DEVELOPMENT AND PSYCHOMETRIC TESTING OF THE IBADAN LOW BACK PAIN DISABILITY SCALE

BY

NOTTIDGE, BOLANLE ABIODUN

(M.Sc PHYSIOTHERAPY) (Ibadan)

A Thesis in the Department of Physiotherapy Submitted to the Postgraduate School,

University of Ibadan, in partial fulfilment of the requirements for the degree of

DOCTOR OF PHILOSOPHY (PHYSIOTHERAPY)

of the

UNIVERSITY OF IBADAN

MARCH 2015

CERTIFICATION

We certify that this research work was carried out by Mrs. B.A. Nottidge in the Department of Physiotherapy, College of Medicine, University of Ibadan, under our supervision.

……………………………………..   …………………………………….

(Main Supervisor)   **(Co-Supervisor)**
Dr. Aderonke O. Akinpelu   Dr. Olufunmilayo I. Fawole
Reader/Consultant Physiotherapist   Reader/Consultant Epidemiologist
Department of Physiotherapy   Department of Epidemiology
College of Medicine   Medical Statistics and Environmental
University of Ibadan   Health
                                                College of Medicine
                                                University of Ibadan

ii

DEDICATION

This work is dedicated to my husband, Dr. T.E. Nottidge, my late father Rev. M. A. Ajagbe, my mother, Mrs. S. Ajagbe and my children, David, James, John-Vidal and Sarah-Blossom for their prayers, love and encouragement throughout the course of this work.

ABSTRACT

Scales which are used for measuring healthcare outcomes are often developed in the context of the culture and environment of the people for which they were originally developed. Scales originally developed for Nigerian culture and environment are few and none exists for Low Back Pain (LBP). The aim of this study was to develop a Nigerian culture and environment-friendly LBP disability scale and investigate its psychometric properties.

Items for the Ibadan Low Back Pain Disability Scale (ILBPDS) were devised through literature review, patients' and experts' interviews. The initial draft with 43 items was taken through content validation, pretesting, factor analysis and series of experts' reviews before the final 18-item scale was produced. Response options were adapted from an existing scale (the Ibadan Knee/Hip Osteoarthritis Outcome Measure) and maximum obtainable score was 72. In this quasi-experimental study, psychometric testing of ILBPDS involved 142 patients with non-specific LBP (Experimental Group [EG]) and 142 age and sex-matched controls (Control Group [CG]) recruited through consecutive sampling from tertiary hospitals across Nigeria. The ILBPDS was completed by EG and CG; participants in EG also completed the Numerical Pain Rating Scale (NPRS) at baseline. ILBPDS was completed again by participants in the EG 48 hours after the initial assessment. Sixty-four participants in EG completed a 5-week physiotherapy programme and were assessed using the ILBPDS and NPRS. ILBPDS scores of EG and CG were compared using Mann Whitney-U test. Internal consistency was analysed using Cronbach's α. Spearman correlation was used to correlate ILBPDS and NPRS scores of EG at baseline and post-intervention; and between changes in ILBPDS and NPRS scores of EG post-intervention. ILBPDS scores taken at baseline and 48 hours later were subjected to Intra Class Correlation (ICC). For EG, pre- and post-

intervention ILBPDS scores and NPRS scores were compared using Wilcoxon signed rank test. Level of significance was set at 0.05.

The age (48.6 ± 12.7 years) of participants in EG was comparable to that (48.2 ±12.3 years) of participants in CG. ILBPDS score of EG [55.2 (45.0-65.6)] was significantly higher than that of CG [21.4(20.0-24.3)](evidence of construct validity) and EG's NPRS score correlated significantly with their ILBPDS score (r = 0.50; p= 0.001)at baseline and post intervention (r =0.35; p= 0.001)(evidence of divergent validity). ILBPDS scores at baseline [55.2 (45.0-65.6)] and 48 hours later [51.8 (41.3-65.7) for participants in EG correlated significantly (ICC=0.80; p= 0.001) (evidence of test re-test reliability). Cronbach's α for ILBPDS was 0.84 (evidence of internal consistency). Post-intervention ILBPDS score [36.5(30.0-49.4)] was significantly lower than the pre-intervention ILBPDS score [55.2 (45.0-65.6)](evidence for responsiveness).Post-intervention changes in ILBPDS and NPRS scores correlated significantly(evidence of responsiveness).

The Ibadan Low Back Pain Disability Scale is valid, reliable, responsive and internally consistent for measuring disability in patients with non-specific low back pain and it is proposed for use in the Nigerian clinical setting.

Keywords: Low back pain, Disability scale, Ibadan, Pain score.

Word Count: 477

ACKNOWLEDGEMENT

I wish to express my sincere gratitude to all who contributed to the successful completion of this work. First of all, I want to thank, the immediate past Head of Department and my supervisor, Dr Aderonke O. Akinpelu, for her hard work, advice and sacrifice without which it would have been impossible to complete this work. I also appreciate my supervisor for getting the University of Ibadan Senate Research Grant to facilitate this work. Thank you so much Ma.I am deeply grateful to my co-supervisor, Dr Olufunmilayo I. Fawole for being available to contribute effectively to the success of this work. May God bless you Ma.

I thank Prof. T.K. Hamzat, the head of Physiotherapy Department for his encouragement, assistance and motivation. I am very grateful to Prof A.O. Sanya for her motherly advice and motivation. I want to appreciate Dr. B.O.A. Adegoke for always being there to advice, encourage and sometimes reading through the work. Thank you Sir. I wish to put on record the effort of Dr. O.M. Akpa of the Department of Epidemiology and Medical Statistics, University of Ibadan in helping with and educating me on factor analysis and other aspects of the data analysis. I am indebted to Dr. Adesola C. Odole and Mrs Nse A. Odunaiya for their constant input, advice, prayer and encouragement. I appreciate both of you.

I want to say a big thank you to the immediate past and current Postgraduate Coordinators – Dr. Adesola Odole and Dr Omoyemi Ogwumike for all their assistance and effort. I am grateful to all the postgraduate lecturers: Dr. Adeniyi, A.F., Dr Fabunmi, Dr Ayanniyi,O., Dr Olubukola Olaleye and Dr Akinremi.

I sincerely appreciate the effort of my colleagues who facilitated data collection from physiotherapy facilities in different parts of the country. They include Mrs. Olanrewaju, from Physiotherapy Department, National Orthopaedic Hospital, Dala Kano, Mr. Akinlade O. and Mrs Temilola Abolarin from National Orthopaedic Hospital, Igbobi Lagos, Mr Akanbi from University of Maiduguri Teaching hospital, Mr Achonwa B. and Mr Salufu P. from University of Nigeria Teaching Hospital Enugu, Dr B. Jonathan from University of Port Harcourt Teaching Hospital, Mr J. Umunna from Nnamdi Azikwe University Teaching Hospital, Nnewi. Mr Boyinde from Jos University Teaching Hospital, Mrs. Uche Oluwatosin from Federal Medical Centre, Umuahia, Mr F. Udoh, Mrs S. Eze, Mr S. Omotunde, Miss E. Umoren and Miss C. Obiako, from University of Uyo Teaching Hospital. I do appreciate the assistance of Mr O. Banjo and Mrs J. Anikpe from University of Uyo Teaching Hospital during the course of this study. I cannot forget the assistance of Mr G. Stephen in the course of this work.May God send help to you in your time of need. I am indebted to Dr. U. Ekripo and Dr Uwem. Ekanem from University of Uyo Teaching Hospital, Mr.Omidiji J. from Finance Department UCH, for their input in analysing the data.

I cannot forget to thank my colleagues in the Physiotherapy Department of the University College Hospital, Ibadan, especially, Mrs Olabisi Akinwande, Dr. Oladapo O. Taiwo, Dr Abiola Atowoju, Mr O. Olasehinde, Mrs A. Akintayo, Dr Oladunni Oshundiya and Mrs Bola Tinubu for their love and encouragement.

I want to say 'a big thank you' to my spiritual leaders for their prayers and encouragement, Pastor Dr & Dr Mrs. U. Obed, Pastor (Engr.) & Mrs. Adeyanju, Pastor Dr & Mrs. Kupolati for their prayers. I sincerely appreciate Mrs. Taiwo Ajagbe for her love, encouragement and

assistance; my siblings – Mr Kunle Ajagbe, Mrs Biola Ikoko, Mr Tope Ajagbe, Mr Olalekan Ajagbe, Mr Tayo Ajagbe for their love, care and prayers that saw me through this work. I am very grateful to my mother, Mrs. S. Ajagbe, for her concern, love, encouragement and prayers. God bless you Mummy.

Lastly, I appreciate my children, David, James, John-V and Sarah-B and my loving husband, Dr. T.E. Nottidge for their unflinching love and support throughout the period of this work.

Finally, I acknowledge my maker, the giver of life and wisdom for seeing me through this programme. Thank you Jesus.

TABLE OF CONTENTS

CHAPTER ONE: INTRODUCTION

CHAPTER TWO: LITERATURE REVIEW

CHAPTER THREE: MATERIALS AND METHODS

CHAPTER FOUR: RESULTS AND DISCUSSION

CHAPTER FIVE: SUMMARY, CONCLUSION AND RECOMMENDATION

LIST OF TABLES

LIST OF PLATES

Page

LIST OF FIGURES

LIST OF ABBREVIATIONS

LBP – Low Back Pain

OM – Outcome Measure

EG – Experimental Group

CG – Control Group

ADL – Activities of Daily Living

NPRS – Numerical Pain Rating Scale

ICC – Intraclass Correlation Coefficient

PROM – Patient Report Outcome Measure

IKHOAM- Ibadan Knee Hip Osteoarthritis Outcome Measure

 ODI – Oswestry Disability Index

RMDQ – Roland-Morris Disability Questionnaire

LBOS – Low Back Outcome Score

QBPDS – Quebec Back Pain Disability Scale

MVAS – Million Visual Analogue Scale

ALBDS – Aberdeen Low Back Disability Scale

NASS LSO – North America Spine Society Lumbar Spine Outcome

LBPRS – Low Back Pain Rating Scale

WDI – Wadell Disability Index

CBSQ – Curtin Back Screening Questionnaire

ILBPDI – Istanbul Low Back Pain Disability Index

GFS – General Function Score

BACKILL – Back Illness and Disability Nine Item Scale

BQ – Bourne Questionnaire

DPQ – Dallas Pain Questionnaire

SPIM – Spinal Pain Independence Measure

FRI – Functional Rating Index

BPFS - Back Pain Functional Scale

DRI – Disability Rating Index

FOE – Frequency of Endorsement

CHAPTER ONE

INTRODUCTION

1.0   Introduction

1.1   Background

Low Back Pain (LBP) is now the leading cause of disability globally, ahead of 290 other conditions (Buchbinder et al, 2013).The latest Global Burden of Disease Study, published at the end of 2012, has highlighted the enormous global burden of LBP (Buchbinder et al, 2013). It is the most prevalent musculoskeletal condition that causes much individual suffering and use of health services (Djiken et al, 2008, Woolf and Pfleger, 2003). Low back pain accounts for a large number of days lost from work and significant economic losses. It is also costly to treat. The total cost of LBP in the United States of America was estimated between 100 and 200 billion dollars annually, two thirds of which were due to decreased wages and productivity (Katz, 2006). Epidemiological studies from North America, Great Britain and other parts of Europe indicate a point prevalence of 12% to 33%, 12 month prevalence rates of 22% to 65% and life time prevalence of 11% to 84% (Walker, 2000). Men and women report about the same prevalence of back pain although some large surveys in UK showed a slightly higher prevalence in women (Walker, 2000).

Many studies have shown that LBP is also a common health problem in Nigeria. Nwuga (1993) found that 88% of a group of Nigerians 60 years and older have had at least one episode of back pain in their lifetime. Omokhodion and Sanya (2003) found the 12 month prevalence of LBP among office workers in Ibadan, South West Nigeria to be 38% and point prevalence to be 20%. Adegoke et al (2008) reported LBP as the most common work –related musculoskeletal disorder

1

among Nigerian physiotherapists with a 12-month prevalence of 69.8%. The report of a study on a small group of Nigerian automobile technicians showed that about 54% had low back pain (Omokhodion, 1996). Akinpelu et al (2011) reported the 12-month prevalence of LBP to be 47% for Igbo-Ora community dwellers.

Outcomes assessment provides a systematic method of measuring treatment effectiveness and efficiency. Therefore, the earlier methods of assessing / measuring outcomes such as laboratory tests, radiographs, physiologic tests, muscle strength, range of motion and spinal mobility have no direct clinical importance to patients and payers and they are not enough to reflect the functional status of the patients (Duruoz, 2013). An individual's functional status has become increasingly important over past decades, reflecting the growing expectation by society of a life without disability or handicap. Moreover, objective findings from these earlier methods of measuring treatment outcomes are not psychometrically sound and are not well correlated with items that concern patients or society. The health insurance companies generally do not care about a patient increased range of motion, instead they want him to perform daily activities and participate in social activities without pain or restriction (Feise et al, 2001).

Self-reported questionnaires or scales (often referred to as outcome measures) of pain, physical performance, functional disability, health status, functional independence and quality of life allow one to evaluate patients before and after a given treatment, and they can be used to detect short term or long term clinical changes of symptoms or disabilities (Grotle et al, 2004). Several scales and questionnaires, otherwise called paper and pencil clinical instruments, have been developed and are available in medical literature. Some of the scales, such as Short-Form 36

Health Survey (SF-36), Nottingham Health Profile (NHP) (Martini and McDowell, 1976), WHO- Quality of Life- Bref (WHO 1996) are generic and others, such as Ibadan Knee Hip Osteoarthritis Outcome Measure (IKHOAM) (Akinpelu et al, 2007), Shoulder Pain and Disability Index (SPADI) (Roach et al, 1995) are disease specific. Low back pain-specific scales include Oswestry Disability Index (ODI) (Fairbank et al, 1980), Roland –Morris Disability Questionnaire (RMDQ) (Roland and Morris, 1982), Quebec Back Pain Disability Scale (QBPDS) (Kopec et al, 1995) and Istanbul Low Back Pain Disability Index (ILBPDI) (Duruoz et al, 2013).

In practice, clinicians and researchers are encouraged to make use of existing outcome measures or scales as often as possible instead of developing new ones. This is because a lot of effort has gone into the development of any standardized outcome measure, no matter how simple it may appear (Streiner and Norman, 2008). Only a few standardized outcome measures originally developed for the Nigerian environment, the Ibadan Knee/Hip Outcome Assessment Measure (IKHOAM) (Akinpelu et al, 2007) and the Stroke Levity Scale (Owolabi and Platz, 2008) are available in medical literature. Most existing standardized outcome measures available were developed in Europe and North America. Outcome measures often reflect the culture and environment of the people for who they were originally developed. Consequently, researchers and clinicians have been provided with guidelines for cross-culturally adapting paper and pencil scales for use in a new environment (Beaton et al, 2000).

A few scales, such as the WHO Quality of Life-Bref, the Stroke Specific Quality of Life Scale and the Visual Analogue Scale, have been found amenable to cross-cultural adaptation or translation into Nigerian indigenous languages (Akinpelu et al, 2006; 2012; Akinpelu and Olowe, 2002; Odole and Akinpelu, 2009). However, the need to develop a new scale for a specific condition may arise when existing standardised scales are found to be deficient in many activities that are considered to be of cultural importance and concern to people in an environment. For patients with low back pain in Nigeria, such activities may include females kneeling down and men prostrating/squatting to greet elders, drawing water from the well, sweeping the floor with short broom and squatting to use the pit or Asiatic toilet. This study was therefore embarked upon to develop a Nigerian culture and environment-friendly scale for assessing treatment outcomes in low back pain and to provide evidence of its psychometric properties.

1.2 Statement of the Problem

Development and integration of standardized outcome measures into clinical practice remains an ongoing research effort all over the world. These scales are important to provision of evidence of effectiveness of healthcare in a way that is meaningful to all stakeholders and to researching evidence-based clinical practice. Standardized Outcome measures originally developed for Nigerian culture and environment are very few (Akinpelu et al, 2007, Owolabi and Platz, 2008). In addition, only few existing scales have been translated and cross-culturally adapted to the Nigerian environment (Akinpelu et al, 2006; 2012). This is probably due the fact that the factors, such as demands of the third party payers and changes in health policies that challenged

4

healthcare professionals in other countries to develop and integrate outcome measures into clinical practice are largely absent in the Nigerian clinical setting. However, Akinpelu et al (2007) observed that the National (Nigerian) Health Insurance Scheme might soon start placing demands on healthcare providers in Nigeria to integrate standardized outcome measures into their practice. In preparation to meeting these possible challenges, it is pertinent for Nigerian researchers to direct effort at developing more Nigerian culture and environment-friendly scales and at cross-culturally adapting others to the Nigerian environment. Such effort would enhance assessment of subjective states which are not amenable to physical measuring instruments in common conditions such as low back pain. This study was intended to contribute to this important social need. The questions arising from the above were: would the clinical instrument for measuring treatment outcomes in low back pain demonstrate adequate degree of:

1. Construct validity?
2. Reliability?
3. Internal consistency?
4. Responsiveness?

1.3 Aim of Study

The aim of this study was to develop a Nigerian culture and environment-friendly low back pain disability scale and to test its psychometric properties.

1.4 Specific Objectives of the Study

The specific objective of this study were:

1. To develop a low back pain disability scale that is Nigerian culture and environment-friendly.

2. To determine the construct validity of the LBP disability scale.

3. To determine the internal consistency of the LBP disability scale.

4. To determine the responsiveness of the LBP disability scale.

5. To determine the reliability of the LBP disability scale.

1.5 Hypotheses

1.5.1 Major Hypothesis

The LBP disability scale that would be developed would not be valid, reliable, responsive or internally consistent.

1.5.2 Sub Hypotheses

1. There would be no significant difference between the scores obtained on the LBP disability scale by the patients with LBP and the age and sex matched controls (construct validity).

2. There would be no significant correlation between the pain intensity score and score obtained on the LBP disability scale for the patients with LBP at baseline (construct [divergent] validity).

3. There would be no significant correlation between the pain intensity score and score obtained on the LBP disability scale for the patients with LBP after treatment (construct [divergent] validity).

6

4. There would be no significant correlation between scores obtained on the LBP disability scale by the patients with LBP on two different occasions   (test –retest reliability).

5. The item-item correlation of the LBP disability scale would not be significant (internal consistency).

6. There would be no significant difference between pain intensity score of the patients with LBP before and after a 5-week physiotherapy programme.

7. There would be no significant difference between the scores obtained on the LBP disability scale by patients with LBP in this study before and after a 5- week physiotherapy programme (responsiveness).

8. There would be no significant correlation between the changes in scores obtained on the LBP disability scale and pain intensity scores by patients with LBP in this study after a 5- week of physiotherapy programme (responsiveness).

1.6   Delimitation

This study was delimited to:

1. Patients diagnosed to have chronic non-specific LBP with or without radiculopathy; without co-morbid conditions, such as symptomatic osteoarthritis of lower limb joints, flaccid or spastic paralysis, spinal tuberculosis, tumour or spinal pathology that may further limit functions and without pregnancy (female patients only)

7

2. Apparently healthy age and sex- matched controls (without LBP) who could speak and read English.

1.7   Limitation

It was difficult for the researcher to be physically present in all the hospitals to collect data. Though the research assistants were trained, there may be some variation in the way the treatment of the patients were carried out. This was reduced by periodic training and retraining of the research assistants.

1.8   Significance of Study

1. When published in a peer reviewed   journal, the outcome of this study would make available a psychometrically sound, Nigerian culture and environment-friendly standardized scale for measuring treatment outcomes in low back pain, a common health problem in Nigeria, as it is elsewhere in the world. This would hopefully enhance objective assessment of functional disability, promote goal setting and facilitate researching evidence- based practice among patients with low back pain in Nigerian clinical setting.

2. The outcome of this study might also help Nigerian Health care providers involved in the care of patients with LBP meet the anticipated demand of Health Insurance companies   in the future as it will provide a scale for measuring disability and outcomes of treatment in LBP.

8

3. The outcome of this study might encourage other Nigerian researchers to implement clinical instrument development studies which will be more culturally relevant.

1.9 Definition of Terms

1. Item: Item is used to refer to individual question in any health measurement. It replaces the more obvious term "question" since not all categories are phrased as a question (McDowell and Newell, 1996).

2. Standardized Outcome Measure: This is a published measurement tool designed for a specific purpose for a given population, with detailed information on administration, scoring, interpretation and psychometric properties of the tool (McKay Lyons, 1998).

3. Chronic non- specific Low Back Pain is defined as pain that is of more than 3 months duration, made worse by activity and relieved by rest.

CHAPTER TWO

LITERATURE REVIEW

2.1   Low Back Pain

Low back pain (LBP) is defined as "activity-limiting low back pain (below the twelfth rib and above the greater trochanter of the femur) (+/- pain referred into one or both lower limbs) that lasts for at least a day (Hoy et al, 2012). Low back pain is a common health problem and a leading cause of disability globally (Buchbinder et al, 2013). It has been described as the most prevalent and the most costly musculoskeletal problem in the world and the first reason for orthopaedic consultation (Melon et al, 2008). It has been observed to be one of the most common complaints of people in the industrialized world, causing major health impairments, limitations in social life and work performance as well as increase in health care cost. (Manchikanti et al, 2014). Low back pain is a common reason for lost work days, estimated at 149 million days of work lost per year (Ricci et al, 2006; Stewart et al, 2003; Guo et al, 1999). The total cost of LBP in the United State of America exceeds 100 billion dollars per year (Katz, 2006). Low back pain can be acute, (0- 6weeks),sub acute (6 to 12weeks) and chronic (more than 3 months) (Solomon et al, 2010).

