms (Wiltse 1991). Today, osteopaths are probably more closely aligned with physiotherapists, in that few osteopaths or physiotherapists propose that manipulation of the spine will cure visceral disease, and consequently both groups are more closely aligned with the medical profession than are chiropractors.

In Australia, spinal manipulative therapy has been used by physiotherapists to treat patients with back pain since the 1930’s (Forster 1975). Much of the early teaching of spinal manipulative therapy to physiotherapists was provided by medical practitioners such as James and John Mennell, and James Cyriax who had been influenced by the teachings of osteopaths. Cyriax (Schiotz and Cyriax 1975) described his attempts to train physiotherapists in the skills of manipulation as follows:

".....By now, I have myself taught these methods to almost a thousand physiotherapists most of them women, who have become as proficient in this part of their work as can be expected of any student - certainly expert enough to obtain good results consistently. It is a proven fact, therefore, that physiotherapists can be taught this work....." (p95)

The manipulative skills that were taught by Cyriax to physiotherapists involved the use of substantial traction forces and poorly localized techniques. However, despite
much criticism of his method, James Cyriax was extremely influential in the emergence of manipulative physiotherapy. Further, Cyriax's use of, and teaching of, manipulative skills at St Thomas's Hospital in London provided spinal manipulative therapy with a credibility in conventional medical circles that it had previously lacked.

Another influential figure in the development of spinal manipulative therapy within the physiotherapy profession was Frederick Kaltenborn, who was a physiotherapist as well as an osteopath and chiropractor (Nwuga 1976). Due to his efforts and charisma he earned the respect of doctors in Norway, who in 1957 supported his attempts to form a group of physiotherapists specializing in spinal manipulative therapy. Kaltenborn advocated the use of many of Cyriax's methods to evaluate the patient, whilst using osteopathic and occasionally chiropractic techniques for treatment (Nwuga 1976). The ideas of Kaltenborn have been taught to many physiotherapists in Australia. In 1985 a special interest group for physiotherapists interested in Kaltenborn's theories and techniques was formed, known as the Australian Institute of Nordic Manual Therapists (AINMT).

Probably the physiotherapist who, more than any other individual, profoundly influenced the development of spinal manipulative therapy within the profession was Geoffrey Maitland. Maitland (1964) proposed a comprehensive system of examination for patients with back pain, as well as other musculoskeletal disorders, and advocated basing treatment decisions on the patient's signs and symptoms rather than on a pathological or structural diagnosis. He also emphasized the importance of
reassessment, suggesting that the usefulness of an intervention could only be determined by measuring before and after the intervention (Maitland 1986).

Physiotherapists use of spinal manipulative therapy to examine patients with spinal pain

The work of Geoffrey Maitland (1964, 1968, 1973, 1981, 1986) provided much of the theoretical basis for the practice of spinal manipulative therapy as it is currently performed by physiotherapists. The present paradigm, within which spinal physiotherapy operates, involves decisions made regarding:

(i) diagnosis, particularly identifying a pathology or structure responsible for the patient’s signs and symptoms, or in the absence of this, providing a clinical diagnosis based on the patient’s signs and symptoms, and,

(ii) management, particularly regarding aims of treatments and techniques to be performed, suitable dosage, and outcome measures to evaluate success of treatment (Refshauge & Gass 1995).

One of the assumptions underlying the practice of spinal manipulative therapy by physiotherapists is that pain and stiffness are related (Mennell 1960, Maitland 1986). Expression of this view comes from the writings of Mennell (1960), Paris (1965), Stoddard (1983), and Bourdillion and Day (1987). Paris (1965 p78) describes the ‘spinal lesion’ suggesting that various pathologies such as disc lesions, facet locking and muscle and ligament strains all interfere with the movement of the spine, thus affecting neighbouring joint tissues and producing pain. Manipulative treatment
directed to these hypomobile joints results in relief of symptoms. Bourdillion and Day (1987) in their chapter entitled ‘What is the Cause of the Pain?’ suggest that stiffness localized to one or more individual spinal joints is nearly always present in patients with back pain and hence is related to the pain:

"...There are relatively few basic physical signs which can be elicited in a sufficient proportion of back pain patients to justify the belief that they are fundamental to the condition. Two objective signs are almost always present, and indeed, if one looks in the right places, it is probably true to say that they are always present. These are loss of mobility and localised changes of tension and texture in the soft tissues." (p213).

Further argument for the existence of a relationship comes from Mennell (1960) who states that:

"...pain causes loss of joint movement, but loss of joint movement also causes pain." (p6)

while Stoddard (1983) suggests that:

"...All spinal joints should be examined for mobility; if restricted, and there are no contraindications, they should be manipulated." (p39).

More recently, Jull et al (1994a) have reported an association between pain and stiffness. In this study a manipulative physiotherapist was able to differentiate between painful and non-painful cervical segments based on the therapist's perception of abnormal motion or stiffness at that level. This study suggested that attending to the stiffness or motion characteristics of different cervical spine levels would enable a therapist to detect the painful spinal levels.
Nordgren et al (1980), who evaluated a sample of military recruits, found that judgements of altered segmental mobility, in conjunction with reports of pain when a PA force was applied to the lumbar spine, provided the best prediction of occurrence of back pain during field service. The predictive value of a negative test was 87%, representing the percentage of recruits with a negative test who did not get back pain during field service, while the predictive value of a positive test was 49%, representing the percentage of recruits with a positive test who developed back pain during field service. The test was found to have moderate sensitivity (55%), and high specificity (85%). A recent study (Phillips and Twomey 1996) also found that the use of manual assessment of PA lumbar movement, in conjunction with verbal subject responses, enabled successful diagnosis of the lumbar level responsible for a subject’s LBP and referred pain. Hence some evidence exists to suggest that PA stiffness may be relevant in patients with spinal pain.

Because of the hypothesized relationship between pain and stiffness, Maitland (1986) and others (eg. Refshauge and Gass 1995) advocate the use of a series of manual tests when examining patients with spinal pain, these being conducted to try and identify abnormalities in segmental range of motion. The information gained from these tests is used to establish a clinical diagnosis for the patient, to select the spinal region to be treated, and to select the most appropriate treatment technique to be used (Refshauge and Latimer 1995). The manual tests usually performed on patients with spinal pain include:
- soft tissue palpation
- passive accessory motion tests
- passive physiological intervertebral movement tests

and these will be considered in turn.

Soft tissue palpation.

Examination of the soft tissues is performed to gain information regarding pathology affecting the muscle, ligament, joint capsule etc., such as the presence of muscle spasm, soft tissue thickenings, local oedema due to inflammation or other causes (Latimer and Maher 1995). Most commonly, evaluation of the soft tissues is performed by palpation, where the examining therapist searches for an indicative alteration in consistency, with the abnormal area becoming harder or softer than the normal tissue (Fischer 1987b). During this examination the therapist may also search for the presence of bony anomalies. It has been demonstrated, however, that when palpating the soft tissues of the spinal region, judgements made regarding the presence of bony anomalies or muscle spasm are not reliable (Keating et al 1990). The manual test does, though, yield reliable information regarding pain and symptom reproduction (Keating et al 1990) and can therefore be used to indicate which vertebral levels should be examined using passive accessory and physiological tests.

In the late 1980’s a handheld mechanical device was developed to objectively measure soft tissue compliance (Fischer 1987a). This device was subsequently used to measure soft tissue compliance in the paraspinal region of a group of children with
juvenile arthritis and a group of healthy children (Hogeweg et al 1995). The results of this study demonstrated no significant difference between the measures of compliance obtained using the device in the juvenile chronic arthritis group and the healthy group. However several studies, discussed later in this thesis, have raised concerns regarding the reliability and accuracy of this hand-held device (Kawchuk and Herzog 1995, Kawchuk and Herzog 1996).

Passive accessory motion tests.

After soft tissue palpation, the physiotherapist typically proceeds to test the ‘passive accessory motion’ of the symptomatic spinal region. One of the tests frequently used to detect increased lumbar PA stiffness, and thereby assist in the formulation of a clinical diagnosis, is a form of ‘springing test’ for the lumbar spine, where a posteroanteriorly directed force is manually applied over each lumbar vertebral segment. This test is commonly called the posteroanterior (PA) pressure test, and was initially described to test the amount of passive accessory intervertebral movement occurring at the segment (Maitland 1986). There is, however, evidence to suggest that the PA pressure test produces relatively little intervertebral movement (Ward 1988, Lee et al 1996), instead causing a complex movement of the spine, pelvis and thorax. The current uses of the PA pressure test are to provoke the patient’s pain, to assess the amount of displacement occurring, and to identify the presence of increased tissue resistance to this displacement (ie: increased stiffness) (Jull 1994b). Figure 1.1 shows the performance of this test.
Fig 1.1 The PA pressure test. The rater uses either their thumbs or the heel of the hand just distal to the pisiform to perform stiffness testing.

The information gained by the therapist from the PA pressure test is then used, in combination with other significant physical findings, to select both the spinal region to be treated and the most appropriate treatment technique. If pain and increased posteroanterior stiffness are detected, then passive mobilization or manipulation techniques may be used at the symptomatic vertebral level in order to increase posteroanterior displacement, reduce stiffness, and relieve pain/symptoms. The test may also be useful in monitoring patient improvement, with a reduced pain provocation on application of the force suggesting an overall reduction in LBP (Latimer and Maher 1995). Although the test is used clinically by some
physiotherapists to document patient recovery, the appropriateness of this test for this purpose has not yet been investigated.

There has been much discussion about what aspect of the vertebral movement, produced when performing the PA pressure test, is of most interest to manual therapists. Maitland (1986) suggests that the absolute movement of the target vertebra is a major consideration. When performing PA pressure testing, the therapist applies a single force to the target vertebra, using a pisiform or thumb grip, and evaluates the movement without reference to the movement induced in the adjacent vertebrae. It is assumed that the movement of the target vertebra reflects the intervertebral movement occurring at the superior and inferior articulations of that vertebra (Maitland 1986). However, there is little evidence to suggest that the PA pressure test measures intervertebral movement. It has been radiographically demonstrated that when a 100 N static load is applied over L3, the vertebrae of the lumbar spine move together into extension, demonstrating little intervertebral shear displacement (Ward 1988). The typical movement pattern of an intact L23 segment (exhibiting Grade 2 disc degeneration) in response to the application of a PA force has been described as extension of approximately 4.99 degrees ± 2.03 degrees when a PA force of 90 N and an extension moment of 9 Nm were applied to the segment (Lee R 1995). This extension was accompanied by superior translation of the anteroinferior corner of the vertebral body of 3.00 ± 1.23 mm, and inferior translation of 0.8 ± 0.59 mm at the posteroinferior corner. Translations in other directions were extremely small. It is difficult however, to use these values to infer the amount of movement occurring to L3 in live subjects when a PA force is applied
to the skin overlying L3, because the skin, subcutaneous tissue and spinal muscles were all removed from the cadaver segments used in this motion study. Lee (1995) therefore conducted an in vivo study where he applied a PA force to the skin overlying L4 and made a series of radiographic measurements of the resultant intervertebral movement. The same pattern of movement was demonstrated in this study as occurred in the in vitro study, although the magnitude of the movements was smaller. For example, when a PA force was applied to L4, the L4/5 segment had mean extension of 1.2 degrees with translations of less than 0.2 mm.

More recently, a series of studies have been conducted to collect data for the development of a computer model designed to predict the movement that occurs when forces are applied to the lumbar spine (Lee M. et al 1995). This model predicts that application of a PA force to the lumbar region will result in an increase in extension, with the target vertebrae also moving anteriorly relative to other vertebrae. However, the magnitude of these predicted intervertebral movements is small, being less than 1 mm translation and 1 degree rotation. In contrast to these small intervertebral movements, the predicted absolute movements of the skin surface overlying the target vertebra are much larger, being around 8mm of movement, again suggesting that use of the PA pressure test provides only a small indication of intervertebral movement and a greater indication of the compressibility of the skin and soft tissues (Lee et al 1995).

Further information regarding the movement produced by a PA pressure in the lumbar spine comes from a study conducted by Lee and Svensson (1993a), who
demonstrated displacement at T₈ and S₁ when a force was applied to L₃. This reinforced the view that the PA pressure test produces movement distant from the vertebral level to which the force is applied. Further support for this view arises from a study that found that for every 100 N of force applied to L₃, an average of 2 degrees of anterior pelvic rotation is produced (Lee et al 1994b).

Passive physiological intervertebral movement tests.

Finally, after using the PA pressure test to assess certain motion characteristics of the spine, physiotherapists often proceed to test the passive physiological movement that occurs between vertebrae to gain more specific information about movement restrictions and hence where best to direct treatment. The studies reported in this thesis are focussed on the problem of stiffness assessment and did not examine the role of passive physiological motion tests in patients with LBP.

**Measurement of stiffness**

**(i) peripheral joints**

Similar to the use of the PA pressure test to measure PA stiffness of the spine, other clinical tests are often performed to measure the stiffness characteristics of peripheral joints. During these tests the therapist attempts to judge stiffness through a range of physiological movement by considering torque and angular displacement, or by judging the stiffness associated with accessory movements by considering force and displacement.
Several mechanical devices have been developed to more reliably measure the stiffness characteristics of different peripheral joints (Fithian et al 1995, Chesworth and Vandervoort 1995, Skalley et al 1993, Yoon and Mansour 1982, Luster et al 1990, Clark et al 1987) but many of these devices have unresolved technical problems. For example, instrumented measurement of patellar mobility has been performed in order to help distinguish among the many possible causes of knee pain, patellofemoral laxity being thought to contribute significantly to symptoms (Fithian et al 1995). One device developed for measuring mobility of the patellar consists of a hand-held force applicator that applies a medially or laterally directed force to the patella, a load cell within the applicator being used to measure the applied force. A displacement sensor is used to measure displacement in the coronal plane. The results of the precision testing of this device indicate that precision in abnormal knees is not enough to discern differences less than 5mm, this representing approximately 25% of the total movement available in an affected (lateral patellar dislocation) knee. Skalley et al (1993), who also developed a device for measuring patellar motion (called the Patellar Pusher) found several technical problems with the device, including inability to control extraneous movement of the entire lower extremity, and difficulty when testing obese subjects.

Tindle (1987) has developed a mechanical device for obtaining force/displacement curves of end-of-range knee extension. Although the author suggests that the device is reliable in its measurement of knee extension, the study examined only ten subjects and determined repeatability only by visually examining the shape of the curves rather than conducting any quantitative analysis.

Several studies have examined the reliability of objective measurement systems for examining passive stiffness of the ankle joint (Kearney et al 1990, Weiss et al 1990,
Chesworth and Vandervoort 1988). One device, used to measure the passive stiffness of the ankle from the neutral position to 10 degrees dorsiflexion, recorded angular displacement using a potentiometer and resistive torque using a strain gauge (Chesworth and Vandervoort 1988). This method of measuring ankle stiffness was shown to be highly reliable and was subsequently used by the authors to compare passive stiffness variables in uninvolved and involved ankle joints of patients following ankle fracture (Chesworth and Vandervoort 1995).

Stiffness of hand joints is a frequent complication following burn injuries. Yet, only this decade has a device been developed for quantifying joint stiffness in the burned hand (Luster et al 1990). This device uses a hydraulic piston system to apply force to the finger and measure the resistance of the joint. A computer calculates the angular displacement of the joint from the motor actuator arm, generating a torque/angular displacement curve for each joint measured. Using this device four subjects with stiff hands were measured before and after dynamic splinting treatments (Luster et al 1990). This study found a significant reduction in joint stiffness [as characterized by measuring the angular displacement at which the largest rate of change (torque per degree) for the curve occurred] following splinting compared to no splinting. Also, the researchers found that those subjects with a higher initial finger stiffness had a significantly smaller reduction in joint stiffness than those with lower initial stiffness values. This study provided some evidence to suggest that measurements made using a mechanical device were sensitive to change in stiffness, and that the values obtained may serve as an indicator of treatment outcome. These findings need to be viewed somewhat cautiously, however, as the authors did not report the reliability or accuracy of their new measurement device.
(ii) lumbar PA stiffness

Recent studies have demonstrated that therapists using the PA pressure test to make judgements regarding the PA stiffness of the spine have poor inter-therapist reliability (Matyas and Bach 1985, Maher and Adams 1994, Binkley et al 1995, Phillips and Twomey 1996). Matyas and Bach (1985), who presented the results of a number of graduate student research projects, found that the reliability coefficient for manual assessment of PA stiffness ranged from $r = 0.09-0.38$. Maher and Adams (1994), who examined the intertherapist reliability of judgements of PA lumbar stiffness on patients, found Intraclass Correlation Coefficient (ICC) values (Shrout and Fleiss 1979) for stiffness judgements ranged from 0.03-0.37 with percent exact agreement scores ranging from 21% to 29%. Both results reflect poor reliability for stiffness judgements made by physiotherapists after manual spinal assessment.

In attempting to identify the cause of this poor reliability, it is necessary to consider factors known to affect measurement of PA stiffness. Such factors can be classified into two groups. The first group includes factors known to affect the physical stimulus value. These include the position of the patient at the time of the PA stiffness test, the subject's breathing pattern, the presence of spinal muscle activity, and the frequency and magnitude of the stiffness testing oscillations used. For example, Herden and Kingsford-Smith (1990) found that lumbar PA stiffness was increased when subjects were measured in an extended lumbar position. Further, it has been demonstrated that PA stiffness at L3 is 12% lower during breath-holding at maximum inspiration, compared to breath holding at the end of normal expiration (Keaveney et al 1989). Also, increased activity of the spinal extensor muscles has
been found to increase PA stiffness (Lee et al 1993b). When testing viscoelastic tissue, the PA stiffness value obtained is dependent on the frequency of force application during testing. If the rate of loading increases, the resistance to deformation of soft tissues also increases. Two studies have demonstrated an increase in PA stiffness with increases in the loading frequency of testing (Lee and Svensson 1993a, Lee and Liversedge 1994a). Finally, it has been suggested that the testing surface on which the PA stiffness test is performed, and the presence of lumbar pain, may be associated with an increase in PA stiffness. These hypotheses can be investigated with a reliable stiffness measuring system.

The second group of variables affecting reported PA stiffness includes factors known to affect perceived stiffness while the physical stimulus remains unchanged. These factors include the manner in which the therapist performs stiffness testing eg: their, visual state at the time of testing and the grip type they use. Maher and Adams (1996a) evaluated whether availability of vision influenced judgements of PA stiffness. This study found that while occluding vision during testing did not affect a subject’s ability to discriminate between stiffness stimuli, the stimuli were judged as significantly stiffer compared to when judged with eyes open. A second study (Maher and Adams 1996b) sought to investigate whether grip type employed during testing affected a rater’s ability to discriminate between stiffness stimuli. The results of this study showed that using either the pisiform grip or the thumb grip gave similar ability to discriminate between stiffness stimuli, however, use of the thumb grip made the stiffness stimuli appear substantially stiffer than when the same stimuli were tested using a pisiform grip.
Therefore, a failure to control either factors known to affect the physical stimulus, or to affect perceived stiffness, will likely result in poor interrater reliability. Currently, the protocol for PA stiffness testing does not attempt to control these variables. In order to improve the interrater reliability of PA stiffness testing it would appear necessary to develop a protocol that clearly controls these factors.

Maher (1995) reviewed the literature on stiffness perception and proposed two further reasons from those mentioned above that may account for the poor interrater reliability. These were, firstly, that humans may simply not possess the ability to accurately discriminate stiffness stimuli in the range represented by human spines, and secondly, that therapists may have idiosyncratic concepts of what constitutes ‘stiffness’. His subsequent research evaluated these hypotheses. One study (Maher and Adams 1995a) investigated the stiffness discrimination threshold of 72 subjects using a mechanical device to provide the stiffness stimuli. This study found that humans have a relatively good ability to discriminate between stiffness stimuli with a change of approximately 11% in the stimulus being needed to be able to detect that a change had occurred. This result suggested that the poor reliability of stiffness judgements was unlikely to be due to difficulty in discriminating linear elastic stiffness. Maher and Adams (1995b) therefore hypothesised that it was each therapist having an individual but different concept of PA stiffness that was a source of the poor reliability. They proposed that individual interpretation of the concept ‘stiffness’ was the source of the disagreement in stiffness judgements between therapists.
As a result of their research Maher and Adams made several suggestions regarding ways to enhance the objectivity and accuracy of stiffness assessments. One of these suggestions was to focus less on manual assessment and to rely more heavily on the use of mechanical devices to measure PA stiffness.

Several mechanical devices designed to measure tissue compliance have been developed in different sites around the world (Fischer 1987a, Kawchuk and Herzog 1995, Horikawa et al 1993, Kawchuk and Herzog 1996). Some of these devices are hand-held instruments, such as the ‘tissue compliance meter’ (Fischer 1987a), which consists of a cylindrical probe that is pressed into the paraspinal tissue, resulting in surface deformation. Displacement of the probe is calculated by measuring the distance the probe pushes into the tissue from a collar which marks the original surface level. The amount of force used to push the probe into the tissue is measured using an analogue gauge similar in mechanism to a bathroom scale. Despite portability and hence potential for clinical use, the reliability and accuracy of this device has been argued to be poor (Kawchuk and Herzog 1995), although this finding was recently challenged by Haas (1996) who suggested that it was a misapplication of the ICC used to determine reliability that had led to the poor interrater reliability results. Reliability involves the comparison of the true variability between subjects with the observed variability, but in Kawchuk and Herzog’s study (1995) the ICC’s were computed from repeated measures of a single stiffness stimulus, thereby having no true between-subject variability to evaluate.
Over the last five years different groups of physiotherapy researchers have attempted to improve the reliability of spinal stiffness measurement by developing several new mechanical devices able to collect both force and displacement data when a PA force is applied to the lumbar spine (Lee & Svensson 1990, Lee and Evans 1992). In 1990, Lee and Svensson developed the 'Spinal Physiotherapy Simulator' to objectively measure lumbar PA stiffness in vivo. In 1992 another group of researchers (Lee and Evans 1992) developed a device that delivered a PA force to a selected lumbar level whilst measuring the displacement of the skin and spinous processes of the adjacent vertebral levels. While both these devices have been shown to be reliable in measurement of PA stiffness in asymptomatic subjects (Lee and Svensson 1990, Lee and Evans 1992), they are large and non-portable and hence have been mainly used for investigating the characteristics of PA stiffness in normal subjects able to attend the university departments where the devices are housed.

In summary, although some progress has been made towards more reliable measurement of PA stiffness in humans, mostly without LBP, there is still difficulty in reliably measuring PA stiffness in the clinic in a symptomatic population. The lack of a portable measurement tool has hindered efforts to investigate whether scientific evidence exists (separate from manual therapists' opinions) to support there being a relationship between LBP and lumbar PA stiffness.

This thesis aims to enable assessment of the relevance of PA stiffness in patients with LBP by designing and evaluating a new mechanical device to measure PA stiffness in
clinical situations, by investigating methods of improving manual assessment of PA stiffness, and by determining whether relevant data support the existence of a relationship between lumbar pain and PA stiffness.
INSTRUMENTED MEASUREMENT OF SPINAL STIFFNESS

The work contained in this chapter has been published as:

ABSTRACT

Background and Purpose: Due to the poor reliability of judgments of spinal stiffness made using manual tests, several mechanical devices have been designed to measure posteroanterior (PA) stiffness. One major limitation of these devices has been the inability to transport them to locations where subjects with spinal pain could be easily accessed and measured. Methods: A portable, mechanical device was designed and manufactured to measure PA lumbar stiffness in subjects suffering low back pain (LBP). The device applies a force to the lumbar spine and measures both the force applied and the displacement which results. The applied force is determined by measuring the change in length of a spring using a linear potentiometer. Displacement is measured using an optical encoder. Results: The manufactured device is easily portable and contains features making it safe for patient use. The device cannot be successfully used in cases where subjects are unable to tolerate the application of a force above 30N or where subjects have hypertrophied lumbar spinal muscles. Conclusion and Discussion: The portable stiffness assessment device has the potential to be useful in research investigating the relevance of altered tissue compliance in patients with LBP.
INTRODUCTION

Altered stiffness of the lumbar spine is thought to frequently occur in subjects with low back pain (LBP) (Maitland 1986, Jull et al 1994a, Jull 1994b, Grieve 1984). In order to identify areas of increased stiffness, physiotherapists perform the central posteroanterior (PA) pressure test which involves the therapist applying an anteriorly-directed force over the spinous process of the prone patient (Maitland 1986). Jull (1994b) clearly describes the nature of the test when suggesting that “...the clinician is measuring the basic load displacement characteristics for a particular segmental direction of motion” (p 520). However, several recent studies have demonstrated that clinical judgments of spinal stiffness made using the PA central pressure test have poor interrater reliability (Matyas and Bach 1985, Maher and Adams 1994, Binkley et al 1995).

A number of mechanical devices have been designed to help improve the reliability and accuracy of stiffness measurement. These devices collect both force and displacement data during a simulated PA central pressure test. One such device, the "Spinal Physiotherapy Simulator", was developed by Lee and Svensson (1990) to objectively measure lumbar PA stiffness in vivo. It has been found to be accurate in measuring PA stiffness, tending to underestimate true stiffness by less than 1%. The device also showed good test-retest reliability, with Lee and Svensson reporting an Intraclass Correlation Coefficient (ICC) (2,1) of 0.88 for same-day repeated measurement of PA stiffness at L₃ in asymptomatic subjects. However, it has proved difficult to evaluate symptomatic subjects using this device since it is large and non-
transportable, as a consequence requiring subjects to travel to the laboratory at the University of Sydney for stiffness testing.

Another mechanical device has been developed to measure PA responses by a different team of researchers, R. Lee and Evans (1992). Rather than measuring the load and absolute displacement of the skin surface at the point of application of the force, as is case with the M. Lee and Svensson device, the R. Lee and Evans apparatus measures the relative displacement of the skin surface at the vertebral levels adjacent to the level to which the PA force was delivered. Lee and Evans found this device to be highly reliable in the measurement of relative displacement at L₃ (ICC = 0.99) and at L₅ (ICC = 0.95) when a PA force was applied to L₄. The measure obtained by the Lee and Evans device has not yet been examined for any relationship to clinically-evaluated posteroanterior stiffness. The way in which the measurement is performed does not appear to be related to the manual judgment of PA stiffness, as advocated by manual therapy authors, where the absolute movement of the target vertebra without reference to the movements induced in adjacent vertebrae is argued to be of primary interest (Maitland 1986, Grieve 1984). This device is also large and non-transportable, making it likewise difficult to examine a symptomatic population.

Therefore, the first aim of the present research was to develop a safe, portable device that could be used clinically to perform instrumented assessment of posteroanterior stiffness in patients with spinal pain. This study evolved out of the collaboration of a clinical researcher (Jane Latimer) who identified a need to improve the accuracy of measurement of PA stiffness in patients with LBP as an important step towards
investigating whether LBP and PA stiffness were related, and a biomedical engineer (Michael Lee), who provided expertise in the engineering and computing areas. Funds for the manufacture of the device were obtained from University of Sydney Research Grants, (University Research Grant and National Staff Development Fund Grant) awarded to Jane Latimer. An electronics technician (Barry Wilkinson) was employed, on the grant, to manufacture the device. The author then identified the performance characteristics required in the device and assessed prototypes. The present chapter reports on the development, characteristics and components of this device.

**DEVELOPMENT OF THE PORTABLE STIFFNESS ASSESSMENT DEVICE**

The principal requirements for the new stiffness assessment device were identified as follows:

(i) the capability to apply an oscillating PA force (of varying frequencies) to the lumbar, thoracic and cervical spines, and to accurately measure the load and displacement of the skin surface at the point of application of the force

(ii) provision of adequate subject safety by enabling a maximum force to be defined and by providing an emergency cut-out switch

(iii) sufficient portability to enable the device to be readily transported and easily assembled, thus enabling testing to be performed in different clinical settings
Figure 2.1 and 2.2 show the finished device ready for testing with subjects.

