

**THE EFFICACY OF AN ADAPTED ROLAND-
MORRIS DISABILITY QUESTIONNAIRE IN
MEASURING FUNCTIONAL STATUS OF
PATIENTS WITH LOW BACK PAIN**

Helen Buchanan

Submitted to the Division of Occupational Therapy, University of Cape Town, in
partial fulfilment of the requirements for the degree of MSc (Occupational Therapy)

August 2001

Supervisor: Thea Coetzer

The copyright of this thesis vests in the author. No quotation from it or information derived from it is to be published without full acknowledgement of the source. The thesis is to be used for private study or non-commercial research purposes only.

Published by the University of Cape Town (UCT) in terms of the non-exclusive license granted to UCT by the author.

TABLE OF CONTENTS

	Page
ABSTRACT	i
DECLARATION	ii
ACKNOWLEDGEMENTS	iii
DEFINITION OF TERMS	iv
LIST OF ABBREVIATIONS	vii
INDEX	viii
LIST OF TABLES	xi
LIST OF FIGURES	xi

**THE EFFICACY OF AN
ADAPTED ROLAND-MORRIS
DISABILITY QUESTIONNAIRE
IN MEASURING FUNCTIONAL
STATUS OF PATIENTS WITH
LOW BACK PAIN**

Helen Buchanan

ABSTRACT

The purpose of this study was to determine whether an adapted Roland Morris Disability Questionnaire (ARMDQ) could be used to determine functional status in subjects with chronic mechanical low back pain, and thereby assist the surgeon in deciding on the appropriateness of spinal fusion surgery for individual patients. The study aimed to determine whether the Roland-Morris Disability Questionnaire (RMDQ) improved its ability to measure function and different aspects of function through the addition of a 26-item scale (AddS).

The study design was retrospective and descriptive. A convenience sample of 42 subjects who consulted one Orthopaedic surgeon working in a private practice in the South Peninsula Municipal area of Cape Town, South Africa, was drawn. All subjects were diagnosed with chronic mechanical low back pain and no other known pathology. Subjects consulted the surgeon between August 2000 and July 2001. All subjects had completed the adapted RMDQ (ARMDQ).

A series of analyses were performed in which the items in the AddS scale were compared with those of the RMDQ to assess the degree to which they enhanced the performance of the RMDQ. Although there was a high correlation between the two scales ($R=.72$; $P<.00$), Cronbach's alpha showed the reliability of the RMDQ to be higher than the AddS and the ARMDQ. This finding was supported in a number of subsequent analyses. There was a low correlation between the rating for surgery and total RMDQ scores ($r=.40$; $P<.01$).

Despite deficiencies in the RMDQ, it is recommended as the basis for comparison in future studies within a variety of South African contexts.

KEYWORDS: LOW BACK PAIN / FUNCTIONAL STATUS MEASURES /
ROLAND-MORRIS DISABILITY QUESTIONNAIRE

DECLARATION

I, HELEN BUCHANAN, hereby declare that the work on which this thesis is based is my original work (except where acknowledgements indicate otherwise), and that neither the whole work nor any part of it has been, is being, or is to be submitted for another degree in this or any other university.

I empower the University of Cape Town to reproduce, for the purpose of research, either the whole or any portion of the contents in any manner whatsoever.

Signed by candidate

SIGNATURE

Signature Removed

DATE

10.09.01

ACKNOWLEDGEMENTS

Many people have helped me in a variety of ways during this research project, and I am thankful to all of them for the contribution they have made towards the end product of this process. The following people deserve particular mention:

- My husband, James, without whose help, I would not have been able to complete this dissertation. He has provided ongoing encouragement, support and practical assistance (sometimes until all hours), and I am indebted to him for this.
- My long-suffering children, Murray and Kirsten, for their understanding.
- My supervisor, Thea Coetzer, who has supported me and provided constructive criticism with the draft reports.
- Professor Ruth Watson who encouraged me to embark on this degree, and provided valuable inspiration along the way.
- Professor Herman Kruissje who assisted with the statistical analyses.
- Dr Guillaim du Toit who prompted this study and, through his meticulous record keeping, made it possible to conduct. His secretary, Rita, for going the extra mile with regards to following up on missing data.
- My sister, Gill O'Carroll, for her many hours of expertise spent with formatting.
- My colleagues in the Division of Occupational Therapy, University of Cape Town, for filling the gaps and for their encouragement throughout the process.
- My fellow masters students gave feedback on the protocol and results, and were available to discuss issues as they arose. Particular thanks go to Suzanne Muller for her part in helping me to formulate the study, and Kathleen Brodrick for reviewing early drafts.

Financial assistance provided by the University of Cape Town is gratefully acknowledged.

DEFINITION OF TERMS

Activity limitations: ‘difficulties an individual may have in executing activities’ (World Health Organization, 2001, pp. 12 & 159).

Activities of daily living: ‘includes those tasks that a person does regularly to prepare for, or as an adjunct to, participating in his/her social and work roles’ (Trombly, 1995; p. 289).

Chronic mechanical low back pain: ‘...ache, pain or discomfort in the lumbar area’ (Ozguler *et al*, 2000, p. 215). Intractable low back pain that continues for more than three months and which results in disability. Backache is mechanical if it varies with physical activity (Waddell, 1998).

Construct validity: ‘the extent to which a test measures the construct (mental representation) variables that it was designed to identify’ (Christiansen & Baum, 1997, p.593).

Content validity: the extent to which the instrument has ‘enough items and adequately covers the domain under investigation’ (Streiner & Norman, 1995, p. 20).

Cross-cultural adaptation: a process that looks at both language (translation) and cultural adaptation issues in the process of preparing a questionnaire for use in other settings (Beaton *et al*, 2000).

Disability: refers to limitations in an individual’s ability to perform activities as well as any restrictions influencing their ability to participate in a life situation (World Health Organization, 2001)

Domain/dimension: ‘...the area of behaviour or experience that we are trying to measure. Domains might include mobility and self-care (which could be further aggregated into physical function) or depression, anxiety, and well-being (which could be aggregated to form an emotional function domain)’(Guyatt, Feeny & Patrick, 1993, p. 623).

Function: ‘an individual’s performance of activities, tasks and roles during daily occupations (occupational performance)’ (American Occupational Therapy Association, 1995a, p. 1016).

Functioning: the person's ability to perform activities and participate in a life situation (World Health Organization, 2001).

Functional limitations: 'restrictions in performance at the level of the individual' (Delitto, 1994, p. 453).

Functional status measures: these assess the ability of an individual to perform particular defined tasks (Wilson & Cleary, 1995).

Health: 'the extent to which an individual or group is able, on the one hand, to realize aspirations and satisfy needs, and, on the other hand, to change or cope with the environment. Health is therefore seen as a resource for everyday life ... a positive concept emphasizing social and personal resources as well as physical capacities' (World Health Organization, 1986, p. 74).

Health-related quality of life / health status: generally refers to 'the combination of measurements of physical, psychological, and social functioning aspects of health' (Boden, 1998, p. 717).

Instrument: 'the actual tool or questionnaire that assesses the health status of the patient' (Boden, 1998; p. 718).

Internal consistency: examines the extent to which the items in an instrument are related to each other (Ottenbacher & Christiansen, 1997; Streiner & Norman, 1995).

Occupation: For the purposes of this study, occupation is used interchangeably with 'work' (refer to the definition for work).

Occupational performance: 'the physical and mental abilities and skills required for satisfactory engagement in a given occupational pursuit'. This is the performance dimension of human occupation. (American Occupational Therapy Association, 1995a, p. 1015). The word 'function' is used interchangeably with 'occupational performance' (American Occupational Therapy Association, 1995b).

Quality of life: a 'concept defined by an individual's perceptions of overall satisfaction with his or her living circumstances, including physical status and abilities, psychological well-being, social interactions, and economic conditions' (Christiansen & Baum, 1997, p.602).

Reliability: ‘Reliability is concerned with the degree to which [measurement scores] can be replicated’ (McDowell & Newell, 1987, p. 31).

Validity: the extent to which an instrument measures what it intends to measure (Streiner & Norman, 1995).

Well-being: ‘... a general term encompassing the total universe of human life domains, including physical, mental and social aspects, that make up the total universe of human life’ (World Health Organization, 2001, p. 157)

Work: ‘any productive activity, whether paid or unpaid, that contributes to the maintenance or advancement of society as well as to the individual’s own survival or development’ (Creek, 1997, p. 34).

LIST OF ABBREVIATIONS

AddS	Added 26-item scale
ADL	Activities of daily living
ARMDQ	Adapted Roland-Morris Disability Questionnaire
BMI	Body mass index
CATPCA	Categorical Principal Component Analysis
CMLBP	Chronic mechanical low back pain
CPN	Continuous back pain
ICIDH-2	International Classification of Functioning, Disability and Health
LBP	Low back pain
O.T.	Occupational Therapy
OCC	Occupation
ODI	Oswestry Disability Index
RMDQ	Roland-Morris Disability Questionnaire
S.A.	South Africa
SF-36	Short-form 36-item Health Survey
SIP	Sickness Impact Profile
TADD	Totals of selected items for the AddS
TORG	Totals of selected items for the RMDQ
U.K.	United Kingdom
U.S.	United States

INDEX

	Page
Chapter 1: Introduction	1
1.1. Prevalence and Costs of Low Back Pain	1
1.2. Treatment for Low Back Pain	4
1.3. Assessment of Low Back Pain	5
1.4. The Study Problem	6
1.5. Purpose	7
1.6. Aims	7
1.7. Objectives	8
Chapter 2: Low Back Pain	9
2.1. Low Back Pain	9
2.2. Measurement of Health Status	10
2.3. The Measurement of Outcomes	11
2.4. Measuring Outcome in Low Back Pain	12
2.5. The Relationship Between Pain, Disability and Function	13
2.6. Measuring Functional Status	13
2.7. Functional Status Measures for Low Back Pain	16
2.8. Generic Measures of Health Status	16
2.9. Back-specific Functional Status Measures	17
2.10. Cross-cultural Adaptation of Self-report Measures	20
2.11. Evaluation of Individual Health Outcome Measures	20
2.11.1. Validity	21
2.11.2. Reliability	22
2.11.3. Sensitivity	22
2.11.4. Practicality	23
Chapter 3: Research Method	24
3.1. Research Design	24
3.2. Study Population and Sample	24
3.3. Instruments	24
3.4. Data Collection	25
3.5. Pilot Studies	27
3.6. Ethical Considerations	27
3.7. Analysis	28

Chapter 4: Results	30
4.1. Sample Description	30
4.1.1. Age and Gender	30
4.1.2. Occupation	30
4.1.3. Body Mass Index (BMI)	32
4.1.4. Continuous Pain	32
4.1.5. Duration of Continuous Pain	32
4.1.6. Rating for Surgery	32
4.1.7. Previous Surgery	32
4.2. Analysis of the Questionnaires	33
4.2.1. Subjects and Items Excluded From Analysis	34
4.2.2. Correlation Between the Item Totals of the RMDQ and the Added Scale (AddS)	36
4.2.3. Comparison of the reliability of the RMDQ and the AddS	36
4.2.4. Item Analysis of the RMDQ and the AddS	37
4.2.5. Correlations Between the RMDQ and the AddS After Selection of Items	41
4.3. Correlations Between the Demographic Variables	42
4.4. Analysis of the Interrelationships Between the Selected Items Measuring Functional Status and Demographic Variables	45
4.5. Analysis of the Relationship Between the Rating for Surgery and Scores on the RMDQ and AddS	49
Chapter 5: Discussion	50
5.1. Sample Description	50
5.2. Analysis of the Questionnaires	52
5.2.1. Subjects and Items Excluded From the Analysis	53
5.2.2. Comparisons between the RMDQ and the AddS	55
5.2.3. Correlations with Demographic Variables	58
5.3. Analysis of the Relationship Between the Rating for Surgery and Scores on the RMDQ and AddS	59
5.4. Problems with the RMDQ	59
5.5. Limitations of the Study	60
Chapter 6: Conclusions and Recommendations	61
6.1. Conclusion	61
6.2. Recommendations	62
References	63

Appendices	74
Appendix 1: The Roland-Morris Disability Questionnaire (RMDQ)	74
Appendix 2: Items Added To The Roland-Morris Disability Questionnaire (Adds – Original Version)	75
Appendix 3: Adapted Roland-Morris Disability Questionnaire Revised Version ARMDQ)	76
Tables	
Table 1: Socio-demographic and Medical Data for Subjects (N=42)	31
Table 2: Items Excluded from the Analysis of the RMDQ	35
Table 3: Items Excluded from the Analysis of the Added Scale (AddS)	35
Table 4: Pearson’s Product-Moment Correlations and Levels of Significance for the RMDQ, AddS and ARMDQ (N=36)	36
Table 5: Component Loadings for the AddS	39
Table 6: Correlations Between Selected Item Totals and Object Scores of the RMDQ and AddS (N=42)	41
Table 7: The Correlations Between the Various Demographic Variables.....	43
Table 8: Correlations With New Scales (Selected Items) and Demographic Variables	44
Table 9: Relationship Between Rating for Surgery and Selected Item Scales	46
Table 10: Relationship Between Rating for Surgery and Selected Item Scales	49
Figures	
Figure 1: Model Depicting the Comparisons in the Analysis	34
Figure 2: Component Loadings for RMDQ Items After Removal of Items not Contributing to Measurement of Functional Status	38
Figure 3: Component Loadings for AddS Items After Removal of Items not Contributing to Measurement of Functional Status	39
Figure 4: Component Loadings of the ARMDQ (Selected Items)	40
Figure 5: Component Loadings for Demographic Variables	44
Figure 6: Relationship Between Demographic Variables and Their Categories	45
Figure 7: Relationship Between Demographic Variables and the Selected Item Scales	47
Figure 8: Component Loadings for Demographic Variables and Items on the RMDQ	47
Figure 9: Relationship of Demographic Variables (and their Categories) to Items 5 and 13 in the RMDQ	48

CHAPTER 1: INTRODUCTION

Disorders causing lower back pain (LBP) are extremely prevalent in all societies (Frymoyer & Cats-Baril, 1991). Impairments of the back and spine have been the most frequently reported cause of activity limitations among chronic conditions in people less than 45 years of age in the United States (U.S.) (Kelsey & White, 1980) as well as in Canada (Kopec & Esdaile, 1998). The impact of back pain has been noted to result in considerable disability and economic loss especially among the working population (Kelsey & White, 1980; Kelsey, Mundt & Goldin, 1992).

1.1. Prevalence and Costs of Low Back Pain

A U.S. survey conducted in a family practice established that 46.3% of all patients aged 18 to 55 years had experienced moderate LBP, and 23.6% had had severe LBP over a three-year period. Only 30.1% had never had LBP (Frymoyer *et al*, 1983). Furthermore, the patients with severe LBP had sought more treatment and lost more work time due to their back pain than other patients seen at the practice (*ibid*). Van Tulder, Koes and Bouter (1995) reported back pain to be one of the most frequent reasons for visits to general practitioners or physiotherapists, and a major cause of work absenteeism and disablement in the Netherlands. On the other hand, a United Kingdom (U.K.) survey by Hillman *et al* (1996) found that while 13.7% of the population between the ages of 25 and 64 years experienced LBP, 20% of these had never consulted anyone about their problem. Of those who had received treatment, 13.7% were treated at primary health care level and 4% at secondary care level. Only 3% had been treated by a complementary therapist such as a physiotherapist (*ibid*). According to Waddell and Hamblen (1983), backache is a common problem as illustrated by 20% to 30% of all new out-patient orthopaedic clinic patients in the U.K. presenting with backache.

Lumbar spine disorders have been reported to affect up to 80% of people in Western countries at some stage of their lives (Kelsey & White, 1980). Hillman *et al* (1996) found the lifetime prevalence of LBP in a U.K. community to be 59% with 6.4% of sufferers taking time off work as a result thereof. Twenty-six percent of this study sample was diagnosed with chronic LBP over a one-year period. In comparison,

Frymoyer *et al* (1983) found a lifetime prevalence of 69.9% amongst men aged 18 to 55 years who were seen in a U.S. family practice over a three-year period. A study conducted on a probability sample of the general population of the Netherlands, reported a 12-month period prevalence of low back problems in the working population to be 44.4% for men, and 48.2% for women, this being even higher among non-workers (Picavet, Schouten & Smit, 1999). Further, more than a third of people with disabilities were disabled as a result of back problems. In France, a study among workers from four occupational sectors by Ozguler *et al* (2000), found that LBP varied in prevalence from 8% to 45% according to the way in which LBP was defined. Their definition for LBP ranged from 'pain for at least one day' to 'sick leave for LBP'. The figures for Southern Africa are relatively unknown. One study of mothers living in a Lesotho district identified that 58.5% of the sample had LBP with 10.1% having severe LBP. Of the mothers with severe LBP, 78.8% were from rural communities and were poor and illiterate (Worku, 2000). There is a paucity of literature concerning the magnitude of the problem of disability related to low back disorders in South Africa (S.A.) (Odendaal, 1999), and no prevalence figures for this country could be located.