2.2   Epidemiology of Low Back Pain

Epidemiological studies from North America, Great Britain and other parts of Europe indicate a point prevalence of 12% to 33%, 12-month prevalence rates of 22% to 65% and lifetime prevalence of 11% to 84% (Walker, 2000). Men and women report about the same prevalence of back pain with a slightly higher prevalence in women (Streenstra et al, 2005, Vingard, 2005).

The prevalence of chronic impairing LBP has risen significantly in North Carolina with continuing high levels of disability and health care use (Freburger et al, 2009). Louw et al (2007) in a review of prevalence of LBP in Africa reported a mean LBP point prevalence of 12% among the adolescents and 32% among the adults. The average one year prevalence of LBP among adolescents was 33% and among adults was 50%. The average lifetime prevalence of LBP among the adolescents was 36% and among adults was 62%.

Many studies have provided the evidence that LBP is common in Nigeria. Nwuga (1993) found that 88% of a small group of Nigerians 60 years and older have had at least one episode of back pain in their lifetime. Omokhodion and Sanya (2003) found the 12-month prevalence of LBP among office workers in Ibadan, Southwest Nigeria to be 38% and point prevalence to be 20%. Adegoke et al (2008) reported LBP as the most common work–related musculoskeletal disorder among Nigerian physiotherapists with a 12-month prevalence of 69.8%. The report of a small study of Nigerian automobile technicians suggested that about 54% had low back pain (Omokhodion and Osungbade, 1996).

More than 80% of the United State population will experience an episode of LBP at some time during their lives (Rubin, 2007). In most cases, the clinical course is benign, with 95% of those affected recovering within a few months of onset (Carey et al, 1995). Some, however, will not recover and will develop chronic LBP (i.e. pain that lasts for 3 months or longer). Recurrence of LBP is common. It is reported that recurrence rate for working populations ranged   from 20%-44% within one year to 85% for lifetime recurrence (Van Tulder et al, 2002).

2.2.1 Aetiology of Low Back Pain

Low back pain can be classified as acute, sub-acute and chronic (Solomon et al, 2010). In addition, the aetiology of low back pain can be classified into specific and non-specific and also into vertebral and non-vertebral causes. Some of the specific causes of low back pain in adults include degenerative conditions like lumbar spondylosis and spondylolisthesis; fractures – both those due to major trauma and insufficiency fractures; infective – spinal tuberculosis, spinal bacterial osteomyelitis, discitis; metabolic – osteoporosis, osteomalacia; neoplasms – benign and malignant (primary and secondary) (Solomon et al, 2010).

2.2.2 Predisposing/Aggravating Factors for LBP

Risk Factors for Low Back Pain: Many risk factors have been associated with LBP. These include Physical, individual, habits and psychosocial factors (Manchikanti and Fellows, 2006).

Physical risk factors: These include occupational risk factors; heavy physical work; static work posture, bending, twisting and lifting; and vibration (Manchikanti and Fellows, 2006). Omokhodion and Sanya (2003) reported that sitting for more than 3 hours was associated with increased severity of low back pain. Whole-body vibration was reported to have a negative effect on the spine in a review of 53 articles on whole body vibration and LBP by Lings and Leboeuf-Yde (1998). They also concluded that the epidemiological studies showed that drivers have an increased prevalence of LBP, postulating that long – term exposure to whole – body vibration can contribute to back disorders.

Individual risk factors: These include genetic predisposition, age, gender, weight, kyphosis, scoliosis and leg length discrepancy (Manchikanti and Fellows, 2006). LBP is associated with

the aging process and some of the related factors include the deterioration of the intervertebral discs from the age of 30 years (Foster, 2001) and the increased incidence of osteoporosis and osteoarthritis in older adults, which for osteoporosis is especially marked in women from the menopause (Manchikanti and Fellows, 2006).

Habits: - smoking, exercise posture, physical activity and alcohol consumption are habits that have been associated with LBP (Huan et al, 2014). Patients with LBP are encouraged to stop smoking, since nicotine levels consistent with those seen in a person smoking 30 cigarettes a day have been shown to result in necrosis of the nucleus pulposus, as well as hypertrophy, cracking and detachment of the annulus (Uematsu et al., 2001). Holm and Nachemson (1988) noted that cigarette smoking also affects metabolite and solute exchange within the disc. Sanya and Omokhodion (2003) found a significant relationship between smoking and low back pain.

Psychosocial Factors: The interplay of psychological and social factors has been found to have great impact on the three phases of LBP namely: onset of pain, pain perception and chronic pain (Yeomans, 2000, Foster, 2001). It has been shown in many people that pre-existing depression and low coping mechanisms may likely predict the onset of pain and later chronic pain (Foster, 2001). The way a patient perceives and copes with his pain at the beginning of an acute attack may actually condition the patient to either recover or develop a chronic condition (Deyo and Weinstein, 2001).

Others risk factors for LBP include genetic factors, pregnancy, infection and low socioeconomic status.

2.3 Management of Low Back Pain

The treatment of low back pain depends on the cause. It has been noted that with minimal intervention, most patients experience full relief of symptoms within 4 to 6 weeks. However , 5% to 10% of patients develop chronic symptoms which require the application of standard treatment protocols, and its management then becomes more focused on pain coping strategies and function activities (Patrick et al, 2014).The goals of treatment are to educate patients, decrease pain, improve function and minimize side effects associated with chosen treatment modalities (Becker and Stumbo,2013). LBP can be managed by medications, bed rest, psychological means, surgery and physiotherapy, in various combinations, in order to address both the subjective and objective aspects (Cailliet, 2003).

Medications: A wide variety of different classes of medications have been used in the management of low back pain. The main goal of therapy is to use the lowest effective dose for the shortest period of time. Non - Steroidal Anti-Inflammatory drugs (NSAIDS) have been used in back pain (Casazza 2012; Chou et al, 2007). A large Cochrane review (Roelofs et al, 2008) supported the use of NSAIDS as the first – line management in the treatment of acute and chronic low back pain without sciatica. This review which included 65 randomized controlled studies found statistically significant results in favour of NSAIDS over placebo for improved functional status, number of patients recovered and decrease in pain intensity from baseline. This review also found moderate evidence that NSAIDS are equally effective as paracetamol/acetaminophen for pain relief and global improvement but NSAIDS were associated with an increased risk of side effects compared with paracetamol.

Tramadol was shown to be more effective than placebo in the management of low back pain (Casazza 2012; Chou et al, 2007; Deshpande et al, 2007). Opioids are typically not considered as a first – line management option in low back pain. They may be considered as a treatment option in patients with severe pain that is not responding to NSAIDS, acetaminophen or other conservative management options. Its use may also be considered in patients that have pain interfering with their sleep. (Carragee 2005; Deshpande et al, 2007). Other medications that have been used in low back pain include anticonvulsants, benzodiazepines. (Casazza 2012; Chou et al, 2007).

Topical analgesics have the advantage of avoiding systemic toxicities but have the limitation of providing treatment to a localized area. Side effects include skin irritation or allergic reactions. They can be used alone or in conjunction with other therapies, an example is Capsaicin (Dalpiazet al.2004; Gagnier et al, 2007). Muscle relaxants when combined with NSAIDS therapy have been shown to be effective for short – term relief of low back pain (Kinkade 2007; Last and Hubbert, 2009). Combinations of the NSAIDS and the opioids have been found to be effective medications and are commonly used by experts in management of LBP. The means of administration is usually oral, supplemented by the intramuscular or intravenous route. The epidural administration of steroids, with or without local anaesthetics, is widely used in the treatment of chronic low back pain (Cohen et al, 2013).

Bed rest: Studies have shown that activity modification is preferred for the treatment of acute low back pain rather than bed rest and immobilization. Bed rest may be recommended for 1 to 2

days if there is severe pain but patients should be educated that longer periods of bed rest can be associated with delayed recovery, joint stiffness and muscle weakness. Reassurance and education will help the patients to know that it is safe to get out of bed and perform activities as tolerated (Deyo and Weinstein, 2011; Shen et al, 2006).

Psychological aspects of pain: This has been recognized as important in the perception of the initial pain episode, its response to treatments and its progression to chronicity (Golob and Wipf, 2014). Cognitive behavioural therapy is advocated as a means of addressing this aspect (Waddell, 1987).

Surgery: Spinal surgery has few indications in the treatment of low back pain and these include persistent pain despite conservative management, and progressive neurological deficits (Becker and Stumbo, 2013). Sphincteric dysfunction in association with acute onset low back pain suggests spinal cord compression and is a surgical emergency (Solomon et al, 2010). The occurrence of post-surgery recurrence of pain is a nuisance and makes for careful case selection.

Physiotherapy: There are many physical methods of treating low back pain including cryotherapy (Gammon and Starr, 1993), heat therapy, traction, transcutaneous electrical nerve stimulation (TENS), interferential current, electrical muscle stimulation, therapeutic ultrasound, spinal manipulation and mobilization, massage, lumbar stabilization exercises, lumbar extensor strengthening exercise, McKenzie method, (Gay and Brault, 2008; Poitras and Brosseau, 2008; Bronfort et al, 2008; May and Donelson, 2008; Imamura et al, 2008; Standaert et al, 2008; Mayer et al, 2008; Brox et al, 2008) . There is no standard protocol for the management of low back

pain. Individualized regimens that include therapist's supervision, stretching and strengthening exercise tend to be associated with the best outcome. The McKenzie method, spine stabilization exercises and home exercise program have also been shown to be beneficial (Petering and Webb, 2011; Kinkade, 2007; Shen et al, 2006).

Heat therapy has been shown to be beneficial in reducing pain associated with acute low back pain. Additional pain relief and improved function are achieved when combined with exercise. Minimal evidence exists for the use of cold therapy in acute low back pain (Casazza, 2012; Chou et al, 2007; Kinkade 2007; Kettenmann et al, 2007; French et al, 2006).

There is no clear evidence for efficacy regarding the use of lumbar corsets in the management

of acute and chronic low back pain. Studies have shown a possible benefit if a lumbar corset is combined with another spinal support, such as a heat – mouldable plastic insert (Million et al, 1981).

Traction therapy: this refers to any method of separating the lumbar vertebrae with the primary force directed along the inferior-superior axis of the spine, in an attempt to treat chronic LBP. The subtypes of traction therapy can be classified based on the duration of the applied force and the direction of the force. It can be applied with the patient supine, prone, lateral decubitus, suspended upright or inverted. The mechanism of action is unclear but thought to revolve around altering the mechanobiology of the spinal motion segment or neural tissues (Gay and Brault, 2008). The ideal patient who is likely to benefit from traction therapy is not well defined, based on the available evidence. In addition, there is more evidence against than for the use of traction therapy for the treatment of chronic LBP (Clarke et al, 2007; Gay and Brault, 2008).

Transcutaneous Electrical Nerve Stimulation (TENS): this delivers an electrical current through superficial electrodes placed on the skin, disrupting the pain signal in surrounding nerves (Poitras and Brosseau, 2008). It may use high or low frequencies and appears to have immediate benefit in reducing chronic LBP but there is no clear evidence of short to long term benefit (Poitras and Brosseau, 2008).

Interferential current, electrical muscle stimulation and therapeutic ultrasound were not found to have a firm basis in the literature, for their use in the treatment of chronic LBP (Poitras and Brosseau, 2008).

Spinal manipulation: the application of high-velocity low amplitude manual thrusts to the spinal joints, slightly beyond the passive range of joint motion. When combined with strengthening exercises, spinal manipulation has been found to be equivalent to prescription of non-steroidal anti-inflammatory medication, in the short to long term (Bronfort et al, 2008).

Spinal mobilization: the application of manual force to the spinal joints within the passive range of joint motion that does not involve a thrust. Flexion-distraction spinal mobilization has been found to be superior to exercise in the short term and superior/similar in the long term (moderate evidence) (Poitras and Brosseau, 2008). Most of the evidence of efficacy and the low risk of serious adverse effects, supports the spinal mobilization and spinal manipulation for the treatment of chronic LBP, as they are at least as effective as other effective interventions (Poitras and Brosseau, 2008).

Massage: This involves soft tissue manipulation using the hands or a mechanical device. There is strong evidence that massage is effective for non-specific chronic LBP (Imamura et al, 2008). In addition, moderate evidence supports the use of massage for patients with chronic LBP in terms of improving symptoms and function (Imamura et al, 2008). These effects of massage are

improved if combined with exercise and education and if delivered by a licensed Physiotherapist; the benefits are long-lasting (at least a year after the end of sessions) (Imamura et al, 2008).

Core stabilization /segmental stabilization / lumbar stabilization exercises are aimed at maintaining stability in the lumbar spine and may be considered a useful tool in the management of patients with chronic non-specific LBP, based on moderate evidence for effectiveness in improving pain and function (Standaert et al, 2008).

The McKenzie school merits special mention, as a modality that combines manipulation, mobilization and exercise (McKenzie, 1981). There is moderate quality evidence to support the claim that exercise regimens administered after treatment, reduce both the number and rate of recurrences of mechanical low back pain (Choi, 2010).

2.4 Outcomes Assessment

Outcomes assessment is a systematic method of assessing the effectiveness of a treatment program (Resnik and Dobrykowski, 2005). The traditional method of carrying out this measurement is not consistent nor standardised and focuses more on variables that are relevant to the clinician, like muscle power, joint range of motion, spinal mobility etc. (McCormick et al, 2013). These are impairment variables and treatment directed at improving them may not necessarily improve function or reduce disability. However, in the last three decades emphasis has shifted to measurement of variables that are important and meaningful to the patients and payers, such as physical performance, functional disability, health status, functional independence and return to work all of which can be assessed using patient-reported outcome measures (PROMs).

Patient-reported questionnaires or scales (often referred to as outcome measures) allow one to evaluate patients before and after a given treatment, and they can be used to detect short term or long term clinical changes in functional disabilities, physical performance, health status and quality of life (Grotle et al, 2004). The assessment of the patient's perception guides the therapist during the course of treatment and can indicate the end point of therapy. These methods of measuring outcomes through patient self-evaluation are not without limitation, as the patient can exaggerate or underrate responses (Feise et al, 2001). Yet, a patient's self–evaluation may be more accurate than the clinical, biochemical or physiological indexes which the traditional method relied on (Feise et al, 2001). For example, an injured employee will regard the ability to return to work as a good outcome. This would also be the primary concern of the person's employer. In addition, payers of healthcare are concerned with cost effectiveness of the treatment modality and are better able to assess this, when the use of an outcome measure can objectively demonstrate an improvement in the patients' function over time (Resnik and Dobrykowski, 2005).

2.5 Outcome Measures

Outcome measures are now more often referred to as Patient Reported Outcome Measures (PROMs). They are now used to measure the health of the patient at different times and not just the outcome of treatment. In addition, they have been used in recent years to assess and compare the outcomes of treatment achieved by different healthcare providers (Black, 2013).

Advantages of considering patients views include:

- Achieving the aims of healthcare – reduce symptoms, minimize disability and improve quality of life; indices only the patient can assess.

- The direct involvement in their care is welcomed by patients and may have health benefits in itself.

- Patient's response rates are invariably better than that of the clinician – the patient completes one questionnaire, while the physician does so for several patients.

- Avoiding observer bias (Black, 2013).

The routine use of PROMs in England, has been motivated by Government policy for public comparisons of HCP performance. In Sweden and the US, the main driver has been the medical profession, with the aim of improving the clinical care of the individual patient.

Healthcare providers must be able to use outcome measures to direct quality improvement initiatives, assist in development of clinical guidelines and make decisions about individual patients (Fritz and Irrang, 2001). This trend is consistent with the final step of an evidence-based approach to practice, which is evaluation of one's clinical practice (Lewis &Latney, 2002). Although, the importance and value of outcome measurement have been recognised by many clinicians, the necessary training and experience to successfully integrate these measures into their practice is lacking (Resnik and Dobrykowski, 2005). Clinicians must be able to evaluate and choose appropriate outcome measures for their population of patients, make inferences about changes in measurement scores that occur during treatment and use this information for clinical decision making and quality improvement (Resnik and Dobrykowski, 2005). Healthcare providers in clinical settings are required to become more familiar with the use of outcome measures for their patients and need to consider using these tools in their daily practice. Providers that actively utilize, analyse, and interpret outcomes information will be able to assess

the severity of the disorder in order to formulate an appropriate treatment plan. The information obtained can also be followed over time to assess treatment efficacy. Ultimately, patients will improve functionally when effective and efficient types of services are identified and offered (McCormick et al, 2013; Lewis and Latney, 2002).

2.6 Classification of Outcome Measures

Outcome measures may be classified using three methods. These are descriptive classification, functional classification and methodological classification (Kirschner and Guyatt, 1985).

2.6.1 Descriptive classification

This method of classification focuses on the scope or the range of topics covered by the scales. Scales are thus classified as generic, disease specific, organ specific and broad syndrome scales. Generic scales may be used to assess treatment outcomes in a variety of conditions and they enable comparison across diseases. Examples of generic outcome measures are the visual analog scale (VAS) (Price, 1994) and the functional independent measure (FIM) (Granger, 1984). Disease specific scales measure treatment outcomes in specific conditions. Examples include the Quebec back pain disability scale (QBPDS) (Kopec et al, 1995) and the LequesneAlgo-functional Index of Knee (Lequesne, 1997). Organ specific scales focus on a particular organ or system, such as vision or hearing. Broad syndrome scales measure broader syndromes, such as health status, emotional well-being and quality of life. Examples of broad syndrome scales are the Sickness Impact Profile (Gilson et al, 1975), the World Health Organisation Quality of Life

Brief (WHOQOL-Brief) (WHO, 1996) and the Short Form-36 Health Survey (SF-36) (Ware and Sherbourne, 1992).

2.6.2   Functional classification: focuses on the specific purpose of the outcome measure and identifies the diagnostic, prognostic and evaluative indices of health measures. Some authors have classified the functional group into discriminative, predictive and evaluative (Kirschner and Guyatt, 1985). This fits broadly into the earlier scheme but provides a wider application in the purpose to which indices of health measures are put. Discriminative index provides a means of distinguishing between individuals or groups, when no external criterion or gold standard is available. They can be used in surveys to determine the burden of illness across different communities and can be said to be diagnostic as in the earlier categories. Examples include the Minnesota Multiphasic Personality Inventory.

Predictive indices classify individuals into specific categories, when a gold standard is already available, such that one can determine the presence of a condition or predict which individuals will likely have the outcome in future. Predictive indices usually offer the advantage of simplicity, lower risk, application earlier in the course of the disease than the gold standard.

Evaluative indices are most common in health rehabilitation and measure the magnitude of longitudinal change in an individual or group on the dimension of interest. These instruments are used to quantify the treatment benefits of interventions and can provide a baseline for determining disability, as in the Oswestry Disability Index.

2.6.3 Methodological classification

This system classifies outcome measures as self/patient reports (subjective measure) and clinician measured (objective measure). Questionnaires are usually patient reports, while test batteries are rated or measured by clinicians. Clinician-Reported Outcomes (CROs) include outcomes either observed by a provider or requiring interpretation. It includes scales completed by a health care provider using information about the patient (Willke et al, 2004). Clinician measured scales can be classified as bio-physiological (e.g. Harvard step test and triple-hop distance test) or observational (e.g. Modified Motor Assessment Scale (MMAS) (Carr and Shepard, 1989) and the Gross Motor Function Measure (GMFM) (Russell et al, 2000).

Patient-reported outcomes (PROs) are reports coming directly from patients about how they function or feel in relation to a health condition and its therapy, without interpretation of a patient's responses by a physician or anyone else (Valderas et al, 2008).The self/patient reported or subjective outcome measures are grouped into:

1. **General Health Status Outcome Measure:** These measures encompass quality of life issues. This group of outcome measures are designed for broad use in a variety of patient populations. They are best combined with specific outcome measures or some other measure with which to compare and contrast. Some of these measures include Short Form 36 survey (SF 36) (Goertz, 1994), Sickness Index profile(SIP) (Bergner et al, 1981), Quality of Well –Being scale (Kanplan and Anderson, 1982), Dartmouth COOP Health Charts (Goertz, 1994), the Health Status Questionnaire (HSQ), (Deyo et al, 1991).

2. **Pain Perception Outcome Measures:** These are used to track the pain perception of individuals experiencing pain, such that the patient matches numbers or words to pain intensity. The accuracy of the assessment is dependent on the efforts of the HCP and the person experiencing pain (Price, 1992). Examples include Visual Analogue Scale (Von Korff et al, 1992, 1993), Numerical Pain Scale (Wallenstein et al, 1980), sMcGill pain scale (Melzack, 1975, 1987).

3. **Condition –specific Outcome Measures:** This group of outcome measure has the ability to measure and track outcomes of specific conditions. They are specifically designed for clinical application to address the patient's specific complaints and are sensitive to change following treatment (McDowell and Newell, 1996). They measure lack of activity tolerance or the inability to perform activities of daily living (ADL) of specific conditions/ diseases. Examples include Oswestry Disability Index (Fairbanks et al, 1980), Ibadan Knee / Hip Osteoarthritis Outcome Measure(IKHOAM) (Akinpelu et al, 2007), Stroke Levity scale (Owolabi and Platz, 2008) and Shoulder Pain and Disability Index (SPADI) (Roach et al, 1995).