Fig 2.1 The final stiffness assessment device being used in testing at $L_3$ by the author.
Fig 2.2 *Line drawing of the stiffness assessment device.*

The stiffness assessment device consists of the following components: an unpadded testing bed, a small rectangular metal plate (indenter) 35mm×20mm×3mm in size, which applies a force to the spine of the subject, a mechanical head (140mm wide, 395mm high, 190mm deep) that both controls the movement of the indenter and measures the force applied and the displacement which results, a small box that houses the central processing unit (CPU) (165mm wide, 85mm high, 140mm deep) containing an analogue-to-digital converter board and memory chips, and a laptop computer. The specially-designed bed consists of a rigid wooden surface (760mm wide, 1240mm long, 230mm deep), hinged in the middle for portability, and detachable aluminium legs. The mechanical head is attached to a steel frame that sits
over the testing bed, and its position, and thus the indenter, can be adjusted in several ways. Height of the frame above the bed surface may be adjusted to enable assessment of patients of varying dimensions. To allow the force to be applied in a cephalad or caudad direction for testing different lumbar levels, the mechanical head also rotates about a transverse bar. The total device weighs approximately 64 kilograms and can be easily transported in a hatch-back car.

The device was developed over an eight month period with several changes in design along the way. In the original design the displacement of the indenter, which was used to calculate stiffness, was to be measured using a linear potentiometer. However, when considering the clinical use of the device, it was apparent that the indenter would be required to move a considerable distance when applying forces to the spine (up to 25 mm), and that few potentiometers would possess sufficient accuracy over this range of travel. Also the design of the mechanical head of the device would need to be large in order to incorporate such a potentiometer. In the final version of the device an optical encoder was used to measure displacement of the indenter.

Earlier in the design of the device, a motor operating in one direction only was to be used to move the indenter, with the upward and downward movement of the indenter being produced via a mechanical arrangement. Such an arrangement would however, have necessitated a much larger mechanical head and also created difficulties in accurately limiting the applied force and changing the cycling frequency. When performing PA stiffness assessment clinically, physiotherapists will sample using
varying frequencies of oscillation up to approximately 5 Hertz (Lee et al 1996). Therefore, it was decided to use a small servomotor that was electronically reversible to provide both upwards and downwards movement of the indenter, and also to enable the force limits and frequency of cycling to be electronically determined.

When designing the rigid testing table to which the mechanical head was attached a surface was needed that would be sufficiently rigid for stiffness testing but light enough to transport. The original design included a testing table made from carbon fibre, however the cost to manufacture such a surface was sufficiently prohibitive that this idea was abandoned. A rigid plywood surface was used in the final device.

All components housed in the mechanical head of the final device are shown in Figure 2.3. A small reversible servomotor\(^1\) (Airpax, model no:9904 120 13316) (M) is stepped down via 2 pulleys to reduce indenter speed and increase motor torque output. This stepdown enables a maximum force production of 300 Newtons at a frequency of 2 Hertz, or lower maximum forces at higher frequencies. An inextensible cable travels around the lower pulley (P) and is attached to the top of the indenter. Movement of this cable produces movement of the indenter pad, the reversibility of the motor producing both upwards and downwards movement.

\(^1\) Philips Electronics, 15 Blue St, North Sydney, NSW, Australia.
Fig 2.3 Schematic diagram of the components of the mechanical head.
Displacement of the indenter is calculated by measuring the rotation of the pulley (P) using an optical encoder² (Hewlett Packard, HEDS-5500) (E), then converting this angular displacement to linear displacement of the indenter. The shaft of the encoder is situated in the centre of pulley (P). One revolution of the pulley/encoder shaft equates to 18.8 mm of linear displacement. The resolution of the encoder (equivalent to 0.009mm) makes it suitable for measuring the small displacements commonly encountered in assessment of PA stiffness (up to 22 mm).

Force is indirectly measured using a conductive plastic linear potentiometer³ (Sakae, Model 15FLP15A) (LP). As the indenter moves downwards, applying a force to the patient, an equal and opposite force is provided upwards by the prone patient. This force causes compression of the spring (S). The amount of spring compression is measured by the linear potentiometer (LP), 1mm of movement corresponding to 100 Newtons of force. Errors in the displacement reading that are brought about by compression of the spring are corrected by using a calibration procedure given below.

The signal from the linear potentiometer is input to the CPU. Here the signal undergoes (i) low pass filtering to remove any frequencies above 20 Hertz, ie. mains interference, and (ii) analogue-to-digital conversion, using a custom-designed card. These data are input to the computer via the serial port, along with the digital displacement data. Five-hundred-and-twelve data points per cycle are collected regardless of the frequency of oscillation of the device.

² Hewlett Packard, 17-23 Talavera Rd, North Ryde, NSW, Australia.
³ R.S. Components Pty Ltd, 129 Beaconsfield St, Silverwater, NSW, Australia.
Operation of the stiffness device is controlled by a laptop computer (Toshiba 386SX Model T3200) via the CPU. The unit is driven using custom-made software written in C++. Using the computer menu it is possible to move the indenter to make contact with the skin, initiate cycling of the indenter, and retract the indenter. The number of cycles, the frequency of cycling and the maximum force and displacement achieved during cycling can also be specified. Movement of the indenter is displacement-controlled, that is, the operator specifies the desired amount of indenter movement via the computer menu. The appropriate displacement for a particular test is established with a preliminary test, wherein the displacement is increased until a selected target maximum force (e.g. 100 Newtons) is achieved. The displacement required to achieve the target force is recorded and, in subsequent testing, the indenter is programmed to move between a minimum value and the target displacement. Should the maximum force be reached prior to the set displacement being achieved, the indenter will stop at this level of displacement until the normal time of indenter reversal. This ensures that a force greater than the maximum force will not be applied.

**Calibration**

Calibration of the device was performed in the following manner. Force data were calibrated by placing an electronic scale (Data Hytech⁴, Australia Post, AP 025, accurate to 0.001 kgs) under the indenter.

---

⁴Data Hytech, Hytech Scales Pty Ltd, 3 Reserve St, Preston, Victoria, Australia.
The stiffness device was programmed to apply a specified force and the corresponding force reading was recorded from the electronic scales. This was performed for forces from 5 up to 120 Newtons. Displacement was calibrated using a dial indicator (Mitutoyo\textsuperscript{5}, Model no: 2050-11), the indenter being moved to a specified displacement and the corresponding dial indicator reading measured to the nearest 0.01mm.

Using a least-squares criterion, regression equations were developed for calibration. These equations were:

Displacement: \[ D = 0.406 + 0.931D_2 \]
where \( D = \) calibrated displacement
and where \( D_2 = \) uncalibrated displacement

Force: \[ F = 2.96 + 0.945F_2 \]
where \( F = \) calibrated force
and where \( F_2 = \) uncalibrated force

Force/displacement data collected using the device were adjusted using these equations prior to data analysis.

**Data analysis**

Throughout each cycle, force and corresponding displacement data are collected, converted using the above calibration equations, and a graphic display of this

\textsuperscript{5}Model No: 2050-11, Mitutoyo Corp, 31-19 Shiba5-chome, Minato-Ku, Tokyo 108, Japan.
converted data provided on the computer screen. A Fortran computer program written for the purpose is then used to analyse the data. This analysis can be performed for an individual force/displacement cycle or for the average of several cycles. The starting point of each cycle is located and least square regression lines fitted to each loading phase between a range of minimum forces and a maximum force of 100 Newtons. By inspecting the $r^2$ values corresponding to the different minimum forces, a non-linear response at low force levels can be detected. The operator then selects the force range and, if more than one cycle is to be used, the desired cycles for averaging. Linear regression is then performed over the selected range of forces using the average force/displacement curve for the loading cycle. Average displacements corresponding to a range of forces up to 30 Newtons are also calculated.

The shape of the force/displacement curve, and hence the analysis, varies depending on the materials being tested. When testing an elastic beam, the force/displacement relationship is completely linear. The stiffness coefficient is obtained by calculating the slope of this force/displacement line. When testing the lumbar spine, the force/displacement curves for the loading cycle usually show a non-linear toe region, where small forces produce relatively large displacements, and a more linear region of higher stiffness for forces above 20-30 Newtons (see Figure 2.4). There is also a difference between the loading and unloading curves not evident when testing an elastic beam, and the curves exhibit time-dependent behaviour.
Fig 2.4 Typical force/displacement loading and unloading curves obtained when testing the lumbar spine.

In an attempt to evaluate the degree of linearity of the loading curve above 30 N, both straight lines [Force = b₀ + b₁(displacement)] and parabolas [Force = a₀ + a₁(displacement) + a₂(displacement)²] were fitted to force/displacement curves obtained from lumbar spine tests. The goodness of fit for both the straight line and parabola, indicated by the coefficient of determination r², was compared to determine which type of function showed the superior fit. Data was analysed using Sigma Plot software to obtain lines and parabolas fitted by the method of least squares to the force/displacement data above 30 N (see Figures 2.5 and 2.6). The results of this

---

6Jandel Scientific Software, PO Box 7005, San Rafael, CA 94912-7005
analysis demonstrated that the data was highly linear at levels of force above 30 N, with a straight line accounting for at least 99.7% of the variance in the data (see Figure 2.5). The parabolic fit (see Figure 2.6) was slightly better than the straight line fit, with greater than 99.9% of the variance accounted for in all cases. However, due to the ease in interpreting the straight line response (rather than the parabola) i.e. calculating the single coefficient ‘K’ (the gradient of the straight line) to represent the stiffness of the movement, fitting a straight line to the data above 30 N appears well-justified. This approach has been adopted by other researchers working on the quantification of stiffness (Sakai et al 1995, Horikawa et al 1993).

Fig 2.5 Force/displacement data for loading curve for typical subject with fitted straight line above 30N.
Fig 2.6. *Force/displacement data for typical subject with fitted parabola above 30N.*

The linear portion of the force/displacement curve relates to the resistance perceived after the initial phase of the PA movement, and this has been the parameter most frequently investigated in studies evaluating PA stiffness in the spine (Lee et al 1993b, Lee et al 1993c). These studies report the stiffness coefficient ‘K’, calculated as the slope of a regression line fitted to this linear region of the curve after 30 N. However, data from the early part of the loading cycle may also be important in judgments of abnormal PA stiffness. Analysis of this region of the force/displacement curve was therefore also performed. This region was quantified by calculating the
displacement of the indenter from 0.5 Newtons to 30 Newtons, and is accordingly called the 'D30' measure. The starting point for measurement of 0.5 Newtons rather than 0 Newtons was chosen as this is the contact force applied prior to testing the selected vertebral level.

Safety features of the device

To ensure that the device always applies a force within the limits of safety, the following features have been incorporated to prevent or minimise the effects of operator or machine error.

To prevent operator error

As mentioned previously, the device moves the indenter up and down a predetermined displacement for a set number of cycles. If the displacement established from the preliminary test is entered incorrectly by the operator then the force may exceed that which was desired. The same effect may occur if the subject moves between the setup and completion of the test. To prevent these errors occurring and to prevent any consequent application of a larger-than-desired force, the maximum force possible is also specified. If this maximum force is reached prior to the set displacement being achieved, the indenter will stop at this level of displacement until the normal time of indenter reversal. When testing the lumbar spine the default setting for maximum force is 105 Newtons, ensuring that if no other value is set, the maximum force delivered will never exceed 105 Newtons, a value well
within the range of forces typically applied by physiotherapists during mobilisation 
(Lee and Moseley 1991).

To prevent machine error

The load that is applied to the subject is measured by a spring and linear 
potentiometer. Should the load be measured incorrectly then the load applied to the 
subject will be likewise incorrect. While the potentiometer is unlikely to fail, the 
wiring between it and the central processing unit (CPU) may. To protect the patient 
from the effect of this machine error, circuitry is included in the device to verify that 
all three wires to the potentiometer are intact. If an error is detected, the local CPU 
shuts down the unit. In this case the indenter is free to move upwards from simply 
the force provided by the prone patient.

A second method of protection against excessive load is provided by the analog-to-
digital converter (ADC). If the digital value of the load is excessive the unit will be 
switched off and the indenter can be pushed upwards by the patient.

The final protection mechanism involves the power supply to the drive motor. The 
main drive motor has a relay in series with it. This relay is closed when the power is 
first applied and, if it is opened due to a machine fault, it cannot be closed again 
without turning the power supply off and on. There is also an 'emergency stop' 
button on the head assembly that directly controls the relay.
Contraindications to testing using the stiffness assessment device

This stiffness assessment device should not be used to test subjects for whom manual posteroanterior stiffness testing is contraindicated. These contraindications have been described in detail elsewhere (Grieve 1984, Maitland 1986) and include metastatic, inflammatory or infective disease processes affecting the lumbar spine, osteoporosis, spondylolythesis and the presence of signs of acute nerve root compromise. In some subjects this device may not be useful for measuring PA stiffness. For example, in subjects with hypertrophied erector spinae muscles, it may be difficult to place the indenter pad over the spinous process due to the fact that the spinous process sits deeply within a muscle gutter, bounded by the erector spinae. Also, some subjects with LBP may be unable to tolerate the application of a sufficient PA force necessary to perform the stiffness analysis. These subjects may be identified prior to testing with the device, when the PA force is manually applied.

DISCUSSION

The portable stiffness assessment device described here has the potential to be useful as a research tool for investigating the significance of altered tissue compliance in subjects with spinal pain, and for evaluating the effect of manual physiotherapy treatments on PA stiffness of the spinal regions. The portability of the device will facilitate research into symptomatic populations, these populations having been
previously unobtainable due to difficulties in encouraging subjects to attend the university departments where the previous generation of mechanical stiffness testing devices were housed.

The factors contributing to the measures of PA stiffness obtained using the device have not been conclusively established, but are likely to include inputs from both spinal and extraspinal structures. Lee (R. Lee 1995) applied a PA force to an L23 cadaver motion segment and examined the effects of sequential dissection on the movements produced. He found that the intervertebral disc provided the greatest resistance to the extension movement produced by a PA force, followed by the inter- and supraspinous ligaments and the zygapophyseal joints. The ligamentum flavum and capsular ligaments provided only slight resistance to the extension moment produced. Extraspinal structures contributing to measures of PA stiffness would include the skin and subcutaneous tissue overlying the spinous process to which the PA force is applied, the abdominal contents, the pelvis and ribcage. There is a paucity of information regarding the relative contribution of these extraspinal structures to PA stiffness at L3, although it has been demonstrated that resistance to anterior pelvic rotation will contribute to PA stiffness measured at L3 (Lee et al 1994b).

Little information is currently available about the effect on the PA response of factors likely to be clinically significant, such as intervertebral joint mobility. Should it be demonstrated that the sensitivity of the PA response to clinically significant variables is low, then measurement of PA responses with a mechanical device such as the one
described here, rather than with the less sensitive manual methods, may be essential to
detect abnormalities.

An entire domain of research has been made possible by the development of this
device. Because it provides a ‘gold standard’ against which manually-determined
estimates of spinal stiffness can be evaluated, methods for improving the accuracy of
spinal stiffness judgments through training and feedback can now be assessed. In
addition, normative data from initial studies has provided critical information about
the range of stiffness values represented by human spines, allowing training and
stiffness discrimination studies (eg. Maher and Adams 1996a,b) to be carried out in
the relevant stiffness region.

In conclusion, although the design of this portable stiffness assessment device makes
it suitable for assessing patients with spinal pain, the real quality of the device can
only be determined by evaluating the accuracy and reliability of the device. The
device appears to have good face validity as it closely replicates the clinically-used,
manual assessment procedure for rating PA lumbar stiffness. The issue of whether
manual or mechanical assessments of PA stiffness actually reflect properties of the
spine relevant to spinal disorders remains unknown, so an investigation of the
relationship between measures of PA stiffness and LBP is needed.

A prior consideration, though, is to evaluate the reliability of the device for measuring
PA lumbar stiffness in a group of subjects with LBP.
APPENDIX A

Circuit diagram of the stiffness assessment device.
CHAPTER 3

RELIABILITY EVALUATION OF A NEW DEVICE FOR MEASURING RESPONSES TO POSTEROANTERIOR FORCES IN A PATIENT POPULATION

The work contained in this chapter has been published as:

ABSTRACT

Background and Purpose: The purpose of this study was to determine the test-retest reliability of measurements of the human spine’s responses to posteroanterior (PA) forces made using a recently-developed, portable stiffness measuring device.

Subjects: Twenty-two subjects with non-specific low back pain (LBP) participated.

Methods: Two indices of the response to PA forces; the length of the low stiffness or ‘toe’ region (D30), and the slope of the linear portion of the force/displacement curve (K) were measured on two occasions. The reliability of repeat measurements was evaluated by intraclass correlation coefficients (ICC) and percent agreement scores. The presence of a systematic change in the response to PA forces was evaluated by a paired-samples t test. Results: The mean ‘K’ values for L2, L3, L4, L5 were 12.78N/mm, 17.46N/mm, 15.94N/mm, 15.14N/mm respectively. The mean ‘D30’ for L2, L3, L4, L5 were 3.99mm, 4.83mm, 5.71mm, 5.54mm respectively. The ICC for ‘K’ was 0.96, with the test-retest error less than 1.8N/mm on 90% of occasions. The ICC value for ‘D30’ was 0.89 with the test-retest error less than 1.1mm on 90% of occasions. For both indices of the response to PA forces, the t test revealed no systematic change with repeated testing. Conclusion and Discussion: In contrast to manual examination, which provides estimates of PA stiffness of low reliability, the new device provides highly reliable measurement of the lumbar PA force/displacement relationship. It is suggested that the device may therefore prove useful in evaluating manipulative physiotherapy theories that have suggested a relationship between spinal symptoms and abnormal PA stiffness.
INTRODUCTION

The cost to the community of low back pain (LBP) has increased dramatically over recent years, at a rate disproportionate to all other illnesses (Frymoyer 1992). Despite the frequency of this condition, there is often great difficulty in establishing the underlying cause of symptoms. Many studies have found a poor correlation between structural abnormalities and LBP. Videman et al (1990), in their study of cadaveric specimens, found no correlation between previous history of LBP and presence of lumbar spine disease while Deyo et al (1994) report that 10-30% of an asymptomatic population may demonstrate pathological findings on radiographic investigation.

Because of the difficulty in formulating an accurate structural or pathological diagnosis, physiotherapists usually treat according to the patient's presenting signs and symptoms. One physical sign which clinicians have associated with LBP is increased lumbar posteroanterior (PA) stiffness (Grieve 1984, Maitland 1986). It has been hypothesized, although not demonstrated, that PA stiffness may be increased in individuals with LBP due to degeneration of articular structures, shortening of joint capsules and adjacent ligaments (Akeson et al 1980), or to the presence of paravertebral muscle spasm (Maitland 1986).

Physiotherapists have traditionally assessed lumbar PA stiffness using manual tests such as the PA mobilization, whereby the therapist applies an oscillating force to the spinous process of the prone patient (Maitland 1986). Stiffness is perceived by the therapist, probably by a combination of estimating the amount of force applied and
estimating the amount of displacement that results. This information is then used in conjunction with information from other examination procedures to select the spinal region to be treated and the most appropriate treatment technique. However, these important clinical decisions are based on manual testing procedures which have been shown to be unreliable (Matyas and Bach 1985, Maher and Adams 1994, Binkley et al 1995) and therefore error prone.

In an attempt to improve the reliability and accuracy of stiffness measurement, two new mechanical devices have been designed to collect both force and displacement data during a simulated PA mobilization (Lee and Svensson 1990, Lee and Evans 1992). One drawback has been that both these devices are large and non-portable, making clinical examination of a symptomatic population difficult. As described in Chapter 2, a portable device was developed to measure PA stiffness. This device may be easily transported, thereby enabling access to symptomatic groups.

This chapter reports on the test-retest reliability of the portable stiffness assessment device in measuring responses to lumbar PA forces in a group of subjects with LBP. A description of the accuracy testing is also included.
METHOD

Subjects

Twenty-two subjects (10 females, 12 males) with non-specific LBP volunteered to participate in the reliability study. Subjects with current LBP were included in the study if they reported pain in the last 24 hours and pain when the examining physiotherapist manually applied a PA force to the lumbar spine. Subjects were excluded if they had any contra-indications to PA stiffness testing such as a known inflammatory, metastatic, or infective disease process, osteoporosis, spondylolysis or signs of nerve root compromise.

Subjects had a mean age of 27.4 years and had experienced pain for an average of 22.4 months (SD=37.95). The mean McGill Pain Score\(^1\), reflected by the pain rating index (PRI) (Melzack 1975) of the group was 11.04 (SD=7.32). This value represents a lower score than that found by Melzack et al (1980) in their investigation of the effect of different treatment strategies on LBP where the mean PRI score was 27.8 (SD not reported). However the subjects used in Melzack's study had a history of chronic LBP unresponsive to conventional treatment and had been referred to the pain clinic where the study was conducted.

\(^1\)The McGill Pain Questionnaire was developed to measure the sensory, affective and evaluative dimensions of pain. The main component of the questionnaire, the pain rating index (PRI), is comprised of 20 groups of pain descriptors. The patient circles those descriptors that describe their pain. Each descriptor is assigned a score depending on the order in which they appear in the group. The pain score or index is calculated by adding the numerical value of the descriptors circled by the patient.
The lower mean score obtained in the current study may reflect the fact that patients were obtained through a private physiotherapy practice rather than a pain clinic. Mean PRI scores for other pain syndromes such as cancer pain, phantom limb pain and post herpetic neuralgia have been reported as 26.0, 25.0, and 22.6 respectively (Melzack 1975).

In this study, some subjects had constant and others intermittent LBP. Subjects are described in detail in Table 3.1. Stiffness assessment was performed at the most painful lumbar spinal level with 15 subjects tested at L₅, 4 subjects at L₄, 1 subject at L₃ and 2 subjects at L₂. Informed consent was obtained prior to testing.
Table 3.1. Subject characteristics.

<table>
<thead>
<tr>
<th>Subject</th>
<th>Gender</th>
<th>Age (yrs)</th>
<th>McGill Pain Score PRI* (0-40)</th>
<th>Low back pain duration (months)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>F</td>
<td>37</td>
<td>3</td>
<td>24.00</td>
</tr>
<tr>
<td>2</td>
<td>F</td>
<td>31</td>
<td>21</td>
<td>1.25</td>
</tr>
<tr>
<td>3</td>
<td>M</td>
<td>31</td>
<td>32</td>
<td>18.00</td>
</tr>
<tr>
<td>4</td>
<td>F</td>
<td>58</td>
<td>7</td>
<td>1.25</td>
</tr>
<tr>
<td>5</td>
<td>M</td>
<td>24</td>
<td>6</td>
<td>4.00</td>
</tr>
<tr>
<td>6</td>
<td>M</td>
<td>46</td>
<td>16</td>
<td>120.00</td>
</tr>
<tr>
<td>7</td>
<td>F</td>
<td>25</td>
<td>8</td>
<td>24.00</td>
</tr>
<tr>
<td>8</td>
<td>M</td>
<td>31</td>
<td>3</td>
<td>96.00</td>
</tr>
<tr>
<td>9</td>
<td>M</td>
<td>20</td>
<td>10</td>
<td>3.00</td>
</tr>
<tr>
<td>10</td>
<td>F</td>
<td>23</td>
<td>16</td>
<td>12.00</td>
</tr>
<tr>
<td>11</td>
<td>F</td>
<td>26</td>
<td>5</td>
<td>3.00</td>
</tr>
<tr>
<td>12</td>
<td>F</td>
<td>29</td>
<td>13</td>
<td>3.00</td>
</tr>
<tr>
<td>13</td>
<td>M</td>
<td>20</td>
<td>2</td>
<td>0.25</td>
</tr>
<tr>
<td>14</td>
<td>F</td>
<td>22</td>
<td>9</td>
<td>6.00</td>
</tr>
<tr>
<td>15</td>
<td>M</td>
<td>25</td>
<td>15</td>
<td>120.00</td>
</tr>
<tr>
<td>16</td>
<td>M</td>
<td>27</td>
<td>10</td>
<td>0.10</td>
</tr>
<tr>
<td>17</td>
<td>F</td>
<td>19</td>
<td>23</td>
<td>0.10</td>
</tr>
<tr>
<td>18</td>
<td>M</td>
<td>24</td>
<td>13</td>
<td>36.00</td>
</tr>
<tr>
<td>19</td>
<td>F</td>
<td>21</td>
<td>6</td>
<td>2.00</td>
</tr>
<tr>
<td>20</td>
<td>M</td>
<td>20</td>
<td>5</td>
<td>0.25</td>
</tr>
<tr>
<td>21</td>
<td>M</td>
<td>20</td>
<td>10</td>
<td>18.00</td>
</tr>
<tr>
<td>22</td>
<td>M</td>
<td>24</td>
<td>10</td>
<td>2.00</td>
</tr>
</tbody>
</table>

| Mean    | 27.40  | 11.04    | 22.46                         |
| SD      | 9.37   | 7.32     | 37.95                         |

* PRI = Pain rating index
Equipment

A portable stiffness device was developed to specifically measure posteroanterior responses in the lumbar spine. Under the control of the experimenter this device, previously described in Chapter 2, applies a PA force to the spine and measures the amount of resultant displacement. The force and displacement data are then used to calculate two indices of the response to PA forces.

Assessment of the accuracy of the device

In order to assess the accuracy of the stiffness assessment device, three aluminium beams of varying stiffness were tested with (i) the new device, and (ii) by applying known forces to the beams and accurately measuring the resulting displacement.

Method

Three regular aluminium channel sections of approximately 1.8 m each were tested. The beams were supported on a steel frame for testing. Stiffness of the beams was firstly established by using 5 dead weights (measured using an electronic scale\(^2\) to 0.001 kg) to apply known forces to the beam, and measuring the resultant displacements as each weight was applied (figure 3.1). Weights were applied in increments to a total of 100 N. The displacements were measured using a dial indicator\(^3\). From these force/displacement data the stiffness of the beams was

\(^2\)Data Hytech, Hytech Scales Pty Ltd, 3 Reserve St, Preston, Victoria, Australia.
\(^3\)Model No: 2050-11, Mitutoyo Corp, 31-19 Shibas-chome, Minato-Ku, Tokyo 108, Japan.
calculated. The stiffness values of the three beams were similar in magnitude to the stiffness values encountered when testing the lumbar spine.

![Image of accuracy testing](image)

**Fig 3.1** *Accuracy testing: measurement of stiffness of aluminium beams using weights and dial indicator.*

The beams were then tested using the new stiffness assessment device (figure 3.2). Force/displacement data were collected for 1 cycle at 0.5 Hertz, to a maximum force of 100 N. The data were adjusted using the regression equations described in the previous chapter. The data were low-pass-filtered with a cut-off frequency of 20 Hertz. Linear regression was then performed on the force/displacement relationship.
The linear stiffness of each beam was calculated by determining the gradient of the force/displacement line.

![Image of a testing rig](image)

**Fig 3.2 Accuracy testing: measurement of PA stiffness using portable stiffness assessment device.**

**Results**

Table 3.2 shows the stiffness values obtained for each beam using the dead weights and dial indicator, the stiffness values obtained using the new stiffness assessment device, and the percentage measurement error. It can be seen that the difference in the stiffness values generated by the two methods is very small. The measurement device marginally underestimates the true stiffness for all measurements, with a
maximum error of 2.5%. Thus a high level of accuracy can be attributed to the stiffness assessment device.

Table 3.2. K values (N/mm) obtained using weights and dial indicator (Method 1) and the new stiffness device (Method 2), and the percentage error.

<table>
<thead>
<tr>
<th>Beam</th>
<th>Method 1</th>
<th>Method 2</th>
<th>%error</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>6.91</td>
<td>6.86</td>
<td>0.7</td>
</tr>
<tr>
<td>2</td>
<td>7.55</td>
<td>7.36</td>
<td>2.5</td>
</tr>
<tr>
<td>3</td>
<td>15.9</td>
<td>15.65</td>
<td>1.5</td>
</tr>
</tbody>
</table>

Procedure

To investigate the test-retest reliability of the stiffness device two measurements of PA responses were performed on each subject on a single day. All subject testing was conducted by the same examiner (JL) who had performed over 20 tests using the new stiffness assessment device prior to commencing this reliability study.