Although the prevalence of back pain has not increased substantially over the past 20 years, the rate of disability has. Reasons for this are unclear but increased awareness of the problem has resulted in escalation in the costs of medical care. According to Odendaal (1999), there is virtually no data available regarding the economic impact of LBP in S.A.. Limited figures are available from insurance disability claims, but this does not fully represent the magnitude of the costs associated with back problems to the country. A study of work-related lumbar spine injuries amongst a group of mineworkers in S.A. suggested that the costs of care were high and that they increased in proportion to the severity of the pathology and the intervention required (Odendaal, 1999). In the U.S., costs were estimated to be more than \$50 billion per year with 75% being attributed to 5% of people who were either temporarily or permanently disabled as a result of back pain (Frymoyer & Cats-Baril, 1991). Spengler *et al* (1986) found that compensation claims related to back injuries in a large industrial company in the U.S. accounted for 19% of all workers' compensation claims, and were responsible for 41% of the total injury costs. They further found that

10% of all back injury claims accounted for 79% of the total back injury costs. A Netherlands study estimated the costs of back pain to be 1.7% of the gross national product in 1991 (Van Tulder, Koes & Bouter, 1995). These figures illustrate the significant economic impact of back injuries on industry (Spengler *et al*, 1986) lending credence to LBP as a health problem of considerable importance (Van Tulder, Koes & Bouter, 1995). They further provide a rationale for affording LBP a high priority in resource allocation for preventive and curative programmes, and research aimed at finding more effective ways of preventing and treating it (Kelsey & White, 1980).

The results of these studies cannot, however, be extrapolated to other societies due to differences in culture and health care systems (Van Tulder, Koes & Bouter, 1995). The picture in S.A. is likely to be similar, if not more exaggerated, due to the high number of unskilled workers employed in jobs as manual labourers. South Africa has a diverse range of people with marked differences in socio-economic status and education levels and great disparities in the health services available. Many people are from disadvantaged backgrounds where poverty is rife. Because many workers are unskilled, their only option is manual labour, an occupation that places workers at risk for back injuries. The prevalence of low back pain could well be higher in South Africa than in more developed countries due to a lack of worker education, prevention programmes and enforceable legislation related to lifting. However, there are no figures to illustrate this. Low back pain in someone who is already disadvantaged leads to further poverty, particularly if the person is unable to work due to back pain. The resulting consequences may be vast in someone who is the only or the primary breadwinner in a family.

Back pain is a chronic and recurring condition for which there is seldom a cure. The main goal of treatment is therefore to improve quality of life by reducing pain and restoring function (Deyo, 1988; Bouter, Van Tulder & Koes, 1998; Deyo *et al*, 1994). With each episode of LBP, the patient becomes increasingly physically and functionally compromised (Cook & Hassenkamp, 2000). It is this group of patients who present the greatest challenge to health care providers and, in lieu of the rising

socio-economic costs, are the greatest cause for societal concern (*ibid*; Frymoyer & Cats-Baril, 1991).

1.2. Treatment for Low Back Pain

Low back pain can present in a variety of ways ranging from acute, mild backache to chronic disabling low back pain. Accordingly, treatment regimes may vary from conservative approaches such as bed rest and physiotherapy, to more invasive approaches such as surgery (Cicinelli, 1997). While most patients with chronic mechanical low back pain (CMLBP) benefit from rehabilitation, a small percentage may require spinal fusion surgery. Odendaal (1999) found that between one and two percent of a group of South African mineworkers with low back problems were treated surgically. The lack of consensus amongst surgeons regarding the choice and utilisation of surgery such as spinal fusion (*ibid*) means that decisions about when it should be performed are not clear-cut. Furthermore, evaluation of the outcomes of spinal fusion have been based on more traditional outcome measures such as radiographic union, return to work or previous activity level, and subjective assessments of pain (Boden, 1998; Bombardier, 2000a). Such measures have disregarded the importance of the patient's perspective concerning the impact of their back problem on everyday function.

Clinical studies have demonstrated that in at least 50% of people who have had extensive investigations for LBP, the exact cause of the problem is unclear (Frymoyer *et al*, 1983). Deyo and Diehl (1983) state that objective physical findings for LBP are often absent and progress can only be assessed in terms of pain resolution and improved function. In the absence of physical findings a diagnosis cannot be reached and clinicians, being reliant on the patient's perceptions of their limitations in everyday activities, may thus document disability from the patient's perspective (Delitto, 1994). Furthermore, while most patients may express satisfaction with the procedure to the treating surgeon, a more reliable method of determining treatment needs and evaluating its effectiveness would enable clinicians and researchers to compare their results more easily. This formed the foundation for the development of outcomes research, which seeks to evaluate treatment effectiveness so that the most

effective forms of treatment may be identified (Andersson & Weinstein, 1994; Cicinelli, 1997). Medical outcomes need to reflect 'a more complete array of measurements, including health status, patient satisfaction, medical costs and quality of life' (Boden, 1998, p. 717).

1.3. Assessment of Low Back Pain

Assessment of health status involves a combination of instruments including those measuring the domains of physical, psychological and social functioning. As a consequence of the practical difficulty of using such a wide array of measurements to determine treatment needs or evaluate treatment effectiveness, Boden (1998) suggested that particular health providers might decide to focus on only one of these areas. Measurement of the effects of treatment for LBP has included, amongst others, functional status questionnaires (Deyo, 1988). These measure the impact of LBP on the person's ability to perform everyday activities. Occupational therapy (O.T.) for people with LBP aims at enabling successful participation in activities that are meaningful to the individual and promote a sense of well-being (Cicinelli, 1997). Occupation is central to the practice of O.T., therefore anything disrupting or preventing an individual from performing their daily occupations is of concern to the therapist. Pain and decreased range may have a marked effect on occupational performance, and may have far-reaching psychological and social consequences. Assessment of the functional status of a person with LBP is thus an area in which occupational therapists should be involved. However, while functional status measures are important to document the effects of treatment, Edwards (1997) recommended that quality of life measures might be better tools for demonstrating O.T. outcomes considering that the profession is based on the assumption that quality of life is achieved through occupational performance or function. Research into functional outcome measures for patients with LBP is a strong focus in present journal publications. Although the use of such measures is currently very limited in South Africa, if we are to evaluate the effectiveness of medical treatment from a health cost-benefit perspective, suitable assessment methods need to be used to determine the most suitable form of treatment and to study treatment outcomes for LBP.

For a functional status measure to be regarded as a worthwhile outcomes measure suitable for use in various health contexts in South Africa, its ability to discriminate between patients with varying degrees of disability (or functional limitations) must be established. To ensure the universal applicability of an instrument, it needs to be standardised. This involves, amongst other things, investigation of its measurement properties, and performance data from a normative population (Keith, 1984). It was expected that through this preliminary study, an improved functional status instrument suitable for use in South Africa might emerge and that it could be tested further in subsequent studies in a variety of contexts. Because only a small percentage of the South African population can afford private health care, it was anticipated that the emerging instrument could ultimately be used to benefit a wider patient population, i.e. those treated for LBP within the public service health sector. To date, the researcher has not located any South African studies in which such an instrument has been tested. The ultimate goal is to develop a functional status measure suitable for use in a variety of South African health service contexts that could also be used as an evaluative measure to determine the outcomes of spinal fusion surgery for LBP.

1.4. The Study Problem

This study originated from the concern of an orthopaedic surgeon in private practice in Cape Town to include the measurement of functional status, both pre- and post-surgery, to capture from the patient's perspective, the impact of their back problem on daily function. The Roland-Morris Disability Questionnaire (RMDQ) (Roland & Morris, 1983a; Roland & Morris, 1983b; Roland & Fairbank, 2000) was introduced as a suitable instrument for this purpose (see Appendix 1). The RMDQ has been extensively validated (Beurskens *et al*, 1995; Beurskens, De Vet & Koke, 1996; Deyo, 1986; Deyo *et al*, 1998; Jensen *et al*, 1992; Leclaire *et al*, 1997; Kopec *et al*, 1995) and is recommended as one of two particularly noteworthy functional outcome measures specific to LBP. It was felt that this measure, together with the clinical examination and the surgeon's expertise, would enable a more effective decision to be made regarding surgery. Furthermore, if surgery was performed, the RMDQ could

be used as a measure of post-operative patient function and would be an indication of the outcome of the procedure.

To broaden the scope of the questionnaire, the surgeon added a further 26 item scale (AddS) to the 24-item RMDQ. In so doing, the properties of the instrument became questionable. Furthermore, according to Underwood, Barnett & Vickers (1999), questionnaires should be validated in the population for which they are intended. Thus, it became necessary to determine whether this adapted instrument (the ARMDQ) was an improvement over the RMDQ in assessing functional status in patients with CMLBP in this particular setting.

The study questions were thus:

- Is the addition to the RMDQ scale an improvement over the original?
- Are the RMDQ and the AddS able to discriminate between patients with severe disability who require surgery and those with minimal to moderate disability who do not require it?

1.5. Purpose

To determine whether the ARMDQ is a functional status measure that could be used to complement the surgeon's clinical examination and to assist the decision-making process regarding the appropriateness of spinal fusion surgery for individual patients.

1.6. Aims

- To establish whether the ARMDQ is an improvement over the RMDQ as a functional status measure for people with CMLBP.
- To determine whether the ARMDQ is able to discriminate between subjects who require spinal fusion surgery and those who do not.

1.7. Objectives

- To describe the demographic profile of the study sample according to age, gender, height, weight (body mass index), duration of continuous pain and type of work (occupation).
- To determine whether there are any associations between the demographic variables as well as their categories.
- To determine whether there is an association between the demographic variables and items in the RMDQ and AddS scales.
- To determine whether there is a correlation between the RMDQ, the AddS and the ARMDQ, i.e. whether all items in the scales relate to the same concept (content validity).
- To establish whether the item reliability in the AddS is at least the same (and preferably higher than) the RMDQ.
- To identify the items in the RMDQ and AddS which make the highest contribution to the measurement of functional status.
- To determine the relationship between responses on the RMDQ and the AddS scales, and the surgery rating of the subjects. It was anticipated that the higher the score obtained on the RMDQ and AddS scales, the more likely it would be that the subject required surgery.

CHAPTER 2: REVIEW OF LITERATURE

2.1. Low Back Pain

According to Waddell and Hamblen (1983), most backache occurs as a result of minor trauma or degenerative changes and is mechanical in nature. Mechanical back pain is related to activity. While some physical activities may make the pain worse, others may relieve it. Pain varies over time and in severity, and is usually classified as chronic when it continues for more than three months (Waddell, 1998; Nordin, 1992). At this stage, the patient's life may be seriously disrupted psychologically, socially and financially (Nordin, 1992). Low back pain may thus have a profound impact on all aspects of a person's life (Leclaire *et al*, 1997).

In most patients with non-specific LBP, no meaningful diagnosis is reached (Delitto, 1994; Waddell, 1992). Doctors are frequently unable to identify the cause of the pain and treatment has a high failure rate. Waddell (1992) found that while up to 15% of patients who have low back surgery will have repeat surgery, the success of subsequent operations has been found to deteriorate progressively. In a study that measured outcome after lumbar disc surgery, Hutchinson *et al* (2000) confirmed that patients who had undergone two operations reported significantly greater disability than those who had only had one operation. In this study, outcome was determined using a generic health status measure and a disease-specific functional status questionnaire. There is now worldwide agreement that the best way to assess low back disability is by basing the assessment on activities of daily living which gives a direct measure of basic activities (Waddell, 1998). Thus, functional status is an important measure of outcome both clinically and in research (Beurskens, De Vet & Koke, 1996; Deyo, 1986).

LBP occurs frequently during adulthood with first episodes commonly occurring between the ages of 20 and 40 years. Being over 40 or 50 years has been shown to be a risk factor for chronic LBP with decreasing occurrence after 60 years of age (Ozguler *et al*, 2000). There appears to be inconsistent evidence concerning gender differences in the frequency of LBP. Some state that women have a higher risk for developing LBP while others report this in men (Hillman *et al*, 1996; Kelsey, Mundt

& Goldin, 1992; Ozguler *et al*, 2000; Waddell, 1998). People in lower socio-economic classes seem to be more frequently affected (Kelsey, Mundt & Goldin, 1992; Waddell, 1998). Many studies show that height and weight do not increase the risk of back pain (Waddell, 1998). Ozguler *et al* (2000) reported that a body mass index (BMI) above 22.6kg/m² was a risk factor in people with back pain who consulted a health professional, had treatment for their back pain, or took sick leave. A number of occupational risk factors have been described in the literature (Kelsey, Mundt & Goldin, 1992) with people in heavy manual jobs reporting more back injuries than those in other jobs (Waddell, 1998). Carrying heavy loads and bending were the most emergent risk factors related to work in people with severe back pain (Ozguler *et al*, 2000). There is strong evidence to support the importance of psychosocial factors over physical factors in the development of chronic LBP and disability (Waddell, 1998; Ozguler *et al*, 2000). Psychological disturbances have been shown to develop secondarily to the physical disorder with depression being a common problem in people with chronic LBP (Waddell, 1992). It is therefore important to look beyond the physical symptoms so that treatment is directed not only at pain relief but also towards restoration of function (*ibid*).

2.2. Measurement of Health Status

Measurement of health status is important in evaluating the outcomes of care or the causes and consequences of differences in health (McDowell & Newell, 1987). It further assists in evaluating the quality of medical care and in planning for the health needs of the population (Ware *et al*, 1981). Benefits to individual patients may include improved decision-making by the surgeon and better timing of interventions. However, measurement of health status has been a much-debated topic and there are no direct measures or standard scales to measure health (McDowell & Newell, 1987). Therefore, to measure health as a total concept, a number of health indicators – each representing a component of the whole – are required (*ibid*). To this end, Ware (1987) recommended the use of instruments to measure the five different dimensions of health. These dimensions are physical and mental health, everyday functioning in social and role activities, and general perceptions of health and well-being. Physical health includes the ability to perform activities of daily living (Boden, 1998) and is

commonly measured in terms of performance limitations in self-care activities, mobility, or more strenuous activities, e.g. participation in sports (Ware, 1987).

2.3. The Measurement of Outcomes

The assessment of treatment outcomes in medical care has become increasingly important due to the demand for evidence to demonstrate the end results of treatment (Deyo *et al*, 1994). Outcome measures are currently the most important tools for clinicians, patients and policy makers to determine the effectiveness of treatment (Andersson & Weinstein, 1994; Boden, 1998; Bombardier, 2000b).

Traditionally, 'hard' measures of outcome have been preferred over 'soft' data in clinical evaluations (Deyo, 1988). However, the boundary between 'hard' and 'soft' data is not always clear and the critical aspect is the reliability of a finding (Deyo, 1988; Deyo *et al*, 1994). Furthermore, many hard measures have no value either to the patient or society (Deyo, 1988). Hard assessment methods provide a skewed impression of the outcome as decisions are based purely on the surgeon's judgement. Outcome is influenced by multiple factors and should therefore be measured comprehensively by including a wider variety of measures such as health status, patient satisfaction, medical costs and quality of life (Boden, 1998; Bombardier, 2000; Deyo *et al*, 1994). The emphasis has thus shifted towards evaluating treatment according to its impact on life satisfaction in social, psychological and physical terms (Bowling, 1991). This is in keeping with the World Health Organization broader definition of health (1986) in which the focus moved away from a disease and illness model towards more subjective measures of health status.

The importance of considering the patient's opinion in decision-making in medical care has been increasingly recognised by clinicians and is now an accepted method of measuring outcomes. (Bowling, 1991; Deyo *et al*, 1994). Such subjective health measures may assess general feelings of well-being, symptoms of illness, or functional ability (Bowling, 1991; McDowell & Newell, 1987). Self-administered (or self-report) questionnaires are subjective measures and are widely used due to their simplicity and low cost (McDowell & Newell, 1987; Ware, 1987). They have been

found to be better at assessing disability (Fairbank & Pynsent, 2000) and have proved to be at least as reproducible, if not more precise than, so-called objective measures (Deyo *et al*, 1994; Ware, 1987). Subjective measures enable practitioners to gain insight into the patient's experiences of treatment thus ensuring that the true outcomes of an intervention become known (Bowling, 1991). The first self-report measure for people with LBP was the Oswestry Disability Questionnaire (Fairbank *et al*, 1980) that was published 20 years ago. Since that time, many other back specific self-report functional status measures have been developed. These will be discussed in more detail later in this section.

2.4. Measuring Outcome in Low Back Pain

Deyo *et al* (1994) indicated that patient outcome in LBP has multiple dimensions. Bombardier (2000a) proposed the use of Wilson and Cleary's (1995) conceptual model of outcomes to demonstrate how different outcomes interrelate. They identified five outcomes namely biological and physiological variables, symptom status, functional status, general health perceptions and quality of life. Each outcome is distinct from the rest but with the possibility of some correlations between them. While all aspects must be considered, one might decide to measure only one dimension of outcome due to the time pressures that make incorporating a number of measures in one study difficult (Deyo *et al*, 1994).

Deyo *et al* (1998, p. 2004) proposed a standardised set of outcome measures to provide 'a common yardstick that is appropriate for use in many types of studies'. Their suggested standard set of measures consisted of instruments to measure symptoms, daily functioning, well-being, work disability and satisfaction. More recently, a panel of experts refined this proposal suggesting that a core set of measures should include five domains, namely, back-specific function, generic health status, pain, work disability, and patient satisfaction (Bombardier, 2000b). Because there is no 'ideal' core set of measures, different instruments may be more or less appropriate depending on the study context and the population involved (Bombardier, 2000b). Findings of poor associations between physiologic measures strengthen the argument that outcomes that are more relevant to patients need to be incorporated

into assessment routines (Deyo, 1988). If the dimension of interest is function, this must be measured directly rather than inferred from more traditional 'objective' methods of assessment (Deyo, 1988; Deyo *et al*, 1998).