4. **Psychosocial Outcome Measures:** Psychosocial factors can influence pain perception, ability to adjust to pain and quality of life (Williams and Feuerstein, 2000). These include Beck's Depression Inventory (Beck, 1967), Modified Zung Depression Index (Zung 1965), Somatic Amplification Ratings Scale (SARS) (Korbon et al, 1989) and Distress and Risk Assessment method (DRAM) (Main et al, 1992).

5. **Disability Prediction Outcome Measures:** These measures can help predict those patients who may prove difficult to manage due to chronic pain, work dissatisfaction,

fear avoidance or combinations thereof. These include Vermont Disability prediction Questionnaire (Hazard et al, 1996), Work APGAR (Bigos et al, 1991), Fear Avoidance Belief Questionnaire (FABQ) (Wadell et al, 1993) and Functional Assessment Screen Questionnaire (FASQ) (Millard, 1989).

6. **Patient Satisfaction Outcome Measure:** These are frequently used as important outcome assessment approach especially by managed care companies when assessing quality assurance issues (Deyo, 1986). These measures yield important information about the quality of the health care service as perceived by the patient by assessing the acceptance of care,perception of the technical competence of the HCP,the setting where care was provided and the effectiveness of the HCP. (Ware and Davies, 1983; Coulter, 1994). Examples of this group of outcome measures include Client Experience survey (Coultier et al, 1994) and Chiropractic Satisfaction Questionnaire (Coulter et al, 1994).

2.7 Development of Measuring Scales and Questionnaires

Researchers and clinicians are encouraged to make use of existing scales as much as possible by adapting them to a new environment and culture, ensuring the steps recommended by Beaton et al (2000) are followed, instead of developing new ones. This is because a lot of effort has been put into the development of any scale no matter how simple it appears (Streiner and Norman, 2008). However, if there is justification for the need to develop a new scale, the seven steps involved are as presented below.

2.7.1 Devising the Items

The first step in developing a scale or an outcome measure is to generate items or questions that will be included in the scale. The sources of items are:

a. **Research findings**: this can be a useful source of items or subscale. For the purpose of scale construction research can be of two types: a literature review of studies that have been done in the area or new research carried out specifically for the purpose of generating items that will be included in the scale being developed. The aim of literature review is to identify what earlier developers deemed relevant to the theme. The items on the scale should be those that have been shown to be the characteristic of a group of people or which differentiates them from other people (Streiner and Norman, 2008).

b. **Focus group discussions:** this involves 6 -12 patients who suffer the disorder for which the scale is being developed. A moderator usually guides the discussion. The patients are to suggest the general themes, which can later be used by the research team to form the items. After this has been done, another focus groups (expert panel) can be used to discuss whether the items are written in terms that can be understood by the target population, clear, relevant, unambiguous and if all the important themes have been included (Streiner and Norman, 2008).

c. **Clinical Observation**: this is one of the most useful sources of generating items. Scales can be regarded as a systematic way of gathering clinical observation to ensure that all the observers are looking for the same thing or all the subjects are responding to the same items (Streiner and Norman, 2008).

d. Key Informant Interview: a small number of people are interviewed because of their distinct knowledge, which could be patients who have had the disorder and can clearly express what they experienced; or clinicians who have extensive experience with the patient and can explain it from their perspective. These interviews are done repeatedly until no new themes emerge.

2.7.2 Content Validation

This is done to ensure that the scale has enough items and adequately covers all the domains under investigation (Streiner and Norman, 2008). The content validity of the generated items is assessed by sending copies of these items with a checklist to experts, asking them to indicate the degree of relevance of each item to the main theme for content relevance and coverage. Content Validity Index (CVI) (which is the number of experts that rate an item as essential or important / total number of experts≤ 10) can be calculated and those items with CVI less than 0.7 should be deleted from the scale. In addition, some items may be eliminated based on the judgement of the experts.

2.7.3 Scaling Responses

A method by which responses will be obtained must be chosen. This will depend partly on the nature of questions being asked. The developer must consider the kinds of responses that may arise, whether categorical or continuous. The different level of measurements must also be put into consideration (nominal, ordinal, interval and ratio).

Many of the variables measured in health care research are continuous rather than categorical, methods must be devised to quantify these judgments. There are three broad categories of quantifying these judgements.

a) **Direct estimation methods:** This is a straightforward method in which subjects are required to indicate their response by a mark on a line or a check in a box. They are designed to elicit from the subject a direct quantitative estimate of the magnitude of an attribute (Streiner and Norman, 2008). Examples include the visual analog scale (a line of fixed length usually 100mm with anchors like 'no pain' and 'pain as bad as it could be' at the extreme ends and no words describing intermediate positions),adjectival scales (use descriptors along a continuum, rather than just labelling the end points, as in visual analog scale, Likert scales and Face scales (for children and those with any type of cognitive disorder). This method is the most commonly used in health measuring scales. The disadvantage of these methods is that data generated from them are usually at ordinal measurement level and this limits the statistical method that can be used in analyzing them.

b) **Comparative methods:** Subjects are required to choose among a series of alternatives that have been previously calibrated by a separate criterion group. The common examples are Thurstone's method of equal appearing intervals, Guttman scaling and the paired-comparison technique (Streiner and Norman, 2008). These methods are intended to correct some of the flaws of the direct estimation methods, but they require training and are time consuming.

c) **Econometric methods:** have roots in economics and have become increasingly popular in the medical literature in applications ranging from clinical trials to decision analysis.

Examples include the Von Neumann-Morgenstern standard gamble and the Time trade off technique (Streiner and Norman, 2008). In time-trade off technique, an example could be a 30 year old man with a chronic disease who has about 40 years of life left is asked how many of the 40 years of life would he be willing to trade off if he could have perfect health.(Russell et al, 2000) The process of restoring him to perfect health may involve a high risk procedure.

2.7.4 Selecting the Items

Selection of the items on a scale under development is done by pretesting the scale on the target population. There are many reasons for pretesting the item on a scale. The items are pretested for comprehensibility, to determine the frequency of endorsement; discrimination ability of each item and homogeneity of the scale.

(i) Comprehensibility Testing: This pretesting for comprehensibility is usually done on a small group of subjects (< 40). The usual rule of thumb is that the scale should not require reading skills beyond that of a 12 year old (Streiner and Norman, 2008).The reasons for this pretesting are to eliminate or re-write items that are not comprehensible to the target population; ambiguous items (have more than one possible meaning); double barrelleditems (ask more than a single question) and items that contain jargon terms or medical terms (Streiner and Norman, 2008).

(ii) Frequency of endorsement: This is the proportion of people who give each response alternative to an item. Items where one alternative has very high or very low endorsement rate are eliminated. Only items with endorsement rates between 0.2 and

0.8 are used (Streiner and Norman, 2008). It is usually done on at least 50 subjects from the target population.

(iii) Discrimination ability of an item demonstrates if a person who has a high total score is more likely to have endorsed the item and if the item discriminates between those who score high (and supposedly have more of the trait) and those who score low (Streiner and Norman, 2008). It differs from endorsement frequency in that it looks at the item in relation to all the other items on the scale, not in isolation (Streiner and Norman, 2008).

(iv) Homogeneity of the scale: occurs when all the items on it are tapping different aspects of the same attribute and not different parts of different traits (Streiner and Norman, 2008). The items should be moderately correlated with each other and each should correlate with the total scale score. These two factors form the basis of the various tests of homogeneity or internal consistency of the scale.

2.7.5 Minimizing Biases in Responding

It is assumed that a respondent will answer honestly to the items on a scale but a number of factors may influence a response making it a less than accurate reflection of reality. Therefore a developer must think of methods of minimizing these biases which may arise from different sources. For instance, the presence of words that have more than one possible meaning in a scale e.g. 'stool' which is intended to mean faeces may be interpreted by a respondent to mean a type of furniture.

Biases in responses can also arise as a result of social desirability (when a person does not deliberately attempt to deceive others but is concerned with giving an expected response), faking good (a deliberate attempt to create a positive false impression), yea – saying acquiescence bias (the tendency to give positive responses to questions) etc. Different methods are used to reduce biases in responding. For example, the aspects of a variable that items tap may be deliberately hidden from respondents, thereby reducing the scale's face validity in order to minimize yea-saying bias (Streiner and Norman, 2008).

2.7.6 Other considerations

The method by which the scale would be administered (self-report or clinician measured) and how scores obtained will be computed and interpreted are important decisions that need to be taken in developing a scale.

2.7.7 Testing Scales or Questionnaires for Psychometric Properties

For an outcome measure to be effectively used in the health sector, the evidence of its psychometric properties must be investigated and published in a peer reviewed journal. Use of outcome measures that are not psychometrically sound will lead to misinterpretation of patient's status and incorrect conclusions regarding any clinical hypotheses under investigation. The psychometric properties include validity, reliability, internal consistency and responsiveness.

Validity: This is generally described as the degree to which a study accurately reflects or assesses the specific concept that the researcher is attempting to measure (Terwee et al, 2007). Nunnally and Bernstein (1994) were of the opinion that the term validity denotes the scientific utility of a measuring instrument, broadly stated in terms of how well it measures what it purports to measure. Validity usually is a matter of degree rather than an all or none property, and validation is an unending process. Validation is the process by which a test developer or test user collects evidence to support the types of inferences that are to be drawn from test scores.

Content Validity: examines the extent to which the concepts of interest are comprehensively represented by the items in the questionnaire. A positive rating is given for content validity if a clear description is provided of the measurement tool, the target population, the concepts that are being measured and the item selection (Mokkink et al, 2010).

Criterion Validity:refers to the extent to which scores on a particular instrument relate to a gold standard. A positive rating is given for criterion validity if convincing arguments are presented that the used standard really is "gold" and if the correlation with the gold standard is at least 0.70 (Mokkink et al, 2010). There are two types of criterion related validity: concurrent criterion validity and predictive criterion validity. In concurrent validity, the new scale and the criterion measure are administered during the same interview or within a short time of each other. In predictive validation, the criterion will not be available until sometime in future. This form of criterion validation is majorly used in college admission tests where the ultimate outcome is the person's performance on graduation four years later or in diagnostic tests where we must await the outcome of an autopsy or the further progression of the disease before we can confirm or disconfirm our predictions (Streiner and Norman, 2008).

Construct Validity:refers to the extent to which scores on a particular instrument relate to other measures, in a manner that is consistent with theoretically derived hypotheses concerning the concepts that are being measured. Construct validity should be assessed by testing predefined hypotheses (e.g. about expected correlations between measures or expected differences in scores between known groups). A positive rating is given when at least 75% of the results are in correspondence with these hypotheses, in (sub) groups of at least 50 patients (Terwee et al, 2007). Construct validity can be further classified as Convergent and Divergent validity.

Convergent validity is demonstrated when scores on the outcome measure being tested are highly correlated to scores on another outcome measure that measures similar or related concepts.

Divergent validity is demonstrated when scores on the measure being tested are not correlated to scores on a measure meant to measure a very different construct.

Reliability: This is a fundamental way to reflect the amount of error, both random and systematic, inherent in any measurement (Streiner and Norman, 2008). The degree to which repeated measurements in stable persons provide similar answers. Reliability concerns the degree to which patients can be distinguished from each other despite measurement error. Test-retest reliability involves instrument self-completion on two occasions separated by a suitable time–period and assuming no change in the underlying health state. It measures the temporal stability of the score (Terwee et al, 2007).

Intra-rater reliability testing is the process by which a measurement tool or method can be shown to give similar results when used by same raters at different times for the same group of subjects while inter-rater reliability is the extent of agreement of two measures by two examiners independent assessment of the same subect (Post et al, 2011, Kurande et al, 2013).

The intraclass correlation coefficient (ICC) is the most suitable and most frequently used method for test-retest for continuous data. The ICC is the variation in the population (inter-individual variation) divided by the total variation, expressed as a ratio between 0 and 1(Terwee et al, 2007). The time period between the repeated administrations should be long enough to prevent recall, though short enough to ensure that clinical change has not occurred.The appropriateness of the time period is not specific, but only require that the time period is described and justified. A minimum of 0.70 is recommended as a standard for reliability.A positive rating is given for reliability when the ICC or weighted Kappa is at least 0.70 in a sample size of at least 50 patients.

Internal Consistency: This is used to assess the homogeneity of the scale. It is an important aspect of the reliability of scales. It is the degree to which answers to individual items in a questionnaire are correlated with the total score or with each of the other items. It suggests how far each item contributes to the overall theme being measured. In internal consistency, it is expected that scores on each item will correlate with scores on all other(Streiner and Norman, 2008).

Responsiveness: The ability of a questionnaire to detect clinically important changes over time, even if these changes are small. lt is a measure of longitudinal validity. The instrument should be able to distinguish clinically important change from measurement error (Terwee et al, 2007, Mokkink et al, 2010).In other words, an outcome measurement with good responsiveness will accurately measure the amount of change in a condition when such change truly occurs. There are two types of responsiveness – internal and external. Internal responsiveness is the ability of an outcome measure to change during a pre-specified time frame. It is determined by administering a measure before and after a treatment of known efficacy. The measure is said to be responsive when the scores before and after the course of treatment differ significantly.

External responsiveness indicates the extent to which changes in a measure relate to changes in another measure, which assesses a different aspect of health. The outcome measure is said to be responsive when there is a significant correlation between the changes in score of the outcome measure being assessed and the other measure (Husted et al, 2000). A number of strategies have been suggested for quantifying responsiveness of evaluative questionnaires,

but there is currently no consensus on the most appropriate method (Davidson and Keating,2002, Beaton et al, 2002). Internal responsiveness can be determined by calculating the effect size (ES) and standardized response mean (SRM). Higher scores are preferred (Terwee et al, 2007). External responsiveness can be calculated with the perceived recovery scale as the external criterion of change. It can be evaluated using the receiver operating characteristic (ROC) curve, which is constructed by calculating the sensitivity (true positive rate) and specificity (true negative rate) of the cut-off point for each of the possible score values. Scores of at least 0.70 are adequate (Terwee et al, 2007).An outcome measure must possess the attribute of responsiveness in addition to validity and reliability.

Providing scientific evidence of psychometric properties of outcome measures: Health measuring instruments with scientific evidence of psychometric properties help register evidence-based practice (Mckay Lyons, 1998). For a measure to be effectively used in the health sector, it has to be standardized with their psychometric properties proven through scientific enquiry. The introduction of new measurement in the clinical setting should be preceded by formal investigation into the meaningfulness of the test or measure. An HCP would do well to expect to see evidence of scientific investigation before engaging in the use of a measuring method and the society is expecting to see that evidence. According to Mckay – Lyons (1998), a standardized outcome measure is a published measurement tool designed for a specific purpose for a given population, with detailed information on administration, scoring, interpretation and psychometric properties of the tool.

2.8 Low Back Pain Specific Scales

The ultimate goals of intervention in low back pain are to relieve pain, restore function and improve the patient's quality of life. Traditional clinician based assessments typically fall short of addressing these important outcomes because they may not describe patients' perception of their state of health (McCormick et al, 2013). A wide variety of patient – reported outcome measure /scales to obtain quantitative data regarding pain, function and general health quality in patients with LBP have been described in the past decades. Each of them evaluates low back performance using both objective and -subjective criteria.

Nineteen low back pain scales have been identified from literature review using OVID, Medline and Pubmed databases. These are:

37

1. Oswestry Disability Index (ODI) (Fairbank et al,1980)

2. Roland-Morris Disability Questionnaire (RMDQ) (Roland and Morris, 1983)

3. Low Back Outcome Score (LBOS) (Greenough and Fraser,1992)

4. Quebec Back Pain Disability Scale (QBPDS) (Kopec et al,1995)

5. Million Visual Analogue Scale (MVAS) (Million et al, 1981)

6. Aberdeen Low Back Disability Scale (ALBDS) (Ruta et al, 1994)

7. The North American Spine Society Lumbar Spine Outcome (NASSLSO) (Daltroy et al, 1996)

8. Low Back Pain Rating Scale (LBPRS) (Manniche et al,1994)

9. Wadell Disability Index (WDI) (Wadell and Main ,1984)

10. Curtin Back Screening Questionnaire (CBSQ) (Harper et al, 1995)

11. Istanbul Low Back Pain Disability Index (ILBPDI) (Duruoz et al, 2013)

12. General Function Score (GFS) (Hagg et al,2001)

13. Back Illness and Disability Scale (BACKILL) (Tesio et al, 1997)

14. Bournemouth Questionnaire (BQ) (Bolton et al, 1999)

15. Dalla Pain Questionnaire (DPQ) (Lawlis et al, 1989)

16. Spinal Pain Independence Measure (SPIM) (Itzkovich et al,2001)

17. Functional Rating Index (FRI) (Fense and Menkie ,2001)

18. Back Pain Functional Scale (BPFS) (Stratford et al, 2000)

19. Disability Rating Index (DRI) (Salen et al, 1994)

The Oswestry Disability Index (ODI) (Fairbank et al, 1980), is a ten item questionnaire that was initiated in 1976 by interviewing a large number of patients with chronic low back pain attending a Specialist clinic by an Orthopaedic surgeon, Occupational Therapist and a Physiotherapist. It was designed as a measure for both assessment and outcome. Version 1.0 was validated in 1980 using a sample of 25 patients with acute low back pain. A larger validation was published in 1994 by Stratford on a population with musculoskeletal LBP (Stratford et al, 1994). The ODI was further developed and validated and is now available in version 2.0 (Fairbank, 1995). A modified ODI is used in the North American Spine Society (NASS) questionnaire (Daltroy et al, 1996). Version 2.0 has been recommended for use and no permission is required for its use (Roland and Fairbank, 2000). The items are on pain, personal care, lifting, walking, standing, sitting, sex life, social life, sleeping and travelling. Item on ability to work is not included. It can be completed in about 5 minutes and scored in less than one minute. The total score is obtained by summing up the scores of all sections, giving a maximum of 50 points. The final score is expressed as a percentage (total score / (5 x number of questions answered) x 100%).

ODI shows good construct validity because it is consistent with some and was used as the standard of comparison for other outcome measures assessing LBP. It correlates with other outcome measures aiming at measuring disability due to LBP (Davidson and Keating, 2002, Fairbank and Pynsent, 2000, Roland and Fairbank, 2000,Fritz and Irrgang, 2001). Validation of the Low back outcome score, Aberdeen score and the Curtin scale was performed using ODI. Test- retest reliability of ODI has been tested (Fritz and Irrgang, 2001), it has an acceptable internal consistency between 0.76 and 0.87 (Kopec et al, 1995, Fisher and Johnson, 1997) and evidence of responsiveness (Davidson and Keating, 2002, Grotle et al, 2004). ODQ has been

compared with RDQ, LBOS, QBPDS (Davidson and Keating, 2002). ODI is a valid, reliable and responsive condition specific assessment tool that has withstood the test of time and scrutiny (Vianin, 2008). The authors of ODI recommend the use of ODI in patients who are likely to have persistent severe disability. ODI has been validated in seventeen languages (English (Fairbank et al, 1980), German (Basler et al, 1997), Finnish (Gronblad et al, 1994), French (Dropsy and Marty, 1994), Iranian (Mousavi et al, 2006), Italian, Greek (Boscainos et al, 2003), Portuguese, Norwegian (Grotle et al, 2003), Danish, Korean, Japanese, Chinese, Thai, Persian, Turkish, and Arabic)

The Roland-Morris Disability Questionnaire (RMDQ)(Roland and Morris, 1983) was derived from the Sickness Impact Profile, of which 24 out of 136 items were selected. It was designed as a self – reporting measure for both assessment and outcome. The 24 items are on pain, walking, bending over, sitting, lying down, dressing, sleeping, lifting, work, stairs, housework and resting, but no question on sex and social life are asked. The questions on RMDQ are straight forward and focus consistently on disabilities related to the back and the answers are dichotomous: yes / no. This might cause subtle changes in the functional abilities of patient to remain undetected. The RMDQ score can be obtained by adding up the number of items checked. The final score ranges from 0 (no disability) to 24 (severe disability). It can be completed in a maximum of 5 minutes and an un-weighted score can be calculated in less than one minute. The original RMDQ also contains a six- point pain rating scale in the form of a pain thermometer (Roland and Morris, 1983). However the authors prefer to use the pain scale of SF-36 instead of the scale described in the original article (Roland and Fairbank, 2000). Despite many published variants of RMDQ – RDQ-23 (Patrick et al, 1995), RDQ-18 (Stratford and Binkley, 1997), RDQ-16 (Dionne et al,

40

1997), RDQ-two (Underwood et al, 1999), RDQ-7p (Walsh and Radcliffe, 2002), RDQ-12 (Atlas et al, 2003), the original version of the RMDQ is favoured by an international expert group (Deyo et al, 1998).

The first validation was done using a LBP population in a general practice treated with pain medication. RMDQ has good construct validity, internal consistency, responsiveness and reliability (Roland and Fairbank, 2000).The RMDQ score correlate well with the data obtained from other physical function score systems such as QBPDS (Kopec et al, 1995) and ODI (Fairbank et al, 1980).It has evidence of internal consistency as demonstrated by Cronbach alpha of 0.93(Hsieh et al, 1992), 0.92(Kopec and Esdaile, 1995) and 0.84(Jarvikoski et al, 1995). Quoted test- retest correlations for RMDQ include 0.91 (same day) (Roland and Morris, 1983), 0.88 (1 week) (Johansson and Lindberg, 1998) and 0.83 (3weeks) (Deyo and Centor, 1986). Data on responsiveness of the RMDQ has been published by a number of authors ((Davidson and Keating, 2002, Grotle et al, 2004). The RMDQ should be used for disability assessment when there is need to detect short term changes in back pain or short term changes in response to treatment in patients with mild to moderate disabilities (Roland and Fairbank 2000). RMDQ has been validated in thirteen languages: English (Roland and Morris, 1983), German (Weisinger et al, 1999), French (Coste et al, 1993), Greek (Buscainos et al, 2003), Portuguese (Nusbaum et al, 2001), Spanish (Kovacs and Llobera, 2002), Swedish (Johansson and Lindberg,1998), Turkish (Kucukdeveci et al, 2001),

Persian, Tunisian, Norwegian (Grotle et al, 2003), Iranian (Mousavi et al, 2006) and Morrocan (Maaroufi et al, 2007).