Prior to testing, anthropometric data and a brief history of LBP was collected from all subjects. Subjects were then asked to undress to their underwear and put on a hospital gown. This precaution was taken to ensure that heavy clothing such as jeans or belts did not produce a more extended lumbar posture than normal and hence alter stiffness between tests. PA testing was then performed as follows. Subjects were asked to lie prone on the testing table while the examiner manually applied a PA force
to each lumbar level. The most painful spinal level was identified and marked. This level was defined as either the level that best reproduced the subject’s presenting pain, or the level with the greatest amount of local pain which could logically refer pain to the symptomatic area. The same markings were used for both the initial and repeat test. The position adopted by the subject during both tests was standardized by having the subject lie with arms by the side and lower legs resting on a pillow.

The mechanical head of the stiffness assessment device was tilted in either a cephalad or caudad direction, depending on the level to be tested, to enable the force to be applied normal to average vertebral body axis directions (Stagnara et al 1982). The reason for directing the force in such a manner was to mimic the manual clinical assessment of PA stiffness as performed by physiotherapists. The directions used for L2, L3, L4 and L5 were cephalad 5.5 degrees for both L2 and L3, and caudad 4.5 degrees and caudad 16 degrees for L4 and L5 respectively. These directions have been employed in previous research (Lee and Liversidge 1994a) as they are within the range of the directions of force application used by experienced physiotherapists (Viner and Lee 1995). The indenter was then moved into contact with the skin at the level to be tested.

At least 5 preliminary cycles were performed to establish the appropriate displacement for the individual subject, to familiarize the subject with the measurement procedure and to precondition the viscoelastic spinal tissues (Loebl 1972) prior to data collection. It has been demonstrated that because of the viscoelastic nature of biological tissue, deformation of tissue in response to an applied
force may be affected by forces previously applied. For example, Loebi (1972), in his study of cyclical abduction loading of the metacarpophalangeal joints, found that stress relaxation of the tissues could not recover before the next cycle was begun. Therefore consecutive cycles commenced from a greater initial abduction displacement value than had the first cycle. Also, the terminal phase slope of the last force/displacement cycles were steeper than the first suggesting that the first cycles serve to increase the elastic stiffness of the metacarpophalangeal connective tissue structures. Loebi (1972) demonstrated that the greatest change in displacement occurred between cycles 1 and 2 and by cycle 5 the force/displacement curves appeared repeatable. A more recent study examining the rheological properties of human lumbar spine ligaments has demonstrated that preconditioning decreases the hysteresis area of the ligament, with the greatest changes occurring within the first 5 cycles (Yahia et al 1991).

Therefore, in the current study the decision was made to precondition the spinal tissues by applying at least 5 cycles of PA forces over the vertebral segment to be tested prior to collecting data.

Following initiation via the computer keyboard, the indenter then proceeded to apply cyclical loading forces, data being collected for five loading cycles to a force of 105 Newtons at a frequency of 0.5 Hertz. This magnitude of force and the frequency of cycling were chosen to try and replicate the manual assessment of PA stiffness where mobilization can be applied at a range of frequencies from 0.5-3 Hz. Cycling at 0.5 Hertz was deemed comfortable by several subjects and hence testing was conducted
using this cycling frequency. During data collection all subjects were asked to hold their breath at functional residual capacity in order to avoid confounding effects due to lung volume (Keaveney et al 1989) and phase of respiration (Beaumont et al 1991).

After a period of 5 minutes the subject was retested. Between tests the mechanical head was removed and the subject lay prone on the testing table. The indenter was then repositioned over the tested level and the procedure repeated exactly.

Data analysis

For each test occasion five loading cycles were obtained. However, the force/displacement data for four loading cycles only (cycles 2-5) were analyzed. Cycle 1 was discarded on the grounds that the data from it were likely to be due to transient behaviour on start-up because the indenter began from a stationary position. The remaining four cycles were averaged and a mean force/displacement curve obtained.

Two measures were then calculated to quantify the main characteristics of the force/displacement curve obtained when measuring the lumbar spine. These characteristics were first, the length of the non-linear toe region, where small forces produce relatively large displacements, and second, the gradient of the linear region of higher stiffness for forces above 30 Newtons. In previous studies (Lee et al 1993b, Lee et al 1993c) evaluating PA stiffness in the spine, only the linear portion of the
force/displacement curve has been investigated. These studies report a coefficient of stiffness ‘K’ calculated as the slope of a regression line fitted to the linear region. However, the non-linear part of the loading cycle may also be important in judgments of abnormal PA stiffness and therefore was also analyzed in this study. The non-linear region was quantified by calculating the displacement of the indenter from 0.5 Newtons to 30 Newtons. Thus the two measures calculated in this study were the stiffness coefficient ‘K’ and the displacement from 0.5 to 30 Newtons (D30).

Statistical analysis

Intraclass Correlation Coefficients ICC (2,1) (Shrout and Fleiss 1979) were calculated for the stiffness measures to determine the test-retest reliability. Due to the small number of subjects tested at L2-4, ICC’s were calculated by combining the data from all levels. The ICC index reaches a maximum of 1.0 when raters are in complete agreement with 1.0 - ICC representing the percentage of variance due to error, some of which is from disagreement between repeat measurement occasions. To detect any systematic changes in stiffness due to repeated testing, a t test for paired samples was carried out to compare the initial and second measurements. The significance level was set at $p < 0.05$. 
RESULTS

The mean ‘K’ values for L₂, L₃, L₄, L₅ were 12.78N/mm, 17.46N/mm, 15.94N/mm, 15.14N/mm respectively (Table 3.3). High test-retest reliability was observed for measurements of the slope of the linear region of the force displacement curve, ‘K’, the ICC (2,1) value being 0.96. The ICC 95% confidence interval was 0.91 - 0.98. On 90% of occasions the test-retest difference was less than 1.8N/mm. No significant difference was found between the initial and second measurements, indicating that there was no systematic effect produced by repeat testing (p=0.82).

The mean displacement ‘D30’ values for L₂, L₃, L₄, L₅ were 3.99mm, 4.83mm, 5.71mm, 5.54mm respectively (Table 3.3). The ICC value obtained was high (ICC (2,1) = 0.89). The ICC 95% confidence interval was 0.76 - 0.95. On 90% of occasions the test-retest difference was less than 1.1mm. No significant difference was found between the initial and second measurements, indicating that there was no systematic effect on the displacement produced by repeated testing (p=0.32).
Table 3.3. K (N/mm) and D30 (mm) values for Test 1 and Test 2 for each subject.

<table>
<thead>
<tr>
<th>Level</th>
<th>K Test 1</th>
<th>K Test 2</th>
<th>D30 Test 1</th>
<th>D30 Test 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>L₂</td>
<td>10.84</td>
<td>10.44</td>
<td>3.49</td>
<td>3.55</td>
</tr>
<tr>
<td></td>
<td>14.72</td>
<td>12.66</td>
<td>4.49</td>
<td>4.35</td>
</tr>
<tr>
<td>Mean</td>
<td>12.78</td>
<td>11.55</td>
<td>3.99</td>
<td>3.95</td>
</tr>
<tr>
<td>L₃</td>
<td>17.46</td>
<td>17.30</td>
<td>4.83</td>
<td>4.45</td>
</tr>
<tr>
<td>Mean</td>
<td>17.46</td>
<td>17.30</td>
<td>4.83</td>
<td>4.45</td>
</tr>
<tr>
<td>L₄</td>
<td>8.87</td>
<td>8.95</td>
<td>8.28</td>
<td>8.53</td>
</tr>
<tr>
<td></td>
<td>26.65</td>
<td>27.71</td>
<td>3.58</td>
<td>3.35</td>
</tr>
<tr>
<td></td>
<td>12.39</td>
<td>12.48</td>
<td>6.04</td>
<td>5.56</td>
</tr>
<tr>
<td></td>
<td>15.83</td>
<td>15.43</td>
<td>4.96</td>
<td>6.04</td>
</tr>
<tr>
<td>Mean</td>
<td>15.94</td>
<td>16.14</td>
<td>5.71</td>
<td>5.87</td>
</tr>
<tr>
<td>L₅</td>
<td>12.94</td>
<td>13.43</td>
<td>5.84</td>
<td>6.13</td>
</tr>
<tr>
<td></td>
<td>15.06</td>
<td>13.77</td>
<td>6.66</td>
<td>6.64</td>
</tr>
<tr>
<td></td>
<td>14.56</td>
<td>14.80</td>
<td>5.05</td>
<td>5.13</td>
</tr>
<tr>
<td></td>
<td>14.12</td>
<td>13.82</td>
<td>4.67</td>
<td>4.98</td>
</tr>
<tr>
<td></td>
<td>12.22</td>
<td>11.46</td>
<td>5.85</td>
<td>5.81</td>
</tr>
<tr>
<td></td>
<td>11.36</td>
<td>11.31</td>
<td>5.55</td>
<td>6.87</td>
</tr>
<tr>
<td></td>
<td>15.26</td>
<td>15.40</td>
<td>5.21</td>
<td>5.02</td>
</tr>
<tr>
<td></td>
<td>19.47</td>
<td>18.77</td>
<td>4.40</td>
<td>3.99</td>
</tr>
<tr>
<td></td>
<td>19.90</td>
<td>18.11</td>
<td>2.67</td>
<td>3.19</td>
</tr>
<tr>
<td></td>
<td>14.15</td>
<td>14.66</td>
<td>5.70</td>
<td>5.50</td>
</tr>
<tr>
<td></td>
<td>11.09</td>
<td>10.86</td>
<td>8.2</td>
<td>8.08</td>
</tr>
<tr>
<td></td>
<td>14.98</td>
<td>14.2</td>
<td>4.61</td>
<td>4.90</td>
</tr>
<tr>
<td></td>
<td>17.12</td>
<td>16.76</td>
<td>5.61</td>
<td>5.81</td>
</tr>
<tr>
<td></td>
<td>22.50</td>
<td>24.07</td>
<td>5.57</td>
<td>5.52</td>
</tr>
<tr>
<td></td>
<td>12.32</td>
<td>16.09</td>
<td>7.58</td>
<td>7.56</td>
</tr>
<tr>
<td>Mean</td>
<td>15.14</td>
<td>15.17</td>
<td>5.54</td>
<td>5.68</td>
</tr>
</tbody>
</table>
There was generally very high consistency in both ‘K’ and ‘D30’ values between successive cycles of loading, that is between cycles 2-5 of the test. Tables 3.4 and 3.5 show typical variations in responses between cycles.

**Table 3.4. K values (N/mm) for cycles 2-5 for typical subjects at each tested level.**

<table>
<thead>
<tr>
<th>Subject</th>
<th>Cycle 2</th>
<th>Cycle 3</th>
<th>Cycle 4</th>
<th>Cycle 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 (L2)</td>
<td>11.79</td>
<td>10.76</td>
<td>10.53</td>
<td>10.37</td>
</tr>
<tr>
<td>2 (L3)</td>
<td>16.95</td>
<td>17.85</td>
<td>16.83</td>
<td>16.85</td>
</tr>
<tr>
<td>3 (L4)</td>
<td>8.81</td>
<td>8.96</td>
<td>8.88</td>
<td>8.81</td>
</tr>
<tr>
<td>4 (L5)</td>
<td>11.14</td>
<td>10.90</td>
<td>10.64</td>
<td>11.73</td>
</tr>
</tbody>
</table>

**Table 3.5. D30 values (mm) for cycles 2-5 for typical subjects at each tested level.**

<table>
<thead>
<tr>
<th>Subject</th>
<th>Cycle 2</th>
<th>Cycle 3</th>
<th>Cycle 4</th>
<th>Cycle 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 (L2)</td>
<td>3.23</td>
<td>3.46</td>
<td>3.64</td>
<td>3.63</td>
</tr>
<tr>
<td>2 (L3)</td>
<td>4.66</td>
<td>4.81</td>
<td>4.94</td>
<td>4.91</td>
</tr>
<tr>
<td>3 (L4)</td>
<td>7.96</td>
<td>8.42</td>
<td>8.31</td>
<td>8.44</td>
</tr>
<tr>
<td>4 (L5)</td>
<td>8.30</td>
<td>8.31</td>
<td>8.26</td>
<td>7.93</td>
</tr>
</tbody>
</table>
DISCUSSION

The results of this study demonstrate that measurements of responses to PA forces in subjects with LBP made using a portable stiffness assessment device (Latimer et al, 1996a) have high reliability when performed by the same examiner, given that the examiner has some experience in operating the device. While several previous studies have demonstrated high reliability using a mechanical device to measure the response to lumbar PA forces in asymptomatic subjects (Lee and Svensson 1990, Lee and Evans 1992) this study is the first to demonstrate reliable measurement in subjects with LBP, where the painful condition might have been thought to make testing less reliable.

The high reliability coefficients obtained for the stiffness coefficient ‘K’ and the displacement ‘D30’ contrast markedly with the poor reliability found for manual assessment of PA stiffness (Maher and Adams 1994, Binkley et al 1995). While therapists have proposed a relationship between spinal symptoms and PA stiffness, the poor reliability of manual assessment has meant that this relationship could not be objectively investigated. The development of a device that reliably measures the response of the lumbar spine to PA forces may enable investigation of this proposed relationship. In addition, the portability of the new device will facilitate testing of a clinically relevant population. Also, the device may be used to provide accurate and immediate feedback when training physiotherapists to manually judge PA stiffness.

Bench testing has demonstrated that the device is accurate in its measurement of PA stiffness when tested on a series of elastic beams. However, the accuracy of the
device in measuring the response of the lumbar region to PA forces is yet to be examined, although the device appears to have good face validity as it closely replicates the manual assessment procedure. The issue of whether either manual or mechanical assessments of PA stiffness actually reflect properties of the spine relevant to spinal disorders is not yet known. Validity issues, such as whether measurements obtained using the device can discriminate between symptomatic and asymptomatic groups, and criterion-related validity, (determined by comparing measurements with those obtained from other established mechanical devices) have yet to be investigated. There is little doubt that, similar to many measures used to infer impairment in LBP sufferers (such as physiological range of movement, muscle strength, endurance, straight leg raise) PA stiffness likewise varies widely in the asymptomatic population, thereby making it difficult to determine whether an individual is symptomatic or not based solely on their PA stiffness result. Information regarding whether PA stiffness changes with the presence of LBP may be one way of differentiating symptomatic from asymptomatic subjects. (Such an investigation is reported later in this thesis).

In this study several decisions were made about how to characterize the force/displacement curve obtained when applying a PA force to the lumbar region. It is currently unclear how physiotherapists interpret force and displacement data to make decisions about stiffness. In this study two parameters were chosen which were believed to best describe the key features of the force/displacement relationship obtained when applying a PA force to the lumbar spine. These two features were the length of the non-linear toe region and the gradient of the higher stiffness linear region. While the stiffness coefficient K has been quantified by previous researchers,
this study also describes a measure of the low stiffness region, the displacement (D30) to an arbitrary low force of 30 Newtons. The good reliability for this parameter (ICC=0.89) together with its value in describing the low stiffness region of the PA response suggest that ‘D30’ may be useful when evaluating the force/displacement relationship during the application of a PA force to the spine.

CONCLUSION

The reliability of measurements of aspects of the PA force/displacement relationship using a new measurement device was determined in subjects reporting LBP. Reliability coefficients were high. Bench testing also demonstrated a high degree of accuracy of measurements made with the device on aluminium beams of varying stiffness. The clinical significance of the PA force/displacement relationship has yet to be established.
CHAPTER 4

TRAINING WITH IMMEDIATE FEEDBACK AND PHYSIOTHERAPY STUDENTS’ ABILITY TO JUDGE LUMBAR STIFFNESS

The work contained in this chapter has been published as:

ABSTRACT

Background and Purpose: One potential explanation for the poor reliability of manual judgments of lumbar posteroanterior (PA) stiffness may be that the conventional method of training students to judge stiffness using feedback provided by physiotherapy tutors is ineffective. The aim of the current study was to investigate whether immediate quantitative feedback, provided from a highly reliable mechanical device, could improve physiotherapy students’ ability to judge lumbar PA stiffness.

Subjects: Four second year physiotherapy student raters and 43 asymptomatic volunteers participated in this study. Methods: The raters assessed the PA lumbar stiffness of the asymptomatic volunteers, performing 75 stiffness judgements in total during pre-test, training and post-test sessions held over a three-week period. During the training sessions raters were provided with accurate and immediate feedback regarding each judgment of PA stiffness. Results: No significant difference in mean absolute error between the pre- and post-tests was found for the raters (p=0.31). Conclusion: Provision of information about the true PA stiffness of each lumbar spine judged did not improve the accuracy of physiotherapy students’ judgments of lumbar PA stiffness.
INTRODUCTION

Prior to treating patients with non-specific low back pain (NSLBP), physiotherapists use a range of manual tests to identify the level of the spine to be treated. One of the most common of these tests is the posteroanterior (PA) central pressure, performed in order to identify spinal segments exhibiting restricted movement or increased PA stiffness. Studies that have investigated the stiffness judgments made by physiotherapists using these tests have found the judgments to be unreliable (Matyas and Bach 1985, Maher and Adams 1994, Binkley et al 1995).

It has been suggested in Chapter 1 that the reason for the poor interrater reliability may reside in the fact that the stiffness assessment protocol taught to physiotherapy students does not attempt to control factors known to affect either the stiffness stimulus or the perception of stiffness. Also, with current teaching methods, the tutor’s perception of PA stiffness is typically regarded as the standard against which students’ judgments are compared (Keating et al 1993, Hickling and Maitland 1970, Maitland 1986). Strategies that rely on tutor feedback are unlikely to be optimal because it is now clear that the tutor’s perception of stiffness may not be accurate. It may be possible to enhance PA stiffness judgments if the feedback provided to the learner was accurate, quantitative and immediate. Such an approach to training has been shown to reduce error in specified force production when using the PA central pressure test (Keating et al 1993). In the case of learning to make stiffness judgments this could be achieved by firstly measuring the stiffness of the spine of a volunteer using a machine (described in Chapter 2) which has been shown to be reliable and
accurate (Chapter 3). The learner could then judge the stiffness and be told how close their perception of stiffness was to the accurately-measured value. Over a series of trials it would be expected that the error associated with each response would reduce, as in studies on motor learning with the provision of trial-by-trial quantitative feedback (Magill 1993).

While many studies have examined the most effective method of providing feedback, or knowledge of results (KR) to improve motor learning (eg. Schmidt et al 1989, Weinstein and Schmidt 1990), few studies have examined the usefulness of feedback in improving perceptual judgments. One perceptual judgment study in which subjects were required to make an estimation about the magnitude of a stimulus was conducted by Annett (1966) in which subjects estimated the number of dots present in a tachistoscopic field. One group estimated with no KR, the second group received KR after every response (ie: was told the correct number of dots after every estimation), and the third group was told the correct number before each trial. At post-test both the second and third groups had reduced their error similarly, however the first group did not improve beyond the initial trial. Therefore this study suggested that when attempting to improve magnitude estimation, feedback (KR) could be given after every response, or could be used to provide anchor information if given prior to each trial. In the current study the decision was made to provide feedback immediately following every trial, enabling the raters’ error to be calculated for all training sessions, as well as for the pre-and post-tests.
The aim of this study therefore, was to investigate whether immediate quantitative feedback, given after every trial, improved physiotherapy students' ability to judge lumbar PA stiffness. The experimental hypothesis was that the absolute error, representing the unsigned difference between manually-based estimates of PA lumbar stiffness made by physiotherapy students, and measurements made by a mechanical device, would decrease following training sessions where each student was provided with immediate feedback on their performance.

**METHOD**

**Research design**

A single case design (ABA) with replication was used in this study. This involved the training of four novice raters by exposing them to 75 asymptomatic lumbar spine stimuli.

The study was approved by the University of Sydney Human Ethics Committee. Informed consent was obtained from all raters and the volunteers who acted as stimuli.
Subjects

The following subjects (SA) and the examiner were involved in the study. A standard setting, calm and the examiner were involved in the study.

Raters

Four second year physiotherapy students (raters) enrolled in the BAppSc (Physiotherapy) at the University of Sydney, Australia, were trained in this study. The raters (1 male and 3 female) were novice raters, not having been previously exposed to tutorial sessions related to spinal examination. Raters who volunteered for the study were asked to provide an assurance that they would be available to complete the training program.

Stimuli

Forty-three volunteers with asymptomatic lumbar spines and with a mean age of 29.1 years (SD=7.4) were used to provide seventy-five lumbar stiffness stimuli to the raters. Thirteen volunteers were presented to the raters once only, twenty-eight volunteers were presented twice and two volunteers were presented three times during the training sessions. All volunteers were fully draped with white sheeting during pre-and post-testing and at training sessions, with only the skin overlying the lumbar spine exposed to ensure that the volunteer could not be recognised and thus remembered.

In the week prior to the commencement of the training study the asymptomatic volunteers underwent PA stiffness testing at the L3 spinal level using the stiffness assessment machine (SAM) described in Chapters 2 and 3.
PA stiffness testing was performed in the following manner. Each volunteer lay prone on a standard testing table and the L₃ spinous process was identified and marked with long-lasting ink by the researcher. They were then instructed to lie with their arms by their side and with their face firmly positioned in the face hole of the plinth. This position was reproduced exactly by the volunteers during the main training study. The stiffness assessment device was then used to measure PA stiffness. Prior to data collection, preconditioning of the viscoelastic tissues was performed by using the device to apply five oscillations up to 100 N at L₃. Immediately following this, data collection commenced, with the device applying loading forces of 100 N to L₃ while measuring the resultant displacement. The data collected by the device were then used to generate a force/displacement curve. A stiffness value at L₃ for each volunteer was obtained by calculating the slope of the regression line fitted to the linear portion of this force/displacement curve. The values obtained for the stiffness coefficient ‘K’ for this group of volunteer stimuli ranged from 9-22 N/mm with a mean of 13.64 N/mm and a standard deviation of 3.34 N/mm. These values, in Newtons per millimetre and rounded to the nearest whole number, were then used to provide feedback for the training sessions. PA stiffness appears to be stable over time in asymptomatic subjects (Latimer et al 1996c) and was therefore deemed unlikely to change from the instrument-measured value during the 2.5 weeks of the study.

**Procedure**

Prior to commencement of the study, student raters attended a one hour session in which the method of performing manual lumbar PA stiffness assessment was
explained and feedback from the researcher was given on their performance. At this session raters were also shown a force/displacement curve and told that in the training study they would be asked to attend only to the linear portion of the force/displacement curve, and that the feedback value they received in the study would be based on the measurement of this portion of the curve. Raters were also told that the mean PA stiffness value obtained in a group of asymptomatic subjects at L₃ was 14.3 N/mm, the values ranging from 7 N/mm to 25 N/mm, these values having been obtained from a pilot reliability study performed on other asymptomatic subjects at L₃.

Two days later the student raters were required to rate the PA stiffness of 10 randomly-ordered volunteers, without feedback, to serve as a baseline measure of performance. Subsequently, five training sessions with feedback were held over 2.5 weeks (ie: 2 per week, Tuesday and Thursday). At each of these sessions volunteers, randomly ordered, were assessed while lying on height-adjustable plinths. Ten stimuli were presented at Session 1, 10 at Session 2, 11 at Session 3, 12 at Session 4, and 9 at Session 5. The L₃ spinous process of each volunteer was clearly marked by the researcher so the rater would not be required to manually identify the level. In the majority of cases the ink mark from the SAM measurement remained and only needed to be highlighted, and in the few cases where it had disappeared L₃ was re-identified and remarked. At this time the researcher also manually applied posteroanterior (PA) mobilisations to precondition the vertebral segment prior to testing by the raters. Each rater was then required to manually assess the PA stiffness of the marked level,
then to tell the researcher their estimate of this value in N/mm (figure 4.1) using only integer values. Raters were able to apply as many PA forces as they believed necessary to enable them to produce the most accurate judgement. Immediately following this, the researcher would show the rater a card displaying the PA stiffness value obtained using the SAM. The rater then moved on to the next volunteer stimulus until all of the spines had been assessed.

Fig 4.1 Performance of PA stiffness testing by rater prior to receipt of feedback. The card with the measured value can be seen face down at the end of the plinth.
To examine whether learning had occurred, two days following the final training session a post-test using 13 volunteers was conducted where the raters had to assess PA stiffness without feedback. None of the stimuli presented to the raters at the post-test had been previously examined by them.

**Data analysis**

The percent exact agreement and the maximum error for the pre- and post-tests for each rater was calculated. The mean absolute error (the absolute error being the unsigned difference between the stiffness value obtained using the instrumented device and the stiffness value judged by the student) for each rater for the pre- and post-tests and for each training session was also calculated. The data for the four raters were then combined, and the mean absolute error and SE for the pre-and post-tests and the training sessions calculated. A paired t-test was then performed to determine whether there was any significant difference in the absolute error between the pre- and post-tests.

**RESULTS**

Table 4.1 shows the percentage of occasions at which the rater was in exact agreement for the pre- and post-tests and the maximum error for each test. The range of stimuli used at the pre- and post-tests ranged from 9-21 N/mm.
Table 4.1. Percentage of occasions at which the rater was in exact agreement (PEA) for the pre- and post-tests and the maximum absolute error for each test for each rater.

<table>
<thead>
<tr>
<th>Rater</th>
<th>PEA (pre-test)</th>
<th>Maximum Absolute Error (N/mm)</th>
<th>PEA (post-test)</th>
<th>Maximum Absolute Error (N/mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>8</td>
<td>16</td>
<td>8</td>
</tr>
<tr>
<td>2</td>
<td>10</td>
<td>8</td>
<td>31</td>
<td>10</td>
</tr>
<tr>
<td>3</td>
<td>10</td>
<td>5</td>
<td>8</td>
<td>7</td>
</tr>
<tr>
<td>4</td>
<td>10</td>
<td>6</td>
<td>10</td>
<td>4</td>
</tr>
</tbody>
</table>

The mean absolute error for each subject for the pre-and post tests and each training session is presented graphically in Figure 4.2. The actual mean scores appear in Table 4.2. It can be seen from the graph that the four raters followed a similar pattern, reducing their absolute error from the pre-test to training session 3, but then demonstrating an increase in error from training session 3 to the post-test. The mean absolute error and SE indicated that the group followed the same pattern (figure 4.3).

A paired t-test revealed no significant difference between the mean absolute error for the pre-and post-tests \((t=1.21, \text{df}=3, p=0.31)\).
Fig 4.2. The mean absolute error (N/mm) at each session for each rater (R1-R4).
Table 4.2. Mean absolute error and standard error (SE) (N/mm) for each rater (R1-R4) for each session. There was a pre-test (Pre), five training sessions (1-5) and a post-test (Post).