2.5. The Relationship Between Pain, Disability and Function

Although pain and disability are related, they are distinct from one another. Pain is a symptom while disability refers to restrictions in function. Each must therefore be measured separately, both relying on input from the patient usually through self-report measures. Self-reports may provide information about the areas of impairment (both physical and psychosocial) and functional limitations, thus assisting in quantifying disabilities (Delitto, 1994). Back pain does not necessarily lead to disability, and the degree of disability may not be proportional to the severity of the pain (Waddell, 1998). Chronic LBP is furthermore not a static phenomenon, and intermittent increases in pain can markedly affect the individual's ability to function in both work and personal spheres of life (McGorry *et al*, 2000). A qualitative study by Cook and Hassenkamp (2000) identified the impact of chronic LBP, finding that the study participants had become increasingly socially isolated because they were unable to partake in their previous leisure activities or go to work. They reported that their quality of life had been compromised mainly as a result of stopping the activities that were important to them.

2.6. Measuring Functional Status

Many indices of physical health build their operational definitions on the concept of functioning. This approach views someone as healthy if they are physically and mentally able to do the things they wish and need to do. Measuring functional ability is a convenient way of comparing the impact of different diseases on different populations and at different times. It is also a common method of assessing the outcome of an intervention (McDowell & Newell, 1987). Functional status measures should be able to discriminate between patients with different severities of back problems, predict prognosis, and evaluate change over time (Kirshner & Guyatt, 1985; Deyo, 1988; Guyatt, Feeny & Patrick, 1993). The construction of instruments

is different for each of these purposes (Guyatt, Feeny & Patrick, 1993; Kirshner & Guyatt, 1985).

Functional status measures had their origin in the field of physical rehabilitation, which emphasised the description and measurement of disability or functional limitations in task performance at the person level (Gresham & Dittmar, 1997). This is in line with the current understanding of 'disability' which has been conceptualised in the International Classification of Functioning, Disability and Health (ICIDH-2) as 'activity limitations' and 'participation restrictions' (World Health Organisation, 2001).

However, the term 'functional status' is confusing in the literature in that it has included measures of spine mobility, muscle strength, employment status and other indicators (Deyo, 1988). Functional status is a measure of the effects of the disease rather than the disease itself. It measures the person's ability to perform everyday activities and is thus distinct from general health status (Beurskens *et al*, 1995; McDowell & Newell, 1987). The terms 'health status' and 'functional status', which are intended to describe aspects of a patient's health, are frequently substituted by 'quality of life' (Gill & Feinstein, 1994). While 'health status' and 'health-related quality-of-life' can refer to different concepts, Wilson and Cleary (1995) argued that they could be used interchangeably. Health-related quality of life encompasses more than adequate ability to function, and includes perceptions of well-being, a basic level of satisfaction and a general sense of self-worth (McDowell & Newell, 1987). McKenna (1993, p. 33) concurred with this by stating that function cannot be equated with quality of life since people 'want more from life than merely to function'. According to Deyo *et al* (1994) and Bowling (1991), functional status is only one component of health-related quality-of-life. Through their interventions, occupational therapists enable people to regain health as well as function (Ottenbacher & Christiansen, 1997).

Although other areas of general health may be incorporated into functional status measures, the primary focus is on the person's level of function and degree of independence in performing various activities, particularly activities of daily living

(ADL) (Stanton, Gresham & Dittmar, 1997). The term ADL has been understood in different ways in the fields of physical rehabilitation and occupational therapy. Occupational therapists thus need to be aware of this when communicating with other professionals or in selecting instruments (Rogers & Holm, 1998). Many functional status instruments measure both basic activities of daily living and instrumental activities of daily living (Gresham & Dittmar, 1997). Basic ADL are universal and necessary to maintain health, however, the way in which they are performed and the relative importance attached to each activity differs culturally. Basic ADL are commonly agreed to be central in determining functional status and include self-care tasks and mobility (Beurskens *et al*, 1995; Gresham & Dittmar, 1997; Keith, 1984; Ottenbacher & Christiansen, 1997; Rogers & Holm, 1998). Instrumental ADL refers to tasks beyond caring for oneself that involve interaction with the physical and social environment, including home management (Ottenbacher & Christiansen, 1997; Trombly, 1995). Keith (1984) argued that while the inclusion of only self-care and mobility tasks provided a narrow focus of function, there was a lack of consensus about what categories of function should be included in a functional status measure. Functional assessment instruments are frequently evaluated using five domains that are relevant to occupational performance assessment. These domains include mobility, communication, self-care, occupation and social relations (Ottenbacher & Christiansen, 1997). According to Wilson & Cleary (1995), the four dimensions of functioning that are commonly measured in a functional status measure are physical function, social function, role function, and psychological function. Although these are not the only domains that may be of interest to a patient or clinician, there is general consensus that they are the minimum required to provide a holistic picture of the individual's abilities. Functional status measures have been criticised for their lack of focus on participation restrictions and environmental barriers both of which may prevent successful rehabilitation outcomes (Stanton, Gresham & Dittmar, 1997).

A wide range of instruments is available to health professionals to measure function. These may either be self-report measures or observations of actual performance. Direct tests of function, such as grip strength assessments, may not indicate ability accurately and subjective feelings or reduced ability to perform ADL may be more important to the patient (Stanton, Gresham & Dittmar, 1997). Thus, many functional

status measures use self-report methods in which respondents report limitations in their own activities (Bowling, 1991). It is important to note, however, that both self-reports and clinician-measured performance tests may be influenced by psychosocial factors (Lee *et al*, 2001). There may thus be a discrepancy between patient's self-reports or clinician's assessments of activity limitations and the individual's actual function.

2.7. Functional Status Measures for Low Back Pain

Since 1980, there has been a proliferation of functional status measures suitable for patients with LBP (Deyo, 1988). Several generic measures of health status and an even greater number of disease-specific measures have been developed. Disease-specific and generic functional status questionnaires have differing advantages and disadvantages. Generic measures may identify complications or side-effects in areas of function or organ systems that are not spine related. They further enable comparisons between conditions and populations and are thus of greatest interest to policy makers (Guyatt, Feeny & Patrick, 1993). Disease-specific functional status measures, on the other hand, focus on back-related problems and are therefore relevant to patients and clinicians interested in determining how the back problem impacts on function (Deyo *et al*, 1994; Guyatt, Feeny & Patrick, 1993). These measures may be more responsive to change in functional status due to back pain than generic instruments. It is thus suggested that both types of measures be included for research purposes (Deyo *et al*, 1994; Boden, 1998).

2.8. Generic Measures of Health Status

Generic health measures are used to measure health status regardless of underlying disease. They broadly assess health, disability, and quality of life (Kopec, 2000; Lurie, 2000). The two major classes of generic instruments are health profiles and utility measures. Health profiles measure health across a number of specific areas such as physical functioning, mental health and role limitations (Garratt, Moffett & Farrin, 2001). They contain items assessing functional limitations related to health in general (Boden, 1998) and therefore provide a more comprehensive picture of health status than back-specific measures (Bombardier, 2000b).

Although these instruments enable the effect of different conditions or treatments to be compared, they may lack the ability to measure changes in certain diseases (Boden, 1998; Lurie, 2000). Numerous generic health status measures have been developed, many of which have been evaluated in patients with LBP. These include the Duke Health Profile (Parkerson, Broadhead & Tse, 1990), the Nottingham Health Profile (Hunt, McEwen & McKenna, 1985), the Short-form 36-item Health Survey (SF-36) (Ware & Sherbourne, 1992), and the Sickness Impact Profile (SIP) (Bergner *et al*, 1976). Deyo *et al* (1998) recommended that the SF-36 or the EuroQol (EuroQoL Group, 1990) be included as measures of well-being in a core set of outcome measures for LBP. According to Lurie (2000), the SF-36 has several advantages over other generic measures, including greater responsiveness, high internal consistency, and good discriminant validity.

2.9. Back-specific Functional Status Measures

Disease-specific measures contain items that identify limitations caused by a specific disease, population, function, or problem (Guyatt, Feeny & Patrick, 1993). They are generally more sensitive to subtle changes in outcome, particularly those that may not affect general health (Boden, 1998). A large number of scales have been used to measure functional disability in patients with back pain (Kopec & Esdaile, 1995). The most widely accepted of these are the Million (Million *et al*, 1982), Oswestry (Fairbank *et al*, 1980), Roland-Morris Disability Questionnaire (Roland & Morris, 1983a), and Waddell Scales (Waddell & Main, 1984). Criticisms of these instruments have been that the methods and criteria of item development and selection have not been fully explained (Kopec & Esdaile, 1995).

More recently new scales have been published including the Aberdeen Back Pain Scale (Ruta *et al*, 1994), Back Pain Functional Scale (Stratford, Binkley & Riddle, 2000), Curtin Back Screening Questionnaire (Harper *et al*, 1995), Dallas Pain Questionnaire (Lawlis *et al*, 1989), Functional Rating Index (Feise & Menke, 2001) and the Quebec Back Pain Disability Scale (Kopec *et al*, 1995; Kopec *et al*, 1996). Kopec and Esdaile (1995) and Kopec (2000) have described additional scales suitable for LBP research. Despite this proliferation in the development of new instruments,

there has been little formal comparison of these measures (Stratford, Binkley & Riddle, 2000).

The two most commonly cited back-specific measures are the RMDQ (Roland & Morris, 1983a) and the Oswestry Disability Index (ODI) (Fairbank *et al*, 1980; Fairbank & Pynsent, 2000). Their widespread use and strong evidence of validity led Deyo *et al* (1998) to recommend their inclusion in a core set of outcome measures for LBP research. More recently, the Quebec scale has been recommended as an acceptable alternative to these two instruments (Kopec, 2000). While the RMDQ is mostly a measure of function, the ODI measures pain as well as function. In practice, however, the difference between these two instruments has been noted to be small, although some floor and ceiling effects have been described for the RMDQ (Bombardier, Hayden & Beaton, 2001; Roland & Fairbank, 2000). Suggestions have been made that the ODI may be more suited to specialty care settings where higher disability is expected while the RMDQ may be more suitable in primary care settings (Bombardier, 2000b; Roland & Fairbank, 2000). However, RMDQ scores have indicated its appropriateness for patients with more severe disability (Roland & Morris, 1983a). Leclaire *et al* (1997) found that the Oswestry and the RMDQ were able to discriminate between two groups of patients with LBP of differing severity. Stratford *et al* (1994) compared the RMDQ to the Oswestry and found them to be similar in terms of sensitivity to change. Both instruments have been criticised for the lack of evidence of content validity and insufficient studies of their internal structure (Kopec & Esdaile, 1995).

The 24-item RMDQ was derived from the SIP (Bergner *et al*, 1976) by Roland and Morris (1983a) (see Appendix 1). Twenty-three items that seemed to be most relevant to patients with back pain were selected from the SIP, and a further item related to the temporal nature of back pain was added (Jensen *et al*, 1992). These represented eight different SIP categories, but no formal sub-scales were developed and the established scoring scale of the SIP was ignored (Delitto, 1994; Deyo, 1988). Furthermore, no rationale was given for the choice of the items (Delitto, 1994). The phrase '*because of my back pain*' was added to each item to distinguish between disability resulting from back pain and disability resulting from other causes. Although the RMDQ

originally included a separate pain scale, this is seldom used (Underwood, Barnett & Vickers, 1999).

Appropriate correlations have been observed between the RMDQ and ordinal pain scales, spine flexion and straight-leg raising (Deyo, 1986; Roland & Morris, 1983a). Short-term test-retest reliability has been demonstrated (Roland & Morris, 1983a) and has been shown to be comparable to that of the SIP (Deyo, 1986; Jensen *et al*, 1992). Other studies have shown the properties of the RMDQ to be better than or equal to that of similar measures (Beurskens *et al*, 1995; Kopec *et al*, 1995; Stratford *et al*, 1994). Responsiveness (Beurskens *et al*, 1996; Patrick *et al*, 1995; Riddle, Stratford & Binkley, 1998; Stratford *et al*, 1994; Stratford *et al*, 1996; Stratford *et al*, 1998) and sensitivity to change have been also determined (Stratford *et al*, 1998; Riddle, Stratford & Binkley, 1998). As a rough guide, Roland and Fairbank (2000) recommended that a change of two to three points on the RMDQ should be considered the minimal clinically important change. A study by Deyo (1986) showed strong correlations between the RMDQ and the physical subscale of the SIP but not with the psychosocial subscale. Although various authors have proposed some modifications to the RMDQ, an international group of experts have recommended the original version because of its wide use in so many countries (Deyo *et al*, 1998).

Bouter, Van Tulder and Koes (1998) and Kopec *et al* (1995) have criticised the method used to develop the RMDQ, stating that it was not based on a conceptual approach or empirical methods of item development, analysis and selection. Roland and Fairbank (2000) argued that while the range of problems covered by the RMDQ was limited, its strength lay in the scoring method, which is easy to understand and interpret. Lee *et al* (2001) identified that patients could interpret some items differently, for example, the item involving 'walking short distances' provided no explanation of what was meant by a 'short distance'. They also suggested that patients might base their responses on a *typical* day rather than the day of assessment. Although there is some evidence favouring the RMDQ above other back-specific functional measures, it is generally felt that further research is needed to study the merits of the different scales under different conditions and with different populations (Kopec & Esdaile, 1995).

2.10. Cross-cultural Adaptation of Self-report Measures

Most health status measures have been developed in English-speaking countries with a Western culture. Adaptation may thus be needed if they are to be used in different contexts (Beaton *et al*, 2000). If the questionnaire is to be used in the same language and culture in which it was developed, then no adaptation is necessary. To enable the RMDQ to be used more widely within the South African context, translations into other of the 11 official languages would be required. Furthermore, because some of the questions may not be applicable to all people living in South Africa, some adaptations may need to be made to the items. The translation and adaptation of the RMDQ to make it culturally appropriate for different groups of people in South Africa is an important future research study.

2.11. Evaluation of Individual Health Outcome Measures

As the process involved in validating instruments is lengthy, it is generally recommended that existing instruments are used (Boden, 1998; Bombardier, 2000a). Deyo *et al* (1994) urged investigators not to develop new measures if existing instruments served the same purpose. Kopec *et al* (1995) however supported the development of new scales arguing that most current measures had no conceptual framework and were deficient in their measurement properties. The fact that many existing measures were designed for use in clinical trials rather than clinical practice where practicality important, further supports Kopec's perception (Deyo *et al*, 1998; Roland & Morris, 1983a).

Shorter scales have been noted to be more acceptable to both patients and clinicians. However, they are less reliable and comprehensive than longer ones (Kopec & Esdaile, 1995). Possible strategies are either to focus on measuring several concepts in one questionnaire, or measuring one concept more accurately in a shorter questionnaire. In making this decision, it is important to realize that for a health status instrument to be considered scientifically acceptable, it must meet the four criteria of validity, reliability, sensitivity and practicality must be met. Moreover, if the instrument is to be used in a different setting than the one in which it was

standardised, it may need to be re-validated (Boden, 1998). There is no literature to suggest that the RMDQ has been validated for use in a South African context.

2.11.1. Validity

Validity is concerned with whether the instrument actually measures the underlying attribute it was designed to measure (Boden, 1998; Bowling, 1991). The assessment of validity usually involves assessment against a 'gold standard', but because there is no gold standard of health against which health status measures can be compared, the most common validation methods used are those of content and construct validity (McDowell & Newell, 1987). A discriminative instrument may be validated by comparing two groups of patients (Guyatt, Feeny & Patrick, 1993), for example, those who require surgery and those who do not. Validation is an on-going process. The more an instrument is used and the greater the number of contexts in which it is used, the greater one's confidence in its validity (Guyatt, Feeny and Patrick, 1993).

According to Deyo *et al* (1998), construct validity is a major concern in selecting appropriate instruments to measure function. Construct validity involves comparisons between measures and examines the logical relationships between a measure and characteristics of patients and patient groups (Guyatt, Feeny & Patrick, 1993). Development of construct validity is a continuous process requiring numerous studies to 'examine various theoretical predications' (Ottenbacher & Christiansen, 1997, p. 115).

Content validity is concerned with whether the components of the scale cover *all* aspects of the attribute to be measured (Boden, 1998; Bowling, 1991). Content validity 'examines the extent to which the domain of interest is comprehensively sampled by the items, or questions, in the instrument' (Guyatt, Feeny & Patrick, 1993, p. 624). Items included in the instrument should thus reflect activities that are important to the population being assessed (Boden, 1998). A valid functional assessment instrument

will include items that examine all areas of ADL with items ranging from easy to difficult (Ottenbacher & Christiansen, 1997).

2.11.2. Reliability

Ottenbacher and Christiansen (1997, p. 113) discussed the importance of reliability as 'a key component of the assessment process'. According to Keith (1984; p. 76), reliability is 'affected by the variability in observers' judgements, in patients' performance, in the assessment situation and by the discriminatory power of the assessment instrument.' Test-retest reliability is more appropriate in establishing reliability in a health status measure used for general populations (Keith, 1984). It is considered to be of particular importance when the purpose of the study is to measure change over time (Ottenbacher & Christiansen, 1997). Test-retest reliability enables the researcher to establish if a change in performance reflects a true improvement rather than reflecting measurement error (*ibid*). In most clinical groups there is potential for improvement and changes may occur between assessments thus affecting the stability of the measure.

Internal consistency examines the extent to which the items in an instrument are related to each other (Ottenbacher & Christiansen, 1997; Streiner & Norman, 1995). It is considered to be more important for questionnaires with discriminative purposes (Kirshner & Guyatt, 1985). Measures of internal consistency are important in determining whether the test is measuring the appropriate construct or whether individual items need to be revised or eliminated (Ottenbacher & Christiansen, 1997). A test with high inter-item correlations is homogeneous and is also likely to produce consistent responses (McDowell & Newell, 1987).