The Low Back Outcome Score (LBOS) (Greenough and Fraser,1992), is a self-reported measure to assess pain and disability in patients with low back pain. It is a 13- item questionnaire, and it includes weighted questions about current pain, employment, domestic chores, sport activities, resting, medical treatment or consultations, drug use, sex life, sleeping, walking, sitting, travelling and dressing. The pain question is answered with an 11-point VAS ranging from 'no pain' to 'maximum pain possible'. The 11 answer possibilities are reduced to four categories (0-2, 3-4, 5-6, and 7-10). All the other questions offer an answer for each possibility, except the sport activities and resting questions, which provide three different answers. The answering possibilities of each item are scored with a four- point scale, but questions are differently weighted. Three different groups of questions can be identified. Items with a nine- point scoring system (pain, employment, domestic and sport activities) in which the score can be 0, 3, 6 or 9 points. Items with a six-point scoring system (resting, medical treatment or consultations, drug use, sex life) in which the score can be 0, 2, 4 or 6 points. Items with a three - point scoring system (sleeping, walking, sitting, travelling and dressing) in which the score can be 0, 1, 2 or 3 points. The final score can be obtained by summing the score of each item and it ranges from 0 to 75, with lower values representing greater disability.

The questionnaire can be completed in about 5minutes and scored in about one minute. The LBOS correlates well with the ODI (r=0.87), the WDI (r=0.74). The test- retest reliability is high (r=0.92), it has a good internal consistency with Cronbach's alpha of 0.85 (Holt et al, 2002). The LBOS has evidence of responsiveness (Kahtri and Greenough, 2002) and it has been validated in English (Greenough and Fraser, 1992, Holt et al, 2002).

The Quebec Back Pain Disability Scale (QBPDS) (Kopec et al, 1995)questions were designed using a conceptual model. Item selection was done using factor analysis for 46 disability items and twenty items were selected. The QBPDS measures functional disability and the items are on self-care, walking, sitting, standing, lifting, sport, stairs, housework and sleep. Items about social life and sex life are not included. The questions can be filled out in about 5- 10 minutes and scored in about 2minutes. Each of the activities is scored with a six- point difficulty scale ranging from 0 'not difficult at all' to 5 'unable to do'. The item scores are added up in order to obtain the disability score, which ranges between 0 and 100. The higher values represent greater disability and sub scores are not reported. The QBPDS was validated on a back pain population and published in 1995. QBPDS was correlated with the RDQ (r= 0.72), ODI (r =0.77). Test-retest reliability was analysed using an intra-class correlation coefficient (ICC) (0.92), internal consistency was measured using the Cronbach's alpha (0.96) and is satisfactory. The QBPDS has been shown to be responsive (Kopec et al, 1995). Davidson and Keating (2002) found similar responsiveness when comparing the QBPDS with the SF-36, the ODI, the RMDQ and the WDI but the QBPDS showed a better reliability than the RDQ and the WDI. The QBPDS is validated in English, French (Kopec et al, 1995), Iranian (Mousavi et al, 2006) and Dutch (Schoppink et al, 1996).

The Million Visual Analogue Scale (MVAS) (Million et al, 1981) is a 15- item questionnaire about disability and pain intensity in patients with LBP. The 15 questions are on pain, sleep, stiffness, twisting, walking, sitting, standing and work. Questions on self-care, sex life, lifting and housework are not included. Information about item selection process is not available. Score is given on a 100mm VAS. An index of severity is obtained in each question by measuring the

distance of the marked point from the origin of the line. The final score is calculated by adding up the individual scores, with higher values representing greater disability. The MVAS was first published in 1981 and validated on patients with chronic LBP. Its reliability has been demonstrated by Pearson's correlation of 0.84-0.94 and internal consistency shows a high Cronbach's alpha of 0.93 (Million et al, 1981). Zoega (2000) found a good correlation and a similar reliability between the MVAS and the ODI. The MVAS is validated in English (Million et al, 1981).

The Aberdeen Low Back Disability Scale(ALBDS) was first published in 1994 by Ruta et al and was validated using a random sample from the general population. The 19 questions focus on pain, sleep, bending, self-care, walking, sitting, standing, sport, housework, resting, loss of feeling and leg weakness. Questions on lifting, sex life, work and dressing are not included. There are 6 multiple choice questions and 13 single choice questions. Answering possibilities of each question can vary between three and six items. The answer categories to each single choice question are scored in an ordinal manner (e.g. 0,1,2,3 points, etc.), while multiple choice questions' responses are assigned a score of one point. The 'back pain severity score' is calculated by summing the score of the responses to each question, and then it is converted to percentages. The final score ranges between 0 and 100 with the higher values representing greater disability. The questionnaire is easy to administer, can be completed within 5 – 10 min and scored within 3 min. The ALBDS was correlated with the SF-36 (r=0.56) (Ruta et al, 1994). The ALBDS showed an acceptable test-retest correlation of 0.94 and acceptable internal consistency with a Cronbach's alpha of 0.8 (Ruta et al, 1994). The ALBDS is able to detect

clinical change two weeks after surgery (Ruta et al, 1994), and is validated in both English (Ruta et al, 1994) and Chinese (Leung and Lam, 1999).

The North America Spine Society (NASS) Lumbar Spine Outcome assessment instrument (NASS LSO) was first published by Daltroy et al. (1996) and is derived from a consensus of the NASS. It consists of 62 main questions obtained from three different existing questionnaires: the SF 36, a modified ODI and a modified employment assessment published by Bigos et al. (1991). The NASS data are grouped into five categories. The first group consists of demographic data (age, sex, race, education and insurance information). The second group consists of the medical history (diagnosis, past surgeries, co-morbidities etc.) The third group includes pain, neurogenic symptoms and function. These domains are measured by a modified ODI version. The fourth group is represented by employment history, evaluated by a score system published by Bigos et al. (1991). The fifth group consists of data about outcomes of treatment, but it is included only in the follow-up module. The scoring is complex and subscores are extractable (modified ODI, SF 36, pain and disability scale, neurogenic symptoms scale, job exertion scale, expectation and satisfaction scale). The questionnaire is long and it takes 20 min to be filled out. The NASS LSO pain and limitation scale correlates with other measures of the same phenomenon (pain VAS $r = 0.84$, SF 36 pain subscale $r = 0.66$ and SF 36 physical limitation subscale $r = 0.75$) (Daltroy et al, 1996). The reliability tests for the condition – specific parts of the NASS LSO show a test-retest reliability of 0.96 for pain and disability and 0.81 for neurogenic symptoms (Daltroy et al, 1996). Cronbach's alpha is 0.93 in both measures (Daltroy et al, 1996). The NASS LSO is validated in English (Daltroy et al, 1996), German (Pose et al, 1999) and Italian (Padua et al, 2001).

45

The Low Back Pain Rating Scale (LBPRS) (Manniche et al, 1994) is a rating system designed to evaluate the clinical outcome of LBP patients. It includes three different components – pain, disability and physical impairment. The pain components consist of six questions divided into 2 groups – three questions about back pain and three questions on leg pain. Each item is scored with the VAS (0 – 10 points). Therefore the pain components in total give 0 – 60 points. The disability component consists of 15 questions on sleeping, housework, walking, sitting, lifting, working, dressing, driving, running, climbing stairs, getting up from a chair, and contact with people and expectations of future pain. Questions about self-care, lifting, standing, sex life and sport are not included. Each question is answered with a three point Likert scale (0 – 2 points).

The disability component gives a total score of 0 – 30 points. The physical impairment component is evaluated by measuring the back muscle endurance, spinal mobility, patient mobility and use of analgesics. Each of these components is scored on a scale ranging from 0-10, thus the physical impairment component in total gives 0 – 40 points. The three different components are weighted: 60 points for pain scoring, 30 points for disability and 40 points for physical impairment. Therefore, combining them, the final LBPRS score ranges from 0 (in patients without back problems) to 130 (in disabled patients). The questionnaire can be filled out in about 15 min and scored in about 3 – 5 min. The LBPRS was first published and validated on different populations by Manniche et al (1994). Christensen et al (1993) showed a correlation (r = 0.82) between the ODI and the LBRS. Internal consistency was calculated using Cronbach's alpha with values between 0.89 and 0.95 for the sub-scores and 0.98 for the entire LBPRS (Manniche et al, 1994). The LBPRS is available in English and validated in Danish (Manniche et al, 1994).

The first article on the Wadell Disability Index (WDI) was published in 1984 and showed the scale to have been validated on a population with chronic low back pain (Wadell and Main, 1984). It has nine items that are based on disabilities (standing, sitting, walking, sex-life, lifting, dressing and travel), social life and body function (sleep and pain). It does not include questions about sports, self-care and work. It takes 5 minutes to complete and 1 minute to score.The WDI was first published in 1984 and validated on a chronic low back pain patients. The WDI test-retest reliability was evidenced with an ICC of 0.74. It is able to detect clinical changes 4 week after surgery (Ruta et al, 1994). It is validated in English (Waddel and Main, 1984) and Spanish. A French translation is available but has not been validated (Guillemin et al, 1994).

Curtin Back Screening Questionnaire (CBSQ) - developed as a discriminative screening instrument and disability assessment tool in patients with moderately severe work related LBP. The CBSQ was developed following the principles of Kirschner and Guyatt (1985) employing data from 74 subjects with moderately severe work related LBP.CBSQ has 93 items arranged in 3 parts. Part I is the body of the questionnaire comprised of 11 disability categories: Pain (4 items), Self-care and mobility (16 items), Work ((items), Sex (1 item), Services, Finances and Legal aspects (11 items), Home management (3 items), Family life and personal relationships (7 items), Social life (7 items), Hobbies and pastimes (4 items), Thinking and problem solving (3 items) and Emotions(14 items).Part 2 contains general and demographic information, while part 3 is for comments and evaluation of the questionnaire by the respondent. All items are presented

in the form of a 4-point Likert scale ranging from 1 (no problem) to 4 (severe problem). Completion time is between 20 and 30 mins and it is self-administered.The screening function of the questionnaire was developed by selecting 8 questions from the whole questionnaire using regression analysis. The screening score is calculated by adding the specific weights for each response of the 8 screening questions. CBSQ was validated using 150 subjects. Construct validity – CBSQ was correlated against SIP for the subscale Pain (r = 0.56), Self-care and mobility (r = 0.65), home management (r = 0.62), Social life (r = 0.72), Hobbies and pastimes (r = 0.58), thinking and problem solving (r= 0.75).Test-retest reliability for the whole questionnaire was 0.98. The Cronbach's alpha coefficients ranged from 0.70 to 0.90 (Harper et al, 1995). CBSQ has been shown to discriminate between fit individuals and those with LBP, and between individuals with different degrees of LBP (Harper et al, 1995).

The Istanbul Low Back Pain Disability index (ILBPDI) was developed to assess the functional disability in patients with chronic LBP. It was developed following the steps outlined by Kirshner and Guyatt (1985). It has 16 items which are - on stairs, walking, running, travelling, sitting, getting up from chair, bending, self-care, lifting, dressing. Questions about sex life, social life and work are not included. Each question is answered using a 6 – point Likert scale ranging from 0 (no difficulty) to 5 (impossible). The total score was calculated as a sum of scores for each item, ranging from 0 – 90, with higher scores representing higher degree of disability. ILBPDI can be administered in 3 minutes. It was well correlated with other functional disability scales associated with low back pain – QBPDS (r = 0.82), revised ODI (r = 0.76), WDI (r = 0.68). The interrater reliability of the scale was found to be 0.79 and the internal consistency of the scale was 0.90.ILBPDI was validated in English and Turkish (Duruoz et al, 2013).

The General Function Score (GFS) is a disease specific instrument consisting of nine items created to measure physical disability in patients with LBP (Hagget et al, 2001). It was derived from an original questionnaire consisting of 17 items. The final GFS includes 9 items, showing high individual correlations, validity, reliability, responsiveness and feasibility. The items are on walking, stair climbing, standing, lifting, dressing and house chores. Each item can be answered with three possible response alternatives – 'can perform', can perform with difficulty' or 'cannot perform', scored as 0, 1 and 2 points. The total score is obtained by summing each items score and is expressed as a percentage, while 0% means no physical disability and 100% means maximal physical disability. It can be administered in 2 mins and scored in 1 min. The GFS was validated using 297 patients with chronic LBP. The Spearman rank correlation coefficient for the GFS with the ODI was 0.61 and with the MVAS was 0.54. The ICC of the total score of the GFS was 0.87. The GFS also showed good responsiveness (Hagget et al, 2001). The GFS has been validated in English (Hagget et al, 2001).

The Back Illness Pain and Disability Nine-Item Scale (BACKILL) (Tesio et al, 1997) aims to detect disability and response to treatment in chronic low-back pain affected patients. Items are selected from three pre-existing validated instruments: the PAIN-FREE8, which is an 8-item version of McGill Pain Questionnaire (Melzack, 1978), the Functional Assessment Screening Questionnaire with five items (Granger and Wright, 1993, Granger et al, 1995), which is derived from the original 15-item FASQ (Seltzer et al, 1982) and the Oswestry low back pain disability questionnaire with eight items, which is a shorter version of the original ODQ (Fairbank at al, 1980). The BACKILL includes two items for pain (aching and tiring), seven items for mobility (lifting, sitting for 30 minutes, standing for 30 minutes, travelling, getting up from a low seat,

walking, and personal care). Items about pain are scored with a four-point scale: none (4 points), mild (3 points), moderate (2 points), and severe (1 point). Three items about mobility (standing, sitting and getting up from a low seat) are also scored with a four-point scale: easy (4 points), a little difficulty (3 points), a lot of difficulty (2 points) and unable to do without help (1 point). Resting mobility items are scored with a six-point scale in which possible answers are specific for each question. Moreover two additional items can be included (fearful and punishing-cruel). They are scored separately from BACKILL items with a four-point scale. The questionnaire is self-administered and it is easy to complete and to score. BACKILL demonstrated adequate evidence of construct and predictive validity, it has been tested for reliability with ICC values of all the items ranging from 0.42 – 0.89 (Tesio et al, 1997).

The Bournemouth Questionnaire (BQ) (Bolten and Breen, 1999) is a short-form multidimensional questionnaire designed to measure the outcomes in back pain patients. The items included in the questionnaire were obtained by reviewing the literature. Seven aspects of the back pain experience were selected. These aspects were the most commonly measured, and showed significant responsiveness to clinical change. Domains are the following: pain intensity; ability to perform daily activities and social activities; anxiety status; depression status; pain interference with work activities and pain locus control each item sis scored with an 11-point numerical rating scale from 0 to 10. A total score can be obtained by summing the result of each item, although the authors recommend expressing the total score of BQ as a percentage. The questionnaire can be completed and scored quickly. BQ was validated on 90 patients with low back pain (Bolten and Breen, 1999). It has face validity, high internal consistency (Cronbach's

alpha=0.9), good test-retest reliability (ICC=0.95). BQ has acceptable construct validity and has high effect size (1.29) which is comparable with established measures (Bolten and Breen, 1999).

The Dallas Pain Questionnaire (DPQ) (Lawlis et al, 1989) is a 16-item instrument to assess the four aspects of daily living affected by chronic back pain: day-to-day activities, such as pain and intensity, personal care, lifting, standing, sitting, walking and sleeping; work and leisure activities, such as social life, travelling and vocational; anxiety-depression status, including anxiety and mood, emotional control and depression; and social interest, such as interpersonal relationship, social support and punishing responses. Each item is scored with a VAS, divided into five, six, seven or eight small segments (it depends in the question). Scale extremities are labelled with specific words (e.g. 'no pain'/'all the time') and with percentage (0%/100%). For every specific question, the patient marks the point on the scale which represents his/her condition. For scoring, 0 points are assigned to the left segment, 1 point to the next segment, 2 points to the next segment and so on to the last segment. Item scores are added and multiplied by a constant to obtain the percentage of pain interference with each of four daily living aspects evaluated by DPQ. The constant used for daily activities section is 3, while the constant used for work/leisure activities, anxiety/depression and social interest section is 5. The DPQ can be answered in 3-5 minutes and scored in less than 1 minute. DPQ was validated in 1989 on a chronic LBP population. It has adequate construct validity, test-retest reliability and internal consistency (Lawlis et al, 1989). It has a French version. Wilhelm et al, 2010 in a study comparing the sensitivity of QBPDS and that of DPQ showed that DPQ was moderately responsive.

The Functional Rating Index (FRI) is a 10-item scoring system designed to measure both patient's perception of function and pain of the spinal musculoskeletal system (Feise and Menkie, 2001).The instrument includes:eight items focused on daily activities (sleeping, self-care, travel, work, recreation, lifting, walking and standing) that can be affected by a spinal disease and two items focused on two different aspects of pain (intensity and frequency). Each item is scored with a five-point scale ranging from 0 (no pain or full ability to function) to 4 (worst possible pain or unable to perform a specific function at all). The index score is achieved by adding up the equally weighted scores, dividing by the maximum possible score, and multiplying by 100%. When all 10 items are answered, the formula is the following: (total score/40) x 100%. The final score ranges from 0 (representing absence of disability) to 100% (representing severe disability). Therefore, the higher the score the higher the perception of dysfunction and pain.The FRI was validated in 2005 using 131 patients with LBP. The validity was supported by a moderate correlation coefficient ($r = 0.67$) between FRI and ODI. The reliability of FRI was demonstrated by an ICC of 0.78. FRI has been shown to demonstrate evidence of responsiveness (Childs, 2005). FRI has been validated in English (Childs, 2005, Feise et al, 2001,Rebbeck et al, 2007), Portugese (Coasta et al, 2007, Lee et al, 2006).

The Spinal Pain Independence Measure (SPIM) is a 12-item questionnaire designed to assess chronic LBP (Itzkovich et al, 2001). It consists of three sections: activities related to mobility, activities performed in sitting and standing and activities performed in the room and bathroom. The mobility section includes five items: mobility for short distances, mobility for moderate distances, mobility for long distance, stair management and maximal walking speed. The activity in sitting and standing section includes three items: carrying loads, activity in the sitting position

and activity in the standing position. The activity indoors section includes four items: mobility in bed, transfers, washing lower body and dressing lower body. The SPIM was validated using 23 patients with chronic low back pain, criterion validity was demonstrated by a correlation coefficient value of 0.78 when correlated with QBPDS, has good inter rater reliability (r = 0.78) and has evidence of adequate responsiveness (Itzkovichi, 2000).

The Back Pain Functional Scale (BPFS) is a self-report measure evaluating patient's functional status in clinical research settings. Item selection was from existing questionnaires (such as SIP, OLBPD QBPD, Dallas pain questionnaire (DPQ), RMQ, MOS-36, PSFS) and interviews with physical therapists. Items reduction was performed by examining the test-retest reliability, internal consistency, content and construct validity. The final version of the BPFS consists of 12 items, investigating work, hobbies, home activities, bending or stooping, dressing shoes or socks, lifting, sleeping, standing, walking, climbing stairs, sitting and driving. Each item is scored with a six-point scale, in which 0 means unable to perform activity,1 extreme difficulty, 2 quite a bit of difficulty, 3 moderate difficulty, 4 a little bit of difficulty and 5 no difficulty. The total BPFS score can vary from 0, representing the lowest functional level, to 60, representing the highest functional level the questionnaire takes less than 5 minutes to complete and about 30 second to score. The BPFS was validated in English (Stratford, 2000). It has good convergent validity, correlation with RMDQ was 0.79 (Stratford, 2000), it demonstrated good test- retest reliability as shown by ICC values of 0.82 – 0.88, (Stratford, 2000, Stratford and Binkley, 2000) and has adequate evidence of responsiveness (Stratford, 2000, Stratford and Binkley, 2000).

The Disability Rating Index (DRI) is a 12-item questionnaire that allows the user to evaluate physical function (Salen et al, 1994). The DRI includes the following activities: dressing; outdoor walks; climbing stairs; sitting for a long time; standing bent over a sink; carrying a bag; making a bed; running; light work; heavy work; lifting heavy objects; participating in exercise/sports. The 12 items are divided into three categories: basic daily life activities (questions 1-4); physical activities (questions 5-8); work-related/vigorous activities (questions 9-12).Each item is scored with a 100mm VAS. Extremities of the scale are labelled with 'without difficulty' (0 points) and 'not at all' (100 points). Patients mark the point on a line, representing their ability to perform the daily activities included in the question list. For scoring, in each item the distance in mm on the VAS between the zero points and the point marked from the patient is measured. The mean of these measurements is calculated, and the DRI score is expressed as a percentage. The DRI is a very quick self-administered questionnaire and can be scored in less than 2 minutes. The DRI was validated on 366 patients with neck/ shoulder/ low back pain and 1092 healthy subjects. It demonstrated good discriminative validity. Its test-retest correlation was 0.95. Its internal consistency is adequate as shown by a Cronbach's alpha of 0.84. The DRI has evidence of responsiveness (Salen et al, 1994).