<table>
<thead>
<tr>
<th>Subj</th>
<th>Session</th>
<th>Pre</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>Post</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>3.4 (.82)</td>
<td>3.2 (.26)</td>
<td>2.8 (.84)</td>
<td>1.3 (.30)</td>
<td>2.9 (.87)</td>
<td>3.6 (.74)</td>
<td>3.23 (.72)</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>2.7 (.47)</td>
<td>2.9 (.87)</td>
<td>3.2 (.49)</td>
<td>2.09 (.16)</td>
<td>2.75 (.81)</td>
<td>2.8 (.72)</td>
<td>3.3 (.63)</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>4.0 (1.13)</td>
<td>3.4 (.91)</td>
<td>1.9 (.72)</td>
<td>1.8 (.55)</td>
<td>2.7 (.69)</td>
<td>4.4 (1.1)</td>
<td>2.5 (.76)</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>3.3 (.65)</td>
<td>3.5 (1.1)</td>
<td>2.3 (.62)</td>
<td>1.6 (.34)</td>
<td>2.83 (.64)</td>
<td>2.7 (.71)</td>
<td>1.7 (.32)</td>
</tr>
<tr>
<td>Mean</td>
<td></td>
<td>3.4 (.39)</td>
<td>3.3 (.50)</td>
<td>2.6 (.34)</td>
<td>1.7 (.18)</td>
<td>2.8 (.37)</td>
<td>3.4 (.41)</td>
<td>2.7 (.33)</td>
</tr>
</tbody>
</table>

Fig 4.3 The mean absolute error and SE (N/mm) at each session for the group of raters.
DISCUSSION

Over a series of sessions, provision of immediate and accurate feedback about the linear component of PA stiffness of the lumbar spine did not significantly improve physiotherapy students' ability to judge lumbar stiffness. In the early stages of this study, it did appear that improvement in performance was occurring, particularly by Session 3. That the power of the study was sufficient to find an overall learning effect if one existed is demonstrated by the fact that when a paired t-test was performed to compare Session 3 with the pre-test a significant reduction in the mean absolute error was demonstrated (t=4.51, df=3, p=0.02). However at Session 4 and thereafter the absolute error increased.

Several possible explanations may account for this. Firstly, debriefing of the students following the study suggested that the student raters found the task of judging the linear portion of the force/displacement curve difficult, as they were sometimes unable to separate out information they perceived from the non-linear toe region of the curve. If, after Session 3, student raters came across volunteers whose spines possessed toe region values not previously encountered, this may have disrupted the concept the raters were forming about the linear portion of the force/displacement curve.
In this study it was decided not to provide raters with feedback regarding the toe region of the curve (commonly characterised by measuring the amount of displacement occurring up to 30 Newtons (See Chapter 2)), in order to simplify the feedback provided. Studies in learning of perceptual skills have demonstrated that the provision of more complex feedback does not assist perceptual learning. For example, in an experiment on complex sound recognition in which each sound varied in five ways each with five different values, it was found that subjects given KR relevant to each of the five dimensions did less well than those who listened to the sounds and were told their names only (Swets and Sewell 1963).

Consequently a toe region value was not obtained for each volunteer, so it was not possible to determine whether the toe region values varied for different sessions and hence supported the above account. Thus one explanation is that the difficulties experienced by the raters in attempting to ignore the non-linear toe region of the curve made it hard for them to make effective use of feedback related to the elastic stiffness or linear region.

Secondly, it is possible that after Session 3, student raters came across some attributes or dimension in the volunteer spines they were assessing that greatly affected their judgments. In this study feedback was provided regarding the elastic stiffness value of the spine because this is considered to be the most critical motion aspect arising from manual clinical examination. There is support for this view, with Jull (1994b) stating that
“Manual examination can be described as a very basic in vivo test of spinal segmental mechanics, testing particularly the elastic properties of the viscoelastic spinal tissues....” (p520).

However, spines vary along other dimensions besides elastic stiffness. Maher and Adams (1995b) have suggested that the dimensions of viscosity and inertia may be important in physiotherapists’ concepts of ‘spinal stiffness’. Until all the dimensions of spinal stiffness perceived by physiotherapists are known, the spines that student physiotherapists palpate during their training are likely to be random with respect to these other dimensions. In this study it may be that the presence of an outlier value on another dimension, such as viscosity or inertia, in one of the volunteers at Session 4 affected raters’ criteria for judgments of linear stiffness, causing them to make substantial error. This conceptual confusion may also have affected the confidence of the raters adding to the increased error at Session 4 and thereafter.

Additionally, there are likely to be unknown variables that affect stiffness perception that were left uncontrolled in this training study and may have contributed to the students inability to learn. For example, the instrumented stiffness measure was obtained by performing linear regression of the lumbar force/displacement curve between 30 and 90 N. In this training study there was no attempt to measure or standardise the downward forces applied by each rater. Indeed, some of the ‘stimuli’ commented that some of the raters seemed to push quite hard, in a similar manner to that performed by the instrumented device, while others pushed more gently. The
reason no attempt was made to standardise the testing force used was that the raters were asked to assess the linear or elastic component of the force/displacement curve. It was argued that as long as the rater applied a force above 30 N they would be operating in the linear portion of the force/displacement relation. In the teaching session conducted prior to the training study they were shown a force/displacement curve and told that they would be asked to judge the linear part of the curve above 30 N. At this session they had time to practice pushing on a set of bathroom scales and were asked to feel the amount of force required to test between 30 and 100 N although they were not trained to apply this force repeatedly. If in fact non-linearity was present, and the rater used different applied loads between volunteer stimuli, or used different loads to the SAM, some error may have been introduced. (This issue is further investigated in Chapter 7 of this thesis.)

In this study the feedback given to students represented the PA stiffness value obtained when the volunteer stimuli were measured using a mechanical device. During mechanical measurement of PA stiffness, testing is performed on a rigid table to which the sensor head of the mechanical device is attached (for full description see Chapter 2). In the current study student raters assessed the PA stiffness of the volunteer stimuli on height-adjustable, minimally padded plinths. These plinths enabled the rater to vary the height of the stimulus for stiffness testing, in a manner similar to that used in clinical practice. Prior to commencing the study an assumption was made that although the PA stiffness values obtained when testing on a rigid surface versus testing on a padded plinth would be different, there would be a high
correlation between the two. Should the assumption of high correlation be incorrect then the feedback may have contained some error due to the fact that the raters assessed stimuli on a different surface, and therefore may have contributed to the inability of raters to learn. (This assumption about the correlation is tested in Chapter 6 of this thesis.)

This study has been the first to provide accurate and immediate feedback to students on their judgments of PA stiffness of the human lumbar spine. In this study students assessed 75 stimuli, receiving feedback on 52 of these stimuli, with the feedback being given immediately after every trial in the training sessions. This amount of training was chosen as it represented the greatest amount of training in PA stiffness assessment that could feasibly be provided to students in any physiotherapy programme. Despite the large number of varied stimuli used, the ratings did not demonstrate evidence of continued improved performance or learning. Another reason for this may lie in the amount of feedback that was provided. It has been suggested recently that providing feedback at every trial may actually operate to degrade performance and learning by producing excessive variability during practice (Wulf and Schmidt 1994). In other words, providing students with trial-by-trial feedback stimulates them to change their criteria for responding more than is necessary (McNicol 1978). Consequently, if a student is provided with frequent feedback this may cause the student to excessively modify their behaviour from trial to trial and so prevent them from learning a stable relationship between sensation and number value. Alternative ways of providing feedback, such as withholding feedback
on certain trials or providing the student with summary feedback ie: feedback generated from the average of several trials, may be needed to significantly improve performance and learning in manual assessment of PA stiffness.

CONCLUSION

After substantial training using immediate quantitative feedback, physiotherapy students’ accuracy in judging lumbar PA stiffness was not improved. Reasons for this may reside in the complex nature of the stiffness judgment task, involving as it does concurrent perceptions of both the non-linear and linear components of the lumbar force/displacement curve, which may make it difficult to attend to the linear component only. There also remains the issue of other attributes which may distract the attention of judges during stiffness assessment. The next chapter reports on a study that trained the same physiotherapy students to assess only elastic stiffness, in an attempt to identify whether it was possible to train students in a simpler task where judgements were related to a linear force/displacement curve only. Arguably, there continue to be uncontrolled factors, likely to affect PA stiffness measurement, that need to be controlled in subsequent in vivo training studies in order to determine the most effective method of training manual physiotherapists. Chapters 6 and 7 investigate some of these possible factors.
## APPENDIX B

### Raw data for in vivo training study

#### Pre-test

<table>
<thead>
<tr>
<th>Instrumented value (N/mm)</th>
<th>Rater 1 (N/mm)</th>
<th>Rater 2 (N/mm)</th>
<th>Rater 3 (N/mm)</th>
<th>Rater 4 (N/mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>15</td>
<td>13</td>
<td>16</td>
<td>12</td>
</tr>
<tr>
<td>21</td>
<td>16</td>
<td>19</td>
<td>14</td>
<td>16</td>
</tr>
<tr>
<td>13</td>
<td>12</td>
<td>17</td>
<td>14</td>
<td>18</td>
</tr>
<tr>
<td>10</td>
<td>11</td>
<td>12</td>
<td>18</td>
<td>18</td>
</tr>
<tr>
<td>17</td>
<td>9</td>
<td>20</td>
<td>16</td>
<td>12</td>
</tr>
<tr>
<td>14</td>
<td>8</td>
<td>19</td>
<td>20</td>
<td>18</td>
</tr>
<tr>
<td>17</td>
<td>16</td>
<td>21</td>
<td>19</td>
<td>20</td>
</tr>
<tr>
<td>15</td>
<td>13</td>
<td>15</td>
<td>17</td>
<td>16</td>
</tr>
<tr>
<td>16</td>
<td>18</td>
<td>17</td>
<td>18</td>
<td>15</td>
</tr>
<tr>
<td>19</td>
<td>17</td>
<td>16</td>
<td>19</td>
<td>19</td>
</tr>
</tbody>
</table>

#### Training Session 1

<table>
<thead>
<tr>
<th>Instrumented value (N/mm)</th>
<th>Rater 1 (N/mm)</th>
<th>Rater 2 (N/mm)</th>
<th>Rater 3 (N/mm)</th>
<th>Rater 4 (N/mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>21</td>
<td>18</td>
<td>17</td>
<td>15</td>
<td>17</td>
</tr>
<tr>
<td>9</td>
<td>22</td>
<td>19</td>
<td>18</td>
<td>15</td>
</tr>
<tr>
<td>17</td>
<td>16</td>
<td>15</td>
<td>19</td>
<td>10</td>
</tr>
<tr>
<td>10</td>
<td>15</td>
<td>9</td>
<td>15</td>
<td>18</td>
</tr>
<tr>
<td>14</td>
<td>8</td>
<td>13</td>
<td>14</td>
<td>12</td>
</tr>
<tr>
<td>15</td>
<td>15</td>
<td>18</td>
<td>19</td>
<td>14</td>
</tr>
<tr>
<td>16</td>
<td>16</td>
<td>15</td>
<td>21</td>
<td>16</td>
</tr>
<tr>
<td>17</td>
<td>18</td>
<td>21</td>
<td>18</td>
<td>19</td>
</tr>
<tr>
<td>19</td>
<td>18</td>
<td>17</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>13</td>
<td>11</td>
<td>14</td>
<td>12</td>
<td>12</td>
</tr>
</tbody>
</table>
### Training Session 2

<table>
<thead>
<tr>
<th>Instrumented value (N/mm)</th>
<th>Rater 1 (N/mm)</th>
<th>Rater 2 (N/mm)</th>
<th>Rater 3 (N/mm)</th>
<th>Rater 4 (N/mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>13</td>
<td>8</td>
<td>15</td>
<td>11</td>
<td>12</td>
</tr>
<tr>
<td>13</td>
<td>14</td>
<td>16</td>
<td>13</td>
<td>14</td>
</tr>
<tr>
<td>10</td>
<td>19</td>
<td>14</td>
<td>12</td>
<td>11</td>
</tr>
<tr>
<td>15</td>
<td>17</td>
<td>19</td>
<td>13</td>
<td>16</td>
</tr>
<tr>
<td>10</td>
<td>14</td>
<td>15</td>
<td>18</td>
<td>17</td>
</tr>
<tr>
<td>11</td>
<td>9</td>
<td>9</td>
<td>12</td>
<td>10</td>
</tr>
<tr>
<td>16</td>
<td>16</td>
<td>11</td>
<td>17</td>
<td>12</td>
</tr>
<tr>
<td>12</td>
<td>9</td>
<td>9</td>
<td>11</td>
<td>10</td>
</tr>
<tr>
<td>12</td>
<td>13</td>
<td>16</td>
<td>14</td>
<td>15</td>
</tr>
<tr>
<td>19</td>
<td>18</td>
<td>19</td>
<td>19</td>
<td>17</td>
</tr>
</tbody>
</table>

### Training Session 3

<table>
<thead>
<tr>
<th>Instrumented value (N/mm)</th>
<th>Rater 1 (N/mm)</th>
<th>Rater 2 (N/mm)</th>
<th>Rater 3 (N/mm)</th>
<th>Rater 4 (N/mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>12</td>
<td>12</td>
<td>14</td>
<td>13</td>
</tr>
<tr>
<td>11</td>
<td>11</td>
<td>8</td>
<td>16</td>
<td>10</td>
</tr>
<tr>
<td>13</td>
<td>13</td>
<td>11</td>
<td>17</td>
<td>14</td>
</tr>
<tr>
<td>17</td>
<td>16</td>
<td>19</td>
<td>17</td>
<td>15</td>
</tr>
<tr>
<td>16</td>
<td>15</td>
<td>18</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>15</td>
<td>13</td>
<td>13</td>
<td>14</td>
<td>14</td>
</tr>
<tr>
<td>21</td>
<td>19</td>
<td>18</td>
<td>20</td>
<td>17</td>
</tr>
<tr>
<td>17</td>
<td>15</td>
<td>15</td>
<td>18</td>
<td>18</td>
</tr>
<tr>
<td>12</td>
<td>9</td>
<td>14</td>
<td>12</td>
<td>11</td>
</tr>
<tr>
<td>14</td>
<td>14</td>
<td>16</td>
<td>17</td>
<td>16</td>
</tr>
<tr>
<td>12</td>
<td>11</td>
<td>13</td>
<td>13</td>
<td>14</td>
</tr>
</tbody>
</table>
### Training Session 4

<table>
<thead>
<tr>
<th>Instrumented value (N/mm)</th>
<th>Rater 1 (N/mm)</th>
<th>Rater 2 (N/mm)</th>
<th>Rater 3 (N/mm)</th>
<th>Rater 4 (N/mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>13</td>
<td>13</td>
<td>13</td>
<td>15</td>
<td>17</td>
</tr>
<tr>
<td>14</td>
<td>15</td>
<td>15</td>
<td>10</td>
<td>15</td>
</tr>
<tr>
<td>21</td>
<td>20</td>
<td>17</td>
<td>22</td>
<td>16</td>
</tr>
<tr>
<td>11</td>
<td>16</td>
<td>11</td>
<td>12</td>
<td>13</td>
</tr>
<tr>
<td>16</td>
<td>10</td>
<td>18</td>
<td>15</td>
<td>12</td>
</tr>
<tr>
<td>10</td>
<td>14</td>
<td>16</td>
<td>12</td>
<td>13</td>
</tr>
<tr>
<td>9</td>
<td>19</td>
<td>18</td>
<td>18</td>
<td>14</td>
</tr>
<tr>
<td>13</td>
<td>13</td>
<td>14</td>
<td>18</td>
<td>13</td>
</tr>
<tr>
<td>15</td>
<td>19</td>
<td>15</td>
<td>18</td>
<td>15</td>
</tr>
<tr>
<td>13</td>
<td>12</td>
<td>10</td>
<td>15</td>
<td>11</td>
</tr>
<tr>
<td>22</td>
<td>20</td>
<td>17</td>
<td>21</td>
<td>15</td>
</tr>
<tr>
<td>11</td>
<td>12</td>
<td>9</td>
<td>10</td>
<td>10</td>
</tr>
</tbody>
</table>

### Training Session 5

<table>
<thead>
<tr>
<th>Instrumented value (N/mm)</th>
<th>Rater 1 (N/mm)</th>
<th>Rater 2 (N/mm)</th>
<th>Rater 3 (N/mm)</th>
<th>Rater 4 (N/mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>17</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>16</td>
</tr>
<tr>
<td>10</td>
<td>15</td>
<td>14</td>
<td>17</td>
<td>15</td>
</tr>
<tr>
<td>22</td>
<td>16</td>
<td>19</td>
<td>14</td>
<td>17</td>
</tr>
<tr>
<td>15</td>
<td>16</td>
<td>17</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>12</td>
<td>13</td>
<td>16</td>
<td>18</td>
<td>14</td>
</tr>
<tr>
<td>13</td>
<td>15</td>
<td>19</td>
<td>22</td>
<td>11</td>
</tr>
<tr>
<td>15</td>
<td>21</td>
<td>15</td>
<td>17</td>
<td>10</td>
</tr>
<tr>
<td>11</td>
<td>16</td>
<td>11</td>
<td>13</td>
<td>15</td>
</tr>
<tr>
<td>13</td>
<td>14</td>
<td>14</td>
<td>12</td>
<td>13</td>
</tr>
</tbody>
</table>
### Post Test

<table>
<thead>
<tr>
<th>Instrumented value (N/mm)</th>
<th>Rater 1 (N/mm)</th>
<th>Rater 2 (N/mm)</th>
<th>Rater 3 (N/mm)</th>
<th>Rater 4 (N/mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>17</td>
<td>12</td>
<td>18</td>
<td>17</td>
<td>17</td>
</tr>
<tr>
<td>11</td>
<td>15</td>
<td>10</td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td>9</td>
<td>16</td>
<td>9</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>12</td>
<td>11</td>
<td>14</td>
<td>14</td>
<td>13</td>
</tr>
<tr>
<td>17</td>
<td>16</td>
<td>19</td>
<td>19</td>
<td>15</td>
</tr>
<tr>
<td>10</td>
<td>18</td>
<td>16</td>
<td>10</td>
<td>14</td>
</tr>
<tr>
<td>11</td>
<td>13</td>
<td>17</td>
<td>21</td>
<td>14</td>
</tr>
<tr>
<td>14</td>
<td>14</td>
<td>20</td>
<td>19</td>
<td>17</td>
</tr>
<tr>
<td>16</td>
<td>20</td>
<td>23</td>
<td>21</td>
<td>19</td>
</tr>
<tr>
<td>10</td>
<td>14</td>
<td>16</td>
<td>13</td>
<td>11</td>
</tr>
<tr>
<td>13</td>
<td>18</td>
<td>15</td>
<td>10</td>
<td>12</td>
</tr>
<tr>
<td>21</td>
<td>21</td>
<td>18.5</td>
<td>22</td>
<td>19</td>
</tr>
<tr>
<td>11</td>
<td>12</td>
<td>13</td>
<td>13</td>
<td>13</td>
</tr>
</tbody>
</table>
CHAPTER 5

TRAINING WITH IMMEDIATE FEEDBACK AND PHYSIOTHERAPY STUDENTS' ABILITY TO JUDGE ELASTIC STIFFNESS.

The work contained in this chapter has been published as:

ABSTRACT

Background and Purpose. Recent unsuccessful attempts to train physiotherapy students to accurately judge lumbar posteroanterior (PA) stiffness may be due to an inability to train judgements on a basic component, elastic stiffness. This study investigated whether feedback training improves students’ ability to judge elastic stiffness. Subjects. Four third-year physiotherapy students were trained. Methods. Students rated the posteroanterior stiffness of seventy-five mechanical stiffness stimuli during pretest, training and post-test sessions held over a 3.5 week period. At each training session trial by trial feedback was given regarding the accuracy of their judgements. Results. A paired t-test demonstrated a significant reduction in absolute error between the baseline and the post-test (t=5.36, df=3, p=0.013). Conclusion and Discussion. Immediate quantitative feedback training improves physiotherapy students’ ability to judge elastic stiffness. The inability of student raters to learn to judge PA lumbar stiffness is not due to an incapacity to learn to judge elastic stiffness.
INTRODUCTION

A possible explanation for the poor reliability of manual judgements of PA stiffness may relate to the methods used to teach undergraduate students to judge PA stiffness. In Australia the current method of teaching involves using the tutor's perception of PA stiffness as the standard against which student judgements are compared (Hickling and Maitland 1970, Maitland 1986). Given the consistently reported low reliability of such stiffness judgements, it is likely that the tutor's perception of stiffness may not be accurate, and so not assist students in learning the skill of judging PA stiffness.

Latimer et al (1996c) (Chapter 4 of this thesis) described an in vivo training study that was performed to investigate whether it was possible to improve the accuracy of manual stiffness judgements by providing physiotherapy students with accurate and immediate feedback. Measures of stiffness used for feedback were derived from a mechanical device rather than a tutor's stiffness perception. The results of this study were that even after substantial training, physiotherapy students' ability to accurately judge PA stiffness in vivo was not significantly improved.

Part of the explanation for these results may reside in the fact that the student raters found the task of judging the elastic stiffness component only, out of the total perceptual complex, a difficult one. To investigate this hypothesis another study was needed that reduced the complexity of the task and established whether the same physiotherapy students could learn to accurately judge elastic PA stiffness. If similar results to the in vivo training study were obtained, then it could be concluded that
training on purely elastic stiffness discrimination with feedback was ineffective. However, should this training study prove successful, it would suggest that exposure to other rheological dimensions, such as viscosity and plasticity (Wright and Johns 1961) may influence manual measurement of in vivo PA stiffness. If such dimensions do influence judgments of the stiffness of the lumbar spine, modifications would need to be made to the mechanical stiffness device to better assess the human back. This may include modification of the way the data is collected and analysed.

The aim of the current study therefore, was to evaluate the effect of immediate feedback training on physiotherapy students’ ability to judge the stiffness of a series of springs. This study used the same four student raters from the original in vivo study described in Chapter 4, but six months later, and a set of springs of stiffness as similar as possible to the backs in the previous study (Latimer et al 1996c).

The experimental hypothesis stated that the mean absolute error, obtained by comparing manual estimates of PA stiffness of a series of springs made by physiotherapy students, and mechanical measures of the stiffness of the same springs, would decrease following training sessions where the students were provided with immediate feedback on their performance.
METHOD

Research design

As before, a single case design (ABA) with replication was used, which involved the training of four third year physiotherapy students.

Approval was received from the University of Sydney Human Ethics Committee and informed consent was obtained from all raters prior to participation in the study.

Subjects

Raters

Four third year physiotherapy students (raters) agreed to participate in this study. The student raters (3 female, 1 male) had a mean age of 23.2 yrs and had completed the seventy hour physiotherapy undergraduate subject related to examination of patients with spinal pain. During this subject approximately twelve hours is spent practicing and receiving tutor feedback on PA stiffness assessment of the lumbar, cervical and thoracic regions. In addition to this, six months previously, these 4 physiotherapy student raters had also received 2 hours of additional tutor training in PA stiffness assessment, and then been asked to estimate the PA stiffness of 75 asymptomatic volunteer spines, receiving accurate and immediate feedback regarding each stiffness estimate on 52 of these. Despite this, at the post-test they were no more accurately able to judge PA stiffness in vivo than at the pre-test.
Stimuli

The elastic stiffness stimuli to be assessed by the raters were generated by using a series of springs housed in a mechanical device. This method of providing stiffness stimuli was described by Maher and Adams (1995a) and has been used in a series of studies (Maher and Adams 1996a,b and 1995b). The mechanical device enables different stiffness stimuli to be produced by varying the length of the lever arm of the resisting spring. The rater uses the heel of their hand, just distal to the pisiform, to contact a bevelled aluminium pad overlying a metal lever, and by applying a posteroanterior force at the point of hand contact, attempts to judge the stiffness associated with the downwards movement (Figure 5.1). The metal lever on which the rater pushes rotates freely about a bearing at one end of the device, and rests on a compression spring at the free end of the device (see Figure 5.2). Changing the position of the spring in the base plate of the device effectively alters the lever arm for the resisting spring and thereby creates a different stiffness.
Fig 5.1 Mechanical device used to generate stiffness stimuli (from Maher C 1996, pp 107). The bottom photograph shows a rater judging the stimulus using a pisiform grip.
Fig 5.2 Plan views of the elastic stiffness generating device (from Maker C. 1996, pp 93).
In the original form of the device the mushroom-shaped aluminium contact pad on the device was uncovered. In earlier studies (Maher and Adams 1995a) subjects commented that the cold feel of the pad was unlike the feel of a human spine. To counteract this in the present study, a 7 mm foam cover was attached to the contact pad. Pilot testing indicated that the foam material made pressing on the stiffness stimuli more like pressing on the human spine, in that it felt warmer and also provided a non-linear toe region similar to that encountered when judging human spines. Figure 5.3 shows the force/displacement curves obtained when testing a low stiffness stimulus (9.14 N/mm) and a high stiffness stimulus (23.86 N/mm). When performing these stiffness tests the stiffness assessment device described previously (Chapter 2) was used to apply a PA force via the indenter placed over the foam cover. It can be seen that the force/displacement curves obtained are similar in shape to the force/displacement curves obtained from human lumbar spines (c.f. Figure 2.5).

In order to determine whether the foam contact pad would change or deform following repeated testing, the stiffness assessment device was used to measure the K and displacement, D30, values of the lowest stiffness stimulus used in the study (9.14 N/mm) and the highest stimulus used (23.86 N/mm) prior to, and following, 100 posteroanterior (PA) oscillations. It was decided to examine the deformation of the pad following 100 PA oscillations as this was within the range of oscillations that would be applied to the pad during the training study by each rater. The stiffness tests were performed by positioning the indenter over the foam contact pad. The results of this testing demonstrated that the K and D30 values obtained were very
similar for the pre and post tests indicating that the foam contact pad would not be expected to vary during the repeated pressing in the training sessions and thereby confound the feedback given to the raters (see Table 5.1).

**Fig 5.3** Force/displacement curves for low stiffness stimulus (solid triangles) and high stiffness stimulus (closed circles) obtained by testing with the stiffness assessment device over the foam covered contact pad.
Table 5.1. K and D30 values obtained when testing stimuli generated by the spring device prior to performing PA oscillations (pre-test) and after the application of one hundred PA oscillations (post-test).

<table>
<thead>
<tr>
<th></th>
<th>K value $N/mm$</th>
<th>D30 value mm</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre-test</td>
<td>Post-test</td>
</tr>
<tr>
<td>Low stiffness</td>
<td>9.14</td>
<td>9.27</td>
</tr>
<tr>
<td>stimulus</td>
<td></td>
<td></td>
</tr>
<tr>
<td>High stiffness</td>
<td>23.86</td>
<td>24.59</td>
</tr>
<tr>
<td>stimulus</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

During this study, seventy-five stiffness stimuli were presented to each rater. These stimuli were selected so that the elastic stiffness ‘K’ would be similar to the stiffness stimuli used in the earlier study described in Chapter 4. This was done to ensure that direct comparisons could be made between the two studies, so that if subjects showed a training effect in the current study, it could be argued that the effect was due to the simpler discrimination task, rather than due to the fact that the training was more intensive, the stimuli presented differently, or that the feedback was given in some more optimal manner.

The instrumented stiffness assessment device (described in Chapter 2) was used to measure the series of stimuli generated by the mechanical spring device. Stimuli that had the same measured value as those human stimuli used in the training study reported in Chapter 4 were used to provide the training stimuli for the current study.
The instrumented value of the stimulus (rounded to the nearest whole value in N/mm) was used to provide feedback to the raters during the training sessions.

**Procedure**

Prior to commencing the study all student raters were shown how to depress the lever arm of the device using the pisiform grip and told that they would be asked to estimate the stiffness of the resulting movement in N/mm. Raters wore a pair of welding goggles with a visor that occluded vision while the researcher changed the spring position for the next stimulus. Raters were instructed to lift the visor with its opaque lens, thereby enabling free vision, when they were rating the spring stiffness (Figure 5.4), and to drop the visor thereafter.
At the first session, student raters were asked to rate the posteroanterior stiffness of 10 stimuli to serve as a baseline measure of performance. Over the next 2.5 weeks, five training sessions were held where the raters received feedback on their
performance. At these sessions stimuli were presented so as to match the previous study (10 stimuli being presented at Session 1, 10 at Session 2, 11 at Session 3, 12 at Session 4, and 10 at Session 5), with each rater being asked to judge the PA stiffness of the stimulus in N/mm. Raters were able to depress the lever arm as many times as was necessary for them to reach a judgement. Immediately following each judgement the rater was shown a card displaying the measured stiffness value of the stimulus in N/mm which had been obtained using the instrumented stiffness assessment device. The researcher then changed the spring position and the rater was asked to repeat the procedure.

To examine whether learning had occurred, a retention test was conducted at 2 days following the final training session (Post-test). At this retention test, raters was asked to assess the PA stiffness of 13 stiffness stimuli without feedback.