2.11.3. Sensitivity

This refers to the ability of an instrument to detect changes or differences that are clinically important, and is important when an instrument has an evaluative purpose (Deyo & Diehl, 1983).

2.11.4. Practicality

The practicality of an instrument is measured according to the rate of missing data, the time of completion, and the complexity of the scoring (Boden, 1998).

CHAPTER 3: RESEARCH METHOD

3.1. Research Design

The study was based on retrospective evidence and was descriptive in its design.

3.2. Study Population and Sample

The study population included all patients with CMLBP, with no other known pathology, who were seen by one Orthopaedic surgeon working in a private practice in the South Peninsula Municipal area of Cape Town. The study population included all patients seen from August 2000 to July 2001. All subjects in the study population were included in the study sample that constituted 42 subjects.

3.3. Instruments

The ODI was initially used as a measure of functional status but was found to be unsuitable. The surgeon reported that patients had difficulty completing it as they found the number of options for each item confusing. It was not, therefore, practical to use. Feise and Menke (2001) had similar criticisms of the ODI. The RMDQ being the other instrument recommended by a panel of experts as suitable for research (Deyo *et al*, 1998) was considered as an alternative (see Appendix 1).

The RMDQ appeared to be more suited to the study for the following reasons:

- As an assessment of activity limitation (disability) due to LBP, it had been extensively used in research as well as for monitoring patients in clinical settings (Roland & Fairbank, 2000).
- The questionnaire was a self-report and was quick to complete, easy to understand and simple to score.
- The items in the questionnaire appeared to be appropriate for the patients seen at this particular private practice both in terms of face validity and content.
- Language was not a problem as patients seen by the surgeon could generally understand English even if it was not their first language.
- As already stated, studies have indicated that it acts as a discriminative outcome measure in LBP.

- Its validity and high internal consistency has been demonstrated in various studies (Beurskens *et al*, 1995; Patrick *et al*, 1995; Roland & Morris, 1983), and its properties have been reported to be equal to or better than other back-specific LBP measures (Beurskens *et al*, 1996; Deyo *et al*, 1998; Kopec & Esdaile, 1995; Stratford *et al*, 1994).

However, on scrutinising the RMDQ for its applicability within a private practice setting in S.A., the surgeon felt the coverage of information was insufficient. The instrument was therefore adapted by including a further scale of 26 items (hereafter called the 'AddS') as a second part to the questionnaire (see Appendix 2). The AddS items were developed by the surgeon from comments that had been made by patients as well as his own experience of living with back pain. These resembled the format used in the RMDQ.

The RMDQ consists of 'yes'/'no' items with one point awarded for each positive response. Subjects complete the questionnaire by ticking (or ringing) the items that apply to them 'today'. Scoring is accomplished by adding the number of positive responses. On the 24-item RMDQ, scores could vary between zero (no disability) and 24 (severe disability). The RMDQ together with the AddS constituted the adapted RMDQ (ARMDQ) with the scores for this new instrument ranging from zero to 50 (see Appendix 3).

3.4. Data Collection

Data were collected from clinical records. Demographic (age, gender, type of work) and anthropometric (height, weight) characteristics were retrieved from the database. This selection of variables has been previously substantiated by literature. Body mass index was calculated from the weight and height measurements by dividing the weight (in kilograms) by height (in metres) squared for each subject to determine whether subjects were overweight or not. This index is a frequently used standard to estimate obesity and is regarded as providing a better estimate of obesity than does relative weight (Wilmore & Costill, 1994). All subjects had completed an ARMDQ during their consultation with the surgeon. The clinical examination consisted of

history taking, examining radiographs and magnetic resonance images of the spine as well as a physical examination that included, amongst other aspects, straight-leg raising and spine flexion tests. Based on these findings as well as information from the patient and professional expertise, the surgeon reached a decision about whether spinal fusion surgery was required or not. This decision was made without the surgeon having seen the completed ARMDQ. The decision regarding surgery (hereafter referred to as the rating for surgery) was documented in the following way:

- Y = yes (definitely needs surgery)
- A = more than 50% chance that surgery is needed
- P = less than 50% chance that surgery is needed
- N = no (surgery definitely not needed).

Completed ARMDQ forms were processed for the researcher by the surgeon's receptionist. Folder numbers were used to retrieve demographic data from the database.

The initial trial to record the data from the completed ARMDQ forms yielded a number of problems with the AddS questionnaires (refer to Appendix 2 for the original AddS):

- Four items were related to the duration for which subjects had experienced continuous pain and were arranged as an ordinal scale. These items differed vastly from the other nominal items, and were therefore excluded from the scoring of the scale. It was also not scientifically correct to allocate a point for each time period that was represented. Responses to these items were used and analysed separately to determine if there was any association between the duration of continuous pain, scores on the RMDQ and AddS, and the other demographic variables.
- There were various numbering errors in the AddS with some numbers having been used twice. The numbering was amended to correct this and the total number of questions remained 50.

Refer to Appendix 3 for the revised version of the AddS.

In recording the responses to the questionnaires, some problems were encountered concerning the extent to which subjects completed the questionnaire according to the instructions. Stratford *et al* (1994) reported similar problems. The particular problems found in this study were dealt with in the following ways:

- One patient placed a dashed circle around three items – these were excluded as this subject had checked other items.
- There were eight instances where patients wrote ‘*sometimes*’ next to an item as well as endorsing the item number. In accordance with Stratford *et al* (1994) these were considered to be positive responses. That patients felt it necessary to write such comments was taken as a reflection of the need for a greater number of response categories for each item.
- Nine subjects, who had circled one or more items for the duration of continuous pain but had not circled item 41 (‘I used to get back pain from time to time but now I have the pain continuously’), were regarded as having checked this item.
- If a subject had selected more than one option under item 41, the longest time period was taken to be the correct option for analysis purposes.

An *Excel* spreadsheet was used to record data from the completed ARMDQ forms together with the decision regarding surgery and demographic details for each patient.

3.5. Pilot Studies

Due to the retrospective nature of this study, a pilot study was not conducted. All items were included in the questionnaire so that, through the analysis, the questions that were most valuable would be identified.

3.6. Ethical Considerations

Ethical approval was granted by the Research Ethics Committee of the Faculty of Health Sciences of the University of Cape Town.

Confidentiality was maintained by ensuring that patients could in no way be identified or linked to the data. Subject’s names were removed from the completed

ARMDQ forms by the surgeon's receptionist prior to photocopying. The researcher worked only with the photocopied forms thus ensuring the anonymity of the subjects. Each form contained the subject's folder number but the researcher had no access to their personal files.

3.7. Analysis

The first step in the analysis involved calculating totals and frequencies for each demographic variable and their categories. Descriptive statistics are presented according to age, gender, occupation, BMI, continuous pain, rating for surgery and previous surgery. Occupation was classified using a simple scale that categorised the occupations according to the amount of physical effort involved. These were: 'sitting and light physical', 'sitting and medium physical' and 'medium to heavy physical'.

Totals for the RMDQ, AddS, and the combined scale (ARMDQ) were calculated to determine the mean scores for each instrument. Item totals were determined to indicate the items that had been most frequently endorsed. Data were analysed using the computer programme *Statistical Package for the Social Sciences (SPSS)*.

The Pearson product-moment correlation coefficient (r) was used to reveal insight into the association between the totals on the RMDQ, AddS and the ARMDQ. The correlation coefficient is a point on the scale between -1.00 and +1.00 – the closer it is to either of these limits, the stronger the relationship between the two variables (Howell, 1995).

For reliability, inter-rater agreement was not applicable in this study as the questionnaires had been self-administered. Due to the retrospective nature of the study, test-retest reliability could not be established, as the ARMDQ had not been administered to subjects a second time. Internal consistency in this study was determined using Cronbach's coefficient alpha, which 'examines the correlation among all individual items and the correlation among individual items and subgroups of items' (Ottenbacher & Christiansen, 1997, p. 113). Although a high Cronbach's

alpha shows that there is consistency amongst the items, it does not explain which items may be better at measuring the phenomenon.

To reveal the inter-relationships between the items in each scale, a categorical principal component analysis (CATPCA) was performed. This enabled the researcher to determine which dimensions of function were being measured by the RMDQ, the AddS and the ARMDQ, and in exactly what way the different items were related. For the ARMDQ to be a discriminative test of functional status, people with chronic disabling back pain would need to attain higher scores (Keith, 1984) reflecting lower functional ability (or greater disability). By inspecting the component loadings for each item, the items that were more discriminative in measuring functional status were identified. Subsequent correlations and CATPCA analyses were performed on the selected items in each scale to determine whether item selection improved the reliability of the scales.

Spearman's rho was used to examine the relationship between the demographic variables. A CATPCA was performed to determine the inter-relationship between the demographic variables.

Pearson product-moment correlation coefficients were calculated for the totals of the selected item scales (RMDQ and AddS) and the demographic variables (continuous pain, occupation, and BMI). A CATPCA was performed to show the inter-relationship between the demographic variables and the total RMDQ and AddS (selected items).

Spearman's correlation coefficient for ranked data (r_s) was used to examine the scores obtained on the RMDQ, AddS and the ARMDQ with the surgery rating of the subjects to establish whether there was any association between them.

CHAPTER 4: RESULTS

4.1. Sample Description

A total of 42 subjects, all of whom had completed the ARMDQ, were included in the study sample. Table 1 (overleaf) provides details concerning demographic, anthropometric and medical information. All patients were regarded to be from middle or upper socio-economic groups as they were either members of a medical aid scheme, or were able to pay for their consultation and treatment. Selected characteristics of the sample will be discussed further.

4.1.1. Age and Gender

The majority of the subjects (50.0%; N=42) were from the age categories of 41 to 54 years, with fewer being 55 years or above (19.1%; N=42). The mean age was 49.9 years. There were more female than male subjects.

4.1.2. Occupation

Thirty-six subjects (85.7%; N=42) were either employed or were housewives while four (9.5%; N=42) had retired. Three of the retired subjects were women and one was a man. Two of these subjects were of retirement age, i.e. 60 years or above for a woman and 65 years or above for a man. The other two subjects had retired early and were aged 55 and 57 years respectively. No information was available regarding their reasons for retiring early. The missing data resulted from one subject being self-employed but with no details pertaining to the type of work performed, while the other subject had not completed the section concerning their occupation. As the analysis was concerned with the type of work performed by subjects rather than whether they were in paid employment or not, housewives were included in the classification. Eleven subjects (26.2%; N=42) were housewives and were classified as performing 'medium to heavy physical activity'.

Table 1: Socio-demographic and Medical Data for Subjects (N=42)

Variable	Categories	No. of subjects			% of total sample
		Male	Female	Total	
Age	25–35	0	6	6	14.3
	36–40	3	4	7	16.7
	41–45	5	4	9	21.4
	46–54	5	7	12	28.6
	55–65	1	5	6	14.3
	66+	1	1	2	4.8
	Total:		15	27	42
Occupation	Sitting and light physical	13			34.2
	Sitting and medium physical	9			23.7
	Medium to heavy physical	14			36.8
	Missing data	2			5.3
	Retired	4			9.5
Total:		38			100.0
Body mass index	Acceptable	20			47.6
	Overweight	13			31.0
	Obese	6			14.3
	Missing data	3			7.1
Total:		42			100.0
Continuous pain	Yes	34			81
	No	8			19
Total:		42			100.0
Duration of continuous pain	Less than 3 months	5			14.7
	More than 3 months up to 6 months	5			14.7
	More than 6 months	8			23.5
	More than a year	16			47.1
Total:		34			100.0
Rating for surgery	Yes	7			16.7
	More than 50% chance	7			16.7
	Less than 50% chance	11			26.2
	No	17			40.5
Total:		42			100.1 *
Subjects who had had previous back surgery	Yes	4			9.5
	No	38			90.5
Total:		42			100.0

* Percentages total more than 100% due to rounding off of decimal places

4.1.3. Body Mass Index (BMI)

According to Lambert (Email; 23.7.01), a subject with a BMI score over 25 is considered 'overweight' while one with a score over 30 is considered 'obese'. Using this as a guide, subjects were recorded as having a satisfactory BMI, or as being overweight or obese. Nineteen subjects (45.3%) were calculated as being overweight or obese, nine of whom were men and 10 were women. There were three subjects (7.1%) for whom BMI's could not be calculated due to missing values for either weight or height or both.

4.1.4. Continuous pain

This variable was derived from responses to item 41 in the AddS (see Appendix 2). Most subjects (81%; N=42) reported having continuous pain.

4.1.5. Duration of Continuous Pain

Of the subjects who had checked item 41, only five subjects had experienced continuous pain for less than three months, while a high number (47.1%; N=34) reported continuous pain for more than a year.

4.1.6. Rating for Surgery

According to the surgeon's rating, only seven subjects (16.7%) required surgery, while 17 (40.5%) definitely did not need it. There was less certainty concerning the need for surgery in the remaining subjects. All subjects who were rated as needing surgery reported continuous back pain, five of whom had experienced pain for more than a year.

4.1.7. Previous Surgery

Only four subjects (9.5%) had had previous surgery. All of these subjects reported continuous pain for either more than six months or more than a year. Three subjects were rated as not needing surgery while one was rated

as having a less than 50% probability of needing it. All subjects reported continuous pain for more than six months.

4.2. Analysis of the Questionnaires

For the ARMDQ to be considered an improvement over the RMDQ, the AddS should relate collinearly to the RMDQ. Furthermore, the AddS should measure the same concept as the RMDQ in addition to at least some new aspects of back pain.

This hypothesis was tested using the following criteria:

- the correlation between both scales should be high but not perfect
- the item reliability of the AddS should reach at least the same level as the item reliability of the RMDQ
- item reduction should improve the canonical correlation and select the most discriminative (disability versus no disability) items.

The various comparisons are represented in Figure 1.

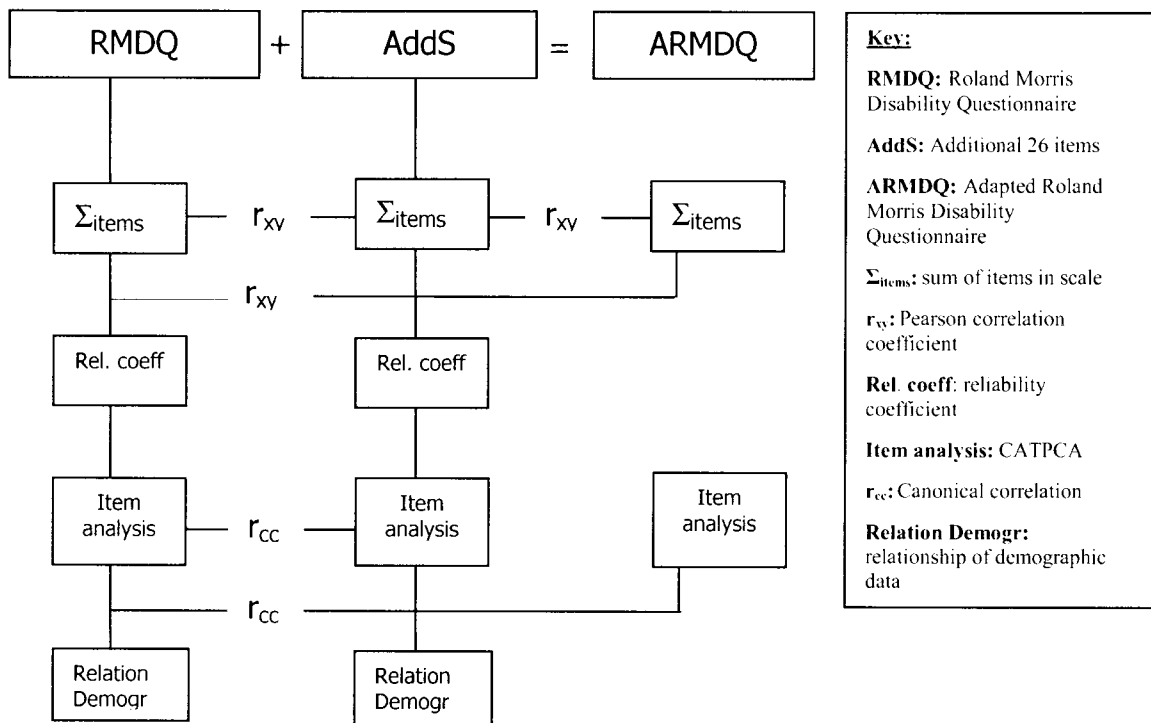


Figure 1: Model Depicting the Comparisons in the Analysis

4.2.1. Subjects and Items Excluded From Analysis

Thirty-six subjects (85.7%; N=42) were included in the analysis. Six subjects were excluded as their response patterns differed markedly from the rest. As stated previously, the total number of responses was calculated per item for each scale (RMDQ and AddS). Items in which there were no or less than 10 responses were excluded from the analysis as there were too few responses to analyse their effects systematically. Similarly, items endorsed by more than 33 subjects (approximately 80 percent of the sample) were also excluded. This was done to ensure that there was enough variation in responses. The mean score for the RMDQ was 8.6 (median = 9.0; range = 21) while that for the AddS was 9.4 (median = 9.0; range = 22).

The number of items analysed in the RMDQ and the AddS was 19 and 14 respectively. Items that were excluded from the analysis of each scale are shown in Tables 2 (RMDQ) and 3 (AddS).

Table 2: Items Excluded from the Analysis of the RMDQ

RMDQ item
1. I stay at home most of the time because of my back.
15. My appetite is not very good because of my back pain.
19. Because of my back pain, I get dressed with help from someone else.
20. I sit down for most of the day because of my back.
24. I stay in bed most of the time because of my back.