Table 1 summarises the nineteen scales / outcome measures described above. The conceptual bases of these outcome measures focused on components of outcomes that are relevant to the patients. Six scales focused on assessment of disability (WDI, ILBPDI, GFS, QBPDS, DRI, SPIM), three are focused on functional pain assessment (ODI, MVAS, DPQ), four focused on disability and functional pain assessment (RMDQ, LBOS, ALBDS, LBPRS), three focused on pain and disability assessment (BACKILL, NASS LSO, BQ), one focused on pain and functional

assessment (FRI), one focused on functional status assessment (BPFS) and one focussed on disability assessment and screening (CBSQ). Seven of these scales have items on impairment and activity limitation (QBPDS, MVAS, ALBDS, LBOS, RMDQ,FRI, BACKILL), four scales have items on activity limitation alone (GFS, DRI, ILBPDI, SPIM) and eight scales have items on impairment, activity limitation and participation restriction (ODI, WDI, NASS LSO, LBPRS, BPFS, BQ, DPQ, CBSQ).The activities of daily living common to many of the scales include walking, standing, sitting, lifting, self-care. The impairment variable common to many of the scales include pain and sleep. They have items ranging from 7 (BQ) to 93 items with CBSQ being the longest with 93 items. The time it takes to complete the scales ranges from 1 – 30 minutes, the GFS requires the shortest time and CBSQ takes the longest time for completion.

Table 1: Low Back Pain Specific Scales

Serial No.	Name of outcome measure	Conceptual Basis	No of Items	Scale Response	Method of Administration/ Type	Utility(administration/scoring time)/ Languages	Validity	Reliability	Responsiveness
1.	Oswestry Disability Index Fairbank et al (1980)	Functional Pain Assessment	10	Scaled text	Self LBP Specific	5 Minutes/>1 min 17 languages	Criterion Validity	Test-retest reliability Internal consistency	X
2.	Roland-Morris Disability Questionnaire Roland and Morris (1983)	Disability and functional pain assessment	24	Yes/No	Self LBP Specific	10 Minutes/>1 min 13languages	Criterion Validity	Test-retest reliability Internal consistency	X
3.	Low Back Outcome Score Greenough and Fraser (1992)	Disability and functional pain and general assessment	13	Scaled text 11-point VAS for pain	Self LBP Specific	5 minutes/>1 min English	Criterion Validity	Test-retest reliability Internal consistency	X
4.	Quebec Back Pain Disability Scale (QBPDS) Kopec et al (1995)	Disability and Functional pain assessment	20	6-point Likert scale (0-5)	Self LBP Specific	5-10 Minutes/>2 mins 4 languages	Criterion Validity	Test-retest reliability Internal consistency	X
5.	Million Visual Analogue Scale (MVAS) Million et al (1981)	Pain and pain intensity locator, functional pain assessment combined with other instruments	15	100mm VAS	Self LBP Specific	10 Minutes/2-3 mins English	Criterion Validity	Test-retest reliability Internal consistency	NI
6.	Aberdeen Low Back Disability Scale (ALBDS) Ruta et al, 1994	Disability and functional pain assessment	19	Scaled text and Multiple choice	Self LBP Specific	10 Minutes/3 mins 2 languages	Criterion Validity	Test-retest reliability Internal consistency	X
7.	The North American Spine Society Lumbar Spine Outcome (NASSLSO) Dultroy et al (1996)	Pain and pain intensity locator, functional pain assessment combined with other instruments	62	Pain locator and 6-point Likert scale	Self LBP Specific	21 Minutes /- 3 languages	Criterion Validity	Test-retest reliability Internal consistency	NI

Low Back Pain Specific Scales contd.

Serial No.	Name of outcome measure	Conceptual Basis	No of Items	Scale Response	Method of Administration / Type	Utility(administration/scoring time)/ Languages	Validity	Reliability	Responsiveness
8.	Low Back Pain Rating Scale (LBPRS), Manniche et al (1994)	Disability and functional pain assessment	21	3-point Likert scale and 11-pont VAS	Self & Clinician measured LBP Specific	15 Minutes/3-5 mins 2 languages	Criterion Validity	Test-retest reliability Internal consistency	X
9.	Wadell Disability Index (WDI) Wadell& Main (1984)	Disability assessment	9	Yes/No	Self LBP Specific	5 Minutes/>1 min 2 languages	Criterion Validity	Test-retest reliability Internal consistency	X
10.	Curtin Back Screening Questionnaire Harper et al, 1995	Disability assessment and Screening tool	93	4-point Likert scale	Self LBP specific	20 – 30 minutes/- English	Criterion Validity	Test-retest reliability Internal consistency	NI
11.	Istanbul Low Back Pain Disability Index (ILBPDI) Duruoz et al, 2013	Disability assessment	18	6-point Likert scale (0-5)	Self LBP specific	3 mins/- 2 languages	Construct, Convergent and Divergent Validity	Inter-rater reliability internal consistency	NI
12.	General function score Hagg et al, 2001	Physical Disability assessment	9	3-point Likert scale (0-2)	Self LBP specific	1-2 minutes/-	Construct, Criterion	Internal consistency, test-retest	X
13.	Back illness and disability Nine item scale (BACKILL) Tesio et al, 1997	Disability assessment and Pain assessment	9	4-point Likert scale (1-4) Scaled text	Self LBP specific	Easy to complete and score	Content Construct Predictive	Test-retest	X
14.	Bournemouth questionnaire Bolton et al, 1999	Pain assessment Disability assessment Psychosocial assessment	7	11-point Numerical rating scale (0-10)	Self LBP specific	Completed and scored quickly	Face validity Construct	Test-retest Internal consistency	X
15.	Dallas pain questionnaire Lawlis et al, 1989	Functional pain assessment	16	VAS (0-8)	Self LBP specific	3-5 minutes/1 min	Concurrent Construct	Test-retest Internal consistency	X

Low Back Pain Specific Scales contd.

Serial No.	Name of outcome measure	Conceptual Basis	No of Items	Scale Response	Method of Administration / Type	Utility(administration/scoring time)/ Languages	Validity	Reliability	Responsiveness
16.	Spinal Pain Independence Measure Itzkovich et al, 2001	Disability assessment	12	8-point Likert scale (0-7)	Self LBP specific	NI	Criterion validity	Inter-rater	X
17.	Functional Rating Index Feise and Menke,2001	Pain and functional assessment	10	5-point Likert scale (0-4)	Self Spine specific	1.15 mins/ >1 min 2 languages	Criterion validity	Test-retest Internal consistency	X
18.	The Back pain Functional Scale (BPFS) Strafford et al, 2000	Functional status	12	6-point Likert scale (0-5)	Self	> 5mins/>1 min	Convergent	Test-retest Internal consistency	X
19.	Disability Rating Index Salen et al, 1994	Disability assessment	12	100mm VAS	Self	2.7 mins/ 1.2 mins	Discriminative	Test-retest Internal consistency	X

NI: Not indicated in literature   X: Responsive

All the scales / outcome measures are self- reported (i.e. rated by patients only) except the LBPRS which is rated by both patients and clinicians. The response option of the scales differ but they all make use of the direct estimate method. Two scales (RMDQ & WDI) have dichotomous answers (Yes / No), one has scaled text (ODI), while six (QBPDS, ILBPDI, GFS, DRI, FRI, BPFS) have Likert scale ranging from 0 to 7. The response scale for BQ is Numerical rating scale, two of the OMs (MVAS, DPQ) make use of Visual Analogue Scale (VAS) as the response scale. Low back pain rating scale makes use of both Likert and Visual analogue scale while LBOS makes use of scaled text and VAS. The response options for ALBDS are scaled text and multiple choice and CBSQ makes use of Likert scale and weighted score.

Eight of these scales have been translated and validated in other languages apart from English (ODI, RMDQ, QBPDS, ALBDS, NASS LSO, LBPRS, WDI, ILBPDI and FRI) to enhance their utility across many cultures with ODI having the highest number of validated versions in other languages (seventeen languages), followed by RMDQ (thirteen languages). The scoring formats of all these scales are available. Eleven of the scales (RMDQ, WDI, QBPDS, LBOS, LBPRS, MVAS, CBSQ, BACKILL, BPFS, SPIM and ILBPDS) express the overall score as the sum of all the individual scores while the remaining nine scales (DRI, ODI, FRI, DPQ, BQ, GFS, ALBDS and NASS LSO) express the overall score as a percentage. All these scales have evidence of one or more psychometric properties proven. None of these scales have contribution from Nigerian authors, therefore culturally relevant items to Nigerian patients like 'sweeping the floor', Squatting to use pit toilet' were not included.

CHAPTER THREE

MATERIALS AND METHODS

3.0 Phases of the Study

The aim of the study was to develop a Nigerian culture and environment friendly low back pain disability scale and to investigate its psychometric properties. The study was conducted in two phases. Phase 1 was the development of the disability scale and phase 2 was the psychometric testing of the disability scale.

3.1 Phase 1— Development of the Low Back Disability Scale

The steps involved in development of the disability scale are finding justification for development of a new scale, conceptual basis of the scale, devising the items, content validation, naming the scale, response scale and scoring, selection of items and minimizing the face validity of the scale.

3.1.1 Finding Justification for Development of a New Scale

The first step in development of the low back pain disability scale was finding justification for the need to develop the scale.Literature was searched through OVID and PubMed databases. The first set ofsearch terms were low back pain, back pain, scoliosis, spinal stenosis andankylosing spondylitis from year 2000 to 2010. These terms were combined with 'OR' Then the following were requested and also combined with 'OR' – questionnaires, outcome assessment and outcome measures. Both sets were combined with 'AND' yielding 1,381,110 articles. These results were combined using 'and' with low back pain, outcome assessment, outcome measure and questionnaires. This yielded 288 articles from which 120 articles were selected after reviewing them by titles and abstracts. A total of 120 full text

60

articles that utilized LBP outcome measures were reviewed. From these articles, nine standardized LBP scales with evidence of one or more psychometric properties that were each used in at least three studies were selected. The content, utility and psychometric properties of the scales were reviewed and the Oswestry Disability Index (ODI) was selected to be used for the pilot study due to its wide utility and its psychometric properties. The pilot study involved administering the ODI on 16 patients with low back pain attending the out-patient Physiotherapy clinic of the University College Hospital (UCH), Ibadan through interview. The patients were also interviewed to find out other activities of daily living they had difficulty performing which are of concern to them and are not included in ODI. The common activities reported by the patients include:

1. Drawing water from a well.

2. Sweeping with a short broom.

3. Sitting on the heel (during the Islamic prayer).

4. Forward bending from the trunk (during the Islamic prayer).

5. Kneeling to greet elders.

6. Attendance at worship (Jumat or Church services).

7. Squatting to use pit latrine

The following were therefore the justifications for development of a new scale:

1. None of the available standardized low back pain scales include any of the seven activities of daily living the 16 patients interviewed above reported they had difficulty performing

2. Some of the existing low back pain scales also contain some terms/phrases, such as walking a few blocks, putting on 'pantyhose' which are not commonly used in Nigeria and which members of the target population may have difficulty understanding.

3. Some of the commonly utilized LBP scales contain some activities (e.g. putting on panty hose, watering flower) that an average Nigerian may not perform in a lifetime.

3.1.2 Conceptual Basis of the Scale

The conceptual basis of the scale that was developed was that patients with LBP experience difficulties in performing certain activities of daily living and the extent of difficulty experienced in these activities represent the degree of activity limitations.

3.1.3 Devising the Items

Items on the scale were devised through the following methods:

a) **Review of Literature**: The primary goal of this literature review was to identify activities deemed relevant by earlier developers of standardized LBP scales. The nine scales that were earlier found through literature search were reviewed to determine common activities included in them. Table 3.1 shows the 28 activities included in the nine scales. It was decided that activities that were included in at least four scales would be selected for inclusion on the list of items for the scale that was being developed.

Twelve items that were devised using these common activities from literature were:

1. Putting on stockings / sandals / shoes.

2. Sleep through the night.

3. Stand up for 20 – 30 minutes.

4. Walk for 1km (15- 20minutes).

Table 3.1 Activities on the Nine Scale Reviewed from Literature

Activity	ODI	RMDQ	LBOS	QBPDS	MVAS	ALBDS	NASS LSO	LBPRS	WDI
Pain	√	√	√	-	√	√	√	√	-
Sleep	√	√	√	√	√	√	√	√	√
Self-care	√	√	-	√	-	√	-	-	-
Walking	√	√	√	√	√	√	√	√	√
Sitting	√	√	√	√	√	√	√	√	√
Standing	√	√	-	√	√	√	√	-	√
Lifting	√	√	-	√	-	-	√	√	√
Sex life	√	-	√	-	-	-	√	-	√
Travelling	√	-	√	-	-	-	-	-	√
Social life	√	-	-	-	√	-	√	√	√
Work	-	√	√	-	√	-	-	√	-
Dressing	-	√	√	-	-	-	√	√	√
Sport	-	-	√	√	-	√	-	-	-
Stairs	-	√	-	√	-	-	-	√	-
Housework	-	√	√	-	-	-	-	√	-
Resting	-	√	√	-	-	-	-	-	-
Appetite	-	√	-	-	-	-	-	-	-
Need of help	-	√	-	-	-	-	-	-	-
Psychological factors	-	-	-	-	-	-	-	√	-
Need of treatment	-	-	√	-	-	-	-	-	-
Need of medication	-	√	-	-	√	√	-	-	-
Car driving	-	-	-	√	-	-	-	√	-
Throwing	-	-	-	√	-	-	-	-	-
Stiffness	-	-	-	-	√	-	-	-	-
Twisting	-	-	-	-	√	-	-	-	-
Bending	-	-	-	-	-	√	-	-	-
Loss of feeling	-	-	-	-	-	√	√	-	-
Leg weakness	-	-	-	-	-	√	√	-	-

ODI - Oswestry Disability Index, **RMDQ** – Roland-Morris Disability Questionnaire, **LBOS** – Low Back Outcome Score, **QBPDS** – Quebec Back Pain Disability Scale, **MVAS** – Million Visual Analogue Scale, **ALBDS** – Aberdeen Low Back Disability Scale, **NASS LSO** – North American Spine Society Lumbar Spine Outcome, **LBPRS** – Low Back Pain Rating Scale, **WDI** – Wadell Disability Index.

5. Putting on stockings / sandals / shoes.

6. Sleep through the night.

7. Stand up for 20 – 30 minutes.

8. Walk for 1km (15- 20minutes).

9. Carrying out house chores (e.g. cooking).

10. Having sex.

7. Travelling for a short distance

8. Sit for 4hours in a chair / sit on an upright hard chair.

9. Doing your usual job.

10. Lifting a moderate weight (10 litre container [keg] of oil/water).

11. Getting up from a low armchair / settee.

12. Attending social gatherings.

b) **Interview of Patients**: A total of 231 patients with LBP attending the outpatient clinics of the Physiotherapy Departments of the following hospitals - University College Hospital, Ibadan(16), National Orthopaedic Hospital, Kano(58), Aminu Kano Teaching Hospital, Kano(21), University of Maiduguri Teaching Hospital, Maiduguri(18), National Orthopaedic Hospital, Enugu(23), University of Enugu Teaching Hospital, Enugu(6), Nnamdi Azikwe University Teaching Hospital, Nnewi(45), University of Uyo Teaching Hospital, Uyo(9), Enugu State University Teaching Hospital(11) were interviewed by the researcher and research assistants/ facilitators, who were all physiotherapists in the various centres. Prior to this, the detailed information on the purpose of the study and the purpose of the interview was given to the research assistants. They were also provided with an interview guide. Copies of the list of 12 items devised through literature were sent to the research assistants as interview guide. The research assistants were asked to find out from each patient if they had difficulty performing the activities. The research assistants were also asked to find out from the patients if there were other activities they had

difficulty performing that were not included in the list. Based on the feedback from the interviews at the various centers, eleven (11) additional items were generated:

1. Drawing water from a well.

2. Sweeping with a short broom.

3. Sitting on the heel and forward bending from the trunk during the Islamic prayer.

4. Prostrating / kneeling/ squatting to greet elders.

5. Attendance at worship (Jumat or Church services).

6. Squatting to use pit toilet.

7. Farming / gardening.

8. Getting up from sitting on a chair.

9. Kneeling to pray.

10. Sitting on the mat.

11. Turnover in bed.

c) **Experts' Interview**: Four Orthopaedic Surgeons and eightPhysiotherapists (with at least Master's degree in physiotherapy or related disciplines) who had experience managing patients with low back pain were interviewed. They were selected based on their seniority and years of experience. Each expert was provided with a list of items already devised through literature review and patients' interview and was asked to suggest additional items that should be included in the list of items. Twenty additional items suggested by the experts were:

1. Sitting on and getting up from a low stool.

2. Washing clothes at floor level.

3. Washing clothes at upright position.

4. Climbing stairs.

5. Descending stairs.

6. Sleeping (lying on the back, lying on the left side, lying on the right side, lying face down).

7. Dressing (putting on trousers, putting on underclothes, removing underclothes).

8. Driving a car.

9. Travelling for a moderate distance.

10. Travelling for a long distance.

11. Getting in and out of low vehicles.

12. Getting in and out of high vehicles.

13. Cooking using standard gas cooker.

14. Cooking with low stove, coal pot or firewood.

15. Sweeping with a long broom or brush.

16. Washing leg / feet during a shower.

17. Getting on and off the water closet toilet.

18. Lifting a light weight (4-5 litre plastic container [keg] of oil/water).

19. Lifting a heavy weight (20-25 litre plastic container [keg] of oil/water).

20. Standing for one hour.

The total number of items devised was 43 (12+ 11 +20 =43).

3.1.4 Content Validation of the Items by First set of Experts

Copies of the list of 43 items were sent to the first set of 10 experts. The experts comprised three experienced Physiotherapists working in the Orthopaedic Physiotherapy clinic of the UCH, Ibadan, and two Nigerian Physiotherapists working in Canada and the USA who are familiar with outcomes assessment and are research oriented, four Orthopaedic Surgeons, and a Physician who was familiar with the development and use of health measuring scales/ questionnaires. They were asked to rate the degree of relevance of each item of the scale, using a 5-point Likert scale shown below:

5. Essential (Item is essential and must be included in the scale).

4. Important (Item is important and should be included in the scale).

3. Acceptable (Item is acceptable and may be included in the scale).

2. Marginally relevant (Item is only marginally relevant and doesn't need to be included in the scale).

1. Not relevant (item should not be included in the scale).

The experts were also asked to assess the coverage of the scale and suggest items that should be included in the scale. No additional items were suggested, but it was suggested that each of two items considered to be double barrelled should be split into two, increasing the number of items to 45.

Based on the feedback from experts, the Content Validity Index (CVI) for each item was calculated as number of respondents that rated an item as essential or important divided by total number of respondents. Three items with CVI less than 0.7 were deleted from the list, bringing down the number to 42.

3.1.5 Naming the Scale

It was decided that the scale would be named Ibadan Low Back Pain Disability Scale (ILBPDS). The word Ibadan was intended to indicate that the scale was developed at the University of Ibadan. The words 'Low Back Pain' make the scale specific to low back pain. The word disability indicates that the scale measures the extent of activity limitations experienced by patients with low back pain. Other instruments such as the Oswestry Disability Index, the Dallas Pain questionnaire and the Ibadan Knee/ Hip Osteoarthritis Outcome Measure have been named likewise (Fairbank et al, 1983, Lawlis et al, 1989, Akinpelu et al, 2007).

3.1.6 The Response Scale and Scoring

The response scale adopted for the ILBPDS was a 5 point (0-4) adjectival scale adapted from a health measuring scale developed by Akinpelu et al, (2007).Respondents were asked to rate

their degree of difficulty as not difficult at all (0), slightly difficult(1), moderately difficult (2), very difficult (3), unable to carry out activity (4). Scores were calculated by adding the responses together. The total possible score was calculated by multiplying the total number of items responded to by participants by 4. The percentage score was then calculated for each participant as follows:

<u>Respondent's score</u>   x   100
Total possible score   (Adapted from Fairbank et al, 1980).

3.1.7 Selection of Items

This involved first pretesting of the scale for comprehensibility, first experts panel meeting for reviewing the scale items, second pretesting of the scale for frequency of endorsement, factor analysis of 43 items, review of scale items by a second set of experts, second experts panel meeting for final item selection and factor analysis of the 18 items.

3.1.7.1 First Pretesting for Comprehensibility: The scale being developed was pretested for comprehensibility on 35 patients with non-specific LBP receiving physiotherapy at the Physiotherapy outpatient clinic, University of Uyo Teaching Hospital, Uyo (18) and University College Hospital, Ibadan (17) respectively. After reading through the scale, the patients were involved in cognitive debriefing interview in order to test their understandability of the scale. (Appendix 3). The aims of pretesting were to determine if the items on the scale:

1. Were understood by the target population.

2. Did not include ambiguous words (i.e. having more than one possible meaning).

3. Did not include double–barrelled items (items asking more than one question).

4. Were homogeneous, i.e. to ensure that the items were tapping the same area or construct.

3.1.7.2   First Experts' Panel Meeting for Reviewing Scale Items

The first experts panel meeting comprising of tenexperts (two Orthopaedic Surgeons, six Physiotherapists, a Physician and a layman with a history of treated low back pain) was held at the Physiotherapy department, University College Hospital, Ibadan on the 8th of February, 2012.The objective of the meeting was to review the 42 items based on findings from comprehensibility testing. The outcome of the meeting was modification of   thirty- six items,merging of six items into three, deletion of one item was and addition of five new items (Appendix 4). At the end of the meeting, the number of the items on the scale was 43 (Appendix 4).