**Data analysis**

The percent exact agreement and the maximum error for the pre- and post-tests for each rater was calculated. The mean absolute error for each rater for the pre-test, post-test, and for the five training sessions was calculated. The data for the four raters were then combined and the mean absolute error and SE for the pre-and post-test and the training sessions calculated. A paired t-test was then performed to identify whether there was any difference in the absolute error between the pre-and post-test.
RESULTS

Table 5.2 shows the percentage of occasions at which the rater was in exact agreement for the pre- and post-tests and the maximum error for each test. The range of stimuli used at the pre- and post-tests ranged from 9-21 N/mm.

Table 5.2. Percentage of occasions at which the rater was in exact agreement (PEA) for the pre- and post-tests and the maximum absolute error for each test for each rater.

<table>
<thead>
<tr>
<th>Rater</th>
<th>PEA (pre-test)</th>
<th>Maximum Absolute Error $N/mm$</th>
<th>PEA (post-test)</th>
<th>Maximum Absolute Error $N/mm$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>7</td>
<td>16</td>
<td>4</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>10</td>
<td>16</td>
<td>6</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>7</td>
<td>47</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>7</td>
<td>54</td>
<td>2</td>
</tr>
</tbody>
</table>

The mean absolute error for each rater for the pre-test, post-test and each training session are presented graphically in Figure 5.5. The actual mean scores can be seen in Table 5.3.
Table 5.3. Mean absolute error and SE (N/mm) for each rater (R1-R4) for each session. There was a pre-test (Pre), five training sessions (1-5) and a post-test (Post).

<table>
<thead>
<tr>
<th>Rater</th>
<th>Pre</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>Post</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4.7(.50)</td>
<td>2.3(.58)</td>
<td>1.5(.34)</td>
<td>1.4(.41)</td>
<td>1.0(.30)</td>
<td>1.8(.41)</td>
<td>1.8(.37)</td>
</tr>
<tr>
<td>2</td>
<td>5.8(.77)</td>
<td>2.3(.67)</td>
<td>1.5(.45)</td>
<td>1.5(.28)</td>
<td>1.5(.51)</td>
<td>1.4(.33)</td>
<td>2.2(.54)</td>
</tr>
<tr>
<td>3</td>
<td>4.2(.57)</td>
<td>2.5(.73)</td>
<td>1.5(.30)</td>
<td>1.5(.42)</td>
<td>1.5(.31)</td>
<td>0.8(.25)</td>
<td>0.9(.31)</td>
</tr>
<tr>
<td>4</td>
<td>1.9(.64)</td>
<td>1.4(.31)</td>
<td>1.5(.27)</td>
<td>0.7(.36)</td>
<td>0.9(.25)</td>
<td>1.1(.38)</td>
<td>0.6(.21)</td>
</tr>
<tr>
<td>Mean</td>
<td>4.2(.37)</td>
<td>2.1(.29)</td>
<td>1.5(.16)</td>
<td>1.3(.19)</td>
<td>1.2(.18)</td>
<td>1.3(.18)</td>
<td>1.4(.21)</td>
</tr>
</tbody>
</table>

Fig 5.5 Mean absolute error (N/mm) at each session for each rater (R1-R4) judging elastic stiffness.
It can be seen from the graph that three of the four raters (R1, R2, R3) followed a similar pattern, reducing their absolute error from the pre-test to training session five. At the post-test the raters varied, some increasing slightly their error while others decreased their error. Rater 4 followed a different pattern due mainly to the small absolute error obtained at pre-test. Figure 5.6 shows the mean absolute error and SE for the group.

![Graph showing mean absolute error and SE (N/mm) at each session for the group of raters judging elastic stiffness.]

**Fig 5.6** *Mean absolute error and SE (N/mm) at each session for the group of raters judging elastic stiffness.*

A paired t-test demonstrated a significant difference between the mean absolute error for the pre-and post-tests. (t=5.36, df=3, p=0.013) with the raters improving the accuracy of their judgements over the training period.
DISCUSSION

Provision of accurate and immediate feedback regarding the stiffness of a series of springs significantly improved the accuracy of judgements of their elastic PA stiffness. The reduction in mean absolute error from the pre- to the post-test demonstrated in this study is similar to that previously described in motor learning studies where subjects were immediately provided with knowledge of results (KR) after every trial (Marteniuk 1986, Swinnen 1990). When KR was withdrawn at the post-test the group mean absolute error increased slightly but still remained significantly lower than at the pretest. Again, this pattern is similar to that seen in other successful motor skills training studies (Magill 1993).

The ability to learn to judge elastic stiffness is in direct contrast to the previous findings of Chapter 4 which demonstrated that training did not improve the ability of physiotherapy students to accurately judge PA lumbar stiffness (Latimer et al 1996c), suggesting that training in vivo stiffness measurement is more complex than training the measurement of mechanical elastic stiffness. The mean error made by the group when rating lumbar PA stiffness, and the mean error made when rating elastic stiffness are shown in Figure 5.7.
Fig 5.7 Mean absolute error and SE (N/mm) at each session for the group of raters judging elastic stiffness (open circles) and judging lumbar stiffness (closed circles).

Interestingly, at the pre-test raters demonstrated a smaller mean error when judging lumbar PA stiffness (group mean error = 3.35 N/mm) than when judging spring stiffness (group mean error = 4.15 N/mm) indicating that the task of judging PA stiffness in vivo may initially not be any more difficult, but accuracy is not improved by providing feedback related to the elastic stiffness dimension. The results of this study suggest that what the raters perceived as stiffness for springs is closely related to what the stiffness assessment device is measuring. What the raters perceived as stiffness for humans, however, does not seem to be as closely related to what the device measures.
Hence, this study has provided further information regarding the several explanations proposed in Chapter 4 to account for the findings reported in the in vivo study. In a general sense these explanations suggest that something distracts or diverts the attention of raters, and thereby masks out the relevant elastic stiffness information. For example, one of the explanations was that the student raters in the in vivo study had difficulty attending to the elastic stiffness component of the force/displacement relations in that they were unable to separate out information they perceived from the non-linear toe region of the curve. That is, variations in the toe region between the spines may have confused the students’ judgements about the linear region. In contrast, when the students were presented with pure elastic stiffness and a standardized toe region (the foam pad) in the current study, training with feedback improved their ability to judge elastic stiffness. A further study, examining whether the ability to judge elastic stiffness is maintained when varying toe regions are applied to the spring device trial by trial, is needed to determine this.

Another explanation for the previous results relates to whether measurement of PA stiffness in vivo involves perception of elastic stiffness only, or whether there are concurrent percepts from the other dimensions along which spines may vary. If this were so then feedback regarding only elastic stiffness may have proved unhelpful for human spines because the raters could not relate the feedback to what they felt. This study provides little support for the view that spines vary along an elastic stiffness dimension only. Were this so, then the student raters examined in the study described in Chapter 4 should arguably have been able to improve their ability to accurately
judge PA stiffness in vivo. The findings of the current study suggest that manual measurement of PA stiffness in human backs may be affected by dimensions other than elastic stiffness. It has been suggested that inertia and viscosity may be two other stiffness dimensions influencing judgements of spinal stiffness (Maher and Adams 1995b) and it is to these component dimensions and their combinations that spinal stiffness assessment researchers must now attend.

In addition two other explanations need to be considered when determining why the student raters used in Chapter 4 and 5 could not accurately judge PA stiffness in vivo but could improve their ability to judge elastic stiffness. These explanations are first, that the feedback provided to the raters in Chapter 4 was obtained by measuring volunteer stimuli on a rigid plinth, whilst the raters performed their manual judgements of stiffness on minimally padded, height adjustable plinths. This may have produced some error in the feedback provided to the raters in Chapter 4. The second explanation is that the in vivo PA response found in subjects above 30 N may be less linear than previously thought, and that the raters failure to learn in Chapter 4 was due to the fact that they were reporting stiffness from different force intervals to that used to determine the instrumented value. As mentioned in Chapter 2, analysis of typical force/displacement relations indicates good linearity of the curve above 30 N, a straight line fit accounting for at least 99.7% of the variance in the data. However, even small degrees of non-linearity may produce changes in the PA stiffness value obtained. Both of these hypotheses requires investigation, and it is to these issues that the next chapters turn.
CONCLUSION

Immediate quantitative feedback training was found to improve physiotherapy students’ ability to judge mechanical elastic PA stiffness. This finding suggests that the failure of students to learn to accurately judge lumbar PA stiffness was not due to difficulty in assessing the elastic stiffness component per se, but may reside in either the fact that lumbar PA stiffness varies along many dimensions which have yet to be clearly defined or that the variety of toe regions encountered in human backs makes it difficult to attend to elastic stiffness. Also, confounding factors such as the plinth padding on which stiffness testing is performed, and the degree of linearity of the force/displacement relation above 30 N need to be investigated to adequately account for the results of the in vivo study reported in Chapter 4.
### APPENDIX C

**Raw data for elastic stiffness training study**

#### Pretest

<table>
<thead>
<tr>
<th>Instrumented value (N/mm)</th>
<th>Rater 1 (N/mm)</th>
<th>Rater 2 (N/mm)</th>
<th>Rater 3 (N/mm)</th>
<th>Rater 4 (N/mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>12</td>
<td>16</td>
<td>18</td>
<td>16</td>
</tr>
<tr>
<td>21</td>
<td>19</td>
<td>20</td>
<td>24</td>
<td>19</td>
</tr>
<tr>
<td>13</td>
<td>20</td>
<td>17</td>
<td>20</td>
<td>15</td>
</tr>
<tr>
<td>10</td>
<td>14</td>
<td>16</td>
<td>16</td>
<td>11</td>
</tr>
<tr>
<td>17</td>
<td>21</td>
<td>21</td>
<td>27</td>
<td>18</td>
</tr>
<tr>
<td>14</td>
<td>20</td>
<td>19</td>
<td>19</td>
<td>16</td>
</tr>
<tr>
<td>17</td>
<td>23</td>
<td>20</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>15</td>
<td>21</td>
<td>20</td>
<td>21</td>
<td>15</td>
</tr>
<tr>
<td>16</td>
<td>20</td>
<td>21</td>
<td>19</td>
<td>16</td>
</tr>
<tr>
<td>19</td>
<td>24</td>
<td>21</td>
<td>25</td>
<td>18</td>
</tr>
</tbody>
</table>

#### Training Session 1

<table>
<thead>
<tr>
<th>Instrumented value (N/mm)</th>
<th>Rater 1 (N/mm)</th>
<th>Rater 2 (N/mm)</th>
<th>Rater 3 (N/mm)</th>
<th>Rater 4 (N/mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>21</td>
<td>19</td>
<td>21</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>9</td>
<td>9</td>
<td>15</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>17</td>
<td>18</td>
<td>23</td>
<td>24</td>
<td>19</td>
</tr>
<tr>
<td>10</td>
<td>11</td>
<td>14</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>14</td>
<td>20</td>
<td>16</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>15</td>
<td>17</td>
<td>16</td>
<td>15</td>
<td>17</td>
</tr>
<tr>
<td>16</td>
<td>12</td>
<td>20</td>
<td>18</td>
<td>16</td>
</tr>
<tr>
<td>17</td>
<td>15</td>
<td>16</td>
<td>13</td>
<td>17</td>
</tr>
<tr>
<td>19</td>
<td>15</td>
<td>18</td>
<td>22</td>
<td>20</td>
</tr>
<tr>
<td>13</td>
<td>12</td>
<td>13</td>
<td>13</td>
<td>15</td>
</tr>
</tbody>
</table>
### Training Session 2

<table>
<thead>
<tr>
<th>Instrumented value (N/mm)</th>
<th>Rater 1 (N/mm)</th>
<th>Rater 2 (N/mm)</th>
<th>Rater 3 (N/mm)</th>
<th>Rater 4 (N/mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>13</td>
<td>15</td>
<td>15</td>
<td>17</td>
<td>16</td>
</tr>
<tr>
<td>13</td>
<td>12</td>
<td>12</td>
<td>14</td>
<td>15</td>
</tr>
<tr>
<td>10</td>
<td>9</td>
<td>9</td>
<td>11</td>
<td>10</td>
</tr>
<tr>
<td>15</td>
<td>15</td>
<td>12</td>
<td>15</td>
<td>16</td>
</tr>
<tr>
<td>10</td>
<td>8</td>
<td>9</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>11</td>
<td>14</td>
<td>10</td>
<td>14</td>
<td>13</td>
</tr>
<tr>
<td>16</td>
<td>17</td>
<td>16</td>
<td>16</td>
<td>17</td>
</tr>
<tr>
<td>12</td>
<td>15</td>
<td>9</td>
<td>15</td>
<td>13</td>
</tr>
<tr>
<td>12</td>
<td>10</td>
<td>13</td>
<td>14</td>
<td>10</td>
</tr>
<tr>
<td>19</td>
<td>19</td>
<td>17</td>
<td>19</td>
<td>17</td>
</tr>
</tbody>
</table>

### Training Session 3

<table>
<thead>
<tr>
<th>Instrumented value (N/mm)</th>
<th>Rater 1 (N/mm)</th>
<th>Rater 2 (N/mm)</th>
<th>Rater 3 (N/mm)</th>
<th>Rater 4 (N/mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>10</td>
<td>9</td>
<td>9</td>
<td>11</td>
</tr>
<tr>
<td>11</td>
<td>13</td>
<td>11</td>
<td>13</td>
<td>11</td>
</tr>
<tr>
<td>13</td>
<td>15</td>
<td>12</td>
<td>14</td>
<td>13</td>
</tr>
<tr>
<td>17</td>
<td>18</td>
<td>17</td>
<td>16</td>
<td>17</td>
</tr>
<tr>
<td>16</td>
<td>18</td>
<td>16</td>
<td>18</td>
<td>16</td>
</tr>
<tr>
<td>15</td>
<td>19</td>
<td>17</td>
<td>17</td>
<td>15</td>
</tr>
<tr>
<td>21</td>
<td>18</td>
<td>18</td>
<td>17</td>
<td>17</td>
</tr>
<tr>
<td>17</td>
<td>17</td>
<td>18</td>
<td>16</td>
<td>17</td>
</tr>
<tr>
<td>12</td>
<td>12</td>
<td>13</td>
<td>13</td>
<td>11</td>
</tr>
<tr>
<td>14</td>
<td>14</td>
<td>19</td>
<td>15</td>
<td>13</td>
</tr>
<tr>
<td>12</td>
<td>11</td>
<td>14</td>
<td>11</td>
<td>13</td>
</tr>
</tbody>
</table>
# Training Session 4

<table>
<thead>
<tr>
<th>Instrumented value (N/mm)</th>
<th>Rater 1 (N/mm)</th>
<th>Rater 2 (N/mm)</th>
<th>Rater 3 (N/mm)</th>
<th>Rater 4 (N/mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>13</td>
<td>13</td>
<td>10</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>14</td>
<td>17</td>
<td>14</td>
<td>14</td>
<td>15</td>
</tr>
<tr>
<td>21</td>
<td>18</td>
<td>19</td>
<td>16</td>
<td>20</td>
</tr>
<tr>
<td>11</td>
<td>11</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>16</td>
<td>15</td>
<td>19</td>
<td>19</td>
<td>15</td>
</tr>
<tr>
<td>10</td>
<td>9</td>
<td>9</td>
<td>10</td>
<td>9</td>
</tr>
<tr>
<td>9</td>
<td>8</td>
<td>10</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>13</td>
<td>13</td>
<td>15</td>
<td>17</td>
<td>13</td>
</tr>
<tr>
<td>15</td>
<td>16</td>
<td>18</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>13</td>
<td>14</td>
<td>12</td>
<td>16</td>
<td>11</td>
</tr>
<tr>
<td>22</td>
<td>22</td>
<td>21</td>
<td>22</td>
<td>19</td>
</tr>
<tr>
<td>11</td>
<td>12</td>
<td>11</td>
<td>11</td>
<td>11</td>
</tr>
</tbody>
</table>

# Training Session 5

<table>
<thead>
<tr>
<th>Instrumented value (N/mm)</th>
<th>Rater 1 (N/mm)</th>
<th>Rater 2 (N/mm)</th>
<th>Rater 3 (N/mm)</th>
<th>Rater 4 (N/mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>17</td>
<td>18</td>
<td>16</td>
<td>14</td>
<td>18</td>
</tr>
<tr>
<td>10</td>
<td>12</td>
<td>9</td>
<td>13</td>
<td>9</td>
</tr>
<tr>
<td>22</td>
<td>20</td>
<td>20</td>
<td>22</td>
<td>20</td>
</tr>
<tr>
<td>15</td>
<td>19</td>
<td>15</td>
<td>17</td>
<td>19</td>
</tr>
<tr>
<td>12</td>
<td>14</td>
<td>11</td>
<td>13</td>
<td>11</td>
</tr>
<tr>
<td>13</td>
<td>17</td>
<td>14</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>15</td>
<td>16</td>
<td>17</td>
<td>16</td>
<td>15</td>
</tr>
<tr>
<td>11</td>
<td>12</td>
<td>11</td>
<td>10</td>
<td>11</td>
</tr>
<tr>
<td>13</td>
<td>13</td>
<td>13</td>
<td>15</td>
<td>13</td>
</tr>
<tr>
<td>9</td>
<td>8</td>
<td>9</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>Instrumented value (N/mm)</td>
<td>Rater 1 (N/mm)</td>
<td>Rater 2 (N/mm)</td>
<td>Rater 3 (N/mm)</td>
<td>Rater 4 (N/mm)</td>
</tr>
<tr>
<td>--------------------------</td>
<td>----------------</td>
<td>----------------</td>
<td>----------------</td>
<td>----------------</td>
</tr>
<tr>
<td>17</td>
<td>18</td>
<td>14</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>11</td>
<td>14</td>
<td>10</td>
<td>13</td>
<td>11</td>
</tr>
<tr>
<td>9</td>
<td>10</td>
<td>9</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>17</td>
<td>20</td>
<td>16</td>
<td>19</td>
<td>16</td>
</tr>
<tr>
<td>10</td>
<td>12</td>
<td>10</td>
<td>11</td>
<td>10</td>
</tr>
<tr>
<td>11</td>
<td>13</td>
<td>11</td>
<td>13</td>
<td>12</td>
</tr>
<tr>
<td>14</td>
<td>15</td>
<td>15</td>
<td>20</td>
<td>14</td>
</tr>
<tr>
<td>16</td>
<td>20</td>
<td>17</td>
<td>22</td>
<td>16</td>
</tr>
<tr>
<td>10</td>
<td>11</td>
<td>10</td>
<td>12</td>
<td>10</td>
</tr>
<tr>
<td>13</td>
<td>17</td>
<td>16</td>
<td>17</td>
<td>15</td>
</tr>
<tr>
<td>21</td>
<td>21</td>
<td>21</td>
<td>22</td>
<td>20</td>
</tr>
<tr>
<td>11</td>
<td>13</td>
<td>13</td>
<td>13</td>
<td>13</td>
</tr>
</tbody>
</table>
CHAPTER 6

PLINTH PADDING AND MEASURES OF POSTEROANTERIOR LUMBAR STIFFNESS

The work contained in this chapter has been published as:

ABSTRACT

Purpose. To investigate whether measurement of posteroanterior (PA) lumbar stiffness is affected by the presence of padding on the testing plinth. Design: Within a repeated measures design, measurements were made of lumbar PA stiffness on a rigid and on a padded plinth surface in subjects without low back pain. Subjects: Nineteen subjects with no history of any low back pain requiring treatment over the preceding twelve months participated in this study. Methods: PA stiffness was measured at L3 on two occasions under two different conditions. The first condition involved measurement of lumbar stiffness on a rigid plinth surface while the second involved measurement on a padded plinth surface. A reliable mechanical device was used to obtain the PA stiffness measures. Results: Mean lumbar PA stiffness values obtained when testing on a padded plinth were 2.86 N/mm less than those values obtained when testing the same lumbar spines on a rigid plinth. A paired t-test showed a significant difference between the PA stiffness measures at L3 obtained on the padded plinth and those obtained on the rigid plinth (t=6.66, df=18, p=<0.0001). Conclusions: These findings suggest that to improve the reliability of lumbar PA stiffness assessment, testing should be performed under the same plinth surface conditions.
INTRODUCTION

Altered stiffness of the lumbar spine is thought to frequently occur in patients with low back pain (LBP). PA stiffness testing is argued to be useful in identifying the symptomatic region of the spine, and in helping the clinician to select treatment and to monitor outcome (Maitland 1986).

Recently, several studies have found judgements of PA stiffness to have poor inter-rater reliability (Maher and Adams 1994, Binkley et al 1995). Many factors are thought to contribute to this poor reliability and can be considered in two groups, firstly, those factors thought to affect the physical stimulus, and secondly, those factors thought to affect perception of stiffness.

Factors known to affect the physical stimulus include factors directly related to the patient, for example, the position of the patient during conduct of the test. It has been demonstrated that PA stiffness is increased in an extended position compared to prone lying, while more flexed positions do not appear to change PA stiffness (Herden and Kingsford-Smith 1990, Cornell and Karlovic 1990). Also the patient’s breathing position during stiffness testing appears to alter the physical stimulus, Beaumont et al (1991) demonstrating that stiffness values obtained while breath-holding at the end of normal expiration were approximately 8% lower than when measured during normal tidal breathing.
Factors known to affect the perception of stiffness include the type of grip used to perform the test, and the visual state under which the test is performed (Maher and Adams 1996a, Maher and Adams 1996b). It appears that use of a thumb grip to perform PA stiffness testing makes the same stimuli appear substantially stiffer than when tested using a pisiform grip. Maher and Adams (1996b) suggest that in areas of the spine where both versions of the test are used to assess PA stiffness, for example, the upper thoracic spine, therapists may incorrectly perceive that the upper thoracic is stiffer than the mid thoracic spine if assessing the upper thoracic with a thumb grip and the mid thoracic spine with a pisiform grip.

Knowledge of the factors affecting stiffness measurement will assist the development of a more reliable protocol for judging PA stiffness, and therefore facilitate development of more effective training programs where confounding variables are controlled. One factor that has not been previously discussed, but seems likely to have some effect on measured stiffness, is the type of plinth surface on which the testing is performed. It has been recognized for some time that the measurement of lumbar PA stiffness as performed by clinicians treating spinal pain is likely to be influenced by the force/displacement characteristics of structures remote to the spine. As yet the exact factors contributing to PA stiffness have not been detailed, but are likely to include not only anatomical structures within and adjacent to the lumbar region, but also factors unrelated to the subject, such as the stiffness of the plinth surface on which the testing is performed. To date there have been no studies that have investigated the role of plinth surface on measurement of PA stiffness.
The aim of this study was to investigate whether posteroanterior stiffness is affected by the nature of the plinth surface on which stiffness testing is performed.

**METHOD**

**Research design**

A repeated measures design was used in this study. Lumbar PA stiffness was measured at L₃ in subjects without LBP on both a rigid plinth surface and on a padded plinth surface. The order of testing was randomized.

**Subjects**

Nineteen subjects (8 male, 11 female) with a mean age of 21.5 yrs (SD=5.3) participated in this study. Subjects had no history of current LBP or of LBP sufficient to require treatment during the preceding twelve months. Subjects were excluded if they were pregnant or had a known contraindication to PA stiffness testing.

The population sampled was one of convenience in that the sample comprised staff and students at The University of Sydney, Australia. Informed consent was obtained
from all subjects prior to their participation in the study. The stiffness testing protocol was approved by the Human Ethics Committee, The University of Sydney, Australia.

**Measurement instruments**

A mechanical device was used to measure PA stiffness in this study. This stiffness assessment device has been described in Chapters 2 and 3, including details of its operation, safety features and reliability. Some simple modifications were performed to the device prior to its use in this study. These modifications involved the use of a load cell (foil strain gauge\(^1\)), rather than a spring to measure the force applied to the patient in order to reduce the height of the mechanical head (see Appendix at the end of this chapter for schematic diagram). This resulted in a height reduction of approximately 50mm making the device easier to transport.

From the data obtained by the device, force/displacement curves were generated for four cycles which were then averaged. Linear regression was then performed on the linear part of the mean curve, between 30-90 Newtons, and the slope of this regression line defined as the stiffness coefficient ‘\(K\)’. It was this ‘\(K\)’ value that was used as the measure of stiffness for each subject at \(L_3\).

---

\(^1\)RS Components No. 308102, 129-137 Beaconsfield St, Silverwater. NSW, Australia
Procedure

While subjects lay prone on a plinth, the skin overlying the L3 spinous process was identified and marked. A manual PA force was then applied to the marked vertebral level to familiarise the subject with the stiffness testing procedure. Subjects were then instructed to first lie on either the rigid unpadded testing table which forms part of the mechanical device, or on the padded plinth surface which was positioned on top of the rigid testing table, depending on which order they were randomly allocated.

The padded plinth surface was created by unscrewing the top of a commercially available, height-adjustable plinth (Australian Medical Couches2), and placing this surface on top of the rigid testing table and securing it. The padded surface was comprised of a layer of foam covered with vinyl. The thickness of this padding (including covering) was 5.5cm. An estimate of the stiffness of the padding was obtained by placing a 54cm × 28cm piece of rigid particle board (similar in dimensions to the surface area of a subject’s torso) on the padded surface and performing the PA stiffness testing by applying the PA force through the rigid board. Figure 6.1 shows the loading curve of the PA stiffness response obtained during this test. Although the response is clearly non-linear a regression line was fitted from 30 to 90 Newtons, as was done for the subjects’ force displacement curves, in order to provide some indication of the padding stiffness. The value obtained for the stiffness coefficient ‘K’ was 77.9 N/mm.

258 East Concourse, Beaumaris, Victoria 3193
Fig 6.1 Force/displacement characteristics of the loading curve (dotted) obtained when performing PA stiffness testing of the padded plinth surface using a torso-sized, rigid board. The regression line used for stiffness analysis is shown (solid).

Mechanical PA stiffness testing was then performed on subjects using the device. The indenter was moved into contact with the skin at the marked level. At least five preliminary cycles where the indenter applied a PA force of 100 Newtons were performed to precondition the viscoelastic spinal tissues prior to data collection to improve the reproducibility of the data (Loebl 1972). Data were then collected for five loading cycles up to a force of 100N at a frequency of 0.5 Hertz. This magnitude
of force and frequency of cycling were the same as used in the study that demonstrated high reliability of measurement (See Chapter 3).

Following this test the subject moved off the testing surface and the padded plinth surface was either removed or applied to the rigid testing table, depending on which test remained to be performed. The head of the indenter was then relocated and a second PA stiffness test was performed as described above.

Data analysis

Mean values for the stiffness coefficient ‘K’ were calculated for PA stiffness testing on both surfaces. A t-test for paired samples was carried out to compare the ‘K’ values obtained from each test. The effect of repeated testing on stiffness values was also examined by using a t-test for paired samples.

Intra-class correlation coefficients, the ICC (2,1) & the ICC (3,1) as defined by Shrout and Fleiss (1979), and Pearson’s ‘r’ were calculated to determine whether any relationship existed between the ‘K’ value obtained when testing on the rigid surface versus the padded surface and to help define the nature of any bias. It has been suggested by Rey et al (1987) that comparisons between the ICC (2,1), ICC (3,1) and Pearson’s ‘r’ can indicate the nature of the relationship between data sets. In particular, the ICC (2,1) is sensitive to both multiplicative and additive bias between
sets of measures, whereas the ICC (3,1) is sensitive only to multiplicative bias, and Pearson’s ‘r’ is sensitive to neither.

RESULTS

The mean ‘K’ value obtained for testing at L3 on the rigid testing surface was 14.87 N/mm (SD=3.21) and for testing on the padded surface, 12.01 N/mm (SD=2.51) with a mean difference of 2.86N/mm. The number of subjects who demonstrated a difference in stiffness when tested on the rigid compared with the padded plinth surface and the value of that difference can be seen in the Table 6.1. All the differences were positive with the maximum difference observed being 6.2 N/mm.