Table 3: Items Excluded from the Analysis of the Added Scale (AddS)

AddS item
25. I avoid going to the cinema because of my back.
27. Because of my back I avoid dinner parties and similar social occasions.
30. When getting out of bed in the morning my back is stiff and sore but eases a bit after being up for a while.
33. My back is better walking briskly rather than slowly.
36. I try to avoid walking uphill because of my back.
37. I try to avoid walking downhill because of my back.
40. Straining on a toilet aggravates my back pain.
41. I used to get back pain from time to time but now I have the pain continuously.
43. I can feel something moving in my back.
46. I feel nervous in a crowd for fear of being bumped.
49. I rely on other people to do things for me because of my back.
50. My back pain is no longer a nuisance it has become a disability.

4.2.2. Correlation Between the Item Totals of the RMDQ and the Added Scale (AddS)

Pearson's product-moment correlation (r) was used to determine the association between the totals for the sub-scales (RMDQ and AddS).

Table 4: Pearson's Product Moment Correlations and Levels of Significance for the RMDQ, AddS and ARMDQ (N=36)

Scale	Correlation	AddS	ARMDQ
RMDQ	Pearson	$r=.722 (**)$	$r=.935 (**)$
	Significance (2-tailed)	$P<.000$	$P<.000$
AddS	Pearson		$r=.921 (**)$
	Significance (2-tailed)		$P<.000$

** Correlation is significant at the .01 level (2-tailed). Nonparametric correlations.

4.2.3. Comparison of the reliability of the RMDQ and the AddS

Reliability was tested using a Cronbach's alpha reliability test and was based on the optimised item response categories (CATPCA)*. Items were treated nominally. Cronbach's alpha for the RMDQ was .92 suggesting that all the items are related and contribute in measuring functional status. The Cronbach's alpha for the AddS was 0.86. Although this suggests that the items are closely related, a comparison shows that the AddS items are not as closely related as those in the RMDQ. The added scale explains 11% less of the variance in the responses. Furthermore, the Cronbach's alpha for the ARMDQ was .93. Thus adding the AddS items to the RMDQ did not improve the reliability of the original questionnaire.

* CATPCA performs a non-linear Principal Component analysis. Responses, treated non-numerically, are optimally transformed and a dimension reduction reveals the inter-relationship between the variables.

4.2.4. Item Analysis of the RMDQ and the AddS

Although Cronbach's alpha showed that there was consistency amongst the items, it did not explain which items were better at measuring functional status. In order to reveal the interrelationship between the items a Categorical Principal Component Analysis (CATPCA) was performed on both scales.

RMDQ. The CATPCA identified two dimensions, suggesting that responses referred to two different aspects of back pain. The eigenvalue (proportion of variance explained by each factor) was 7.6 and explained 40% of the total variance. The first dimension explained 28% of the variance (eigenvalue = 5.3), while the second explained 12% (eigenvalue = 2.3).

Closer inspection of the component loadings for each item identified those that contributed highly to either dimension 1 or to dimension 2. Since items with low component loadings seemed unrelated to the remaining items (further analysis had not revealed further clustering of the items), a subsequent CATPCA was performed in which only the items that were explicitly measuring the phenomenon were included.

The analysis of the selected items showed an improved eigenvalue (4.7), and explained 59% of the total variance. Figure 2 (overleaf) illustrates the interrelationship between the items that were selected by previous analysis and considered as most contributive to measuring functional status in the study sample.

An angle of 90° implies that the items are unrelated. The relationship between items is illustrated in the following way: the smaller the angle between the vectors, the higher the relationship between them, and vice versa.

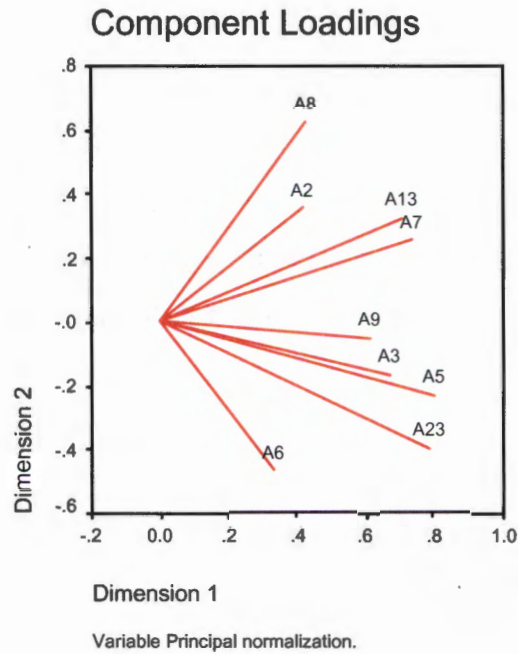


Figure 2: Component Loadings for RMDQ Items After Removal of Items not Contributing to the Measurement of Functional Status

Items formed two clusters – those between A8 and A7, and those between A9 and A6. On examination of these items, it appeared that the former cluster seemed to be related to ‘*compensations made because of back pain*’, while the latter appeared to relate to ‘*reduction in function resulting from LBP*’.

AddS. The results of the CATPCA of the AddS showed a different pattern. Firstly, fewer items were removed from the analysis (refer to Table 2). The eigenvalues, however, were relatively lower than those for the RMDQ, being 2.8 for dimension 1 and 2.2 for dimension 2. The total eigenvalue was 5.0 and explained 36% of the total variance. The first dimension explained only 20% of the variance and the second 16%. Thus, there was a considerable drop in performance compared to the responses to the items of the original scale – even compared to the initial analyses before the item selection. The interrelationship between the items is presented in Figure 3. The corresponding component loadings are presented in Table 5.

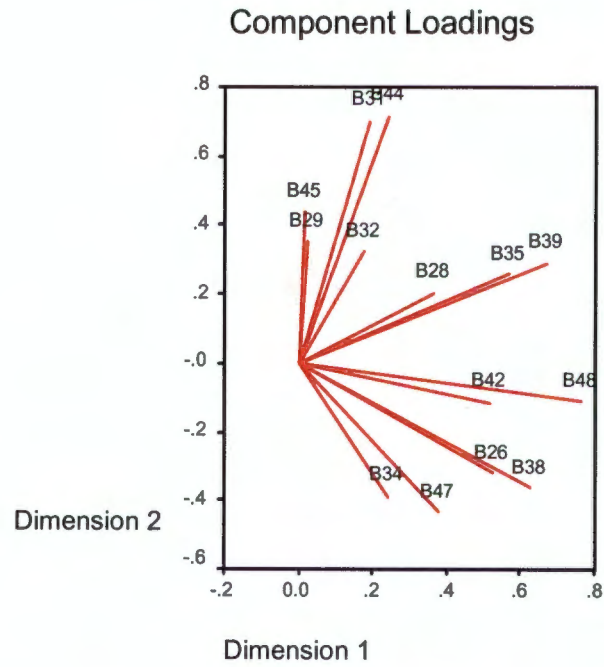


Figure 3: Component Loadings for AddS Items After Removal of Items not Contributing to Measurement of Functional Status

Table 5: Component Loadings for the AddS

Item number	Dimension 1	Dimension 2
B26	.522	-.321
B28	.362	.204
B29	.028	.351
B31	.189	.698
B32	.180	.321
B34	.239	-.391
B35	.571	.257
B38	.623	-.364
B39	.673	.288
B42	.519	-.114
B44	.245	.715
B45	.017	.437
B47	.378	-.436
B48	.766	-.111

Figure 3 shows three clusters of items – B48 to B34, B28 to B39, and B45 to B32. The cluster bordered by B48 and B34 seems to relate to ‘*activities adapted/avoided to reduce back pain*’, while those bordered by B28 and B39 appear to involve ‘*activities that aggravate back pain*’. The cluster of items from B29 to B32 seem to relate to the ‘*negative consequences of performing a specific activity*’.

ARMDQ. Although the above analyses suggest that the AddS does not reach the level of consistency in measuring the related aspects of functional status as the original scale, a CATPCA was used to analyse the combined scales (22 selected items). The eigenvalues were 4.8 and 2.6 for dimensions 1 and 2 respectively explaining 34% of the total variance. This is somewhat less than the 35% explained by the additional scale only, but is a considerable decrease (25%) compared to the RMDQ.

Close scrutiny of the component loadings showed that except for A2, all RMDQ items load high on dimension 1, whereas six out of the sixteen items in the additional scale load high on dimension 2 and therefore seem to measure something different than the RMDQ (refer to Figure 4).

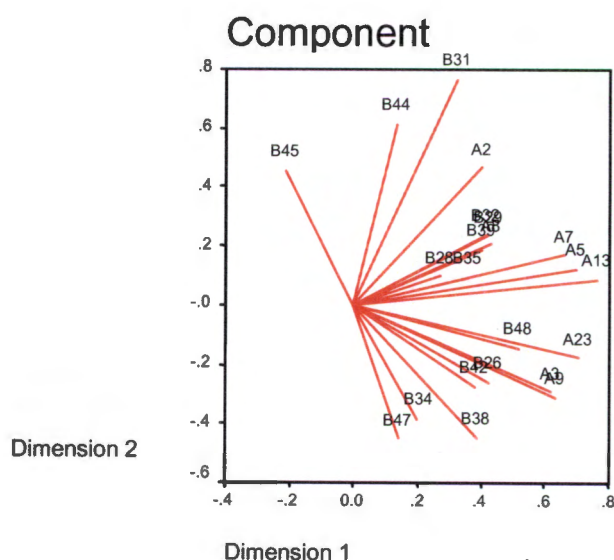


Figure 4: Component Loadings of the ARDMQ (Selected Items)

It is evident here that ‘B’ items (items included in the AddS) seem to be different from the ‘A’ items (items in the RMDQ), and therefore, the added scale seems to measure a different aspect of functional status.

4.2.5. Correlations Between the RMDQ and the AddS After Selection of Items

A non-linear Canonical analysis was performed to show the correlation between the set of selected items of the RMDQ and the set of selected items of the AddS. This revealed a canonical correlation of .86 for dimension 1 and .79 for dimension 2. Thus, a slight increase in the correlation is observed compared to the Pearson correlation ($r = .72$) of the item totals of the RMDQ with the AddS (see Table 6). The correlation of $r = .72$, however, was an overestimation since separate analysis of both scales revealed the need to select items. The selection of items improved the consistency between the items and thus the construct validity of the scales. The correlation between the selected item totals, however, decreased towards $r = .52$.

Table 6: Correlations Between Selected Item Totals and Object Scores of the RMDQ and AddS (N=42)

	Correlation	Total AddS	Selected AddS items (dimension 1)	Selected AddS items (dimension 2)
Total RMDQ	Pearson	$r=.524 (**)$	$r=.502 (**)$	$r=.020$
	Significance	$P<.000$	$P<.001$	$P<.900$
Selected RMDQ items (dimension 1)	Pearson	$r=.511 (**)$	$r=.488 (**)$	$r=-.002$
	Significance	$P<.001$	$P<.001$	$P<.991$
Selected RMDQ items (dimension 2)	Pearson	$r=.091$	$r=.125$	$r=.298$
	Significance	$P<.568$	$P<.431$	$P<.056$

** Correlation is significant at the 0.01 level (2-tailed)

Further examination of the relationship between the item totals (sum of selected items for each scale) and the object scores (selected items based on CATPCA) per dimension revealed significant correlations in the same range for only dimension 1. This suggests that the item totals of the selected scales are representative of the performance on the selected items.

The correlations between the RMDQ and the AddS explain 25% of the variance. Although this is statistically significant ($P < .000$), 75% is not explained. The assumption that 75% of what is measured by the AddS is something unique to the RMDQ is strong.

Furthermore, the item totals of the RMDQ and AddS correlate highly with the first dimension and very low with the second dimension. This supports the results of the CATPCA where only a few items defined the second dimension. It suggests that the totals are fairly representative as a summary of the performance of the scale.

4.3. Correlations Between the Demographic Variables

Spearman's rho (r_s) was used to examine the relationship between the demographic variables. Occupation is a categorical variable and age and the other variables are always considered in terms of ordinal categories. The categorisation of the variables is presented in Table 7.

The results of this analysis showed poor associations between age and occupation (OCC), continuous pain (CPN) and BMI. The results are presented in Table 8. As was expected, there was a high correlation between height and weight and consequently with the BMI. The number of subjects varies for some of the correlations due to missing values for data.

Table 7: Categorisation of Demographic and Anthropometric Variables

BMI = Body Mass Index		OCC = Occupation	
1	= 18.00 – 24.98	1	= sit and light physical
2	= 24.99 – 29.00	2	= sit and medium physical
3	= 29.01 upwards	3	= medium to heavy physical
CPN = Continuous Pain		Weight	
1	= <3 months	1	= 47 – 65 kilograms
2	= 3 months ≤6 months	2	= 66 – 79 kilograms
3	= >6 months and up to one year	3	= 80 – 99 kilograms
4	= > one year	4	= 100 kilograms upwards
HGT = Height			
1	= 154 – 165 meters		
2	= 165 – 174 meters		
3	= 175 – 180 meters		
4	= 181 meters upwards		

A CATPCA was performed to examine the relationship between demographic variables more closely. The number of subjects in the analysis was 42 of which there were 30 active cases and 12 with missing values. Missing values were either as a result of subjects not having continuous pain (of which there were 8 subjects) or because data was not available. The latter applied to the occupation of one subject and the BMI of three subjects. Age had a very low correlation with the other demographic variables and was therefore removed from this analysis. Occupation was treated as a nominal variable while the others were treated as ordinal variables.

The CATPCA yielded an eigenvalue of 2.2 for dimension 1 and 1.7 for dimension 2 with a total value of 3.9 that explained 78 percent of the total variance. Further analysis of the component loadings showed that the demographic variables clustered in two groups as presented in Figure 5.

Table 8: Correlations between the Various Demographic Variables

Demographic Variable	Correlation Co-efficient	OCC	CPN	Height	Weight	BMI
Age	Spearman's rho (r_s)	-.005	.260	-.084	-.097	-.117
	Significance (P)	.973	.137	.613	.553	.459
	N	41	34	39	40	42
OCC	Spearman's rho (r_s)		.314	-.253	-.139	.085
	Significance (P)		.075	.125	.399	.596
	N		33	38	39	41
CPN	Spearman's rho (r_s)			.174	.065	.001
	Significance (P)			.350	.722	.997
	N			31	32	34
Height	Spearman's rho (r_s)				.748 **	.182
	Significance (P)				.000	.268
	N				39	39
Weight	Spearman's rho (r_s)					.635
	Significance (P)					.000
	N					40

** Correlation is significant at the .01 level (2-tailed)

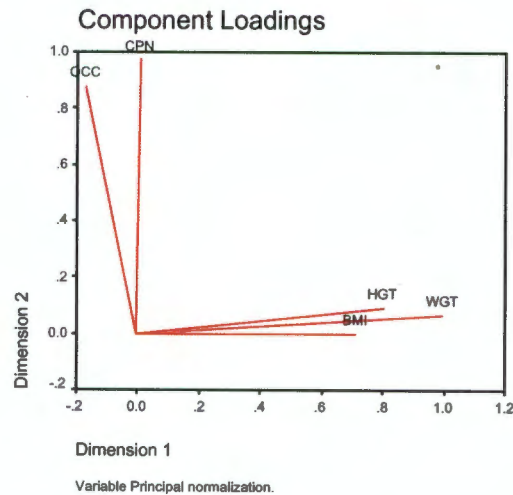


Figure 5: Component Loadings for Demographic Variables

The plot clearly illustrates that occupation (OCC) and continuous pain (CPN) are closely related and determine dimension 2 while height (HGT), weight (WGT) and BMI are strongly related and largely determine dimension 1. This seems to suggest that continuous pain and occupation are not necessarily related to BMI, height and weight.

Figure 6 illustrates how the variables are related on a categorical level. For example, higher categories of continuous pain are associated with higher categories of occupation. Thus, pain experienced over a period of more than a year (Category 4) is associated with occupations of ‘medium to heavy physical work’ (Category 3). Categories 1 and 2 of CPN (pain less than 3 months, or more than 3 months up to 6 months) fall on the same point, and are associated with Category 1 of occupation (sitting and light physical work). Furthermore, the greater the height, the greater will be the weight and the BMI. This is evident as BMI is based on height and weight.

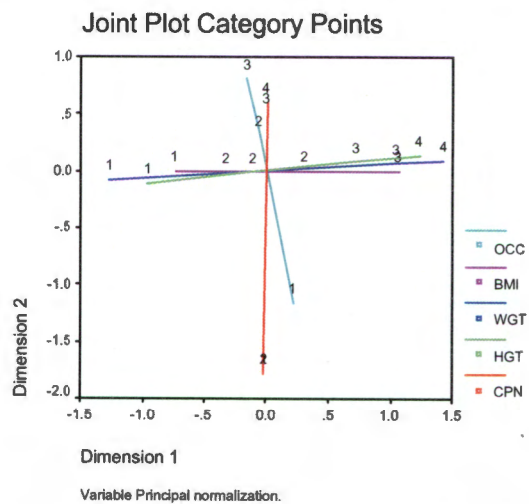


Figure 6: Relationship Between Demographic Variables and Their Categories

4.4. Analysis of the Interrelationships Between the Selected Items Measuring Functional Status and Demographic Variables

Totals of selected items for the RMDQ (TORG) and the AddS (TADD) were correlated with CPN, OCC and BMI. The correlation matrix (Table 9) shows that the TORG correlates better with CPN and BMI than the TADD.