3.1.7.3   Second Pretesting for   Frequency of Endorsement: This scale containing 43 questions was used for the second stage of pretesting [testing for frequency of endorsement – (FOE)] , in order to reduce / select the number of items on the scale, This was carried out on a total of 114 patients with LBP   attending the Physiotherapy out- patient clinics of the following hospitals.

   i.   University of Uyo Teaching Hospital, Uyo (19 patients)

  ii.   University College Hospital, Ibadan (14 patients)

 iii.   National Orthopaedic Hospital Kano (20 patients)

 iv.   National Orthopaedic Hospital Igbobi (19 patients)

  v.   University of Nigeria Teaching Hospital Enugu (11 patients)

 vi.   University of Port Harcourt Teaching Hospital (9 patients)

vii.   University of Maiduguri Teaching Hospital, Maiduguri (9 patients)

viii.   Federal Medical Centre, Umuahia (7 patients)

 ix.   Lagos University Teaching Hospital (4 patients)

  x.   Nnamdi Azikwe University Teaching Hospital, Nnewi. (2 patients)

Fifteen percent of the respondents complained that the scale was too lengthy and that there were repetitions of questions.

Plate 3.1. First Experts' Panel Meeting

Plate 3.2. First Experts' Panel Meeting

Frequency of endorsement of the response options was calculated for all the items. The recommended value for FOE is 20 – 80 %. Items with FOE not within this range should be eliminated from the scale (Streiner and Norman, 2008). Items could not be selected, (that is no item could be eliminated from the list) based on frequency of endorsement because the FOE did not fall within 20 – 80% for all the items on the scale. It was then decided that the items would be subjected to factor analysis in order to select items that would be retained on the scale.

3.1.7.4   Factor Analysis of 43 items

Factor analysis of the 43 items was performed and the items were loaded under 4 factors. Twenty one items loaded on the 1st factor, 7 items loaded on the 2nd factor, 3 items on the 3rd factor and 5 items on the 4th factor. Three items loaded on more than one factor while four items did not load on any factor. It was again difficult to select items using factor analysis because of the distribution of the items loading on the four factors. In addition, two of the four items that did not load on any factor were of cultural importance.

3.1.7.5   Review of Scale Items by the Second set of Experts

In order to finalise the process of item selection, a second set of ten experts was constituted for the purpose of item selection. They were made up of five Physiotherapists, three Orthopaedic Surgeons and two Physicians with history of LBP). They were informed about the need to reduce the number of items on the scale. Each expert was provided with the list of 43 items and was asked to rate them according to their degree of importance on a 1-5 Likert scale.(5 – Extremely Important, 4 – Very Important, 3 – Important, 2 – Slightly Important, 1 – Not Important).

3.1.7.6 Second Experts' Panel Meeting for Final Item Selection

The Second Experts' Panel (comprising of five Experts) Meeting was held at the Physiotherapy department, College of Medicine, University of Ibadan in August, 2012. The objective of the meeting was to select the items that would be on the final scale. The final selection of items on the scale was done by input from these five experts based on:

a. Findings from second set of experts' review and comments: Items that were rated as extremely important or very important by at least 70% of the experts were selected to be retained on the scale. This was assigned one point. In addition, experts suggested merging and deletion of some items. Items devised on the same activities (such as sitting on different types of chair for different duration were merged into one. Three items on lifting were also considered to be repetitions and lifting heavy weight (20-25L keg of water) was selected.

b. Relevance of items to the Nigeria culture and environment: this was assigned two points since this was the justification for developing the scale. Items of cultural relevance (such as farming / gardening, bending to wash clothes at floor level, sweeping with a short broom, drawing water from the well) were scored two points each and were selected to be on the final scale.

c. Review of items on existing standardized LBP scales: this was assigned one point. Some items were retained on the scale because they were devised from activities that were deemed relevant to assessment of low back pain by early developers. Such items included walking outdoors for 15-20 mins (1km), lifting heavy weight, standing for 15 -20mins, sleeping through the night and climbing stairs.

The items that scored at least two points were selected to be retained on the final scale. Some items scored two points but they were deleted from the scale because they were regarded as

repetitions of another item that had earlier been selected to be retained on the scale. Such items included walking for one hour (walking for 15 – 20mins was earlier selected), getting up from an office or dinning chair (getting up from low chair or stool had been chosen) and turning in bed during sleep (sleeping during the night had been selected).

At the end of this exercise, an 18 item scale that is Nigerian culture and environment friendly was produced (Appendix 9). It takes about 5-7 minutes to administer and about 2-3 minutes for scoring.

3.1.7.7 Factor analysis of the 18 items on the Final Scale

Factor analysis of the 18 items was performed and 17 items loaded on two factors. One item on 'having sexual intercourse' did not load on any factor. It was decided that this item should be included on the scale based on experts' opinion, patients' comments and review of the existing standardized LBP scales.

3.1.8   Minimizing the Face Validity of the Scale

Deliberate effort was made to reduce the face validity of the scale by mingling together items assessing different domains of activity limitations in self-care, walking, standing and sitting. This is to prevent yea- saying bias or participants' tendency to give the same response to items that are under the same domain.

PHASE 2.0

3.2   Investigating Psychometric Properties of the ILBPDS

3.2.1 Participants

1. One hundred and forty two patients diagnosed to have chronic non-specific LBP with or without radiculopathy, sourced from the out-patient clinic of the physiotherapy

departments of University of Uyo Teaching Hospital, Uyo (60), University College Hospital, Ibadan(8), National Orthopaedic Hospital Kano (25), National Orthopaedic Hospital Igbobi (24), University of Nigeria Teaching Hospital Enugu (14), University of Maiduguri Teaching Hospital, Maiduguri (8) and Jos University Teaching Hospital (JUTH) (3).

2. One hundred and forty two apparently healthy age and sex matched controls without symptoms of low back pain. They were staff and relatives of patients' from the various hospitals listed above.

3.2.2 Inclusion Criteria

The study included:

1. Patients on first appointment referred for physiotherapy with the diagnosis of chronic non-specific LBP with or without radiculopathy, who could read and write English in the experimental group.

2. Participants who were age and sex peers of the patients in EG without symptoms of LBP, without neuromuscular/ musculoskeletal disorder of the lower limbs who could read and write English in the control group.

3.2.3 Exclusion Criteria

This study excluded:

1.Patients with LBP caused by an underlying pathology (such as tumour or infection), or those with co-morbid conditions such as neuromuscular or other musculoskeletal diseases.

2. Pregnant women with low back pain.

3. Pregnant women without low back pain.

3.2.4 Data Collection Instruments.

The following instruments were used for data collection in the study:

I.   A demographic data collection form (Appendix 7).

II.   The Numeric Pain Rating Scale (Appendix 7): This was used to measure the
      intensity of pain experienced by the patient (Wallenstein et al, 1980). The scale is a
      horizontal line with number 0 -10 written on it, zero written at one end (left) and
      10 written at the other end (right). It has been shown to be a reliable scale for
      assessing intensity of pain (Price et al, 1994).

III. A body Chart for marking the site of pain (Mckenzie 1981) (Appendix 8).

IV.   The Ibadan Low Back Pain Disability Scale (Appendix 9).

3.2.5 Venue of Study

The study was a multicentre study conducted in the Physiotherapy Department of seven
hospitals. They were University of Uyo Teaching Hospital, Uyo, University College
Hospital, Ibadan, National Orthopaedic Hospital, Kano, National Orthopaedic Hospital
Igbobi, University of Nigeria Teaching Hospital Enugu, University of Maiduguri Teaching
Hospital, Maiduguri and Jos University Teaching Hospital, Jos.

76

3.2.6 Sample Size

The sample size was calculated using the formula:

$$n = \frac{Z_{1-\alpha/2}^2 P(1-P)}{d^2}$$

$Z_{1-\alpha/2}$=1.96 is the number of standard errors away from the mean

P= 20% is the prevalence for LBP (Omokhodion & Sanya 2003)

d=7% is the precision

$$n = \frac{1.96^2 \times 0.20 \times (1-0.20)}{0.07^2}$$

$$= \frac{1.96^2 \times 0.20 \times 0.80}{0.07^2}$$

$$= 125.44$$

Adding 10% attrition rate gives n= 139 (Lemeshow et al, 1993).

To enhance precision, the computed minimum sample size was rounded up to 150. The subjects that could however not be matched with their age and sex peers were dropped bringing the number of subjects to 142.

3.2.7 Research Design

The research design was quasi-experimental comprising of three phases: baseline, intervention and post intervention phases.

Baseline – ILBPDS was administered on both the EG and CG, NPRS was administered on the experimental group only.

Intervention phase – a 5-week physiotherapy programme at the rate of two times per week.

Post – Intervention -   the NPRS and ILBPDS scores of the EG were evaluated.

3.2.8 Sampling Technique

Participants were recruited consecutively into the study which was carried out concurrently in all the centres. All participants that met the inclusion criteria were recruited consecutively into the study from September 2012 to September 2013.The patients with LBP were in the Experimental Group (EG) while their age and sex matched apparently healthy individuals without symptoms of LBP were in the Control Group (CG). The age of the control was +/– 2 years that of the EG.

The convenience sampling was used to select the controls. Staff and relatives of patients were approached in their offices and outpatient clinics respectively. Those who met the inclusion criteria were informed about the purpose of the study and informed consent was obtained from those who were willing to participate. This was done until the sample size was achieved.

3.2.9   Training of Research Assistants

Six Physiotherapists with a minimum of 10years clinical experience were trained as research assistants by the researcher. They all have postgraduate training and are research oriented. Training was done individually and it involved teaching the research assistants on how to administer ILBPDS and NPRS through interview and that explanation should be provided to

participants when needed. They were informed about the inclusion and exclusion criteria for recruitment of patients with low back pain and the age and sex matched controls. They were also taught on how to compute score of each participant. Items that were not applicable to participants were to be marked as NA (Not Applicable). They were informed that the physiotherapy should be twice weekly for five consecutive weeks. The treatment protocol was written and made available for each research assistant and adherence to this protocol was emphasized to ensure standardization of the protocol.

They were informed not to include patients in the EG who may not complete the physiotherapy programme in the post intervention administration of ILBPDS and NPRS. They were also taught on how to explain the protocol to participants in simple languages and to obtain their informed consent by asking them to sign on the appropriate place on the form. The research assistants were instructed not to coarse patients to participate in the study. Patients who may refuse to participate should not be penalized. They were also instructed that data about participants should be kept secret. The training session took about 45mins to 1 hour for each research assistant. The questions asked were answered by the researcher to further educate the research assistants on the procedure. Intra observer error was minimised by periodic re- training via phone calls and email.

3.2.10 Ethical Consideration

The approval of the University of Ibadan/University College Hospital Research Ethics Committee and the Ethical Committee of the University of Uyo Teaching Hospital was obtained before commencement of the study. Permission from the Heads of Physiotherapy departments of the hospitals where data were collected was also obtained. Information on the

rationale and the procedure for the study was adequately explained to the participants and their informed consent was obtained.

Confidentiality of data: All data collection instruments, materials, and documentation developed during this project were treated with utmost secrecy and confidentiality. The data collected from the respondents was used for the purpose of this research. The questionnaires were identified with numbers and every data collected from the participants was safeguarded using a password protected computerized system and protected from a third party. The research assistants were instructed not to divulge information about participants' data.

Translation of protocol to local language: This study involved participants that can read or write English therefore no translation of protocol was done.

Beneficence to participants: The interviews were conducted in a friendly manner that enabled participants to communicate their true disability status easily. The study results and recommendations would be communicated properly in a way that would enhance planning for interventions that will help educate participants with LBP. Participants were also trained on how to prevent reoccurrence.

Non-maleficence to participants: The research posed no harm, risks or injury to the respondents, as no new procedure was being tested and the results obtained was used for the purpose of the study only.

Voluntariness: The participants were free to choose whether or not to participate in the study. A voluntary consent form was attached to the data collecting materials, every patient approached to participate in the study carefully read through with the aid of a research assistant, and voluntarily decided to participate after understanding all the procedures

involved in the study. There was no penalty attached to those who decided not to take part in the study.

3.2.11 Procedure for Data Collection

This was done by the principal investigator and trained research assistants from September 2012 to September 2013.

Demographic Data Form: demographic data (age, sex, occupation, marital status, religion) and clinical data (duration of LBP, site of the pain / paraesthesia and the present treatment) were collected through interview.

Body Chart: Each patient with LBP was asked to mark the site where he has pain and paraesthesia if present on the body chart.

Numerical Pain Rating Scale: Each patient in the EG was asked to choose a number from 0 to 10 on the horizontal line (NPRS) which represents his /her level of perceived pain intensity at the low back during the first visit. The number chosen by the patient was recorded as the patient's pain intensity score. This was taken and recorded again after five weeks of physiotherapy.

Administration of ILBPDS: The ILBPDS was administered to participants in both EG (142) and CG (142) through interview in order to assess the ILBPDS for construct validity. The ILBPDS was re- administered to 113 participants in EG during the second visit, two days after (48 hours) (Streiner and Norman, 2008) for the purpose of testing ILBPDS for test-retest reliability. This time interval was meant to forestall carrying over effect. The participants in the EG after completing ILBPDS during their second visit then started physiotherapy at the frequency of two times a week for five consecutive weeks.

The physiotherapy included the following:

Infrared radiation (IRR) for 10-15 minutes for superficial pain − Patient was comfortably positioned in the prone or side lying depending on the position comfortable for the patient. The area to be radiated was cleansed using cotton wool and methylated spirit in order to remove oil from the skin. IRR is then applied to the area such that the distance between the IRR machine and patient's skin is at least 50cm from the skin (Robertson et al,2006). The eyes are covered with a small face towel to prevent IRR getting to the eyes. The intensity was regulated so that patient experience the heat that was comfortable for him/her (Casazza, 2012; Chou et al, 2007).

Cold/hot packs: Patients was comfortably positioned on the treatment couch. The cold/hot pack was wrapped with a towel with 4 layers in between the pads and patient's skin. The cold/hot pack was then placed over the painful area for applied for 10 − 15 minutes (Gammon and Starr 1993;Chou et al, 2007; Kinkade 2007).

Therapeutic Ultrasound for localized deep pain points: Patients was comfortably positioned on the treatment couch. Ultrasonic gel was applied over the painful point. Ultrasound was then applied via the treatment head of the ultrasound machine which was moved in a circular manner for 10 minutes (Robertson et al, 2006; Poitras and Brosseau, 2008).

Back flexion and extension exercise based on assessment using the MacKenzie diagnosis protocol. This was performed repeatedly either in the standing or lying position. The exercise is carried out 10times each time (MacKenzie 1981).

Lumbar Stabilization exercises: Transversusabdominis and multifidus muscles were re-activated with patient in the supine and prone lying position respectively (Standaert et al, 2008).

Transversus abdominis re-activation: The patient was placed in the supine lying position. He / She was then instructed to tuck in the tummy by drawing the navel in towards the back and to hold it for 10 seconds. This was repeated for 10 times.

Multifidus re-activation: The patient was instructed to swell up the multifidus muscle which is located between the two dimples at the low back region while in the prone lying position. This was held for 10 seconds and repeated for 10 times.

Strengthening exercises: Based on assessment, any weak group of muscles were worked through high resistance and low repetition using appropriate weights / sandbag (Choi et al, 2010).

Stretching exercises: Based on assessment, tight muscle groups were stretched according to the recommended stretching procedure for each muscle group (Choi et al, 2010).

Soft tissue massage: This was done to the painful areas of the back with analgesic topical gel [an example is Capsaicin (Dalpiazetal.2004; Gagnier et al, 2007)].Patient was comfortably positioned on the treatment couch and was instructed to relax before this procedure was carried out. Finger and palmar kneading, friction massage, hacking and soothing stroking were employed in treating the patients.(Imamura et al, 2008).

Patient Education: Each patient was educated on good standing and sitting posture. They were instructed to use appropriate/correct type of chair for sitting and firm (not sagging / soft) mattress for sleeping. They were advised to avoid prolonged standing, sitting or walking which may aggravate pain. They were also educated on lifestyle modification and proper lifting technique.

Postural correction / Retraining: Patients with postural deformity such as lateral shift and exaggerated lordosis were trained on how to prevent further deformity by avoiding abnormal posture. Postural correction exercises were taught in order to correct established deformity.

Home programme: This included stabilization, stretching and strengthening exercises. The participants that were managed with Mackenzie protocol were also instructed to treat themselves by carrying out the back flexion or extension exercise at home ((Petering and Webb, 2011; Kinkade, 2007; Shen et al, 2006).

It was ensured that the above protocol was used in all the centres by periodic calling of the research assistants. Ibadan low back pain disability scale and NPRS were re-administered to 64 participants in the EG who completed the five consecutive weeks of physiotherapy to assess responsiveness of ILBPDS.

3.3   Data Analysis

Multiple method of data analysis was used.

1. Data was summarized using mean, median, range and standard deviation for continuous variables, proportion and percentage for categorical variables.

2. Mann – Whitney U-test was used to compare the scores obtained on the ILBPDS for the EG and CG (construct validity).

3. Spearman's correlation coefficient was used to determine the correlation between the NPRS and ILBPDS scores of the EG before treatment. (Divergent (construct) validity), and also between changes in ILBPDS and NPRS scores of the EG post intervention.

4. Wilcoxon Signed Rank Test was used to compare the ILBPDS and NPRS scores of the EGrespectively before and after 5 weeks of physiotherapy (responsiveness).

5. Cronbach's alpha was used to determine the item-to-item correlation of the ILBPDS (internal consistency).

6. Intra-class correlation coefficient (ICC) was used to determine the correlation between scores obtained on ILBPDS by the EG on two different occasions (48 hours interval) (Test-retest reliability).

The significance level was set at 0.05.

4.1 RESULTS

This study was conducted in two phases – development of the ILBPDS (phase 1) and psychometric testing of the ILBPDS (phase 2), hence results of the two phases will be presented separately.

Phase 1 Results

4.1.1 First Pretesting for Comprehensibility

Table 4.1 summarises responses to cognitive debriefing interview following the pretesting. Thirty three out of the 35 individuals who participated in pretesting of the scale for comprehensibility reported that the introductory letter was easy to understand, however, two participants reported that they had difficulty understanding two words. Their normal understanding of the word "instrument" was equipment rather than scale or questionnaire and of "administering" was giving a course of treatment, e.g. injection, rather than giving or distributing copies of a questionnaire. These participants also reported that the term 'usual work' in question 11 of the scale was vague and might be misunderstood as house chores rather than vocation.

The participants understood the items on the scale but five respondents complained that the scale was too long and that many questions were repeated. Four respondents also complained that each of question 12 and 38 was   asking more than one question because more than one activity are involved. [question 12 (Participating in social functions e.g. wedding, naming, birthday and burial ceremonies) and question 38 (Participating in religious gatherings - Jumat or Church services)].

Emphasis on relating time to distance was encouraged by the participants. It was also suggested by one of the participants that two questions should be modified – 'Hand washing of clothes at the floor level' and 'putting on socks/sandals/shoe' was to read 'bending down to wash clothes at floor level 'and 'bending down to put on socks /sandals /shoe'. The participants commended the expression of specific volumes of water in kilogrammes for ease of understanding.

4.1.2 First Experts' Panel Meeting

The outcomes of the expert panel's meeting for item selection based on findings from pretesting for comprehensibility were as follows: modifications of 36 items; merger of six items (items 12&38, 20 &41,25&39) into three; deletion of one item (item 10) and addition of five new items. At the endthe of meeting, a 43 item scale (Appendix 5) instead of 42 item scale (Appendix 3) was produced.

4.1.3 Second Pretesting for Frequency of Endorsement

The proportion of people who gave response alternatives to all items on the ILBPDS was 0% to 50.5%. The acceptable frequency range for item to be retained on the scale should be 20% to 80 % (Streiner and Norman, 2008). Items with any of the response option less than 20% or greater than 80% should be eliminated from the scale (Streiner and Norman, 2008). All the items on the scale had at least one of the response options less than 20%. (See Table 4.2).

The item on 'greeting elders (prostrating, kneeling or squatting) was not endorsed by participants in the east and South-South of Nigerian.It was observed that majority of the respondents did not endorse the response option 'unable to carry out the activity' for any of the items on the scale (Table 4.2).