Table 6.1. The number of subjects who demonstrated an increase in stiffness when tested on the rigid compared with the padded plinth surfaces and the value of that increase.

<table>
<thead>
<tr>
<th>Increase in stiffness</th>
<th>N/mm</th>
</tr>
</thead>
<tbody>
<tr>
<td>-1-0</td>
<td>0</td>
</tr>
<tr>
<td>0-1</td>
<td>3</td>
</tr>
<tr>
<td>1-2</td>
<td>3</td>
</tr>
<tr>
<td>2-3</td>
<td>3</td>
</tr>
<tr>
<td>3-4</td>
<td>1</td>
</tr>
<tr>
<td>4-5</td>
<td>6</td>
</tr>
<tr>
<td>5-6</td>
<td>1</td>
</tr>
<tr>
<td>6-7</td>
<td>2</td>
</tr>
</tbody>
</table>

Number of subjects
Figure 6.2 shows a scatterplot (and the regression line) of the relationship between $K_{unpadded}$ and $K_{padded}$. A paired t-test demonstrated a significant difference between the two stiffness tests, indicating a higher measured value of stiffness when testing on the rigid surface ($t=6.66$, df=18, $p<0.0001$). Another paired t-test was conducted on the stiffness values from test 1 and test 2, irrespective of padding, to evaluate the effect on stiffness of a second test. Repeat testing did not significantly affect the stiffness value ($t=1.34$, df=18, $p=0.196$).

Fig 6.2 Scatterplot showing the relationship between $K_{unpadded}$ and $K_{padded}$ and the regression line.
The difference between the ICC(2,1) at 0.532 and the ICC (3,1) at 0.788 indicated the presence of an additive bias, whereas the smaller rise to the Pearson’s ‘r’ value of 0.812 revealed little multiplicative bias (Rey et al 1987). The effect of plinth padding was therefore to cause a relatively constant reduction in measurement values across all ranges of spinal stiffness, not a reduction dependent on the size of the stiffness value.

**DISCUSSION**

PA lumbar stiffness values obtained on a padded plinth surface were 2.86 N/mm or approximately 20% less than those obtained when the same subjects were tested on a rigid surface. This finding may explain some of the disagreement observed between therapists in previous reliability studies of manual judgements of PA stiffness. If the surface on which the PA stiffness test was performed was not controlled between therapists it may have been a source of considerable error. Logically, the plinth surface on which PA stiffness testing is performed must be controlled in the development of a reliable protocol for judging PA stiffness. The degree to which padding stiffness varies among different commercial treatment couches is not known.

A simple theoretical analysis of the experimental situation can be considered in order to attempt to predict the effect of the addition of plinth padding on PA movement stiffness. The subject and the padding can be simplified to become two springs in
series, with stiffness values $K_{\text{subject}}$ and $K_{\text{padding}}$ respectively. The overall PA
movement stiffness can then be predicted, based on the effect of placing 2 springs in
series, to be $K_{\text{total}}$, where:

$$\frac{1}{K_{\text{total}}} = \frac{1}{K_{\text{subject}}} + \frac{1}{K_{\text{padding}}}$$  \hspace{1cm} \text{(Hall 1973)}$$

which becomes:

$$K_{\text{total}} = \frac{K_{\text{subject}}}{1 + \frac{K_{\text{subject}}}{K_{\text{padding}}}}$$

From this form of the equation it can be seen that flexible padding in series will cause
a decrease in PA stiffness, with the ratio of subject stiffness to padding stiffness being
the critical variable in determining the reduction in stiffness. The stiffness data
obtained from each of the 19 subjects in this study was subsequently substituted in the
above equation to give an estimated mean value for the padded plinth of 61.6 N/mm,
based on the model of padding forming series stiffness. This value was then
compared with the stiffness value obtained by actually testing the padded plinth using
the stiffness assessment device. During this test the PA force was applied to a piece
of rigid particle board (similar in dimensions to the surface area of a subject’s torso)
that rested on the padded surface. The force/displacement curve shown in Figure 6.1
provides an approximation of the stiffness of the padded plinth, the ‘K’ value obtained
being 77.9 N/mm. This value was obtained by performing a linear regression, over
the same force limits as was used for the subjects, in order to give an approximate
indication of the contribution of padding stiffness to the measured values. From this
comparison it can be seen that the value obtained from the experimental testing is similar to that predicted by the equation.

In effect, a padded surface contains a large number of non-linear springs in parallel, with the actual number of springs which are involved in providing the resistance to the PA force, and each spring’s stiffness, being dependent on the surface area of the subject and the pressure distribution over the contact area between the patient and the plinth. These factors are likely to vary substantially among subjects, making it difficult to exactly determine the value of $K_{padding}$ for each individual subject. The above analysis would therefore explain why all our subjects showed some reduction in stiffness, but not be able to accurately predict the size of this change.

It is also worth considering whether the difference in measured PA stiffness caused by the testing surface would be clinically detectable. This study found a 20% difference in the value obtained when PA stiffness was measured on a rigid surface compared to a padded surface. Maher and Adams (1995a) demonstrated that humans have a good ability to discriminate between stiffness stimuli, with a shift in the stimulus value of approximately 11% being needed for subjects to detect that a change had occurred. Based on this data, it is reasonable to suggest that a 20% difference in PA stiffness would be detectable to most clinicians examining patients with spinal pain. However, the Maher and Adams study (1995a) used a mechanical spring device to provide the stiffness stimuli, and it is likely therefore that when trying to perform the more
difficult task of assessing PA stiffness in vivo, where both linear and non-linear components of stiffness occur, the discrimination threshold may be greater than 11%.

This study supports the previously held view (Lee and Svensson 1993a) that structures other than those in the lumbar region may affect the results obtained from PA stiffness testing. Lee et al (1995) have demonstrated that application of a PA force to the lumbar spine results in a small amount of intervertebral movement, and produces physiological extension of the whole lumbar spine. In addition, the pelvis tilts forward and the rib cage and abdominal contents are compressed. However, our study also suggests that the compressibility of the surface on which the subject is lying will alter the measured magnitude of lumbar PA stiffness.

It is unclear whether plinth surface will affect the PA stiffness values obtained equally in all regions of the spine. Lee et al (1994b) have shown that lumbar PA forces produce anterior rotation of the pelvis, and have predicted that lumbar PA stiffness will be sensitive to the resistance to pelvic rotation. Plinth surface padding may be an important factor in determining this resistance to pelvic rotation. In the thoracic spine the lesser effect of this mechanism would be likely to produce a different relationship between PA stiffness and plinth padding.
CONCLUSION

Measurement of PA stiffness is significantly affected by the plinth surface on which testing is performed. A padded plinth surface was found to reduce the PA stiffness value obtained by approximately 20% compared to when testing on a rigid surface. Development of a reliable protocol for judging PA stiffness must therefore ensure that repeat testing is performed on the same plinth surface.
APPENDIX D

Schematic diagram of modified stiffness assessment device.

- Strain Gauge
- Measure of Applied Force using Strain Gauge
- Angular Displacement measured by Optical Encoder (E) located on pulley shaft
- Motor
- Pulley (P)
- Subject
The work contained in this chapter has been published as:

ABSTRACT

Background and Purpose. A potential explanation for the poor reliability of manual judgements of posteroanterior (PA) stiffness may be that physiotherapists are reporting stiffness values after using different PA testing forces. Should non-linearity exist in the force/displacement relation above 30 Newtons, then the PA stiffness values obtained will be dependent upon the testing forces applied. The purpose of this study was to determine the reliability of instrument measurements of PA stiffness obtained using different applied loading forces. Subjects. Twenty-five subjects with a mean age of 23.3 years and no history of LBP were measured in this study. Methods. Instrumented measures of the stiffness coefficient $K$ were performed using a portable, mechanical stiffness assessment device. A testing force of 200 N was applied to the skin overlying the L$_3$ spinous process of all subjects. Six stiffness coefficients were then determined from the force/displacement curve obtained from each subject, by performing linear regressions from 30-80N, 30-150N, 30-200N, and from 30-83.3N, 30-136.7N, and 136.7-200N. Intraclass correlation coefficients were used to determine the extent to which the first three stiffness tests were measures of the same property, and likewise whether the second three stiffness tests were measures of the same property. A repeated measures one-way analysis of variance was conducted to examine whether there was a significant difference between any of the first three stiffness measures, and between any of the second three stiffness measures. Results. Moderate reliability was demonstrated for the stiffness measures obtained by regressing between 30-80N, 30-150N and 30-200N, with an ICC (2,1) of 0.67 being obtained. A repeated measures, one-way analysis of variance demonstrated that there was a significant difference between the ‘$K$’ values obtained
for each of the three, successively larger, regression intervals (p<0.05), with the ‘K’ value increasing as the maximum force increased. Poor reliability was demonstrated for the stiffness measures obtained by regressing over the discrete non-overlapping intervals from 30-83.3N, 83.3-136.7N, 136.7-200N, with an ICC (2,1) of 0.39 being obtained. A repeated measures, one-way analysis of variance demonstrated that there was a significant difference between the ‘K’ values obtained for each of the three discrete regression intervals (p<0.05), with the obtained ‘K’ value increasing with increasing force. **Conclusion and Discussion.** This study suggests that instrumented measurement of spinal stiffness must be performed by calculating ‘K’ values from the same force intervals. Also, revised protocols for manually judging PA stiffness should ensure that stiffness is assessed by sampling specified force intervals rather than the rater determining his or her own force limits. Future training studies examining ways of improving the accuracy of manual judgements of PA stiffness need to consider training raters to sample PA stiffness within specified force limits.
INTRODUCTION

The need to develop a new protocol for PA stiffness testing has been identified by many researchers (Maher and Adams 1994, Lee et al 1996, Latimer et al 1996c). Development of such a protocol would foster the growth of more effective training programs for physiotherapy students learning the skill of stiffness testing. However, when this idea was first proposed there was little information available regarding factors which might affect instrument measurement of stiffness or manually derived perception of stiffness. Since this time, research (Maher and Adams 1996a,b, Lee et al 1996) has identified many factors associated with changes in stiffness and which are therefore important to control in a revised protocol for PA stiffness testing.

One factor not previously considered, but arguably likely to affect the stiffness measure obtained, is the maximum amount of force applied to the skin overlying the target vertebra during the stiffness test. Much of the past research related to PA stiffness testing has performed the test using loading forces up to 100 N (Lee et al 1993c, Latimer et al 1996c, Lee and Liversidge 1994). Yet a review of the manual therapy literature suggests that physiotherapists use a wide variety of applied forces (Lee and Moseley 1991, Matyas and Bach 1985, Simmonds et al 1995). For example, Lee and Moseley (1991) reported on a group of experienced physiotherapists who applied an average maximum force of 240 N, with one individual applying a maximum of 428 N. They also examined a group of postgraduate manipulative therapy students and found that the maximum forces applied by these students ranged from 30 N to 280 N (Lee and Moseley 1991). Matyas and Bach (1985), who reported the results of a number of studies in which the force applied to human subjects was measured, found the largest force used to be 350 N. Finally, Petty (1995) reported on the range of forces used by a group of
It is possible that, if physiotherapists are applying different loading forces, the perceived stiffness values obtained during testing may also be different. In the training study reported in Chapter 4 (Latimer et al. 1996c) no attempt was made to train the raters to apply a standardized force when judging PA stiffness. It was argued that as long as the rater applied some force above 30 N the force/displacement relation would be linear and the stiffness value obtained should therefore be the same. However, if in fact some non-linearity was present and the raters applied different loads between volunteer stimuli, error would have been introduced.

Recently, in a pilot study, Lee et al. (1997) examined the PA stiffness values obtained when testing the lumbar spine in two subjects using testing forces from 30-100 N and 30-275 N. The results obtained suggested that, due to a small degree of non-linearity, there was a tendency for stiffness to moderately increase as the load increased. The stiffness value obtained when calculated between 30-275 N was 25% greater than when calculated between 30-100 N. This study suggested, therefore, that the PA stiffness measure obtained was dependent on the testing force used and by implication this would be a necessary factor to control in any revised PA stiffness testing protocol. However, a larger study was needed to determine whether the effect seen in this pilot research would be reproduced in a larger sample and what degree of consistency there would be between stiffness values obtained with different force intervals remained to be ascertained.

Patients with spinal pain who have seen more than one physiotherapist frequently report that the previous physiotherapist pushed harder or more gently during stiffness
assessment or mobilisation. This raises the question of whether the same stiffness judgement can arise from testing forces made to different maximum limits. It would be interesting, therefore, to examine the interrater reliability of a group comprising physiotherapists who used the same minimum forces but different force intervals (the force interval being defined as the maximum force minus the minimum force used). To simulate this, investigation of the interrater reliability of measures from the stiffness assessment device using a small force interval from 30 to 80 N (representing a physiotherapist pushing lightly when PA stiffness testing), from the stiffness assessment device using a medium force interval from 30 to 150 N (representing a physiotherapist pushing moderately), and from the stiffness assessment device using a high force interval from 30 to 200 N (representing a physiotherapist pushing hard) may provide useful information about the poor reliability of stiffness judgements.

An alternative set of assessment strategies may involve physiotherapists performing PA stiffness assessment by estimating the elastic stiffness over a fixed force interval with variable maximum and minimum forces. For example, of a group of physiotherapists who used a similar force interval, one physiotherapist, who assesses using low force, may apply an initial force of 30N and then judge elastic stiffness by sampling between 30 and approximately 80N. A physiotherapist who assesses using moderate force over a similar interval may apply an initial force of 80N and assess between 80 and approximately 135N, and a physiotherapist who assesses using higher force may judge stiffness from the interval between 135 and 200N. Currently, there is no information available about the assessment strategies therapists use to judge PA stiffness, only information indicating the range of maximum forces applied (discussed earlier). Therefore, in the present study, the decision was made to investigate the reliability of stiffness measurements made using two types of strategies. Firstly, assessing PA stiffness when using the same minimum force but up to different maximum forces was considered, and secondly, the strategy of sampling over the
same-width force intervals but with different minimum and maximum forces. In this study the stiffness assessment device described previously was used to perform the stiffness measures, being programmed to represent real physiotherapists employing the different assessment strategies outlined above.

Specifically, the current study examined the reliability of instrumented measures of PA stiffness when stiffness testing was done by (i) using the 30-80N, 30-150N, and 30-200N force ranges, and (ii) using a fixed force interval with different maximum and minimum values, that is, low forces from 30-83.3N, moderate forces from 83.3-136.7N, and high forces from 136.7-200N.

**METHOD**

**Research design**

A repeated measures design was used in this study, wherein all the stiffness estimates based on different forms of sampling from the data obtained for each subject could be compared.

**Subjects**

Twenty-five subjects (17 male, 8 female), with a mean age of 23.5 yrs (SD=4.98) with no history of LBP requiring treatment over the preceding 12 months, participated in this study. Subjects were excluded if they had any contraindication to PA stiffness testing such as a known malignancy, inflammatory or infective disease
process affecting the spine, known osteoporosis, history of spinal surgery or were pregnant.

Ethics approval was obtained from the University of Sydney Human Ethics Committee, and subjects were free to withdraw from the study at any time simply by indicating their wish to do so.

**Procedure**

Anthropometric data were collected from all subjects prior to stiffness testing. Subjects were then asked to undress to their underwear and put on a hospital gown. This precaution was taken to ensure that heavy clothing such as jeans or belts did not produce a more extended lumbar posture than normal. The L3 spinal level was then palpated and marked by the researcher and several manual PA pressures were applied to ensure that no pain would be elicited during the stiffness testing procedure. PA stiffness was measured using the modified instrumented device (fully described in Chapters 2 and 3 of this thesis, with modifications as described in Chapter 6). Prior to commencement of data collection, five oscillations were applied to the subject’s skin overlying the spinous process of L3 using the device. During these 5 oscillations the force was slowly incremented from 100 to 200 N to precondition the viscoelastic tissues prior to testing, and also to ensure that the subject would be able to tolerate the applied force.

It was decided to test to a maximum force of 200N rather than 275N as had been used by Lee et al (1997) in their pilot study where the researchers themselves were
used as subjects, due to the fact that a larger number of subjects were to be tested in the current study. During instrumented stiffness testing the force is applied through a rectangular metal indenter with dimensions 35mm×20mm. These dimensions are approximately 50% of the surface area used to deliver a PA force when performing stiffness testing using a pisiform grip, and as a consequence the average surface pressure generated when using the indenter to apply the force is greater (although the peak pressure may be lower). Instrumented stiffness testing using forces above 200N may become uncomfortable for the subject, even though substantially higher loading forces have been documented in clinical practice. The decision was therefore made to not exceed a force of 200N.

Subjects were asked to hold their breath at the end of normal expiration and the stiffness device was then used to measure PA stiffness at L₃ by applying 5 testing oscillations to 200 N at a frequency of 0.5 Hertz. The relative change in force and displacement was then used to calculate the stiffness coefficient ‘K’ (in N/mm) for the tested spinal level.

**Data extraction**

The data was extracted in the following manner. A mean force/displacement relation was generated using the loading curves of cycles 2, 3, 4 and 5 collected by the device. Cycle one was discarded from the analysis as being more likely to be due to transient behaviour on startup. After averaging cycles 2 to 5, the resulting mean curve characterized the lumbar spine’s response to the application of up to 200N of force. Six stiffness values (measures of the stiffness coefficient ‘K’) were then extracted. The first three stiffness values were obtained by performing linear regression over three force intervals with the same minimum force, 30N, but having a different
maximum force, i.e. 30-80N, 30-150N, 30-200N, and calculating the gradient of the linear region. The last three stiffness values were obtained by performing the linear regression over three equal force intervals with increasing minimum force i.e. 30-83.3N, 83.3-136.7N, 136.7-200N. This was conducted to determine the reliability of measurement obtained from different discrete force intervals on the curve.

**Data analysis**

The group mean and SD for each of the 6 stiffness measures was calculated. Intra class correlation coefficients, ICC (2,1) and ICC (3,1) and their associated confidence intervals (Shrout and Fleiss 1979) were calculated using software developed at the University of Sydney. These statistics were used to determine the extent to which the first three stiffness measures could be interpreted as measures of the same property, whether the second three stiffness measures were likewise able to be viewed measures of the same property, and the nature of any association between them.

Finally, a repeated measures, one-way analysis of variance was conducted to see if there was a significant difference between any of the first three stiffness measures, and any of the second three stiffness measures. A Student-Newman-Keuls method was used to determine where these differences lay.
RESULTS

At the sampling rate employed, 256 points were sampled for each loading curve from 0-200 N. The 'K' values and $r^2$ values (the latter representing the goodness of fit of the regression line to the force/displacement curve) for each subject were calculated using regression over forces from 30-80N, 30-150N, 30-200N and are shown in Table 7.1. The resulting functions have a high degree of linearity, and this is reflected in the $r^2$ values being close to 1.0 (Table 7.1).

The group mean 'K' values, calculated by performing linear regression from 30-80 N, 30-150 N and 30-200 N were 15.8 (SD=3.42,) N/mm, 18.6 (SD=4.04) N/mm, and 20.1 (SD=4.26) N/mm respectively. Moderate reliability was observed between these measures, with the ICC (2,1) being 0.67 (95% CI =0.14 -0.87), and the ICC (3,1) being 0.88 (95% CI =0.79-0.94). Because there is a difference of greater than 0.1 between these two forms of the ICC, this is evidence suggesting the presence of an additive bias, as indeed can be observed across the three measures.

A repeated measures analysis of variance demonstrated a significant difference existed between the three measures ($p=0.00861$). The results of the Student-Newman-Keuls test demonstrated that 'K' values obtained by regressing between 30-80N (representing a physiotherapist pushing lightly) were significantly lower than values obtained by regressing between 30-150N (representing a physiotherapist pushing more firmly), which in turn were significantly lower than values obtained by regressing between 30-200N (representing a physiotherapist pushing very firmly) ($p<0.05$).
Table 7.1: K and $r^2$ values for each subject calculated using regression forces from 30-80 N, 30-150 N, 30-200 N.

<table>
<thead>
<tr>
<th>Subject</th>
<th>30-80</th>
<th>$r^2$</th>
<th>30-150</th>
<th>$r^2$</th>
<th>30-200</th>
<th>$r^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>9.37</td>
<td>0.999</td>
<td>11.18</td>
<td>0.996</td>
<td>12.59</td>
<td>0.992</td>
</tr>
<tr>
<td>2</td>
<td>19</td>
<td>0.999</td>
<td>20.37</td>
<td>0.999</td>
<td>21.11</td>
<td>0.999</td>
</tr>
<tr>
<td>3</td>
<td>14.84</td>
<td>0.999</td>
<td>17.02</td>
<td>0.998</td>
<td>17.77</td>
<td>0.998</td>
</tr>
<tr>
<td>4</td>
<td>14.63</td>
<td>0.997</td>
<td>17.21</td>
<td>0.997</td>
<td>18.68</td>
<td>0.996</td>
</tr>
<tr>
<td>5</td>
<td>11.23</td>
<td>0.999</td>
<td>12.16</td>
<td>0.998</td>
<td>13.32</td>
<td>0.996</td>
</tr>
<tr>
<td>6</td>
<td>22.52</td>
<td>0.999</td>
<td>25.61</td>
<td>0.998</td>
<td>26.64</td>
<td>0.999</td>
</tr>
<tr>
<td>7</td>
<td>12.63</td>
<td>0.997</td>
<td>18.1</td>
<td>0.986</td>
<td>20.76</td>
<td>0.987</td>
</tr>
<tr>
<td>8</td>
<td>16.64</td>
<td>0.994</td>
<td>17.68</td>
<td>0.999</td>
<td>18.76</td>
<td>0.998</td>
</tr>
<tr>
<td>9</td>
<td>11.56</td>
<td>0.998</td>
<td>15.54</td>
<td>0.991</td>
<td>17.86</td>
<td>0.988</td>
</tr>
<tr>
<td>10</td>
<td>11.96</td>
<td>0.996</td>
<td>14.84</td>
<td>0.995</td>
<td>16.49</td>
<td>0.994</td>
</tr>
<tr>
<td>11</td>
<td>17.86</td>
<td>0.997</td>
<td>21.22</td>
<td>0.997</td>
<td>21.01</td>
<td>0.999</td>
</tr>
<tr>
<td>12</td>
<td>15.82</td>
<td>0.995</td>
<td>20.4</td>
<td>0.995</td>
<td>22.14</td>
<td>0.995</td>
</tr>
<tr>
<td>13</td>
<td>17.18</td>
<td>0.993</td>
<td>19.73</td>
<td>0.998</td>
<td>21.01</td>
<td>0.998</td>
</tr>
<tr>
<td>14</td>
<td>16.03</td>
<td>0.997</td>
<td>18.28</td>
<td>0.996</td>
<td>20.75</td>
<td>0.992</td>
</tr>
<tr>
<td>15</td>
<td>20.72</td>
<td>0.999</td>
<td>22.33</td>
<td>0.999</td>
<td>23.82</td>
<td>0.997</td>
</tr>
<tr>
<td>16</td>
<td>12.55</td>
<td>0.993</td>
<td>18.13</td>
<td>0.986</td>
<td>20.85</td>
<td>0.987</td>
</tr>
<tr>
<td>17</td>
<td>20.64</td>
<td>0.994</td>
<td>27.11</td>
<td>0.993</td>
<td>29.87</td>
<td>0.994</td>
</tr>
<tr>
<td>18</td>
<td>14.79</td>
<td>0.991</td>
<td>20.59</td>
<td>0.99</td>
<td>23.18</td>
<td>0.991</td>
</tr>
<tr>
<td>19</td>
<td>13.48</td>
<td>0.999</td>
<td>15.16</td>
<td>0.998</td>
<td>16.14</td>
<td>0.998</td>
</tr>
<tr>
<td>20</td>
<td>20.79</td>
<td>0.997</td>
<td>22.81</td>
<td>0.999</td>
<td>23.66</td>
<td>0.999</td>
</tr>
<tr>
<td>21</td>
<td>12.63</td>
<td>0.999</td>
<td>13.59</td>
<td>0.999</td>
<td>14.84</td>
<td>0.996</td>
</tr>
<tr>
<td>22</td>
<td>18.8</td>
<td>0.994</td>
<td>24.4</td>
<td>0.994</td>
<td>26.77</td>
<td>0.994</td>
</tr>
<tr>
<td>23</td>
<td>15.43</td>
<td>0.999</td>
<td>14.23</td>
<td>0.999</td>
<td>14.57</td>
<td>0.999</td>
</tr>
<tr>
<td>24</td>
<td>17.33</td>
<td>0.994</td>
<td>20.33</td>
<td>0.998</td>
<td>21.37</td>
<td>0.998</td>
</tr>
<tr>
<td>25</td>
<td>15.66</td>
<td>0.999</td>
<td>17.08</td>
<td>0.999</td>
<td>18.16</td>
<td>0.998</td>
</tr>
</tbody>
</table>
The ‘K’ values and $R^2$ values for each subject calculated using regression forces from 30-83.3N, 83.3-136.7N, 136.7-200N are shown in Table 7.2.

The group mean ‘K’ values, calculated by performing linear regressions from 30-83.3N, 83.3-136.7N, and 136.7-200N were 15.9N/mm (SD=3.46), 20.7N/mm (SD=4.87), and 24.5N/mm (SD=5.32) respectively. Lower reliability was observed between these measures, with the ICC (2,1) being 0.39 (95% CI=0.002-0.70), and the ICC (3,1) being 0.72 (95% CI=0.53-0.85) with the difference between these two forms of the ICC also suggesting the presence of an additive bias.

A repeated measures analysis of variance demonstrated a significant difference existed between the three measures ($p=0.00103$). The results of the Student-Newman-Keuls test demonstrated that ‘K’ values obtained by regressing between 30-83.3 N/mm were significantly lower than values obtained by regressing between 83.3-136.7 N/mm, which in turn were significantly lower than values obtained by regressing between 136.7-200 N/mm ($p<0.05$).
Table 7.2: K and $r^2$ values for each subject calculated using regression forces from 30-83.3 N, 83.3-136.7 N, 136.7-200 N.