Table 9: Correlations With New Scales (Selected Items) and Demographic Variables

Scale	Correlation	Occupation (N=41)	CPN (N=34)	BMI (N=42)
TORG	Pearson	R=.228	R=.303	R=.275
	Significance	P<.151	P<.082	P<.077
TORG – dimension 1	Pearson	R=.217	R=.289	R=.248
	Significance	P<.174	P<.097	P<.114
TORG – dimension 2	Pearson	R=.029	R=.187	R=.208
	Significance	P<.856	P<.288	P<.186
TADD	Pearson	R=.281	R=.193	R=.107
	Significance	P<.075	P<.274	P<.499
TADD – dimension 1	Pearson	R=.290	R=.176	R=.080
	Significance	P<.066	P<.319	P<.615
TADD – dimension 2	Pearson	R=-.021	R=.019	R=.076
	Significance	P<.895	P<.917	P<.633

Occupation, however, shows an opposite picture and seems more related to TADD. Thus, the AddS might measure some aspects that are more occupation related whereas the RMDQ relates more to pain sensation and physiology (see Figure 7). Furthermore, correlations for dimensions 1 and 2 for both scales show that dimension 1 contributes more highly while dimension 2 makes very little contribution. Furthermore, the total Cronbach's alpha for dimension 1 was .68 while that for dimension 2 was .38 showing that the items in dimension 2 are not closely related.

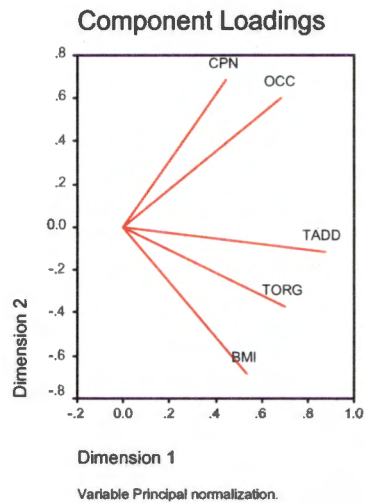


Figure 7: Relationship Between Demographic Variables and the Selected Item Scales

Since the above results indicated that the RMDQ with selected items was most consistent in measuring functional status, a CATPCA of these items with continuous pain, occupation and BMI was conducted. The CATPCA revealed eigenvalues of 3.77 for dimension 1, and 1.59 for dimension 2. The total eigenvalue was 5.36 and explained 49% of the total variance. A closer inspection of the component loadings showed that the relationship between the demographic variables follows the previous pattern in that CPN and OCC seem more related to item 13 and item 8 whereas items 5, 23 and others (items 2, 3, 7 and 9) are more related to BMI (Figure 8).

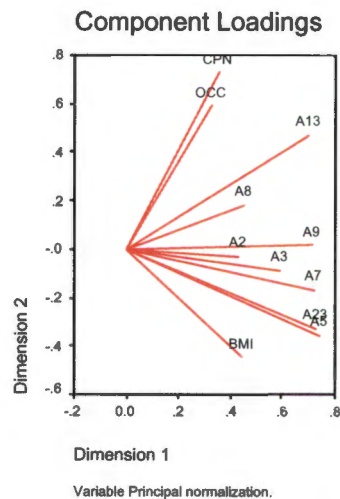


Figure 8: Component Loadings for Demographic Variables and Items on the RMDQ

Figure 9 illustrates the relationship between the demographic variables and items 5 and 13 through the categories.

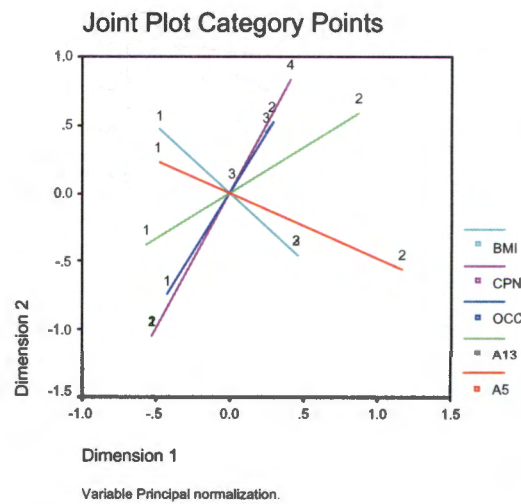


Figure 9: Relationship of Demographic Variables (and their Categories) to Items 5 and 13 in the RMDQ

Item 5 refers to *'Because of my back, I use a handrail to get upstairs'* whereas item 13 refers to *'My back is painful almost all the time'*. A positive answer to *'Because of my back, I use a handrail to get upstairs'* (item 5) is related to a BMI of 25 or more. In contrast a BMI less than 25 is related to a negative answer to this item.

Likewise, a positive answer to *'My back is painful almost all the time'* is related to *'medium and heavy physical activity'* and complaints about CPN of more than 6 months. A negative response to this item is related to *'sitting and light physical activity'* and continuous pain less than and including 6 months.

To compare how the AddS performs in this sense to the RMDQ an additional CATPCA was performed with the selected AddS items and CPN, OCC and BMI. Consistent with the previous insight, the results showed a decrease in performance: The total eigenvalue was 5.38 and explained 32% of the total variance.

4.5. Analysis of the Relationship Between the Rating for Surgery and Scores on the RMDQ and AddS

Spearman's rho (r) was used to examine the relationship between the responses on the RMDQ and the AddS scales and the surgery rating of the subjects (see Table 10).

	RMDQ	AddS
Rating for surgery	.40	.53
Significance	P<.01	P<.01
Variance explained	16%	25%

Table 10: Relationship Between Rating for Surgery and Selected Item Scales

The relationship between the total RMDQ scores and the rating for surgery was low ($r=.40$; $P<.01$). Although this is statistically significant, the correlation only explains 16% of the variance. The correlation between the total AddS scores and the rating for surgery was slightly higher than this ($r=.53$) but only explained 25% of the variance. It was therefore not opportune to execute further analyses to examine the predictive power of the ARMDQ and the subscales (RMDQ and AddS) in discriminating between subjects who needed surgery and those who did not.

CHAPTER 5: DISCUSSION

5.1 Sample Description

The high number of subjects between the ages of 41 and 54 years (50%; N=42) concurs with the pattern reported by Hillman *et al* (1996) who found a significantly higher prevalence of LBP in people between the ages of 45 and 54 years. This is further supported by Roland and Morris (1983b) who reported that the highest consultation for LBP was for people between these ages. Although the present study sample was small, the trend followed that described by Ozguler *et al* (2000) with the majority of subjects being older than 40 years of age (69.1%; N=42) and fewer being over 60 years of age (11.9%; N=42).

The higher number of women (64.3%) than men (35.7%) in the study sample further corresponds with the findings of Ozguler *et al* (2000) who reported a greater number of women with chronic LBP. Hillman *et al* (1996) on the other hand, found no significant differences in the prevalence of back pain between men and women in general, while Kelsey, Mundt and Goldin (1992) reported approximately equal frequency of LBP in men and women in younger age groups. The latter finding differed from that of the present study where there were more women under the age of 40 years than men. Moreover, the number of women in the present study sample increased with age with the greatest number being over 45 years of age. This increased prevalence of LBP amongst women has been supported by other research findings (Kelsey, Mundt & Goldin, 1992).

The small percentage of subjects (2.6%; N=42) who were involved in heavy work can be explained by the fact that the study was conducted in a private practice. Subjects were therefore from middle to upper socio-economic groups and were thus more likely to have studied further, enabling them to be employed in a professional capacity where work tends to be of a more sedentary nature. The under-representation of subjects performing heavy occupations is problematic in that other studies have shown a higher incidence of reported back injuries amongst people in this group compared to those in other types of work (Kelsey, Mundt & Goldin, 1992; Waddell, 1998). Moreover, Kelsey, Mundt and Goldin (1992) and Waddell (1998) found that

people employed in jobs requiring heavy manual labour tended to be from a lower socio-economic group. This is representative of the situation in S.A. where unskilled people from lower socio-economic income groups usually perform heavy work. The relatively large number of housewives in the present study (26.2%; N=42) is similar to the findings of Picavet, Schouten and Smit (1999) whose study sample also constituted a large number of housewives with low back problems.

Classification of some of the occupations of subjects within the present study was difficult as there were insufficient details regarding the requirements of each subject's job. For example, a nurse could work in an administrative position that is sedentary, or in a ward where heavy physical activity is required. Thus, occupations of some subjects may have been categorised incorrectly thereby affecting the results to some extent. This also applies to retired subjects who could not be categorised for occupation, as there was no information related to the type of occupations they performed. They could thus not be included in the analysis. As work was classified in terms of 'any productive activity' (refer to the definition of terms) all subjects should have been classified according to the occupations they currently performed.

The high percentage of subjects (44.3%; N=42) who were classified as overweight or obese in this study has been supported by Ozguler *et al* (2000) who found that subjects with a BMI greater than 22,6kg/m² were more likely to visit a health professional or have treatment for LBP. Other researchers have reported conflicting results pertaining to the relationship of weight and height to LBP. For instance, Croft *et al* (1999) found an association between high weight and LBP in women but in men, neither height nor weight predicted LBP. According to Waddell (1998), most studies show that body weight, and even obesity, does not increase the risk of LBP.

The number of subjects in the sample who complained of continuous pain (81%; N=42) is higher than that reported by Kopec and Esdaile (1998). These researchers found that 28% of their study sample (N=23) reported having back pain 'always' or 'almost always'. A possible explanation for this difference is that subjects in the present study had *chronic* mechanical LBP while in Kopec and Esdaile's study the only stipulation was that subjects were seeking professional help for back pain.

Furthermore, subjects in the latter study were mainly recruited from physiotherapy and psychiatry clinics and may have had less severe back pain than those in the present study.

The finding that 16.7% (N=42) of the sample were rated as definitely requiring spinal fusion surgery is higher than that reported in other studies. For instance, Odendaal (1999) reported that one to two percent of the subjects included in his study were treated surgically. Katz *et al* (1997) however, found that the rate of surgery varied between nine and 56% depending on the surgeon involved in making the decision. The relatively high figure for surgery in the present study could again be due to a greater severity of LBP experienced by the subjects. Furthermore, because the surgeon involved in the study is regarded as an expert in the field of back problems, the subjects could have been consulting him for a second opinion if they had not been treated successfully elsewhere. That subjects who had undergone previous surgery were rated as either definitely not needing surgery or having a less than 50% probability of needing it, has been supported by Waddell (1992) who reported a progressively deteriorating success rate in subsequent spine operations. Hutchinson *et al* (2000) produced further evidence of this occurrence in their findings that subjects who had undergone two operations for lumbar discectomy had significantly greater disability than those who had only had one operation.

5.2 Results of the Questionnaires

The mean RMDQ score of 8.6 in this study is lower than that reported by Roland and Morris (1983a). The latter study found a mean score of 11.4 in a sample of 230 LBP subjects seen at a group practice. In a study of patients with mechanical back pain Deyo (1986) reported a slightly lower mean score of 10.1, while Lee *et al* (2001) found a mean of 10.4 amongst patients with LBP from an orthopaedic spine clinic. In contrast to these studies, Hillman *et al* (1996) reported a mean RMDQ score of 5.4 amongst subjects who had experienced LBP during the previous year and reported LBP on the day of the study. The lower mean RMDQ score displayed in this study was not expected considering that all subjects had chronic back pain and would therefore be expected to have had higher RMDQ scores.

The range of RMDQ scores in previous studies differed slightly from that of the present study in which the range was 0 to 21. Roland and Morris (1983a) reported a range of 0 to 23 while Lee *et al* (2001) calculated a range of 0 to 24. Patrick *et al* (1995) reported a range of 0 to 23 for a modified version of the RMDQ in a sample of subjects with sciatica. The reason for these variations could be due to the differences in the sample sizes for each study as well as the differences in the severity of the low back problem.

5.2.1 Subjects and Items Excluded From the Analysis

The nature of the study was explorative, and therefore, the matter of the most interest was the overall pattern of responses rather than individual responses to items. This reasoning has been supported by Streiner and Norman (1995) who advocated the importance of having enough variation in responses as well as a sufficient number of observations between items to avoid floor and ceiling effects. Furthermore, in a dichotomous scale, such as the RMDQ and AddS, items where one alternative has a very high (or very low) endorsement rate are usually eliminated (*ibid*). An endorsement rate over 95% (or under 5%) suggests that most people are responding in the same direction. It follows therefore that such items do not improve the psychometric properties of a scale (*ibid*). Streiner and Norman (1995) recommended using only items with endorsement rates between 20% and 80%. Stratford and Binkley (1997) used similar response frequencies (less than 20% or greater than 90%) in considering items for deletion when they developed the Back Pain Functional Scale.

Items included in the analysis for the present study have been supported by previous research. Stratford *et al* (1993) conducted a study to identify the most relevant items from the SIP for patients with back pain. Twenty items were found to have been checked most frequently of which only seven appeared in the RMDQ. These seven items were also included in the analysis for the present study, with three of these (items 9, 16 and 21) having been endorsed by more than 50% of the study sample. It is

interesting to note that only 50% of the items identified by Stratford *et al* (1993) were from the physical subscale of the SIP while the others were from the psychosocial domain. This contrasts strongly with the RMDQ, which comprises, almost entirely, items from the physical subscale of the SIP.

Item 21 (*'I avoid heavy jobs around the house because of my back'*) was the most frequently endorsed item (69.1%; N=42) in the present study. Lee *et al* (2001) similarly found this item to be the most highly endorsed having been checked by 78.3% of their study subjects. The AddS item that was checked most frequently was item 41 (*'I used to get back pain from time to time but now I have the pain continuously'*). This item was endorsed by 81% of the subjects (N=42) and was therefore excluded from the analysis. Examination of the item reveals that it seems to be concerned with pain rather than function.

Other studies have identified items in the RMDQ that seem to be less discriminative in measuring functional status. In a study conducted by Stratford and Binkley (1997), results suggested that six items could be deleted from the scale (namely items 2, 15, 17, 19, 20 and 24). Reasons for excluding these items included few endorsements by subjects, low item-total correlation, high correlation with a similar item and poor face validity. The present study supports these findings in that four of these items (items 15, 19, 20 and 24) were eliminated from the analysis due to being endorsed by fewer than 10 subjects. In a study by Hillman *et al* (1996) to establish the prevalence of symptoms of LBP, 10 RMDQ items were endorsed most frequently. Five of these items (numbers 2, 10, 11, 16 and 21) were checked by at least 50% of the subjects in the present study. In a modification of the RMDQ, Patrick *et al* (1995) deleted five items and replaced them with four others from the SIP. Unfortunately, due to inconsistencies in the article, it is not clear exactly which items were deleted. Thus, there appears to be some similarities in the pattern of responses between the present study and other published study findings.

5.2.2 Comparisons Between the RMDQ and the AddS

Comparison of the correlations between the RMDQ and the ARMDQ ($r=.94$; $P<.00$), and the AddS and the ARMDQ ($r=.92$; $P<.00$), showed that there was a slightly higher correlation between the RMDQ and the ARMDQ. The decrease in the correlation coefficient between the RMDQ and the AddS ($r=.72$; $P<.00$) shows that there is less agreement between the two scales than between the ARMDQ and each scale. This was predictable considering that each scale is included in the ARMDQ. Other studies have compared the RMDQ with back-specific measures in a similar way. For instance, correlations between the RMDQ and the ODQ have been reported as .66 and .72 (Leclaire *et al*, 1997), .77 (Beurskens *et al*, 1995) and .79 (Stratford *et al*, 1994). Correlation coefficients have been reported to be .68 for the Aberdeen Back Pain Scale (Garratt, Moffett & Farrin, 2001), .79 for the Back Pain Functional Scale (Stratford, Binkley & Riddle, 2000), and .77 for the Quebec Back Pain Disability Scale (Kopec *et al*, 1995) compared with the RMDQ. Deyo (1986) and Jensen *et al* (1992) reported correlations for the RMDQ and the SIP to be .85 and .78 respectively.

The Cronbach's alpha of .92 for the RMDQ showed a high internal consistency among the items. This compares favourably with that of other studies where Cronbach's alpha was determined as .84 (Jarvinkoski, Mellin & Estlander, 1995), .87 (Stratford, Binkley & Riddle, 2000), .90 (Kopec & Esdaile, 1995; Kopec *et al*, 1995) and .93 (Hsieh *et al*, 1992). Underwood *et al* (1999) reported a Cronbach's alpha of .96 for a modified version of the RMDQ. The internal consistency of the items in the AddS was lower than that of the RMDQ with a Cronbach's alpha of .86. Furthermore, Cronbach's alpha for the ARMDQ did not increase showing that the added items did not improve the internal consistency of the scale. This indicates that the items in the AddS are not well related to, and measure something different from those in the RMDQ. Closer examination of the AddS items shows that the AddS items appear to be more related to pain while those in the RMDQ are more related to function.

The identification of two dimensions in the RMDQ by the CATPCA is a unique finding and other studies in which similar analyses have been performed on the RMDQ are not apparent. Some researchers have used factor analysis (CATPCA belongs to the group of factor analyses) in the development of functional status measures for LBP. A factor analysis of the Dallas Pain Questionnaire showed factor loading on two major components – ‘functional’ and ‘emotional’ (Delitto, 1994). Ruta *et al* (1994) used this same method to identify separate health ‘factors’ within a questionnaire designed to evaluate outcome on patients with LBP. A factor analysis performed by Stratford & Binkley (1997) demonstrated that all the items included in the Back Pain Functional Scale loaded on a common factor. Kopec *et al* (1995) also used factor analysis to develop the Quebec back pain scale.