Table 4.1. Frequency Distribution of Responses to Cognitive Debriefing Interview

Section		Yes	Percentage	No	Percentage	Total
1. Introduction	Easy to understand.	33	94.3	2	5.7	35
	Purpose of the study clearly indicated.	31	88.6	2	11.4	33
2. Overall layout	Easy reading	33	94.3	2	5.7	35
	Pages clearly numbered and stapled.	30	96.8	1	3.2	31
3. Main body of questionnaire	There are words not understood	3	6.2	30	93.8	33
	The questionnaire is of appropriate length	30	85.7	5	14.2	35
	Irrelevant/ unnecessary questions	3	8.8	31	91.2	34
	There are words that have more than one meaning	3	9.4	29	90.6	32
	Question asking more than one question	4	12.5	28	87.5	32
	Response options easy to apply	28	93.3	2	6.7	30

Table 4.2. Frequency of Endorsement of 43 Items

Item		ND (%)	SD (%)	MD (%)	VD (%)	UC (%)	NA (%)
1	Walking (15 - 20mins)	3.4	30.5	22.9	**13.3**	**0.9**	1.9
2	Sitting (30 minutes or less).	45.2	25.0	25.0	**4.8**	**0.0**	1.0
3	Lifting weight (20-25 kg)	**7.1**	25.2	26.3	34.3	**7.1**	6.6
4	Bending at floor level (wash clothes)	**8.2**	32.0	23.7	29.9	**6.2**	5.8
5	Getting up from an office /dining chair	34.0	34.0	20.7	**11.3**	**0.0**	0.0
6	Standing (1 hour or more)	**8.9**	21.8	**19.8**	43.6	**5.9**	1.9
7	Bending to put on underclothes	26.7	30.7	23.8	**17.8**	**1.0**	1.9
8	Sitting (1 hour or longer)	29.3	27.3	22.2	20.2	**1.0**	6.6
9	Getting on a motor bike /motor cycle	35.3	**18.8**	27.1	**18.8**	**0.0**	13.
10	Performing your routine job activities	23.1	29.8	30.8	**14.4**	**1.9**	1.0
11	Religious/social functions	37.3	23.5	23.5	**12.8**	**2.9**	2.9
12	Driving (30mins or less)	48.7	22.4	21.0	**7.9**	**0.0**	24.6
13	Sitting at floor level	24.7	33.7	20.8	**17.8**	**3.0**	3.8
14	Getting up from a modern toilet	39.2	37.2	**16.7**	**6.9**	**0.0**	0.0
15	Sitting on a low chair (30mins or less)	32.7	27.7	21.8	**16.8**	**1.0**	1.0
16	Getting off a high vehicle (Jeep/Bus)	30.2	25.0	28.1	**15.6**	**1.0**	6.8
17	Lifting weight (10kg)	28.1	25.0	21.9	**18.8**	**6.2**	5.9
18	Walking for one hour (3 - 4km)	**15.6**	**16.7**	22.9	35.4	**9.4**	6.8
19	Bending to put on shoes or wash feet	33.7	20.2	**16.3**	25.0	**4.8**	1.9
20	Getting up from a low chair/ stool	24.3	34.9	20.4	20.4	**0.0**	1.0
21	Stand (15-20 minutes)	34.3	28.4	20.6	**14.7**	**2.0**	1.0
22	Getting into a low vehicle (saloon cars)	48.0	27.0	21.0	**3.0**	**1.0**	3.9
23	Having sexual intercourse	33.8	29.6	21.1	**11.3**	**4.2**	26.8
24	Greet (prostrating/ kneeling squatting)	31.3	**16.9**	27.7	20.5	**3.6**	19.4
25	Travelling (1 hour or more)	30.3	23.2	27.3	**16.2**	**3.0**	4.8
26	Night sleep	50.5	20.8	20.8	**6.9**	**1.0**	3.8
27	Driving for 1 hour or more	34.3	25.4	25.4	**13.4**	**1.5**	33.0
28	Drawing water from a well	**13.8**	29.3	22.4	22.4	**12.1**	43.1
29	Getting off a low vehicle (a saloon car)	39.8	23.6	28.0	**8.6**	**0.0**	4.1
30	Going down the stairs	35.0	35.0	21.0	**9.0**	**0.0**	3.9
31	Travel (less than one hour)	47.9	26.0	**19.8**	**5.2**	**1.0**	6.8
32	Getting into a high vehicle (Jeep / Bus)	35.8	26.3	21.0	**15.8**	**1.1**	9.5
33	Squatting (pit toilet/ latrine)	**19.3**	24.1	24.1	28.9	**3.6**	18.6
34	Turning in bed during sleep	34.3	37.2	**16.7**	**10.8**	**1.0**	1.9
35	Sweeping with a broom	**14.8**	32.9	**11.4**	37.5	**3.4**	12.0
36	Farming/ gardening activities	**8.4**	26.8	**18.3**	38.0	**8.5**	33.0
37	Riding a motorbike/ motorcycle	30.2	22.6	22.6	**17.0**	**7.6**	41.1
38	Climbing stairs	**19.4**	30.6	26.5	22.5	**1.0**	3.2
39	Getting up from floor level	25.5	20.6	23.5	30.4	**0.0**	1.9
40	Lifting weight (4-5 kg)	45.4	**18.6**	**15.4**	**16.5**	**4.1**	5.8
41	Sitting on low stool (1 hour or more)	29.6	22.4	27.6	**18.4**	**2.0**	3.0
42	Prolonged kneeling	**17.3**	23.5	23.5	25.5	**10.2**	4.9

ND-Not Difficult at all   **NA**- Not Applicable

VD- Very Difficult   **UC**- Unable to carry out activities

MD- Moderately Difficult

Bolden values are those with frequency of endorsement < 20%

4.1.4 Factor Analysis of 43 Items

Factor analysis showed that the scale had four factor groups. Factors with eigenvalue of ≥ 1 according to the scree test result (Figure 1) were considered to be scale factors. The eigenvalue of first, second, third and fourth factors were 12.97, 3.02, 2.25 and 1.96 respectively.

Factor analysis of the 43 items identified four factors which explained 47% of the total variance (Table4.3).Factor 1 explained 30.2% of the total variance, factor 2 explained 7.0% of the total variance, factor 3 explained 5.2% of the total variance and factor 4 explained 4.6% of the total variance, According to Liau et al, 2011, an item is said to load on a particular factor if the factor loading is ≥ 0.40. The degree of loading of each item on a factor represents its contribution to that factor. Factor 1 had 21 items (postural activities involving lower limb movement) loaded on it, factor 2 had 7 items (activities requiring upper extremity movement) loaded on it, Factor 3 had only three items (activities relating to transportation) loaded on it, while factor 4 had five items (mainly house chore activities requiring bending and use of the upper extremities) loaded on it. Four items (walking for one hour (3 - 4km), having sexual intercourse, greeting elders (prostrating / kneeling /squatting etc.), drawing water from a well) did not load on any of the factors. Three items [participating in religious/social functions (e.g. Church/ Jumat service or wedding/ birthday/ funeral), driving for a short period (30mins or less)], carrying out farming activities or gardening] loaded on more than one factor, two of these items are complex tasks that require more than one elementary activity (Table 4.4). The communalities were all above 0.4.

Table 4.3. Factor Analysis of the 43 items on the Scale after Pretesting

Item		1	2	3	4	Communality(h^2)
1	Walking (15 - 20mins)	**.51**	.09	.18	.09	.71
2	Sitting (30 minutes or less).	**.47**	.00	.34	.33	.76
3	Lifting weight (20-25 kg)	.16	.05	.08	**.41**	.74
4	Bending at floor level(wash clothes)	.27	-.16	-.02	**.62**	.70
5	Getting up from an office /diningchair	**.67**	.07	.15	.07	.60
6	Standing (1 hour or more)	**.42**	.35	.07	-03	.73
7	Bending to put on underclothes	**.66**	-.15	.03	.28	.79
8	Sitting (1 hour or longer)	.36	-.14	.02	**.46**	.78
9	Getting on a motor bike /motor cycle	.15	.01	**-.44**	.34	.69
10	Performing your routine job activities	**.55**	.02	.19	.28	.63
11	Religious/social functions*	.46*	.06	.13	.41*	.73
12	Driving (30mins or less)*	.04	.67*	.42*	.15	.83
13	Sitting at floor level	**.68**	-.01	-.19	.03	.73
14	Getting up from a modern toilet	**.77**	-.00	.21	-07	.67
15	Sitting on a low chair (30mins or less)	**.57**	.02	-.13	.24	.66
16	Getting off a high vehicle (Jeep/Bus)	**.51**	.15	-.30	.06	.81
17	Lifting weight(10kg)	.02	**.51**	-.22	.20	.85
18	Walking for one hour (3 - 4km)	.34	.33	-.19	.17	.74
19	Bending to put on shoes or wash feet	**.73**	-.04	.11	.08	.73
20	Getting up from a low chair/ stool	**.73**	-.06	-.21	-02	.72
21	Standing (15-20 minutes)	**.56**	.34	.12	-04	.64
22	Getting into a low vehicle (salooncars)	.15	**.65**	-.04	-11	.64
23	Having sexual intercourse	.12	.35	.32	.26	.65
24	Greeting (prostrating / kneeling)	.22	.15	-.37	.11	.66
25	Travelling (1 hour or more)	.32	**.53**	.01	-21	.69
26	Night sleep	**.45**	.15	-.21	.02	.66
27	Driving for 1 hour or more	-.02	**.63**	.32	.20	.77
28	Drawing water from a well	-.04	.16	-.19	.39	.55
29	Getting off a low vehicle (a saloon car)	.04	**.72**	-.17	-.06	.80
30	Going down the stairs	**.43**	.30	-.19	-.11	.64
31	Travelling (less than one hour)	.27	**.61**	-.17	-.03	.74
32	Getting into a high vehicle (Jeep / Bus)	**.55**	.04	-.33	-.02	.77
33	Squatting (pit toilet/ latrine)	.23	.04	**-.48**	.16	.68
34	Turning in bed during sleep	**.53**	.16	-.08	.90	.60
35	Sweeping with a broom	.23	-.06	-.18	**.63**	.75
36	Farming/ gardening activities*	-.23	.45*	.03	.45*	.74
37	Riding a motorbike/ motorcycle	-.26	.25	-.31	**.53**	.69
38	Climbing stairs	**.76**	-.01	-.11	-.13	.77
39	Getting up from floor level	**.72**	-.00	-.26	-.03	.77
40	Lifting weight (4-5 kg)	-.12	**.62**	-.21	.15	.87
41	Sitting on low stool (1 hour or more)	**.59**	-.05	-.23	.17	.71
42	Prolonged kneeling	**.51**	.17	.10	-.00	.73
43	Getting off a motorbike/ motorcycle	.01	.28	**-.61**	.19	.81
	% Variance Explained	30.2	7.03	5.20	4.60	
	Eigen value	12.97	3.02	2.25	1.96	

*Items loaded on more than one factor

(h^2) – Communality: the amount of variability in the item explained by the four factors.

Bolden items are those loading on the respective factors under which they appear

Table 4.4. Items Loading on the Four Factors

Item		1	2	3	4	Communality(h²)
				Factors		
1	Walking (15 mins or less)	**.51**	.09	.18	.09	.71
2	Sitting (30 mins or less)	**.47**	.00	.34	.33	.76
5	Getting up from an office/ dining chair	**.67**	.07	.15	.07	.60
6	Standing (1 hour or more)	**.42**	.35	.07	.03	.73
7	Bending to put on underclothes	**.66**	-.15	.03	.28	.79
10	Performing your routine job activities	**.55**	.052	.19	.28	.63
13	Sitting at floor level	**.68**	-.02	-.19	.03	.73
14	Getting up from a modern toilet	**.77**	-.00	.21	-.07	.67
15	Sitting on low chair (30 mins or less)	**.57**	.02	-.13	.24	.66
16	Getting off a high vehicle (jeep/bus)	**.51**	.15	-.29	.06	.81
19	Bending to put on shoes or wash feet	**.73**	-.04	.11	.08	.73
20	Getting up from a low chair or stool	**.73**	-.06	-.21	-.02	.72
21	Standing (15-20 mins)	**.56**	.34	.12	-.04	.65
26	Night sleep	**.45**	.15	-.21	.02	.66
30	Going down the stairs	**.43**	.30	-.19	-.11	.64
32	Getting into a high vehicle (jeep/ bus)	**.57**	.04	-.33	-.02	.77
34	Turning in bed during sleep	**.53**	.16	-.08	.09	.60
38	Climbing stairs	**.76**	-.01	-.12	-.13	.71
39	Getting up from floor level	**.72**	-.00	-.26	-.03	.77
41	Sitting on low stool (1 hour or more)	**.59**	-.05	-.23	.17	.71
42	Prolonged kneeling	**.51**	.17	.10	-.00	.73
17	Lifting weight (10 kg)	.02	**.51**	-.22	.20	.85
22	Getting into a low vehicle saloon cars	.15	**.61**	-.04	-.12	.64
25	Travelling (1 hour or more)	.32	**.53**	.01	-.21	.69
27	Driving for 1 hour or more	-.02	**.63**	.32	.20	.77
29	Getting off a low vehicle (saloon car)	.04	**.72**	-.17	-.06	.80
31	Travelling (less than 1 hour)	.27	**.61**	-.17	-.03	.74
40	Lifting weight (4-5 kg)	-.12	**.62**	-.21	.15	.87
9	Getting on a motorbike/ motorcycle	.15	.01	**-.44**	.34	.69
33	Squatting (pit toilet/ latrine)	.23	.02	**-.48**	.16	.68
43	Getting off a motorbike/ motorcycle	.01	.28	**-.61**	.19	.81
3	Lifting weight (20-25kg)	.16	.05	.08	**.41**	.74
4	Bending at floor level (wash clothes)	.27	-.16	-.02	**.62**	.70
8	Sitting (1 hour or longer)	.36	-.14	.02	**.46**	.60
35	Sweeping with a broom	.23	-.06	-.18	**.63**	.75
37	Riding a motorbike/ motorcycle	-.21	.25	-.31	**.53**	.69
11	Religious/ social functions*	.46*	.06	.13	.41*	.73
12	Driving (30 mins or less)*	.04	.67*	.42*	.15	.83
36	Farming/ gardening activities*	-.23	.45*	.03	.45*	.74
18	Walking for one hour (3-4 km)+	.34	.33	-.19	.17	.74
23	Having sexual intercourse+	.12	.35	.32	.26	.65
24	Greeting (prostrating/ kneeling/ squatting)+	.22	.15	-.37	.11	.66
28	Drawing water from a well+	-.04	.16	-.19	.39	.55

* Items loaded on more than one factor

+ Items not loaded on any factor

Bolden items are those loading on the respective factors under which they appear
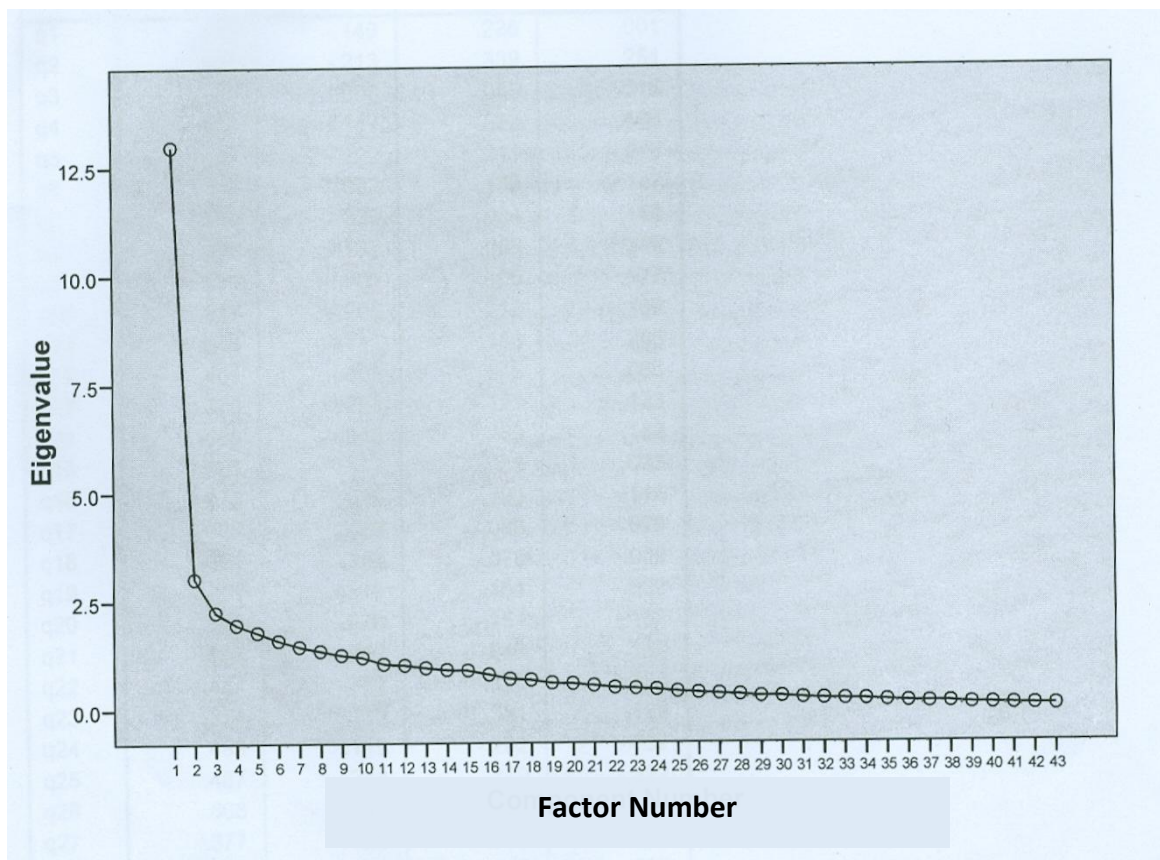
Figure 4.1. Scree plot of initial 43 items on ILBPDS

4.1.5 Second Experts' Panel Meeting for Final Item Selection

Items that scored at least two points out of the four points with the exception of those that were regarded as repetition were retained on the scale. The outcome of this meeting was deletion of 25 items from the scale and selection of 18 items which were retained on the scale (Table 4.5)

4.1.6 Factor Analysis of the 18 items on the Final Scale.

Factor analysis showed that the scale had two factor groups. Factors with eigenvalue of ≥ 1 according to the scree test result (Figure 4.2) were considered to be scale factors. The eigenvalue of first and second factors were 9.3 and 1.6 respectively. The two factors represent 55.2% of the total variance. Factor 1 explained 49.1% of the total variance while factor two explained 6.2% of the total variance. Items with factor loading of ≥ 0.5 were considered to load under a factor (Liau et al,2011). The first group had 13 items loaded on it (Items1-6, 8, 10, 13, 15 – 18) ; they are Common Activities of Daily Living (CADL). The second factor group had 4 items loaded on it (Items 7, 9, 11, 12); these are mainly Culture Specific Activities of Daily Living (CSADL) (Table 4.6).

Item 14 did not load on any factor. However, it was decided that item 14 (having sexual intercourse) should be part of the scale based on experts' opinion and participants comments. The final copy of the ILBPDS is shown in Appendix 9.

Table 4.5 Second Experts Panel Meeting for Final Item Selection

Item	Decision	Reason
1.Walking (15 - 20mins)**(2pt)** 3. Lifting weight (20-25 kg)**(2pt)** 6. Standing (1 hour or more)**(2pt)** 7. Bending to put on underclothes**(2pt)** 21. Standing (15 - 20 mins)**(2pt)** 26. Night sleep**(2pt)** 38. Climbing stairs**(2pt)**	Retained	They were ADL affected by LBP
2. Sitting (30 mins or less)**(2pt)** 8. Sitting (1 hour or longer)**(2pt)** 15. Sitting low chair (30 mins or less)**(2pt)** 41. Sitting on low stool (1 hour or more)**(2pt)**	Merged into one item on sitting	They were all activities related to sitting
4. Bending to wash clothes (floor level) **(2pt)** 13. Sitting at floor level**(3pt)** 20. Getting up from a low chair or stool**(3pt)** 24. Greeting (prostrating/ kneeling/ squatting)**(2pt)** 28. Drawing water from a well**(2pt)** 33. Squatting to usepit toilet/latrine**(2pt)** 35. Sweeping with a broom**(2pt)** 36. Farming / gardening activities**(2pt)** 39. Getting up from floor level**(3pt)**	Retained	They were considered as activities of cultural relevance
23. Having sexual intercourse**(2pt)**	Retained	Considered important to the target population
5. Getting up from an office/dinning chair**(2pt)** 14. Getting up from a modern toilet**(1pt)**	Deleted	They were considered as repetition of item 20 by target population
17. Lifting weight (10 kg)**(2pt)** 40. Lifting weight (4-5 kg)**(2pt)**	Deleted	They were considered as repetition of item 3 by target population
18. Walking for 1 hour (3-4 km)**(1pt)**	Deleted	It was considered as repetition of item 1
19. Bending to put on shoes or wash feet**(1pt)**	Deleted	It was considered to require similar functional movement as item 7
9. Getting on a motorbike / motorcycle**(0pt)** 16. Getting off a high vehicle (jeep/bus)**(0pt)** 22. Getting into a low vehicle saloon cars**(1pt)** 29. Getting off a low vehicle (saloon car)**(1pt)** 30. Going down the stairs**(1pt)** 32. Getting into a high vehicle**(0pt)** 37. Riding a motorbike / motorcycle**(0pt)** 42. Prolonged kneeling**(0pt)** 43. Getting off a motorbike/motorcycle**(0pt)**	Deleted	They were considered not relevant to assessment of patients with LBP
25. Travelling (1hour or more)**(2pt)** 31. Travelling (less than 1 hour)**(1pt)**	Deleted	Considered to be a repetition of items on sitting
10. Performing your routine job activities**(2pt)** 11. Religious/social functions**(1pt)**	Deleted	They were considered to be vague/ambiguous
34. Turning in bed during sleep**(2pt)**	Deleted	Considered as a repetition of item 26
12. Driving (30mins or less)**(1pt)** 27. Driving for 1 hour or more**(1pt)**	Deleted	They were considered not applicable to all participants

The bolden figures in bracket are the scores of each item.