<table>
<thead>
<tr>
<th>Subject</th>
<th>30-83.3</th>
<th>$r^2$</th>
<th>83.3-136.7</th>
<th>$r^2$</th>
<th>136.7-200</th>
<th>$r^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>9.42</td>
<td>0.999</td>
<td>13.03</td>
<td>0.998</td>
<td>17.22</td>
<td>0.999</td>
</tr>
<tr>
<td>2</td>
<td>19.13</td>
<td>0.999</td>
<td>21.13</td>
<td>0.999</td>
<td>23.14</td>
<td>0.999</td>
</tr>
<tr>
<td>3</td>
<td>15.02</td>
<td>0.997</td>
<td>17.91</td>
<td>1</td>
<td>19.64</td>
<td>0.999</td>
</tr>
<tr>
<td>4</td>
<td>14.77</td>
<td>0.999</td>
<td>19.23</td>
<td>0.999</td>
<td>23.25</td>
<td>0.999</td>
</tr>
<tr>
<td>5</td>
<td>11.27</td>
<td>0.999</td>
<td>12.97</td>
<td>0.998</td>
<td>17.05</td>
<td>0.999</td>
</tr>
<tr>
<td>6</td>
<td>22.8</td>
<td>0.997</td>
<td>27.45</td>
<td>0.999</td>
<td>29.44</td>
<td>0.999</td>
</tr>
<tr>
<td>7</td>
<td>12.88</td>
<td>0.993</td>
<td>24.1</td>
<td>0.999</td>
<td>28</td>
<td>0.97</td>
</tr>
<tr>
<td>8</td>
<td>16.77</td>
<td>0.998</td>
<td>17.99</td>
<td>0.999</td>
<td>22.48</td>
<td>0.999</td>
</tr>
<tr>
<td>9</td>
<td>11.76</td>
<td>0.996</td>
<td>19.01</td>
<td>0.999</td>
<td>25.9</td>
<td>0.999</td>
</tr>
<tr>
<td>10</td>
<td>12.11</td>
<td>0.997</td>
<td>16.94</td>
<td>0.999</td>
<td>21.24</td>
<td>0.999</td>
</tr>
<tr>
<td>11</td>
<td>18.16</td>
<td>0.995</td>
<td>22.76</td>
<td>0.999</td>
<td>20.46</td>
<td>0.999</td>
</tr>
<tr>
<td>12</td>
<td>16.13</td>
<td>0.993</td>
<td>23.23</td>
<td>0.999</td>
<td>26.34</td>
<td>0.999</td>
</tr>
<tr>
<td>13</td>
<td>17.36</td>
<td>0.997</td>
<td>20.87</td>
<td>0.999</td>
<td>24.58</td>
<td>0.999</td>
</tr>
<tr>
<td>14</td>
<td>16.06</td>
<td>0.999</td>
<td>20.76</td>
<td>0.998</td>
<td>29.27</td>
<td>0.998</td>
</tr>
<tr>
<td>15</td>
<td>20.76</td>
<td>0.999</td>
<td>23.55</td>
<td>0.999</td>
<td>28.73</td>
<td>0.997</td>
</tr>
<tr>
<td>16</td>
<td>12.83</td>
<td>0.993</td>
<td>24.04</td>
<td>0.999</td>
<td>27.76</td>
<td>0.999</td>
</tr>
<tr>
<td>17</td>
<td>21.03</td>
<td>0.994</td>
<td>32.29</td>
<td>0.999</td>
<td>37.31</td>
<td>0.997</td>
</tr>
<tr>
<td>18</td>
<td>15.13</td>
<td>0.991</td>
<td>25.07</td>
<td>0.999</td>
<td>30.96</td>
<td>0.999</td>
</tr>
<tr>
<td>19</td>
<td>13.51</td>
<td>0.999</td>
<td>16.79</td>
<td>0.999</td>
<td>18.59</td>
<td>0.999</td>
</tr>
<tr>
<td>20</td>
<td>20.99</td>
<td>0.997</td>
<td>23.24</td>
<td>0.999</td>
<td>26.43</td>
<td>0.999</td>
</tr>
<tr>
<td>21</td>
<td>12.66</td>
<td>0.999</td>
<td>14.18</td>
<td>0.999</td>
<td>19.17</td>
<td>0.999</td>
</tr>
<tr>
<td>22</td>
<td>19.15</td>
<td>0.994</td>
<td>28.05</td>
<td>0.999</td>
<td>32.92</td>
<td>0.999</td>
</tr>
<tr>
<td>23</td>
<td>15.36</td>
<td>0.999</td>
<td>13.54</td>
<td>0.999</td>
<td>16.54</td>
<td>0.999</td>
</tr>
<tr>
<td>24</td>
<td>17.62</td>
<td>0.994</td>
<td>21.57</td>
<td>0.999</td>
<td>24.83</td>
<td>0.999</td>
</tr>
<tr>
<td>25</td>
<td>15.71</td>
<td>0.999</td>
<td>18.4</td>
<td>0.999</td>
<td>20.88</td>
<td>0.999</td>
</tr>
</tbody>
</table>
DISCUSSION

The results of this study suggest that physiotherapists who judge PA stiffness by sampling from similar minimum forces but up to different maximum forces are predominantly measuring the same property, however, those using high maximum forces will be likely to report higher stiffness. On the other hand, physiotherapists who judge PA stiffness by sampling with the same-width force intervals but with different minimum and maximum forces will demonstrate lower reliability in the values obtained, having less correspondence in terms of the property they are assessing. For example, physiotherapists who push lightly when performing PA stiffness assessment will demonstrate a poor reliability with physiotherapists who push moderately firmly, and demonstrate a poor reliability with physiotherapists who push very firmly (assuming that therapists are able to judge the elastic stiffness dimension). Again, the application of higher forces will be likely to result in higher evaluated stiffness.

It appears therefore, that the gradient of the line fitted to the force/displacement curve obtained when applying 200 N of force to the lumbar spine may vary depending on the limits of force considered. For example, the current study found a 24% change in the gradient obtained by extending the limits of force reached from 30-80N to 30-200N. This finding is extremely similar to the 25% difference found by Lee et al (1997) in their pilot study. Interestingly, the current study found that a greater change in the gradient would occur if physiotherapists were reporting stiffness based on evidence from limited-width force intervals with differing minimum and maximum
forces. For example, this study found a 42% difference in the stiffness obtained when regressing between 30-83.3N and 136.7-200N.

The reasons for the difference in the gradient obtained when regressing between different force limits appears to be due to variance between subjects. For example, some subjects show extremely linear force/displacement relations above 30N, and calculation of the gradient varies little even though the limits of the force interval sampled may change (see Figure 7.1).

**Fig 7.1** Force/displacement curve (open circles) for a subject exhibiting predominantly linear force/displacement behaviour. Regression lines have been fitted using different force intervals but same minimum forces (left figure), and using different fixed intervals of the force/displacement curve (right figure).
However, there are some subjects whose functions demonstrate a less linear pattern, and calculating the gradient for different force limits will produce ‘K’ values that vary substantially (see Figure 7.2).

**Fig 7.2** Force/displacement curve (open circles) for a subject exhibiting less linear force/displacement behaviour. Regression lines have been fitted using different force intervals but same minimum forces (left figure), and using different fixed intervals of the force/displacement curve (right figure).
To date there is little information available enabling prediction of whether a subject will exhibit typical linear behaviour above 30N or exhibit a more curvilinear force/displacement relation. One issue for future research in this area is to examine any association between the degree of linearity of the PA response, and factors, yet to be determined, that may produce a more curvilinear response.

The findings of this study go some way to explain the results of the in vivo training study reported in Chapter 4. It is possible that one of the reasons for the raters' inability to improve the accuracy of their judgements of lumbar PA stiffness during training with feedback relates to the fact that they may have been sampling different subjects' PA stiffness using different force limits from those used during mechanical testing. One possibility for improved training studies may be to ensure that raters are first trained to sample stiffness using specified force limits and that these same force limits are used to measure stiffness using the mechanical device. Also, information is currently lacking which would enable researchers to determine which force limits to use. If it is demonstrated that subjects with LBP show changes in their elastic stiffness over time (when calculated over a specific force interval), this would provide some early indication of the force limits to use during PA stiffness-assessment training.

This study also provides some explanation as to why the reliability of manual judgements of PA stiffness has been found to be so poor. It has been often documented that physiotherapists apply a wide range of forces when performing PA stiffness assessment and mobilization (Lee and Moseley 1991, Matyas and Bach 1985, Simmonds et al 1995, Petry 1995). One likely reason for the lack of consistency between physiotherapists when manually judging PA stiffness may relate to the varying forces used to judge stiffness. The findings here with respect to the maximum force used have several precedents in the research literature.
Similar to the pilot study by Lee et al (1997), the current study found that stiffness increased with increasing load, an observation consistent with previous reports of soft tissue behaviour (Markolf 1972, Yahia et al 1991, Taylor et al 1990). Markolf (1972), in an in vitro experiment, examined the load-deformation behaviour of whole intervertebral joints, and demonstrated increasing stiffness with increasing load. Yahia et al (1991), examined the rheological properties of the interspinous-supraspinous ligament complex and demonstrated similar behaviour. Finally Taylor et al (1990), who examined the viscoelastic properties of muscle-tendon units, also found that the application of increased load resulted in increased stiffness.

This study suggests that instrumented measurement of PA stiffness must be performed by calculating ‘K’ values from the same force intervals. The study reported in Chapter 3 demonstrated the high reliability of instrumented measures of PA stiffness when linear regression was performed between 30 and 90N. Therefore, when comparing PA stiffness between individuals or before and after treatment the force intervals used to calculate ‘K’ should remain the same. It would be useful to record the force interval used to calculate the ‘K’ value with the ‘K’ value.

CONCLUSION

Because the data reported here demonstrate that both the maximum force reached, and the limits of the fixed intervals sampled for evidence, affect the measured values of stiffness it can be argued that one of the reasons for the poor reliability of manual judgements of PA stiffness is that physiotherapists perform stiffness assessments by sampling using different force limits. To have improved reliability, a revised protocol for PA stiffness testing would therefore ensure that physiotherapists judge stiffness by sampling between specified force intervals rather than determining their own force
limits. By implication future training studies aimed at improving the accuracy of manual judgements of PA stiffness should train raters to sample PA stiffness over a specified force interval.

Nevertheless, the training of physiotherapists to accurately and reliably assess PA stiffness is only of value if there is demonstrated to be a relationship between PA stiffness and LBP. Chapter 8 reports on a study designed to investigate this hypothesized link.
CHAPTER 8

AN INVESTIGATION OF THE RELATIONSHIP BETWEEN LOW BACK PAIN AND LUMBAR POSTEROANTERIOR STIFFNESS

The work contained in this chapter has been published as:

Objective: To investigate the relationship between low back pain (LBP) and lumbar posteroanterior (PA) stiffness. Design: A repeated measures design was used to measure lumbar PA stiffness on 2 occasions in subjects with and without LBP. Subjects: Twenty-five subjects with acute or subacute LBP and twenty-five painfree subjects participated. Pain subjects reported pain on the application of a manual PA force to the lumbar spine and had no contraindication to PA stiffness testing. Painfree subjects reported no history of LBP requiring treatment, and obtained a score of 0 on the McGill Pain Questionnaire. Methods: Posteroanterior stiffness was measured in subjects with LBP when (i) they first presented with pain, and, (ii) when their pain had resolved by more than 80%. Painfree subjects, matched with pain subjects on gender, age, vertebral level to be tested and time between tests, were also measured on two occasions, in order to control for the effects of repeated stiffness testing and the passage of time. Results: In subjects with LBP, stiffness decreased by 1.21N/mm between test 1 and test 2. A paired t-test found a significant difference between the tests (t=3.04, df=24, p=0.006). In subjects without pain there was an increase in stiffness of 0.74N/mm between test 1 and test 2, a paired t-test finding no significant difference between the tests (t=−1.673, df=24, p=0.107). Conclusions: Subjects with LBP showed increased PA stiffness compared to when they had little or no pain, while painfree subjects showed unchanged PA stiffness over time.
INTRODUCTION

Low back pain (LBP) forms one of the most common reasons for people seeking medical attention. However, despite the prevalence of LBP, there is often great difficulty in establishing its underlying cause. Even after an extensive medical workup, only about 15% of patients can be given a definitive diagnosis (Bigos et al 1994, Kelsey 1982). Because of this, clinicians treating patients with LBP often establish a clinical diagnosis for the patient based on the patient’s presenting signs and symptoms.

Both restricted PA movement and increased PA stiffness of the lumbar region are clinical signs thought to commonly occur in patients with LBP (Maitland 1986, Jull et al 1994a). Inflammation and immobilization may contribute to this loss of mobility. For example, it has been demonstrated that immobilization may produce a loss of extensibility of the connective tissues of the spine, producing shortening of joint capsules and adjacent ligaments and tendons (Akeson et al 1980). Following immobilization there is an alteration in the biochemical composition of the connective tissue involving a decrease in the glycosaminoglycan and water content, and an increase in collagen synthesis. The decrease in water content causes the collagen fibres to lie much closer together, resulting in collagen proliferation producing additional interfibre crosslinks between fibres (Amiel et al 1983). This may inhibit the lengthening of spinal connective tissues such as capsule, ligament and tendon, resulting in decreased movement. In addition, it has been suggested that spinal movement may be decreased due to degeneration of articular structures (Rodnan and
Schumacher 1983), or to the presence of paravertebral muscle spasm (Jull et al 1994a).

To date there has been little investigation of whether patients with LBP do have altered PA stiffness. A pilot study (Shirley and Lee 1993) provides some information supporting the existence of a relationship between LBP and PA stiffness. In this small-\(n\) study it was found that two LBP subjects showed higher lumbar PA stiffness than six unmatched controls. The study by Shirley and Lee (1993) was inconclusive due to the small subject numbers and the use of two unmatched groups of subjects. A more powerful investigation of the relationship between LBP and lumbar stiffness would be achieved by measuring PA stiffness in the same group of subjects under two different conditions; firstly, when they had significant pain and secondly, when their pain had largely resolved.

One of the reasons for the absence of research investigating the effect of the presence of pain on the stiffness behaviour of the spine has been the difficulty in reliably measuring the amount of PA displacement occurring when applying a PA force to the lumbar region, and the resistance to this displacement. Many studies have demonstrated the poor reliability of subjective judgments of PA stiffness (Maher and Adams 1994, Binkley et al 1995), and therefore mechanical devices have been developed to measure PA stiffness in the lumbar region (Lee and Svensson 1990, Lee and Evans 1992, Latimer et al 1996a). More recently, the development of a portable
stiffness assessment device (see Chapter 2) has facilitated access to a population suffering LBP, thereby enabling the current research project.

The aim of this study therefore, was to investigate whether lumbar PA stiffness is different when subjects have LBP compared to when they have little or no pain. The experimental hypotheses were:

(i) lumbar PA stiffness is decreased in subjects when they have little or no LBP compared to when they have LBP.

(ii) there is no change in lumbar PA stiffness over time in subjects not reporting LBP.

METHOD

Research design

A two group, repeated measures design was employed in this study (see Figure 8.1). Instrument-determined spinal stiffness assessments at the symptomatic level were performed on 25 subjects with LBP (Pain Group); firstly, when they presented complaining of lumbar pain, and secondly, when their reported pain had improved by at least 80%. A matched group of 25 subjects without LBP (Painfree Group) were included in the research design, and also measured on two occasions, to control for the effects of repeated stiffness testing and to observe PA stiffness over time in painfree subjects.
Subjects

Twenty-five subjects (15 female, 10 male, mean age 36.5 years) with acute/subacute LBP (mean duration 18.1 days) were recruited for the study. All subjects reported pain on the application of a PA force to at least one level of the lumbar spine, and all were native English speakers enabling them to complete the McGill Pain Questionnaire (MPQ) (Melzack 1975). Subjects were excluded if they were
pregnant, had a history of lumbar spine surgery, or had any contraindication to PA stiffness testing e.g., malignancy, inflammatory or infective process affecting the spine, moderate osteoporosis, spondylolythesis, or signs of altered nerve conduction. Subjects complaining of significant pain, unrelated to their lumbar symptoms, elsewhere in the body were also excluded, as the unrelated pain could distort the measurement of LBP when using the McGill Pain Questionnaire. Subjects are described in detail in Table 8.1.

Twenty-five painfree subjects (15 male, 10 female, mean age 35.1 years) were also recruited. These subjects reported no history of LBP requiring treatment, and obtained a score of 0 on the McGill Pain Questionnaire. Subjects were excluded if they were pregnant or had a known contraindication to PA stiffness testing. Painfree subjects were matched with the pain subjects on gender, age, vertebral level to be tested and time between tests. An independent samples t-test demonstrated that there was no significant difference between the Pain and Painfree Groups, on age (t=-0.4537, df=48, p=0.65) or number of days between tests (t=-0.2401, df=48, p=0.81). Subjects are described in detail in Table 8.2.
Table 8.1. LBP subject characteristics.

<table>
<thead>
<tr>
<th>Subject</th>
<th>Gender</th>
<th>Age</th>
<th>BMI*</th>
<th>Lumbar Level tested</th>
<th>Initial MPQ# Score</th>
<th>Final MPQ# Score</th>
<th>Days between tests</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>F</td>
<td>15</td>
<td>25.25</td>
<td>L2 - L3</td>
<td>5</td>
<td>18</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>M</td>
<td>29</td>
<td>28.72</td>
<td></td>
<td>4</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>F</td>
<td>35</td>
<td>30.84</td>
<td></td>
<td>3</td>
<td>19</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>F</td>
<td>45</td>
<td>33.88</td>
<td></td>
<td>5</td>
<td>22</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>F</td>
<td>18</td>
<td>21.08</td>
<td></td>
<td>4</td>
<td>15</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>F</td>
<td>69</td>
<td>27.35</td>
<td></td>
<td>5</td>
<td>20</td>
<td>3</td>
</tr>
<tr>
<td>7</td>
<td>F</td>
<td>16</td>
<td>17.27</td>
<td></td>
<td>5</td>
<td>8</td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td>F</td>
<td>45</td>
<td>23.80</td>
<td></td>
<td>5</td>
<td>16</td>
<td>0</td>
</tr>
<tr>
<td>9</td>
<td>M</td>
<td>22</td>
<td>26.30</td>
<td></td>
<td>5</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>10</td>
<td>F</td>
<td>41</td>
<td>23.82</td>
<td></td>
<td>5</td>
<td>23</td>
<td>0</td>
</tr>
<tr>
<td>11</td>
<td>M</td>
<td>49</td>
<td>22.84</td>
<td></td>
<td>5</td>
<td>28</td>
<td>0</td>
</tr>
<tr>
<td>12</td>
<td>F</td>
<td>47</td>
<td>22.36</td>
<td></td>
<td>5</td>
<td>24</td>
<td>2</td>
</tr>
<tr>
<td>13</td>
<td>M</td>
<td>36</td>
<td>25.84</td>
<td></td>
<td>5</td>
<td>23</td>
<td>4</td>
</tr>
<tr>
<td>14</td>
<td>M</td>
<td>38</td>
<td>21.91</td>
<td></td>
<td>3</td>
<td>14</td>
<td>0</td>
</tr>
<tr>
<td>15</td>
<td>F</td>
<td>40</td>
<td>24.56</td>
<td></td>
<td>3</td>
<td>35</td>
<td>12</td>
</tr>
<tr>
<td>16</td>
<td>F</td>
<td>26</td>
<td>21.87</td>
<td></td>
<td>5</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>17</td>
<td>M</td>
<td>35</td>
<td>25.93</td>
<td></td>
<td>5</td>
<td>16</td>
<td>0</td>
</tr>
<tr>
<td>18</td>
<td>M</td>
<td>23</td>
<td>25.30</td>
<td></td>
<td>3</td>
<td>14</td>
<td>0</td>
</tr>
<tr>
<td>19</td>
<td>M</td>
<td>36</td>
<td>27.13</td>
<td></td>
<td>3</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>20</td>
<td>F</td>
<td>41</td>
<td>18.06</td>
<td></td>
<td>4</td>
<td>11</td>
<td>0</td>
</tr>
<tr>
<td>21</td>
<td>M</td>
<td>42</td>
<td>25.97</td>
<td></td>
<td>5</td>
<td>8</td>
<td>0</td>
</tr>
<tr>
<td>22</td>
<td>F</td>
<td>44</td>
<td>27.43</td>
<td></td>
<td>3</td>
<td>22</td>
<td>1</td>
</tr>
<tr>
<td>23</td>
<td>F</td>
<td>30</td>
<td>21.78</td>
<td></td>
<td>5</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>24</td>
<td>F</td>
<td>45</td>
<td>32.07</td>
<td></td>
<td>5</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>25</td>
<td>M</td>
<td>46</td>
<td>31.79</td>
<td></td>
<td>5</td>
<td>9</td>
<td>0</td>
</tr>
</tbody>
</table>

| Mean    | 36.52 | 25.32 | 15.4 | 0.92 | 22.64 |
| SD      | 12.25 | 4.12  | 7.98 | 2.53 | 23.65 |

*BMI=Body mass index (mass/height²)

#MPQ=McGill Pain Questionnaire Score (Pain Rating Index)
Table 8.2. Painfree subject characteristics.

<table>
<thead>
<tr>
<th>Subject</th>
<th>Gender</th>
<th>Age yrs</th>
<th>BMI* kg/cm²</th>
<th>Level Tested L₂ - L₅</th>
<th>Days b/t tests days</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>F</td>
<td>19</td>
<td>22.77</td>
<td>5</td>
<td>7</td>
</tr>
<tr>
<td>2</td>
<td>M</td>
<td>34</td>
<td>24.56</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>F</td>
<td>31</td>
<td>23.88</td>
<td>3</td>
<td>21</td>
</tr>
<tr>
<td>4</td>
<td>F</td>
<td>43</td>
<td>28.19</td>
<td>5</td>
<td>28</td>
</tr>
<tr>
<td>5</td>
<td>F</td>
<td>24</td>
<td>22.27</td>
<td>4</td>
<td>57</td>
</tr>
<tr>
<td>6</td>
<td>F</td>
<td>57</td>
<td>24.21</td>
<td>5</td>
<td>13</td>
</tr>
<tr>
<td>7</td>
<td>F</td>
<td>23</td>
<td>18.28</td>
<td>5</td>
<td>56</td>
</tr>
<tr>
<td>8</td>
<td>F</td>
<td>43</td>
<td>27.43</td>
<td>5</td>
<td>23</td>
</tr>
<tr>
<td>9</td>
<td>M</td>
<td>23</td>
<td>25.81</td>
<td>5</td>
<td>7</td>
</tr>
<tr>
<td>10</td>
<td>F</td>
<td>30</td>
<td>21.52</td>
<td>5</td>
<td>53</td>
</tr>
<tr>
<td>11</td>
<td>M</td>
<td>43</td>
<td>31.22</td>
<td>5</td>
<td>26</td>
</tr>
<tr>
<td>12</td>
<td>F</td>
<td>56</td>
<td>22.77</td>
<td>5</td>
<td>55</td>
</tr>
<tr>
<td>13</td>
<td>M</td>
<td>34</td>
<td>21.42</td>
<td>5</td>
<td>7</td>
</tr>
<tr>
<td>14</td>
<td>M</td>
<td>38</td>
<td>27.62</td>
<td>3</td>
<td>9</td>
</tr>
<tr>
<td>15</td>
<td>F</td>
<td>38</td>
<td>25.99</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>16</td>
<td>F</td>
<td>29</td>
<td>19.69</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>17</td>
<td>M</td>
<td>34</td>
<td>24.33</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>18</td>
<td>M</td>
<td>20</td>
<td>21.18</td>
<td>3</td>
<td>9</td>
</tr>
<tr>
<td>19</td>
<td>M</td>
<td>38</td>
<td>24.18</td>
<td>3</td>
<td>28</td>
</tr>
<tr>
<td>20</td>
<td>F</td>
<td>28</td>
<td>22.18</td>
<td>4</td>
<td>14</td>
</tr>
<tr>
<td>21</td>
<td>M</td>
<td>44</td>
<td>22.15</td>
<td>5</td>
<td>7</td>
</tr>
<tr>
<td>22</td>
<td>F</td>
<td>40</td>
<td>27.18</td>
<td>3</td>
<td>63</td>
</tr>
<tr>
<td>23</td>
<td>F</td>
<td>28</td>
<td>33.96</td>
<td>5</td>
<td>14</td>
</tr>
<tr>
<td>24</td>
<td>F</td>
<td>34</td>
<td>27.46</td>
<td>5</td>
<td>11</td>
</tr>
<tr>
<td>25</td>
<td>M</td>
<td>46</td>
<td>22.20</td>
<td>5</td>
<td>7</td>
</tr>
</tbody>
</table>

**Mean**: 35.1 yrs  24.50 kg/cm²  21.16 days

**SD**: 10.1 yrs  3.57 kg/cm²  9.91 days

*BMI=Body mass index (mass/height²)*
The population sampled was one of convenience, and subjects were recruited at locations in both Australia and the United States over a one year period until the required numbers were obtained. Informed consent was obtained from all subjects prior to participation in the study. Ethical approval for the study was obtained from The University of Sydney Human Ethical Review Committee, the South Sydney Area Health Service Research Committee and the Internal Review Board of the Hospital for Bone and Joint Diseases, New York University.

**Measurement instruments**

**Measurement of stiffness**

A portable stiffness assessment device was used to measure the posteroanterior stiffness of the lumbar spine. The lumbar stiffness assessment device has been described in detail elsewhere (see Chapter 2). This device applies a PA force via a metal applicator or probe to a selected vertebral level of the lumbar region and measures the amount of resulting displacement, with the relative change in force and displacement being used to calculate stiffness.

The main measure generated by using the device is the stiffness coefficient, K. This is generated by calculating the slope of the linear portion of the force/displacement curve (from 30N to 90N) obtained when testing the lumbar region.
Another characteristic of the force/displacement curve was also measured in this study. An indication of the length of the toe region was obtained by calculating the amount of displacement (mm) occurring between 0.5N and 30N of force. This measure was termed displacement D30. The starting point for measurement of 0.5N rather than 0N was chosen as this is the contact force applied prior to testing the selected vertebral level. The displacement D30 measure has been only infrequently reported in previous studies that performed mechanical stiffness testing, however, it was argued here that this portion of the force/displacement curve may be affected by the presence of pain and therefore should be quantified.

The stiffness assessment device has been found to be reliable in its measurement of lumbar PA stiffness in subjects with LBP, in that high test-retest reliability has been observed for the stiffness coefficient, K, with the ICC (2,1) value being 0.96 (95% CI=0.91-0.98). For the displacement D30 the ICC (2,1) obtained was only slightly lower at 0.89 (95% CI=0.76-0.95) (Latimer et al 1996b, Chapter 2 and 3).

Measurement of pain

(i) McGill Pain Questionnaire

Administration of the McGill Pain Questionnaire (MPQ) was used to measure the amount of LBP experienced by subjects in the 24 hours immediately prior to each test. It was also used to give certainty with regard to the status of the painfree subject group, since any subject not obtaining a score of zero was excluded from this
group. Administration to subjects was as a pencil and paper test, but subjects were encouraged to ask for clarification of any instruction or word that they did not understand. The reliability of the MPQ when used as a measure of pain over a period of time has been shown to be high (Hunter et al 1979) and the questionnaire also displays acceptable face and construct validity (Reading 1983). The number of LBP studies that include the MPQ as a dependent measure testify to this (Melzack et al 1980, Fox and Melzack 1976).

(ii) Absolute Visual Analogue Scale (AVAS)

An absolute visual analogue scale (shown in Figure 8.2) was used to evaluate the intensity of pain experienced during the performance of each lumbar PA stiffness test. The AVAS was presented to subjects as a vertical scale, 100 mm in length, with anchor points of 'no pain' and 'the worst possible pain' at either end of the scale. When marking the scale, subjects were instructed to differentiate low back and/or leg pain from the slight discomfort produced by the testing procedure. Use of the AVAS has demonstrated high reliability with pain estimates showing good reliability over time (Revill et al 1976).
worst possible pain

no pain

Figure 8.2 Absolute visual analogue scale

Procedure

Pain subjects were asked to complete a MPQ to quantify the amount of low back ± leg pain that they had experienced in the preceding 24 hours. Subjects undressed to their underwear and put on a hospital gown. They were then asked to lie prone on the testing table, and a manual PA force was applied to each lumbar vertebral level in order to identify the vertebral level to be examined using the stiffness assessment device. This level was defined as:

(i) the level that best reproduced the subject’s presenting pain on manual PA stiffness testing, or if no levels reproduced the presenting pain,

(ii) the level with the greatest amount of local pain that could logically refer pain to the symptomatic area.
Subjects were then positioned in prone for PA stiffness testing (Figure 8.3). The force applicator of the stiffness device was brought into contact with the skin overlying the vertebral level to be tested. Five preliminary cycles were then performed using the testing device. The purpose of these cycles was to ensure that the subject could easily tolerate the force required to perform the stiffness test, to help familiarize the subject with the stiffness testing procedure, and to precondition the viscoelastic tissues prior to data collection (Loebl 1972). Stiffness data collection then commenced with five testing cycles being performed to a maximum force of 100N at a frequency of 0.5Hz. Following stiffness testing, subjects marked a point on a 100 mm absolute visual analogue scale to indicate the amount of LBP ± leg pain that they experienced during the stiffness testing procedure.

Fig 8.3 *Lumbar PA stiffness testing.*
Subjects were then asked to contact their treating physiotherapist or the researcher when low back ± leg pain had improved by at least 80%. At that time the McGill Pain Questionnaire was readministered, as on the first occasion, and if the score obtained was at least 80% less than the initial value, stiffness testing was repeated in exactly the same manner as before. Again, subjects were asked to mark an AVAS to indicate the amount of pain that they felt during the stiffness testing procedure.