The analysis performed in the present study identified two dimensions. Dimension 1 items were concerned with *‘function’* while those in dimension 2 seemed to be more related to *‘behaviour’*. Examination of the items included in each cluster revealed that the RMDQ items from A8 to A7 seemed to relate to *‘compensations because of pain’*, while those from A9 to A6 seemed to relate to *‘reduction in function resulting from LBP’* (refer to Figure 2 in the results section).

The three clusters of items identified in AddS scale (refer to figure 3 in the results section) seemed to be related to:

- *‘negative consequences of performing a specific activity’* (B32 to B45).
- *‘activities aggravating LBP’* (B28 to B39).
- *‘activities adapted or avoided to reduce back pain’* (B34 to B48).

While all these clusters seem to be concerned with pain resulting from performing activities, their content differs slightly in emphasis. Items loading highly on dimension 1 seemed to be related to *‘activities or actions increasing/aggravating back pain’* while those loading more highly on

dimension 2 appeared to be concerned with '*behaviours adopted as a result of the pain*'. Thus, although the dimensions appear to be similar to those identified in the RMDQ, the emphasis in the AddS seems to be on pain.

The cluster labels further illustrate the differences between the two scales showing that each scale measures different aspects of functional status related to LBP. The component loadings for the combined scale further illustrate the clear difference between items in the RMDQ and those in the AddS. While almost all items in the RMDQ loaded on dimension 1, a fairly high number of AddS items loaded highly on dimension 2. Furthermore, the plot of the component loadings (refer to figure 4 in the results section) clearly shows that items in the RMDQ (vectors marked 'A') clustered completely differently from those in the AddS (vectors marked 'B'). It is clear that the content of the two scales differs.

The decreasing performance of the AddS and subsequently the ARMDQ is clearly evidenced by the decreasing eigenvalues and the amount of variance explained. When items that were more discriminative in measuring functional status in the study sample were selected, the eigenvalue for the RMDQ increased. Furthermore, removing those items that were not discriminating highly, improved the consistency between the items and thus the construct validity of the scale. The RMDQ therefore performed better than the AddS as well as the ARMDQ. This is not surprising considering that the AddS was developed on an ad-hoc basis and had not been tested prior to this study. Furthermore, items included in the AddS were not based on a theoretical construct of functional status. Thus, the addition of the AddS items added considerable 'noise' to the scale thereby reducing not only its reliability but also its construct validity.

5.2.3 Correlations with Demographic Variables

The poor association between age and OCC as well as CPN and BMI appears to be fairly logical in that age does not usually influence the occupations performed by a person, nor does it necessarily have any bearing on CPN and BMI. The high correlation between height and weight and BMI was expected considering that BMI is calculated using these measurements.

The clustering of the demographic variables showed that OCC and CPN were related to, and determined, dimension 2 (*'behaviour changes to accommodate LBP'*) while BMI (and height and weight) were related to, and determined, dimension 1 (*'changes in function resulting from LBP'*). This implies that a person with CPN will tend to compensate in the way they perform activities in order to reduce their pain. Furthermore, it seems that the higher the BMI, the greater the extent to which the person's occupational performance will be affected. Examination of the relationship of these variables and their categories provided further insight to this relationship. For example, longer duration of CPN is associated with heavier occupations. If this is compared to the 'label' for dimension 2, it seems that the greater the duration of CPN, the greater the behavioural changes will need to be to accommodate the pain.

Pearson correlations showed that the selected RMDQ items correlated with CPN and BMI while those of the AddS correlated more closely with OCC. The RMDQ therefore seemed to be more related to the sensation of pain as well as physiology, while the AddS seemed to be associated with more occupation-related issues. The high correlations between the demographic variables and dimension 1 indicate that *'changes in function resulting from LBP'* are closely related to OCC, CPN and BMI. Comparison of the performance of the AddS with the demographic variables showed a decrease in performance compared with that of the RMDQ. The CATPCA for the AddS and demographic variables showed a decrease in the

eigenvalue and explained 32% of the total variance in comparison to the RMDQ, which explained 49% of the total variance. This confirmed the better performance of the RMDQ in comparison with the AddS.

5.3 Analysis of the Relationship Between the Rating for Surgery and Scores on the RMDQ and AddS

There was a higher correlation between the AddS and the rating for surgery than between the RMDQ and rating for surgery. This could be explained by the fact that the AddS was designed by the surgeon who would have included items which he considered to be important in making a decision regarding surgery. This agrees with Katz *et al* (1997) who found the surgeon to be the most powerful predictor of spinal fusion. While the AddS seemed to be more highly correlated with the rating for surgery, it still only explained 25% of the variance. Thus, it seems that neither the RMDQ nor the AddS can be used to indicate with certainty the need for surgery. Furthermore, since both scales only succeeded in explaining so little of the variance, it is debatable whether either are actually worth pursuing any further.

5.4 Problems with the RMDQ

Problems related to the format of the RMDQ became evident during this study. Firstly, the questionnaire required subjects to ring or check an item if it applied to them 'today'. However, there was no way of knowing whether a subject had mistakenly omitted an item rather than purposely not checking it. This could have resulted in underscoring on both the RMDQ and the AddS which could have affected the results. Secondly, if the scale was to be used to evaluate change over time (for example to determine whether spinal fusion had resulted in an improvement in functional status), the dichotomous nature of the questionnaire may lead to an artificial polarisation of responses, thus skewing the results. Stratford and Binkley (1997) identified this as an apparent deficiency in the RMDQ. This is illustrated by the fact that some subjects wrote 'sometimes' beside an item, suggesting that they may not have completed the scale based on their back pain 'today'. However, dichotomous scales are still considered to be a good option for discriminative measures (Kirshner & Guyatt, 1985).

5.5 Limitations of the Study

The study was limited to subjects seen within one private practice by one surgeon and was not representative of the diversity of people in S.A. in terms of ethnic group, level of education, type of occupation, language or socio-economic level. The results cannot, therefore, be generalised. Furthermore, because only subjects with CMLBP were included in the study (for the sake of homogeneity) the results cannot be generalised to all patients with LBP.

The fact that one surgeon was involved reduced the variables involved in making decisions about the need for surgery. Furthermore, as data capture had been geared towards one instrument, it precluded the possibility of comparing it with another measure, such as the Back Pain Functional Scale (Stratford, Binkley & Riddle, 2000).

CHAPTER 6: CONCLUSIONS AND RECOMMENDATIONS

6.1 Conclusion

Most functional status measures have been developed in the U.K., Canada and the U.S.. However, these instruments are currently not frequently used in South Africa to determine the extent of disability experienced by individuals with LBP. This may be due, in part, to the fact that their validity and reliability have not been tested within the multiplicity of South African contexts; they are not generalisable to the people (and contexts) in South Africa due to wide variations in the culture and language of its people. Low levels of education and a high level of illiteracy may furthermore render self-report functional status measures difficult to use in some contexts. Occupational therapists, with their understanding of function and the impact health problems may have on performance, should be more familiar with a variety of functional status measures and can moreover make a valuable contribution to research in this field in S.A.. Functional status measures may also assist clinical decision-making for individual patients (Stratford *et al*, 2000).

This study examined an addition to the RMDQ in a homogeneous group of subjects with CMLBP in an effort to demonstrate whether this scale was an improvement over the original. Results showed that the internal consistency of the scale decreased with the addition of the AddS. Item reduction improved the internal consistency and thus the construct validity of the RMDQ. Subsequent analysis revealed that items in the RMDQ and AddS measured two different dimensions of function. Furthermore, this analysis identified two clusters of items in the RMDQ and three clusters in the AddS. Findings showed that the selected AddS items were more closely related to OCC and CPN, and the RMDQ to BMI. The low correlations between scores on the RMDQ and AddS and the rating for surgery showed that the surgeon's evaluation is at this stage more reliable in determining whether surgery is needed than the RMDQ or AddS scores. The CATPCA items identified the items that were most discriminative in measuring functional status. This improved the reliability of the scale.

The ARMDQ was found to be a suitable instrument for completion by the subjects included in this study, as it was quick and simple to complete. It was also easy to

score. Problems identified with the scale were related to its design, particularly, the use of a nominal scale and the method of completion (checking the items that were applicable). Modifications could alleviate these problems. It was noted that many items included in the AddS scale might not be universally applicable to the general population of S.A., for example, 'going to the cinema'. Some items in the RMDQ may similarly be inappropriate within a South African context.

6.2 Recommendations

- The AddS did not improve the reliability of the RMDQ and also seemed to be measuring a different aspect of function. There seem to be two possibilities. Either the AddS items should be removed from the scale to preserve the integrity of the original RMDQ, or the performance of the RMDQ and AddS items selected by the CATPCA should be used as a means of comparison with that of another functional status scale such as the Back Pain Functional Scale (Stratford, Binkley & Riddle, 2000) together with a quality of life scale.
- Future studies should involve larger sample sizes in order to improve reliability and to establish cut-off points for disability.
- While the AddS may be of use to the surgeon, it does not correlate highly with the rating for surgery and should therefore not be used to predict which patients will require surgery. Further studies should be done to investigate whether other functional status scales could be used for this purpose.
- Some improvements to the layout of the RMDQ are suggested: instead of checking items, the options 'yes', 'no', and some intermediate option(s) should be presented in a check-box format to enhance accuracy.
- Further research to investigate the appropriateness of specific functional status measures should include patients treated in State hospitals and health services to ensure a more representative sample. This will require scales to be translated and tested for cultural appropriateness.

REFERENCES

Andersson, G.B.J. & Weinstein, J.N. 1994, 'Health outcomes related to low back pain', *Spine*, vol. 15, no. 19 (18 Suppl), pp. 2026S-2027S.

American Occupational Therapy Association. 1995a, 'Position paper: Occupation', *American Journal of Occupational Therapy*, vol. 49, no. 10, pp. 1015–1018.

American Occupational Therapy Association. 1995b, 'Position paper: Occupational performance: Occupational Therapy's definition of function', *American Journal of Occupational Therapy*, vol. 49, no. 10, pp. 1019–1020.

Beaton, D., Bombardier, C., Guillemin, F. & Ferraz, M. 2000, 'Guidelines for the process of cross-cultural adaptation of self-report measures', *Spine*, vol. 25, no. 24, pp. 3186–3191.

Bergner, M., Bobbitt, R., Kressel, S., Pollard, W., Gilson, B., & Morris, J. 1976, 'The Sickness Impact Profile: Conceptual formulation and methodology for the development of a health status measure', *International Journal of Health Services*, vol. 6, no. 3, pp. 393–415.

Beurskens, A., De Vet, H., Koke, A., Van der Heijden, G. & Knipschild, P. 1995, 'Measuring the functional status of patients with low back pain. Assessment of the quality of four disease-specific questionnaires', *Spine*, vol. 20, no. 9, pp. 1017–1028.

Beurskens, A., De Vet, H. & Koke, A. 1996, 'Responsiveness of functional status in low back pain: A comparison of different instruments', *Pain*, vol. 65, pp. 71–76.

Boden, S. 1998, 'Outcome assessment after spinal fusion. Why and how?', *Orthopedic Clinics of North America*, vol. 29, no. 4, pp. 717–728.

Bombardier, C. 2000a, 'Outcome assessments in the evaluation of treatment of spinal disorders', *Spine*, vol. 25, no. 24, pp. 3097–3099.

Bombardier, C. 2000b, 'Outcome assessments in the evaluation of treatment of spinal disorders. Summary and general recommendations', *Spine*, vol. 25, no. 24, pp. 3100–3103.

Bombardier, C., Hayden, J. & Beaton, D.E. 2001, 'Minimal clinically important difference. Low back pain: Outcome measures', *Journal of Rheumatology*, vol. 28, no. 2, pp. 431-438.

Bouter, L., Van Tulder, M. & Koes, B. 1998, 'Methodologic issues in low back pain research in primary care', *Spine*, vol. 23, no. 18, pp. 2014–2020.

Bowling, A. 1991, *Measuring Health. A Review of Quality of Life Measurement Scales*, Open University Press, Milton Keynes.

Christiansen, C. & Baum, C. (eds) 1997, *Occupational Therapy. Enabling Function and Well-being*, 2nd edn, SLACK Incorporated, New Jersey.

Cicinelli, L. 1997, 'Back Pain', in *Occupational Therapy and Physical Dysfunction. Principles, Skills and Practice*, eds A. Turner, M. Foster, & S. Johnson, Churchill Livingstone, New York, pp. 667–690.

Cook, F.M. & Hassenkamp, A. 2000, 'Active rehabilitation for chronic low back pain', *Physiotherapy*, vol. 86, no. 2, pp. 61–67.

Creek, J. 1997, *Occupational Therapy and Mental Health*, 2nd edn, Churchill & Livingstone Medical Division of Pearson Professional Limited, United States of America.

Croft, P.R., Papageorgiou, A.C., Thomas, E., MacFarlane, G.J. & Silman, A.J. 1999, [Abstract of 'Short-term physical risk factors for new episodes of low back pain. Prospective evidence from the South Manchester Back Pain Study', *Spine*, vol. 24, no. 15, pp. 1556-1561], [Electronic], Available: Silver Platter File: MEDLINE/99386025 [2001, June 26].

Delitto, A., 1994, 'Are measures of function and disability important in back care?', *Physical Therapy*, vol. 74, no. 5, pp. 452–460.

Deyo, R. 1986, 'Comparative validity of the Sickness Impact Profile and shorter scales for functional assessment in low-back pain', *Spine*, vol. 11, no. 9, pp. 951-954.

Deyo, R. 1988, 'Measuring the functional status of patients with low back pain', *Archives of Physical Medicine and Rehabilitation*, vol. 69, no. 12, pp. 1044-1053.

Deyo, R. & Diehl, A. 1983, 'Measuring physical and psychosocial function in patients with low-back pain', *Spine*, vol. 8, no. 6, pp. 635-642.

Deyo, R., Andersson, G., Bombardier, C., Cherkin, D., Keller, R., Lee, C., Liang, M., Lipscomb, B., Shekelle, P., Spratt, K. & Weinstein, J. 1994, 'Outcome measures for studying patients with low back pain', *Spine*, vol. 19, no. 18, pp. 2032S-2036S.

Deyo, R., Battie, M., Beurskens, A., Bombardier, C., Croft, P., Koes, B., Malmivaara, A., Roland, M., Von Korff, M. & Waddell, G. 1998, 'Outcome measures for low back pain research', *Spine*, vol. 23, no. 18, pp. 2003-2009.

Edwards, D.F. 1997, 'The effect of occupational therapy on function and well-being', in *Occupational therapy. Enabling function and Well-being*, 2nd edn, SLACK Incorporated, New Jersey, pp. 557-574.

EuroQoL Group. 1990, 'EuroQoL: A new facility for the measurement of health-related quality-of-life', *Health Policy*, vol. 16, pp. 199-208.

Fairbank, J., Couper, J., Davies, J. & O'Brien, J. 1980, 'The Oswestry Low Back Pain Disability Questionnaire', *Physiotherapy*, vol. 66, no. 8, pp. 271-273.

Fairbank, J. & Pynsent, P. 2000, 'The Oswestry Disability Index', *Spine*, vol. 25, no. 22, pp. 2940-2953.

Feise, R. & Menke, J. 2001. 'Functional Rating Index. A new valid and reliable instrument to measure the magnitude of clinical change in spinal conditions', *Spine*, vol. 26, no. 1, pp. 78-87.

Frymoyer, J.W., Pope, M.H., Clements, J.H., Wilder, D.G., MacPherson, B. & Ashikaga, T. 1983, 'Risk factors in low-back pain', *Journal of Bone and Joint Surgery*, vol. 65A, no. 2, pp. 213–218.

Frymoyer, J.W. & Cats-Baril, W.L. 1991, 'An overview of the incidences and costs of low back pain', *Orthopedic Clinics of North America*, vol. 22, no. 2, pp. 263–271.

Garratt, A., Moffett, J. & Farrin, A. 2001, 'Responsiveness of generic and specific measures of health outcome in low back pain', *Spine*, vol. 26, no. 1, pp. 71–77.

Gill, T. & Feinstein, A. 1994, 'A critical appraisal of the quality of Quality-of-Life measurements', *Journal of the American Medical Association*, vol. 272, no. 8, pp. 619–625.

Gresham, G.E. & Dittmar, S.S. 1997, 'Instruments used to assess function and measure outcomes in physical rehabilitation', in *Functional Assessment and Outcome Measures for the Rehabilitation Health Professional*, S.S. Dittmar & G.E. Gresham, Aspen Publishers, Gaithersburg, pp. 31–36.

Guyatt, G., Feeny, D. & Patrick, D. 1993, 'Measuring health-related quality of life', *Annals of Internal Medicine*, vol. 118, no. 8, pp. 622–629.

Harper, A., Harper, D., Lambert, L., De Klerk, N., Andrews, H., Ross, F., Straker, L. & Lo, S. 1995, 'Development and validation of the Curtin Back Screening Questionnaire (CBSQ): A discriminative disability measure', *Pain*, vol. 60, no. 1, pp. 73–81.

Hillman, M., Wright, A., Rajaratnam, G., Tennant, A. & Chamberlain, M. 1996, 'Prevalence of low back pain: Implications for service provision in Bradford, UK', *Journal of Epidemiology and Community Health*, vol. 50, no. 3, pp. 347–352.

Howell, D.C. 1995, *Fundamental Statistics for the Behavioral Sciences*, 3rd edn, Duxbury Press, Belmont.

Hsieh, C.Y., Phillips, R.B., Adams, A.H. & Pope, M.H. 1992, 'Functional outcomes of low back pain: comparison of four treatment groups in a randomized controlled trial. *Journal of Manipulative and Physiological Therapeutics*, vol. 15, no. 1, pp. 4-9.