In the Painfree Group of subjects the stiffness testing procedure was performed as for the Pain Group, except for the identification of the level to be tested and determining the time of retest. For Painfree subjects, the vertebral level to be tested corresponded to the level tested in their age-and gender-matched pain subject, and the second test was performed following a time interval as close as possible to the time which had elapsed between tests for the matched pain subject.

Data extraction

Stiffness data

Lumbar PA stiffness data were analyzed in the following manner. For each test occasion, force/displacement data for five loading cycles were obtained. The force/displacement data for four loading cycles only (cycles 2-5) were analyzed, cycle 1 being discarded as representing transient behaviour on start-up. The remaining four cycles were averaged and a mean force/displacement curve obtained. From this mean curve the main measure, the stiffness coefficient, K, was calculated. An indication of
the length of the toe region, the displacement, D30, was also obtained by measuring the amount of downwards displacement in millimeters occurring between forces of 0.5 and 30 Newtons.

**Pain data**

AVAS values were obtained by measuring the distance in millimeters from the anchor point 'no pain' to the mark on the page made by the subject.

**Statistical analysis**

Although the design of the study was a $2 \times 2$ (ie. two groups $\times$ two repeated measures) design, neither a main effect of groups averaged across repeats (tests) nor a main effect of tests averaged across groups was hypothesized to occur. Accordingly, data were first considered descriptively, then t-tests conducted within groups to assess the statistical significance of any changes in the measures of interest between test 1 and test 2.

Group means and standard errors (SE) for the stiffness coefficient, K, the displacement, D30, and the AVAS score, were calculated for each test for the pain and painfree groups. A paired t-test was then used to compare the stiffness coefficient, K between test 1 and test 2 in both groups. Similarly, a paired t-test was used to compare differences between test 1 and test 2 for the displacement, D30, and the AVAS scores, in both pain and painfree groups.
For the LBP group Pearson Product-moment correlations were performed to examine whether the change in the stiffness coefficient, K, and the change in the displacement D30, between the two tests was correlated with subject age, height, weight, body mass index (BMI), initial McGill score, initial AVAS score, change in McGill score between tests, change in AVAS between tests, or days between tests. Pearson Product-moment correlations were also performed to examine whether there was a correlation between the change in AVAS scores and subject age, initial McGill score, final McGill score and the change in McGill score between the two test occasions. Due to the large number of correlations performed a significance level of 0.01 was used. Statistical analysis was performed using the ‘SigmaStat’\(^1\) statistical package (Version 1.0).

In addition, to give some indication of the reliability of the PA stiffness measurement procedure over an extended period of time the test-retest reliability of measures of K and the displacement D30 was investigated by calculating ICC (2,1)’s using test 1 and test 2 for the painfree group. The mean number of days between the two tests was 21.2 (SD=9.91), ranging from 2-63 days. For this reliability calculation, data from all spinal levels (L1-5) was grouped together.

**RESULTS**

**Stiffness coefficient, K**

The mean value obtained for test 1 for the stiffness coefficient, K, in subjects with LBP was 14.96 N/mm (SD=2.74) with the mean decreasing to 13.74 N/mm (SD=2.10) for test 2. Figure 8.4 shows change scores for individual subjects, ranked from left to right in ascending order, according to the magnitude of change in the

\(^{1}\text{Jandel Scientific Software, PO Box 7005, San Rafael, CA 94912-7005.}\)
stiffness coefficient, \( K \). It can be seen that in 68% of subjects there was a decrease in the stiffness coefficient \( K \) between the two test occasions. A paired t-test demonstrated that the mean \( K \) value obtained at test 2 was significantly lower than that obtained at test 1 \( (t=3.04, \text{df}=24, p=0.006) \).

**Fig 8.4** Change in \( K \) for subjects with LBP. Change scores for individual subjects ranked in ascending order from left to right according to the magnitude of change.
No significant correlation was found between any of the following pain subject variables: age, height, weight, BMI, initial McGill score, initial AVAS score, change in McGill score between tests, change in AVAS between tests, or days between tests; and the magnitude of change in the stiffness coefficient ‘K’ (Table 8.3).

Table 8.3: Product-moment correlation between change in K and subject variables for subjects in the LBP group.

<table>
<thead>
<tr>
<th></th>
<th>age</th>
<th>ht</th>
<th>wt</th>
<th>BMI Q(I)</th>
<th>MPQ (C)</th>
<th>AVAS (I)</th>
<th>AVAS (C)</th>
<th>days b/t tests</th>
</tr>
</thead>
<tbody>
<tr>
<td>corr.</td>
<td>0.12</td>
<td>-0.20</td>
<td>-0.02</td>
<td>0.08</td>
<td>-0.05</td>
<td>0.06</td>
<td>-0.36</td>
<td>0.38</td>
</tr>
<tr>
<td>change in ‘K’</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>p level</td>
<td>0.56</td>
<td>0.33</td>
<td>0.93</td>
<td>0.70</td>
<td>0.81</td>
<td>0.80</td>
<td>0.09</td>
<td>0.07</td>
</tr>
</tbody>
</table>

ht = height  
wt = weight  
BMI = body mass index  
MPQ(I) = initial McGill Pain Questionnaire score  
MPQ(C) = change in McGill Pain Questionnaire score  
AVAS(I) = initial AVAS score  
AVAS(C) = change in AVAS score

In the group of painfree subjects the mean value obtained for test 1 for the stiffness coefficient, $K$, was 14.84 N/mm (SD=3.46) increasing to 15.59 N/mm (SD=3.73) for test 2.). Figure 8.5 shows change scores for individual subjects ranked from left to right in ascending order. It can be seen that in 60% of cases there was an increase in the stiffness coefficient, $K$, between the two tests. However, a paired t-test demonstrated that there was no significant difference in the $K$ value between test 1 and 2 ($t$=-1.673, df=24, p=0.107). In addition, there was no significant difference in
K at test 1 between the group of subjects with LBP and the group of painfree subjects (t=0.13, df=48, p=0.89).

Because each subject in the painfree control group had two stiffness measures made at the same lumbar level, it was possible to calculate reliability indices. High test-retest reliability was observed for K, even when measured over an extended time period, the ICC (2,1) value being 0.80. The 95% confidence interval was 0.59-0.91. On 80% of occasions the test-retest difference was less than 2 N/mm.

![Change in K for painfree subjects](image)

**Fig 8.5** Change in K for painfree subjects. Change scores for individual subjects ranked in ascending order from left to right according to the magnitude of change.
Displacement, D30

The mean value obtained for test 1 for the displacement, D30, in subjects with LBP was 4.50 mm (SD=1.43) increasing to 5.15 mm (SD=1.67) for test 2. Figure 8.6 shows change scores for individual subjects, ranked from left to right in ascending order, according to the magnitude of change in the displacement, D30 between test occasions. It can be seen that in 68% of subjects there was an increase in the displacement D30, between the two test occasions. A paired t-test demonstrated that this difference was statistically significant (t=-3.499, df=24, p=0.002).

Fig 8.6 Change in the displacement D30 for subjects with LBP. Change scores for individual subjects ranked in ascending order from left to right according to the magnitude of change.
No significant correlation was found between the pain subject’s age, height, weight, BMI, initial McGill score, initial AVAS score, change in McGill score between tests, change in AVAS between tests, or days between tests, and the change in the displacement D30 (Table 8.4) when the significance level of 0.01 was adopted in order to protect Type 1 error rate against the number of tests.

Table 8.4: Product-moment correlation between change in displacement D30 and subject variables for subjects in the pain group.

<table>
<thead>
<tr>
<th></th>
<th>age</th>
<th>ht</th>
<th>wt</th>
<th>BMI</th>
<th>MPQ (I)</th>
<th>MPQ (C)</th>
<th>AVAS (I)</th>
<th>AVAS (C)</th>
<th>days b/t tests</th>
</tr>
</thead>
<tbody>
<tr>
<td>corr. with change in 'D30'</td>
<td>-0.20</td>
<td>0.44</td>
<td>0.37</td>
<td>0.24</td>
<td>-0.15</td>
<td>0.09</td>
<td>0.16</td>
<td>-0.05</td>
<td>0.20</td>
</tr>
<tr>
<td>p level (\alpha=0.01)</td>
<td>0.33</td>
<td>0.03</td>
<td>0.06</td>
<td>0.23</td>
<td>0.51</td>
<td>0.66</td>
<td>0.63</td>
<td>0.79</td>
<td>0.34</td>
</tr>
</tbody>
</table>

*where:*  
ht = height  
wt = weight  
BMI = body mass index  
MPQ(I) = initial McGill Pain Questionnaire score  
MPQ(C) = change in McGill Pain Questionnaire score  
AVAS(I) = initial AVAS score  
AVAS(C) = change in AVAS score

In the group of painfree subjects the mean value obtained for test 1 for the displacement, D30, was 4.88 mm (SD=1.14) decreasing to 4.45 mm (SD=1.29) for test 2. Figure 8.7 shows change scores for individual subjects, ranked from left to right in ascending order, according to the magnitude of change in the displacement, D30. It can be seen that in 76% of subjects there was a decrease in the displacement.
D30 between the two test occasions. Displacement D30 values obtained at test 2 were significantly lower than those obtained at test 1 ($t=2.269$, $df=24$, $p=0.033$). There was no significant difference in D30 at test 1 between the group of subjects with LBP and the group of painfree subjects ($t=-1.627$, $df=48$, $p=0.11$).

Moderate test-retest reliability was observed for displacement D30 when measured over an extended time period, the ICC (2,1) value being 0.67. The 95% confidence interval was 0.37-0.84. On 80% of occasions the test-retest difference was less than 1.2mm.

![Figure 8.7](image)

**Figure 8.7** Change in the displacement D30 for painfree subjects. Change scores for individual subjects ranked in ascending order from left to right according to the magnitude of change.
AVAS score

The mean value obtained for test 1 for the amount of pain experienced during the stiffness test, as measured using an AVAS, in subjects with LBP was 21.46mm (SD=21.80) decreasing substantially to 2.23mm (SD=4.04) for test 2. A paired t-test demonstrated that this difference was statistically significant (t=4.523, df=24, p<0.001), and indicating that subjects reported significantly less pain at the second stiffness test.

No significant correlation was found between the change in AVAS score and subject age, initial McGill score, final McGill score and the change in McGill score between the two test occasions (Table 8.5).

<table>
<thead>
<tr>
<th></th>
<th>age</th>
<th>MPQ (I)</th>
<th>MPQ (F)</th>
<th>MPQ (C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>corr. with</td>
<td>0.08</td>
<td>-0.01</td>
<td>0.04</td>
<td>0.09</td>
</tr>
<tr>
<td>change in</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>‘AVAS’</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>p level</td>
<td>0.71</td>
<td>0.97</td>
<td>0.84</td>
<td>0.80</td>
</tr>
</tbody>
</table>

MPQ(I) = initial McGill Pain Questionnaire score  
MPQ(F) = final McGill Pain Questionnaire score  
MPQ(C) = change in McGill Pain Questionnaire score
In the group of painfree subjects the mean AVAS score obtained for test 1 was 0.76mm (SD=2.7) increasing to 0.96mm (SD=3.5) for test 2. A paired t-test demonstrated no significant difference between test 1 and test 2 (t=1.0, df=24, p=0.329).

**DISCUSSION**

For a group of subjects with LBP this study demonstrated that their stiffness coefficient, K, decreased as their LBP decreased. This stiffness reduction represents a change of approximately 8.0% of the original stiffness value. No decrease was observed in a group of painfree subjects who were matched with the LBP subjects on gender, age, vertebral level tested and time between tests.

It is likely that in some pain subjects the change in PA stiffness between conditions would be detectable clinically ie: manually. In this study nine subjects demonstrated reductions in PA stiffness in the range of 14 - 37%. Maher and Adams (1995a) have examined threshold levels for discrimination of the stiffness coefficient, K and found that, on average, manipulative physiotherapists are able to distinguish between stiffness stimuli that vary by 11%. Therefore, in subjects whose PA lumbar stiffness decreases by 14 - 37% as their LBP resolves, it is possible that this magnitude of change in PA stiffness could be detected by the treating clinician.
In this study the ‘K’ value was obtained by calculating the gradient of the force/displacement relation between 30N and 90N. Stiffness measured using these force limits decreased as LBP resolved. Therefore, this finding provides some early indication of the force limits to use when training physiotherapy student raters to perform manual PA stiffness testing. If students are trained to determine PA stiffness magnitude over a 30N-90N force interval, this should enable them to clinically detect reductions in PA stiffness in subjects whose LBP is resolving.

The factors contributing to PA stiffness have not yet been conclusively established but are likely to include inputs from both spinal and extraspinal structures. Spinal structures would include the intervertebral disc, the zygapophyseal joint complex, the spinal ligaments and muscles. Extraspinal structures would include the skin and subcutaneous tissue overlying the spinous process to which the PA force is applied, the pelvis and thorax. Information regarding which of these anatomical factors contributes most significantly to the decreasing PA stiffness seen in subjects with LBP as LBP improves, is not available. In this study, subjects varied considerably in the time elapsed between tests, with some subjects demonstrating a quick resolution of their LBP, while other subjects took several months for their LBP to resolve (mean time between tests = 22.64 days, range = 2-105 days). In subjects demonstrating a rapid improvement in pain it is possible that the presence of paravertebral muscle spasm was a source of increased PA stiffness at the first test. Resolution of this muscle spasm may have produced a decrease in PA stiffness at the second test (Shirley and Lee 1993). Future research may need to measure muscle activity in
order to investigate this hypothesis. In subjects measured with several months between tests, it is possible that changes in the mechanical properties of the connective tissue structures of the spine contributed to the change in PA stiffness.

The amount of pain experienced during stiffness testing in subjects with LBP, as measured using an AVAS, markedly decreased as a subject’s LBP decreased. Although there was not a significant correlation between the amount of change in the AVAS and the final McGill score, the direction of change was similar and both pain scores reduced substantially. This result suggests that the amount of pain produced when applying a PA force to the lumbar spine may be a useful indicator as to whether a subject’s LBP, measured using the McGill pain questionnaire, is decreasing or not. Although the PA stiffness test is already used by some clinicians as a measure of outcome following treatment and of outcome over time, the current report is the first to demonstrate its usefulness through its association with resolution of LBP.

Because this is only the second study, (after the study that reported the reliability of the testing instrument, Latimer et al 1996b, see Chapter 3), to investigate changes in displacement, D30, in a patient population, little information is available about how to interpret these changes. In the present study the displacement, D30, generally increased in subjects with LBP as their LBP resolved. This increase represents a mean change of 14.5% from the original value. Although little is known about this measure, this finding is in agreement with the view often put forward by clinicians treating patients with LBP, that the length of the toe region appears to increase as a
patient's LBP resolves. Interestingly, in the group of matched, painfree subjects the amount of displacement D30 decreased over time. To investigate whether this change may simply reflect the existence of an inverse relationship between K and D30, correlations were computed for K and D30 for the LBP and painfree groups at test 1 and test 2 and also for the change in K and D30 between test 1 and 2. For each group at test 1 and test 2 K and D30 were significantly inversely correlated [LBP group: test 1 $r = -0.63$, test 2 $r = -0.42$; Painfree group: test 1 $r = -0.46$, test 2 $r = -0.59$ ($r_{err} = 0.396$ df=23 $\alpha = 0.05$, 2 tailed test)]. This level of correlation was not evident for the change in K in relation to the change in D30 between test 1 and test 2 (LBP group: change scores $r = -0.25$; Painfree group: change scores $r = -0.32$, both not significant at $\alpha = 0.05$). These results provide evidence that the change in K is independent of the change in D30. However, there does appear to be an inverse association between the two measures when obtained on the same test occasion. Future PA stiffness studies which measure both the K and D30 components of the force/displacement curve are needed to provide an adequate account for the observed association between these measures.

The results of this study also provide evidence of test-retest reliability of the instrumented stiffness measure over an extended time period. Chapter 3 describes the high test-retest reliability of PA stiffness measurement made using the stiffness assessment device when performed over a short time period of approximately 5 minutes (ICC 2,1 for K = 0.96, for D30 = 0.89). In the Chapter 3 study subjects remained still on the testing table between tests, requiring only the indenter of the
stiffness device to be correctly repositioned for the second test. In the study reported in Chapter 6 an ICC (3,1) of 0.80 was obtained for measures of K obtained from subjects while lying on two different plinth surfaces. This lower ICC value is likely to reflect not only error introduced by the different plinth surfaces but also the fact that in this study subjects ambulated for approximately 5 minutes between tests while the plinth surfaces were changed, necessitating subject repositioning as well as indenter repositioning for the second test. If the subject was repositioned slightly differently for the second test the amount of subject area contacting the plinth surface at the second test could vary from the first test, hence producing a different stiffness value and a lower ICC value.

Information from the present study suggests that measures of PA stiffness obtained using the stiffness assessment device demonstrate high reliability when tests are performed with durations of 2-63 days between tests (mean days between tests = 21), despite the fact that variations in subject and indenter positioning must surely have occurred. This implies that the stiffness measurement procedure would be useful in evaluating the effects of a course of physiotherapy treatment on PA stiffness. The data here also provides evidence that lumbar PA stiffness (represented by K) is relatively stable over time in painfree subjects.
CONCLUSION

This study found an association between LBP and PA stiffness. Subjects with LBP were found to have increased PA stiffness (characterized by measuring the stiffness coefficient, K, of the force/displacement curve) compared to when they had little or no pain. Future studies investigating whether certain therapeutic interventions, such as manipulation, actually change PA stiffness, and subsequently whether this change in stiffness produces relief of concurrent LBP, will be needed to obtain further information regarding the exact nature of the relationship between LBP and lumbar PA stiffness.
CHAPTER 9

CONCLUDING REMARKS
OVERVIEW OF THE MAIN RESULTS

A major focus of this thesis has been consideration of improving the reliability and accuracy of posteroanterior stiffness measurement in human backs by both instrumented and manual methods. In so doing, several factors have been identified that affect the measured value of posteroanterior (PA) stiffness. In addition, the development of a portable stiffness assessment device has made it possible to examine the relationship between low back pain (LBP) and lumbar PA stiffness.

The design and development of a portable device for measuring PA stiffness was reported in Chapter 2 of this thesis. Chapter 3 reported reliability for PA stiffness measures made using this device, with an ICC (2,1) of 0.96 being found for measurement of the stiffness coefficient ‘K’ (95% CI 0.91-0.98) and an ICC (2,1) of 0.89 being found for measurement of the toe region ‘D30’ (95% CI 0.76-0.95). The device was also found to be highly accurate in its measurement of the stiffness of a series of aluminium beams, with a maximum error of 2.5%. The advantages arising from the portability of the device are substantial, in that the wherewithal to collect data on site in physiotherapy practices, hospital outpatient departments, as well as at the School of Physiotherapy, has made possible assessment of LBP symptomatic groups. Some of the subjects studied in Chapter 8 of this thesis were measured at the Occupational and Industrial Orthopaedic Centre in New York, when the device was transported there to conduct research.

Research into a number of areas has been facilitated by the development of the stiffness testing device described here, with the information provided about lumbar PA stiffness being used in studies examining therapists’ perception of stiffness (Maher and Adams 1996a, Maher and Adams 1996b), in studies investigating the effects of physiotherapy treatments (Goodsell 1996), and in studies investigating muscle factors
contributing to PA stiffness (Shirley and Lee 1993) as well as being used in the studies described in this thesis.

When designing the mechanical device it was believed that the most important parameter of stiffness to measure was the elastic stiffness dimension. This view was derived from reading manual therapy authors (Jull 1994b, Maitland 1986) who have identified elastic stiffness as the most significant component of total stiffness, with Jull (1994b) stating that:

“manual examination can be described as a very basic in vivo test of spinal segmental mechanics, testing particularly the elastic properties of the viscoelastic spinal tissues. What is measured perceptually is the displacement and the segmental tissue resistance to the applied force. In other words, the clinician is measuring the basic load displacement characteristics for a particular segmental direction of motion.” P 520.

However, information gained from the studies reported in Chapters 4 and 5 challenges the notion of elastic stiffness as the only stiffness dimension attended to during manual stiffness assessment. The study in Chapter 4 was the first study ever conducted in which it was possible to provide accurate and immediate feedback to students on their judgements of PA stiffness of the human lumbar spine. Yet, despite the substantial training provided to the students, they were unable to improve the accuracy of their lumbar stiffness judgements between the pre- and post-tests.

It was considered possible that the fact that students had rated PA stiffness on minimally padded plinths rather than a rigid plinth surface may have contributed to their failure to learn. However, subsequent study (Chapter 6) demonstrated that there was a high correlation between measures obtained on the rigid and padded plinths.
It was also suggested that raters might have been reporting PA stiffness values after sampling from different force intervals, and that a degree of non-linearity of the force/displacement curve was sufficient to have contributed to their inaccuracies. Subsequent study (Chapter 7) found this hypothesis to be true, this study demonstrating that sampling from different fixed intervals with varying maximum and minimum forces between 30 and 200N, would produce significantly different ‘K’ values, the higher the force used in testing, the greater the measured stiffness. Also it was demonstrated that testing PA stiffness using the same minimum force of 30N but varying maximum forces would result in values exhibiting moderate reliability, however, significantly different ‘K’ values would be obtained using force limits of 30-80N, 30-150N, and 30-200N. By implication instrumented measurement of spinal stiffness must be performed by calculating ‘K’ values from the same force intervals. Also, it appears that to effectively train physiotherapy students to rate PA stiffness it will be necessary to train them to assess stiffness by sampling a specified force interval rather than choosing their own force limits.

Another possible explanation for the failure to find evidence of learning relates to the fact that manually-based judgement of PA stiffness is a consequence of simultaneously attending to the many stiffness component dimensions. The study conducted in Chapter 5 demonstrated that the same four raters from the study reported in Chapter 4 could be trained to assess elastic stiffness only, if given quantitative feedback. This suggested that their failure to learn to accurately judge lumbar PA stiffness in vivo did not relate to an inability to learn how to judge elastic stiffness per se. However, when asked to judge elastic stiffness in subjects with varying toe regions, and in subjects whose spines would have been exhibiting other stiffness dimensions such as viscosity, or frictional stiffness, the raters were unable to learn to make correct judgements of elastic stiffness.
In order to more effectively train physiotherapy students to judge PA stiffness, it seems likely that it will be necessary to establish, in patients with LBP, what are the other relevant stiffness dimensions, and how they should be measured. In an early study by Wright and Johns (1961), the authors attempted to break total stiffness measurements down into component dimensions, and to quantify subjects with respect to each of these components. The results of this study, evaluating diseased and normal metacarpalphalangeal joints, suggest that the elastic stiffness component contributed most to total stiffness in both normal and diseased joints.

“It was of interest that in arthritic joints the relation between elastic, viscous, and frictional stiffness was the same as in normal joints. Only in severely damaged joints which showed marked clinical and radiographic deterioration could friction be demonstrated. Even in these joints frictional stiffness did not contribute significantly to joint stiffness, the major component being increased elastic stiffness.” p.42

Although this work was conducted some thirty-six years ago, on the metacarpalphalangeal joint rather than the spine, and quantified the stiffness of a physiological movement rather than an accessory movement, the conceptualization of stiffness by these authors may point to the future for instrumented measurement of spinal stiffness, and training for manual raters of spinal stiffness. Developments may enable a new version of the portable stiffness assessment device (or a new protocol of testing and analysis) to simultaneously measure elastic, viscous and frictional stiffness.
Current status of posteroanterior stiffness assessment in human backs

Table 9.1 sets out the advantages and disadvantages as currently understood regarding the two methods of assessing posteroanterior stiffness in human backs.

Table 9.1 Current status of PA stiffness assessment in human backs.

<table>
<thead>
<tr>
<th>Manual Assessment</th>
<th>Mechanical Assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Advantages</strong></td>
<td><strong>Disadvantages</strong></td>
</tr>
<tr>
<td>inexpensive</td>
<td>judgments found to be unreliable</td>
</tr>
<tr>
<td>therapists believe it to provide them with useful information</td>
<td>judgments are likely to be inaccurate re: elasticity, if influenced by other dimensions</td>
</tr>
<tr>
<td>complex information comes as a single percept</td>
<td>individual therapists derive their own concept based on their pattern of experience, rather than on a known set of training stimuli</td>
</tr>
</tbody>
</table>
Directions for future research

The clear implication arising from this thesis is that to solve the complicated problem of PA stiffness assessment, researchers in the area need to establish which stiffness dimensions exist and which are likely to be perceived by physiotherapists performing manual stiffness assessment. Also, there is a need to establish which mechanical parameters of stiffness are relevant markers of pathology and dysfunction in LBP, and whether the presence of deviant values of any given parameter suggest that manually-based treatments such as mobilization or manipulation will provide the most effective treatment for this patient group. Knowledge of these factors would provide useful information about how to reprogram the stiffness assessment device described in Chapter 2, to facilitate measurement of these parameters.

Chapter 8 suggests that elastic stiffness is one dimension which is altered in subjects with LBP, with measured elastic stiffness significantly reducing as LBP resolves. To examine the diagnostic value of other stiffness parameters, epidemiological studies need to be conducted. In order to identify these other stiffness dimensions, research will need to be informed from the domain of rheology, the study of elastic and plastic phenomena which take place upon deformation of matter (Tanner 1985). Study of rheology will likely also provide ideas regarding methods of modifying the mechanical stiffness measuring device described before so that it can be used to measure all stiffness dimensions thought to be important in total lumbar PA stiffness.

Although Chapter 3 has demonstrated that the instrumented measurement protocol produces highly reliable measures of PA stiffness, there remain differences between the instrumented protocol and the strategies adopted by therapists when they perform PA stiffness assessment and these differences may produce a disassociation between
instrumented measures and manual judgements. With the current instrumented protocol, measures are obtained by firstly applying at least five testing cycles to precondition the vertebral segment, and then collecting data using a loading force of 100N at a frequency of 0.5 Hertz. Depending on the vertebral level to be tested the indenter applies an angled force as discussed in Chapter 3. The instrumented stiffness measures are then obtained by generating a mean curve from the loading curves of cycles 2-5, the first cycle being discarded due to transient behaviour on start up.

It is possible that during the preconditioning process used when obtaining instrumented measures, information related to viscosity is reduced. When manually judging PA stiffness, it has been suggested that therapists make their stiffness judgements based on information gained from the very first PA force application (Maitland 1986). Arguably, therapists are obtaining more information regarding viscosity than the mechanical device, and using this to come to a total judgement of PA stiffness. Therefore, future instrumented measurement of PA stiffness could examine data obtained without preconditioning the vertebral segment, the device being reprogrammed to extract data related to the viscosity dimension of stiffness. Also, the loading force used to perform stiffness assessment both manually and mechanically could be standardized, and the frequency of testing specified. Analysis of the force/displacement curves obtained using the stiffness testing device could include generating a mean curve using Cycles 1-5, and including analysis of the unloading as well as the loading curve, as it is not clear whether manual judgements of PA stiffness are based solely on the loading curve.

It appears that for the time being instrumented measurement of PA stiffness will most frequently be used in research studies. It would seem unprofitable to spend time developing a ‘low tech’ version of the current mechanical device suitable for clinical use until some of the investigations suggested above have been conducted and information is available about what parameters (other than elastic stiffness) the device
should be capable of measuring. When the first studies described in this thesis were conducted, there was only one laboratory in Australia, at the University of Sydney, investigating the complex area of spinal stiffness measurement (Latimer et al 1991, Maher and Latimer 1992, Lee and Svensson 1993a, Latimer et al 1993). Since then, researchers in England (Petty 1995), Canada (Kawchuk and Herzhog 1996) and the USA (Simmonds et al 1995) have also begun investigating this area. This increase in research activity worldwide makes it more likely that, with several research teams focusing on the problem of posteroanterior spinal stiffness, its relevance to the management of LBP will become better understood, and patient pain and disability reduced as a result.
REFERENCES