Hunt, S.M., McEwen, J. & McKenna, S.P. 1985, 'Measuring health status: A new tool for clinicians and epidemiologists', *Journal of the Royal College of Practitioners*, vol. 35, pp. 185-188.

Hutchinson, P., Laing, R., Waran, V., Hutchinson, E. & Hollingworth, W. 2000, 'Assessing outcome in lumbar disc surgery using patient completed measures', *British Journal of Neurosurgery*, vol. 14, no. 3, pp. 195-199.

Jarvikoski, A., Mellin, G., & Estlander A. 1995, 'Outcome of two multimodal back treatment programmes with and without intensive physical training', *Journal of Spinal Disorders*, vol. 6, pp. 93-98.

Jensen, M.P., Strom, S.E., Turner, J.A. & Romano, J.M. 1992, 'Validity of the Sickness Impact Profile Roland scale as a measure of dysfunction in chronic pain patients', *Pain*, vol. 50, no. 2, pp. 157-162.

Katz, J.N., Lipson, S.J., Lew, R.A., Grobler, L.J., Weinstein, J.N., Brick, G.W., Fossel, A.H. & Linag, M.H. 1997, 'Lumbar laminectomy alone or with instrumented or non-instrumented arthrodesis in degenerative lumbar spine stenosis. Patient selection, costs, and surgical outcomes', *Spine*, vol. 22, no. 10, pp. 1123-1131.

Keith, R.A. 1984, 'Functional assessment measures in medical rehabilitation: Current status', *Archives of Physical Medicine and Rehabilitation*, vol. 65, no. 2, pp. 74-78.

Kelsey, J.L. & White, A.A. 1980, 'Epidemiology and impact of low-back pain', *Spine*, vol. 5, no. 2, pp. 133-142.

Kelsey, J.L., Mundt, D.J., & Goldin, A.L. 1992, 'Epidemiology of low back pain', in *The lumbar spine and back pain*, 4th edn, M.I.V. Jayson & A.S. Dixon, Churchill Livingstone, Edinburgh, pp 537–549.

Kirshner, B. & Guyatt, G. 1985, 'A methodological framework for assessing health indices', *Journal of Chronic Diseases*, vol. 38, no. 1, pp. 27–36.

Kopec, J.A. 2000, 'Measuring functional outcomes in persons with back pain', *Spine*, vol. 25, no. 24, pp. 3110-3114.

Kopec, J. & Esdaile, J. 1995, 'Functional disability scales for back pain', *Spine*, vol. 20, no. 17, pp. 1943-1949.

Kopec, J.A., Esdaile, J.M., Abrahamowicz, M., Abenhaim, L., Wood-Dauphinee, S., Lamping, D. L. & Williams, J.I. 1995, 'The Quebec Back Pain Disability Scale', *Spine*, vol. 20, no. 3, pp. 341-352.

Kopec, J.A., Esdaile, J.M., Abrahamowicz, M., Abenhaim, L., Wood-Dauphinee, S., Lamping, D. L. & Williams, J.I. 1996, 'The Quebec Back Pain Scale: Conceptualization and development', *Journal of Clinical Epidemiology*, vol. 49, no. 2, pp. 151-161.

Kopec, J. & Esdaile, J. 1998, 'Occupational role performance in persons with back pain', *Disability and Rehabilitation*, vol. 20, no. 10, pp. 373–379.

Lambert, M. 2001, *Standard tables: Weight and height*, E-mail to H. Buchanan (hbuchan@uctgsh1.uct.ac.za), 23 July [2001, July 23].

Lawlis, G.F., Cuencas, R., Selby, D. & McCoy, C.E. 1989, 'The development of the Dallas Pain Questionnaire. An assessment of the impact of spinal pain on behavior', *Spine*, vol. 14, pp. 510-516.

Leclaire, R., Blier, F., Fortin, L. & Proulx, R. 1997, 'A cross-sectional study comparing the Oswestry and Roland-Morris Functional Disability scales in two populations of patients with low back pain of different levels of severity', *Spine*, vol. 22, no. 1, pp. 68–71.

Lee, C.E., Simmonds, M.J., Novy, D.M. & Jones, S. 2001, 'Self-reports and clinician-measured physical function among patients with low back pain: A comparison', *Archives of Physical Medicine and Rehabilitation*, vol. 82, no. 2, pp. 227–231.

Lurie, J. 2000, 'A review of generic health status measures in patients with low back pain', *Spine*, vol. 25, no. 24, pp. 3125–3129.

McDowell, I. & Newell, C. 1987, *Measuring health: A guide to rating scales and questionnaires*, Oxford University Press, New York.

McGorry, R.W., Webster, B.S. & Snook, S.H. 2000, 'The relation between pain intensity, disability, and the episodic nature of chronic and recurrent low back pain', *Spine*, vol. 25, no. 7, pp. 834-840.

McKenna, K. 1993, 'Quality of life: A question of functional outcomes or the fulfilment of life plans', *Australian Journal of Occupational Therapy*, vol. 40, no. 1, pp. 33–35.

Million, R., Hall, W., Nilsen, K.H., Baker, R.D. & Jayson, M.I.V. 1982, 'Assessment of the progress of the back pain patient', *Spine*, vol. 7, no. 3, pp. 204–212.

Nordin, M. 1992, 'Prevention of back pain in industry', in *The Lumbar Spine and Back Pain*, 4th edn, M.I.V. Jayson & A.S. Dixon, Churchill Livingstone, Edinburgh, pp. 551–563.

Odendaal, T. 1999, 'Outcomes and costs in a sample of patients who sustained significant injuries of the lumbar spine in work-related accidents', *South African Bone & Joint Surgery*, vol. 9, no. 2, pp. 80-88.

Ottenbacher, K.J. & Christiansen, C.C. 1997, 'Occupational Performance Assessment', in *Occupational Therapy. Enabling Function and Well-being*, 2nd edn, SLACK Incorporated, New Jersey, pp. 105–135.

Ozguler, A., Leclerc, A., Landre, M., Pietri-Taleb, F. & Niedhammer, I. 2000, 'Individual and occupational determinants of low back pain according to various definitions of low back pain', *Journal of Epidemiology and Community Health*, vol. 54, no. 3, pp. 215–220.

Parkerson, G.R., Broadhead, W.E. & Tse, C.K. 1990, 'The Duke Health Profile: a 17-item measure of health and dysfunction', *Medical Care*, vol. 28, pp. 1056–1072.

Patrick, D., Deyo, R., Atlas, S., Singer, D., Chapin, A. & Keller, R. 1995, 'Assessing health-related quality of life in patients with Sciatica', *Spine*, vol. 20, no. 17, pp. 1899–1909.

Picavet, H., Schouten, J. & Smit, H. 1999, 'Prevalence and consequences of low back problems in The Netherlands, working vs non-working population, the MORGEN-Study. Monitoring project on risk factors for chronic disease', *Public Health*, vol. 113, no. 2, pp. 73–77.

Riddle, D., Stratford, P. & Binkley, J. 1998, 'Sensitivity to change of the Roland-Morris Back Pain Questionnaire: Part 2', *Physical Therapy*, vol. 78, no. 11, pp. 1197–1207.

Rogers, J.C. & Holm, M.B. 1998, 'Evaluation of Occupational Performance Areas', in *Willard and Spackman's Occupational Therapy*, 9th edn, M. Neistadt & E.B. Crepeau, Lippincott Williams & Wilkins, Philadelphia, pp. 185–208.

Roland, M. & Morris, R. 1983a, 'A study of the natural history of back pain. Part I: Development of a reliable and sensitive measure of disability in low-back pain', *Spine*, vol. 8, no. 2, pp. 141–144.

Roland, M. & Morris, R. 1983b, 'A study of the natural history of back pain. Part II: Development of guidelines for trials of treatment in Primary Care', *Spine*, vol. 8, no. 2, pp. 145–150.

Roland, M. & Fairbank, J. 2000, 'The Roland-Morris Disability Questionnaire and the Oswestry Disability Questionnaire', *Spine*, vol. 25, no. 24, pp. 3115–3124.

Ruta, D., Garratt, A., Warlaw, D. & Russell, I.T. 1994, 'Developing a valid and reliable measure of health outcome for patients with low back pain', *Spine*, vol. 19, no. 17, pp. 1887-1896.

Spengler, D.M., Bigos, S.J., Martin, N.A., Zeh, J., Fisher, L. & Nachemson, A. 1986, 'Back injuries in industry: A retrospective study. I. Overview and analysis', *Spine*, vol. 11, no. 3, pp. 241-245.

Stanton, M.P., Gresham, G.E. & Dittmar, S.S. 1997, 'Relationship among measures of disease, general health, and functional status', in *Functional assessment and outcome measures for the rehabilitation health professional*, S.S. Dittmar & G.E. Gresham, Aspen Publishers, Gaithersburg, pp. 31–36.

Stratford, P., Solomon, P., Binkley, J., Finch, E. & Gill, C. 1993, 'Sensitivity of Sickness Impact Profile items to measure change over time in a low-back pain patient group', *Spine*, vol. 18, no. 13, pp. 1723-1727.

Stratford, P., Binkley, J., Solomon, P., Gill, C. & Finch, E. 1994, 'Assessing change over time in patients with low back pain', *Physical Therapy*, vol. 74, no. 6, pp. 528-533.

Stratford, P., Binkley, J., Solomon, P., Finch, E., Gill, C. & Moreland, J. 1996, 'Defining the minimum level of detectable change for the Roland-Morris questionnaire', *Physical Therapy*, vol. 76, no. 4, pp. 359-365.

Stratford, P. & Binkley, J. 1997, 'Measurement properties of the RM-18. A modified version of the Roland-Morris Disability Scale', *Spine*, vol. 22, no. 20, pp. 2416-2421.

Stratford, P., Binkley, J., Riddle, D. & Guyatt, G. 1998, 'Sensitivity to change of the Roland-Morris Back Pain Questionnaire: Part 1', *Physical Therapy*, vol. 78, no. 11, pp. 1186-1196.

Stratford, P., Binkley, J. & Riddle, D. 2000, 'Development and initial validation of the Back Pain Functional Scale', *Spine*, vol. 25, no. 16, pp. 2095-2102.

Streiner, D. & Norman, G. 1995, *Health Measurement Scales. A Practical Guide to their Development and Use*, Oxford University Press, Oxford.

Trombly, C.A. 1995, 'Retraining basic and instrumental activities of daily living', in *Occupational Therapy for Physical Dysfunction, 4th edn*, ed C.A. Trombly, Williams & Wilkins, Baltimore, pp. 289-318.

Underwood, M., Barnett, A. & Vickers, M. 1999, 'Evaluation of two time-specific back pain outcome measures', *Spine*, vol. 24, no. 11, pp. 1104-1112.

Van Tulder, M.W., Koes, B.W. & Bouter, L.M. 1995, 'A cost-of-illness study of back pain in the Netherlands', *Pain*, vol. 62, no. 6, pp. 233-240.

Waddell, G. 1992. 'Understanding the patient with backache', in *The lumbar spine and back pain, 4th edn*, M.I.V Jayson & A.S. Dixon, Churchill Livingstone, Edinburgh, pp. 469-485.

Waddell, G. 1998, *The back pain revolution*, Churchill Livingstone, Edinburgh.

Waddell, G., & Hamblen, D.L. 1983, 'The differential diagnosis of backache', *Practitioner*, vol. 227, no. 1381, pp. 167-1175.

Waddell, G. & Main, C.J. 1984, 'Assessment of severity of low back disorders', *Spine*, vol. 9, no. 2, pp. 204-8.

Ware, J. 1987, 'Standards for validating health measures: Definition and content', *Journal of Chronic Diseases*, vol. 40, no. 6, pp. 473-480.

Ware, J., Brook, R., Davies, A. & Lohr, K. 1981, 'Choosing measures of health status for individuals in general populations', *American Journal of Public Health*, vol. 7, no. 6, pp. 620–625.

Ware, J.E. & Sherbourne, C. 1992, 'The MOS 36-item short-form survey (SF-36): I. Conceptual framework and item selection. *Medical Care*, vol. 30, pp. 473–483.

Wilmore, J.H. & Costill, D.L. 1994, *Physiology of Sport and Exercise*, Human Kinetics, Illinois.

Wilson, I. & Cleary, P. 1995, 'Linking clinical variables with health-related quality of life', *Journal of the American Medical Association*, vol. 273, no. 1, pp. 59–65.

Worku, Z. 2000, [Abstract of 'Prevalence of low back pain in Lesotho mothers', *Journal of Manipulative and Physiological Therapeutics*, vol. 243 no. 3, pp. 147-154], [Electronic], Available: Silver Platter File: CINAHL/2000050193 [2001, April 18].

World Health Organization. 1986, 'A discussion document on the concept and principles of health promotion', *Health Promotion*, vol. 1, pp. 73-78.

World Health Organization. 2001, *ICIDH-2: International Classification of Functioning, Disability and Health. Final draft, Full version*, World Health Organization, Geneva.

APPENDIX 1: THE ROLAND-MORRIS DISABILITY QUESTIONNAIRE (RMDQ)

When your back hurts, you may find it difficult to do some things you normally do. This list contains sentences that people have used to describe themselves when they have back pain. When you read them, you may find that *some* stand out because they describe you *today*. As you read the list, think of yourself *today*. When you read a sentence that describes you today, put a tick against it. If the sentence does not describe you, then leave the space blank and go on to the next one. Remember, only tick the sentence if you are sure it describes you today.

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

___ **Score**

The score is the total number of items checked – i.e., from a minimum of 0 to a maximum of 24.

APPENDIX 2: ITEMS ADDED TO THE ROLAND-MORRIS DISABILITY QUESTIONNAIRE (AddS – Original version)

25. I avoid going to the cinema because of my back.
26. When travelling longer distances, I have to stop and get out from time to time to relieve my back pain.
27. Because of my back I avoid dinner parties and similar social occasions.
28. I avoid cocktail parties because the standing aggravates my back.
29. When rising from a seat my back is stiff and sore.
30. When getting out of bed in the morning my back is stiff and sore but eases a bit after being up for a while.
31. Because of my back I constantly shift around when seated.
32. Walking slowly such as ambling around a shopping centre aggravates my back pain.
33. My back is better walking briskly rather than slowly.
34. I avoid walking on soft sand because of my back.
35. Walking on a level, firm surface initially relieves my back pain but walking too far will again aggravate the pain.
36. I try to avoid walking uphill because of my back.
37. I try to avoid walking downhill because of my back.
38. I brace myself when coughing or sneezing because of my back pain.
39. Stumbling, tripping or in any way jolting my spine aggravates my back pain.
40. Straining on a toilet aggravates my back pain.
41. I used to get back pain from time to time but now I have back pain continuously.
I have had continuous daily back pain for:
42. Less than 3 months.
42. More than 3 months
43. More than 6 months
44. More than 1 year.
45. My back feels unstable, as if it will give way on me.
46. I can feel something moving in my back.
44. My back pain is worse after lying still in one position for any length of time.
45. I can bend without much pain but remaining in a stooped position aggravates my back pain.
46. I feel nervous in a crowd for fear of being bumped.
47. I have stopped playing sport because of my back.
48. I avoid gardening because of my back.
49. I rely on other people doing things for me because of my back.
50. My back pain is no longer just a nuisance - it has become a disability.

___ **Score**

APPENDIX 3: ADAPTED ROLAND-MORRIS DISABILITY QUESTIONNAIRE REVISED VERSION (ARMDQ)

When your back hurts, you may find it difficult to do some things you normally do. This list contains sentences that people have used to describe themselves when they have back pain. When you read them you may find that some stand out because they describe you *today*. As you read the list think of yourself *today*. When you read a sentence that describes you today, ring the number against it. If the sentence does not describe you, then leave the space blank and go on to the next one. Remember, only tick the sentence if you are sure that it describes you *today*.

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

_____ **Subscore**

25. I avoid going to the cinema because of my back.
26. When travelling longer distances I have to stop and get out from time to time to relieve my back pain.
27. Because of my back, I avoid dinner parties and similar social occasions.
28. I avoid cocktail parties because the standing aggravates my back.

29. When rising from a seat my back is stiff and sore.
30. When getting out of bed in the morning my back is stiff and sore but eases a bit after being up for a while.
31. Because of my back, I constantly shift around when seated.
32. Walking slowly, such as ambling around a shopping centre, aggravates my back pain.
33. My back is better walking briskly rather than slowly.
34. I avoid walking on soft sand because of my back.
35. Walking on a level firm surface initially relieves my back pain but walking too far will again aggravate the pain.
36. I try to avoid walking uphill because of my back.
37. I try to avoid walking downhill because of my back.
38. I brace myself when coughing or sneezing because of my back pain.
39. Stumbling, tripping or in any way jolting my spine aggravates my back pain.
40. Straining on a toilet aggravates my back pain.
41. I used to get back pain from time to time but now I have back pain continuously.
 - I have had continuous back daily pain for:
 - (a) Less than 3 months.
 - (b) More than 3 months
 - (c) More than 6 months
 - (d) More than 1 year.
42. My back feels unstable - as if it will give way on me.
43. I can feel something moving in my back.
44. My back pain is worse after lying still in one position for any length of time.
45. I can bend without much pain but remaining in a stooped position aggravates my back pain.
46. I feel nervous in a crowd for fear of being bumped.
47. I have stopped playing sport because of my back.
48. I avoid gardening because of my back.
49. I rely on other people doing things for me because of my back.
50. My back pain is no longer just a nuisance - it has become a disability.

___ **Subscore**

___ **Total score**