pt - point

Table 4.6 – Factor Analysis of the 18 Items on the Final Version of ILBPDS

Item		Factors		Communality (h^2)
		CADL	CSADL	
1	Walking (15 – 20mins)	**.76**	.18	.64
2	Sitting on a chair (1 hour or more).	**.70**	.22	.54
3	Lifting heavy weight	**.75**	.28	.64
4	Standing (15 - 20 minutes)	**.84**	.21	.75
5	Bending (wash clothes floor level)	**.65**	.41	.60
6	Climbing the stairs	**.64**	.40	.57
8	Sitting at floor level	**.63**	.42	.58
10	Standing (1 hour or more)	**.89**	.20	.83
13	Getting up from floor level	**.71**	.43	.68
15	Sleeping at the night	**.53**	.19	.32
16	Sweeping with a broom	**.59**	.45	.55
17	Getting up from a low chair/ low stool	**.64**	.40	.58
18	Bending to put on underclothes.	**.74**	.33	.65
7	Greeting (kneeling/ prostrating)	.14	**.64**	.43
9	Farming/gardening activities	.41	**.53**	.45
11	Drawing water from the well	.18	**.70**	.51
12	Squatting (pit toilet/ latrine).	.21	**.70**	.53
14	Sexual intercourse *	.33	.06	.11
	% Variance explained	51.50	8.83	
	Eigen value	9.27	1.59	

* Item not loaded on any factor

(h^2) – Communality: the amount of variability in the item explained by the two factors

CADL – Common Activities of Daily Living

CSADL– Culture Specific Activities of Daily Living

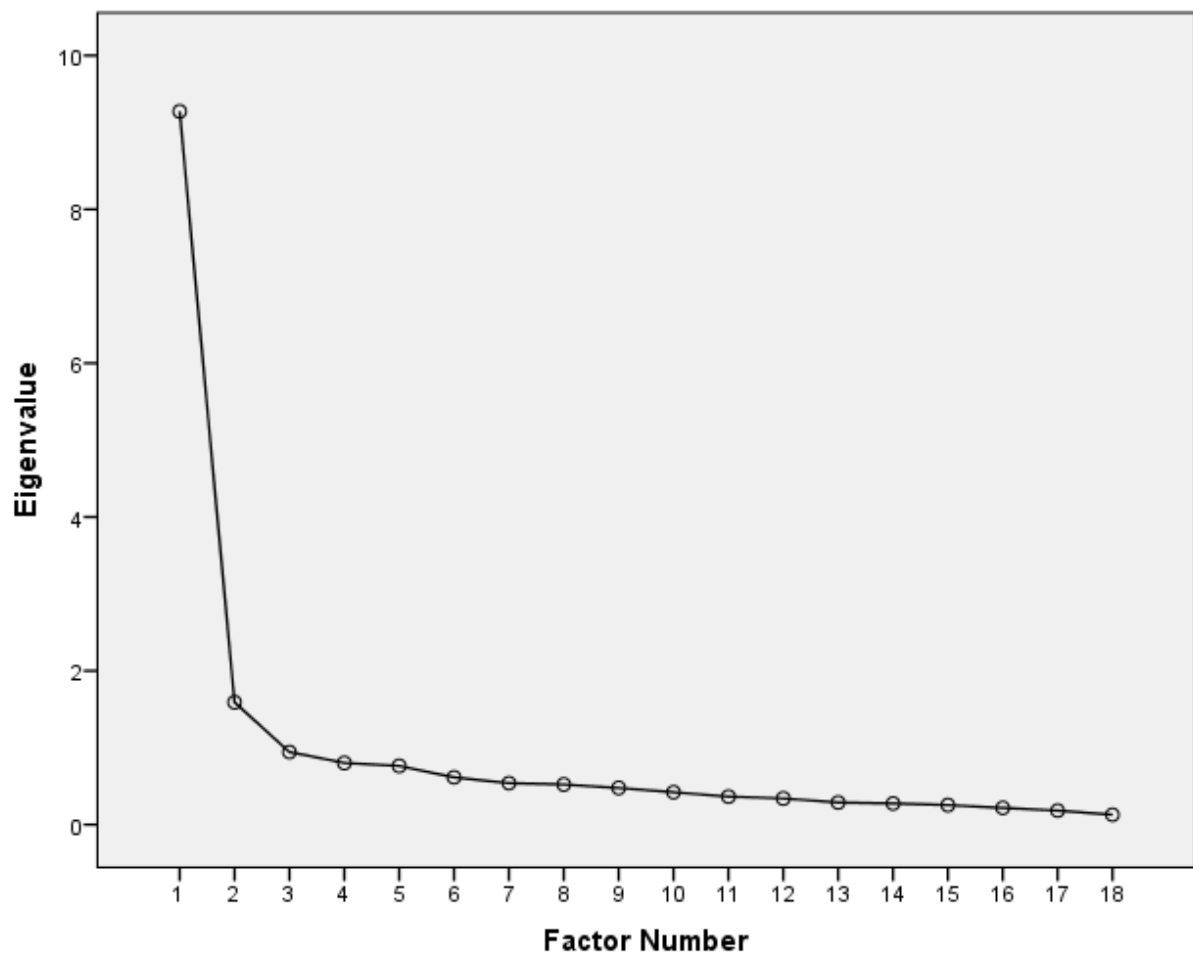Bolden items are those loading on the respective factors under which they appear.

Figure 4.2. Scree Plot of final 18 items on ILBPDS

Phase 2 Results

4.1.7   Demographic Characteristics of Participants

Construct Validity of ILBPDS: This involved one hundred and forty-two patients (69 male, 73 female) with LBP (Experimental Group - EG) and one hundred and forty-two (69 male, 73 female) age and sex matched apparently healthy controls without symptoms of LBP (Control Group - CG).The mean age of the EG(48.6 ± 12.7 years) and that of the CG (48.2 ± 12.3 years) did not differ significantly. The highest percentage of the participants in the EG (34%) were within the age group 50 – 59 years (Figure 4.3).Most participants were married in EG (73%) and CG(66%).Table 4.7 shows that the percentages of participants who were single, divorced and widowed were small. Most participants were semi-skilled in the EG (60%) and CG (41%). The percentages of participants who were skilled, unskilled, pensioner and student were small (Table 4.7).

Reliability of ILBPDS: One hundred and thirteen patients (56 male, 57 female) with LBP in the EG participated in the investigation of ILBPDS for test of reliability. Their mean age was 49.2 ± 12.6 years.

Responsiveness of ILBPDS: Sixty-four patients with LBP (28 male, 36 female) participated in investigation of ILBPDS for evidence of responsiveness. Their mean age was 49.5 ± 12.6 years

4.1.8 Clinical Status of the Participants in the EG

The median duration of LBP for the patients in EG was 24 months (9 – 60 months) and 75.6% of them had pain / paraesthesia radiating to one / two lower limbs while 24.7% had pain on the low back region only. The median NPRS score was 7.0(5.0-8.0) at the time of the first visit. During the first appointment, 81.0% were on medication, 2.8% were on both medication and traditional treatment, 0.7% were on traditional treatment alone while 15.5% were not on medication or traditional treatment.

Table 4.7 Demographic Characteristics of Participants

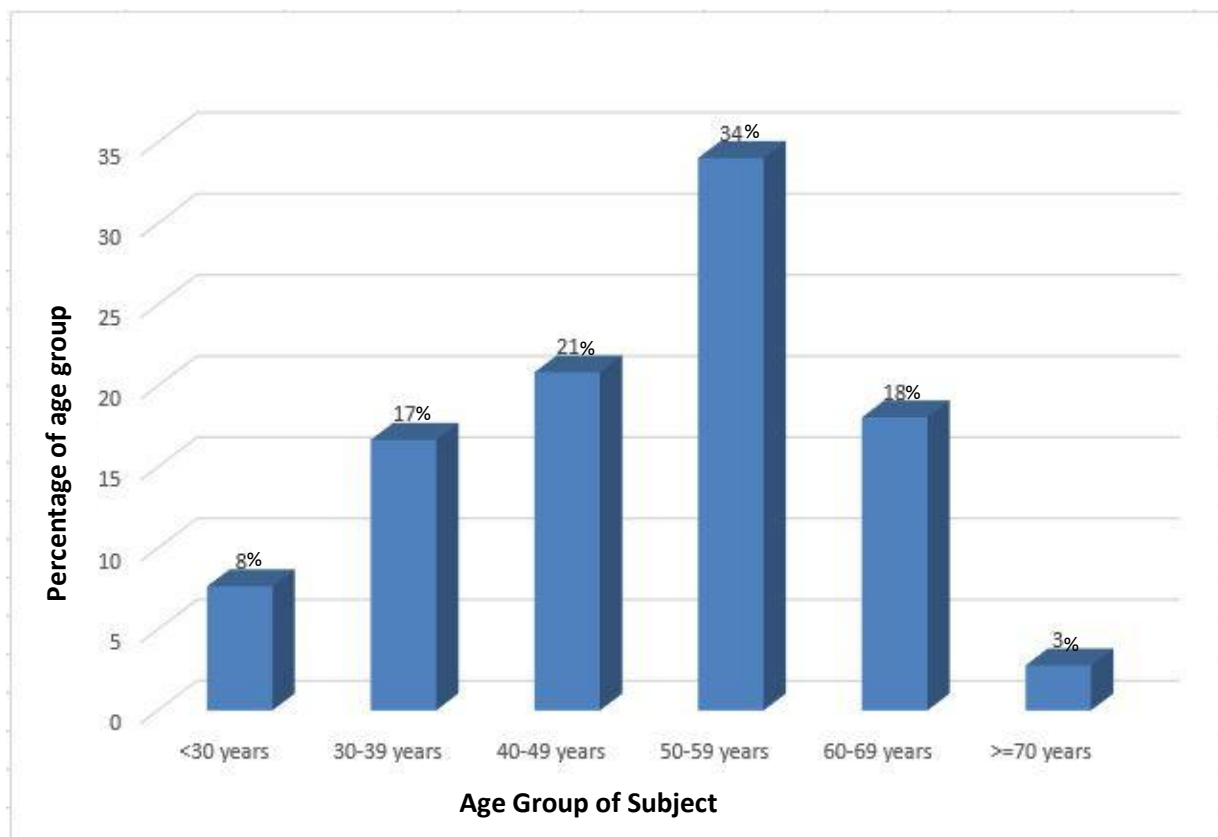| Variable	EG(N=142)		CG(N=142)	
	n	%	n	%
Sex				
Male	69	48.6	69	48.6
Female	73	51.4	73	51.4
Marital Status				
Single	19	13.4	19	13.4
Married	113	79.6	94	66.2
Divorced	5	3.5	2	1.4
Widowed	5	3.5	3	2.1
Unspecified	0	0	24	16.9
Occupation				
Skilled	28	19.7	40	28.2
Semi-skilled	79	55.6	54	38.0
Unskilled	7	4.9	2	1.4
Pensioner	21	14.9	14	9.9
Student	4	2.8	5	3.5
Unspecified	3	2.1	27	19.0

99

Figure 4.3. Age Group of Participants in the EG

4.1.9   Testing ILBPDS for Construct Validity

The median (range) ILBPDS score of the EG [55.2 (45.6-65.6)] was significantly higher (p=0.0001) than that of CG [21.4 (20-24.3)] (Table 4.8). The median NPRS score of the EG [7.0 (5.0-8.0)] correlated significantly with their ILBPDS score (r = 0.50,p = 0.0001)) at baseline (Table 4.8). The median NPRS score of the EG [4.0 (3.0 -5.0)] correlated significantly (r= 0.35, (p= 0.0001) with their ILBPDS score after 5 weeks of physiotherapy (Table 4.9).

4.1.10   Testing ILBPDS for Reliability and Internal Consistency

The median (range) ILBPDS score at first visit [55.7(45.3-65.6)] and 48 hours later [51.8(41.3 − 65.7)] were significantly correlated (ICC = 0.8, p< 0.05).The internal consistency of ILBPDS was demonstrated by a Cronbach's alpha of 0.84.

4.1.11   Testing ILBPDS for Responsiveness

The median (range) ILBPDS score after the 5 weeks physiotherapy 36.5(30.0-49.4), was significantly lower (P < 0.05) than that before treatment 55.2(45.6-65.6) (Table 4.10). There was a significant moderate correlation (r = 0.62, p < 0.05) between change in ILBPDS score and change in NPRS score of the EG after 5 weeks of physiotherapy (Table4.9).

Table 4.8 Comparison of ILBPDS Score of EG and CG using the Mann-Whitney U Test

EG Median (Range)	CG Median (Range)	z-value	p-value
55.2 (45.6-65.6)	21.4 (20-24.3)	-14.49	0.0001

Table 4.9 Spearman's Correlation Coefficient between ILBPDS and NPRS Score of EG.

	Pre – Rx NPRS Score r(p)	Post – Rx NPRS Score r(p)	Post – Rx Change in NPRS Score r(p)
Pre-Rx ILBPDS Score	0.50 (p = 0.0001)		
Post – Rx ILBPDS Score		0.35 (p = 0.0001)	
Post – Rx Change in ILBPDS Score			0.62 (p = 0.0001)

Rx - treatment

Table 4.10 Wilcoxon Signed Rank Test for Pre & Post Treatment of ILBPDS and NPRS Scores of EG

Variable	Pre Treatment Median (Range)	Post Treatment Median (Range)	z- value	p-value
NPRS Score	7.0(5.0-8.0)	4.0 (3.0-5.0)	6.207	0.0001
ILBPDS Score	55.2(45.6-65.6)	36.5(30.0-49.4)	-6.130	0.0001

4.1.12 Hypotheses Testing

Hypothesis 1

Statement: There would be no significant difference between the ILBPDS scores of the EG and CG (construct validity).

Test statistic: Mann − Whitney U-test

Observed values: $z = -14.49$ $p = 0.0001$

Judgement: since $p < 0.05$, the hypothesis was rejected and it was concluded that there was a significant difference between the ILBPDS scores of the EG and CG (construct validity).

Hypothesis 2

Statement: There would be no significant correlation between the NPRS and ILBPDS scores of the EG at baseline (construct [divergent] validity).

Test statistic: Spearman correlation coefficient.

Observed values: $r = 0.50$ $p = 0.0001$

Judgement: since p is < 0.05 the hypothesis was rejected and it was concluded that there was a significant correlation between the NPRS and ILBPDS scores of the EG at baseline (construct [divergent] validity).

Hypothesis 3

Statement: There would be no significant correlation between the NPRS and ILBPDS scores of the EG post intervention (construct [divergent] validity).

Test statistic: Spearman correlation coefficient.

Observed values: $r = 0.35$ $p = 0.0001$

Judgement: since p is < 0.05 the hypothesis was rejected and it was concluded that there was a significant correlation between the NPRS and ILBPDS scores of the EG post intervention (construct [divergent] validity).

Hypothesis 4

Statement: There would be no significant difference between the ILBPDS score of the EG in this study pre and post intervention (responsiveness).

Test statistic: Wilcoxon Signed Rank Test

Observed values: z= -6.13   p = 0.0001

Judgement: since p is < 0.05, the hypothesis was rejected and it was concluded that there was a significant difference between the ILBPDS score of the EG in this study pre and post intervention (responsiveness).

Hypothesis 5

Statement: There would be no significant difference between the NPRS scores of the EG in this study pre and post intervention (responsiveness).

Test statistic: Wilcoxon Signed Rank Test

Observed values: z = 6.21, p =0.0001

Judgement: since p is < 0.05, the hypothesis was rejected and it was concluded that there was a significant difference between the NPRS scores of the EG in this study pre and post intervention (responsiveness).

Hypothesis 6

Statement: The item-item correlation of the ILBPDS would not be significant (internal consistency).

Test statistic: Cronbach alpha

Observed values: α = 0.84, p =0.0001

Judgement: since p is < 0.05, the hypothesis was rejected and it was concluded that the item-item correlation of the ILBPDS was significant (internal consistency).

Hypothesis 7

Statement: There would be no significant correlation between the ILBPDS scores of the EG taken on two different occasions (reliability).

Test statistic: Intraclass correlation

Observed values: ICC = 0.80, p =0.0001

Judgement: since p is < 0.05, the hypothesis was rejected and it was concluded that there was a significant correlation between the ILBPDS scores of the EG obtained on two different occasions (reliability).

Hypothesis 8

Statement: There would be no significant correlation between the changes in NPRS and ILBPDS scores of the EG post intervention (responsiveness).

Test statistic: Spearman correlation coefficient.

Observed values: r =0.62, p =0.0001

Judgement: since p is < 0.05 the hypothesis was rejected and it was concluded that there was a significant correlation between the changes in NPRS and ILBPDS scores of the EG post intervention (responsiveness).

4.2   DISCUSSION

4.2.1 Development of ILBPDS

Standardized low back disabilityscales are not only suitable for routine assessment of treatment outcome but also provide high quality information for research (Durouz et al,2013).Although there are many disability scales that have been developed for evaluation of patients with low back pain (Fairbank et al,1980, Million et al,1981,Roland and Morris, 1983,Wadell and Main,1984, Greenough and Fraser,1992 , Ruta et al,1994, Daltroy et al, 1994,Manniche et al, 1994, Kopec et al,1995, Durouz et al, 2013).Some of these disability scales have been translated and cross culturally adapted into different languages to promote their wide use in different populations. Many of these scales do not include some activities that are important to patients with LBP in Nigeria.In addition, some of them also include activities that an average Nigerian may not perform in a lifetime and terms that he may not easily understand (Manniche et al, 1994, Kopec et al, 1995). None of these scales has been translated or cross culturally adapted into any of the major Nigerian languages.

The initial intention of the researcher therefore was to modify the ODI and then adapt it to the Yoruba culture. This is to comply with the injunction of Streiner and Norman, (2008) that researchers should adapt existing scales to a new environment and culture instead of developing new ones. This is because a lot of effort has gone into the development of any scale no matter how simple it may appear. Fairbank (the developer of ODI) was contacted to obtain permission to modify ODI. The purpose of this modification was to include in ODI the seven activities that the patients with LBP attending out-patient Physiotherapy clinic of the UCH, Ibadan have reported to have difficulty with during an interview. However, Fairbank agreed that the scale could be translated into Yoruba but did not give permission for any modification (Appendix 1). This study was then aimed at developing a new scale that would accommodate

the activities of cultural importance that are not included in ODI or any of the other LBP scales that were reviewed and to test its psychometric properties.

The researcher's experience in developing ILBPDS indicated that pretesting of the initial draft of the scale on a sample of the target population is essential (Streiner and Norman, 2008). For example, although it was thought that sitting and lifting were important to evaluation of treatment outcomes in LBP, the feedback from pretesting indicated that the target population perceived many questions related to sitting on chairs of different heights for different duration as repetitions. They also considered many questions related to lifting loads of different weights as repetitions. This suggests that patients do not perceive issues as healthcare providers do (Feise et al, 2001). It also emphasized the importance of involving patients in scale development. The participants at the stage of pretesting also demonstrated a lot of interest in completing the questionnaire. Consequently, they provided useful suggestions, such as the need to reduce the number of items on the scale, to remove vague words/items, to use less technical words and to add some phrases to some items to make them clearer (Streiner and Norman,2008).

The level of the participants' understanding of the purpose of the scale was impressive. This suggests that patients will appreciate it if the scale is incorporated into assessment of patients with LBP in the Nigeria clinical setting. Many participants at pretesting phase expressed the facts that the addition of time needed to cover distance in a particular item facilitated easy understanding. This suggests that in developing scales for use in the Nigeria clinical setting, researchers should endeavour to add time requiredto a distance because many Nigerians cannot easily conceptualise distance expressed in miles, blocks and kilometres (Akinpelu et al, 2007, Meenan et al, 1980).

In this study, frequency of endorsement indicated redundancy of all the items. Therefore no item could be deleted from the scale based on frequency of endorsement since there can be no scale without items. This finding showed that the usefulness of this method is limited where there are many response options, it can only be used in scales with dichotomous (Yes / No) response options(Buchanan, 2007).The observation that majority of the respondents did not endorse the response option 'unable to carry out the activity' for many of the items on the scale during pretesting for frequency of endorsement showed that Nigerians will not easily accept they are unable to carry out some activities. This may be because endorsing the response option 'unable to carry out the activity' was seen as a personal weakness on their part. The observation that the item on greeting elders (prostrating, kneeling or squatting) was not endorsed by participants in the east and South-South of Nigerian shows that Nigeria is a multi-culture nation. This was put into consideration during the process of developing ILBPDS because the scoring system allows scores to be computed based on the number of items endorsed by the participants (Fairbank et al,1980, Ruta et al,1994, Akinpelu et al,2007).

The results from factor analysis of the 43 items indicated that factor analysis though very useful in guiding item selection should not be used in isolation (McDowell and Newell,1996). The fact that some items loaded on more than one factor and some cultural relevant items did not load on any factor suggested the need to use another method for item selection.   Factor analysis of the 18 items finally selected showed 17 of the items loaded on two factors. Thirteen loaded on factor 1 which could be best described as Common Activities of Daily Living (CADL), four items loaded on factor 2 which could be described as Culture Specific Activities of Daily Living (CSADL) and the last item (on sexual activity) did not load on any factor but was retained because of its importance to the target population. The feedback from the target population and the experts in this study showed that sexual activity is very important to the

target population and supports the fact that it is one of the activities that are affected by low back pain (Fairbank et al, 1983, Wadell and Main, 1984, Greenough and Fraser, 1992, Daltroy et al, 1996).

4.2.2. Demographic Characteristic of Participants

The patients for the validation study were 69 males and 73 females. The distribution of the sexes in this study is in agreement with the findings of Walker 2000 and Vingard 2005, which stated that men and women report about the same rate of prevalence of LBP with a slight increase among women especially for severe low back pain. It was difficult to get individuals above 40 years who did not have symptoms of low back pain especially among the females.

4.2.3   Construct Validity of ILBPDS

The finding that ILBPDS score of the EG was significantly higher than that of CG confirms that the ILBPDS is able to differentiate between patients with symptoms of LBP and those without symptoms of LBP.The significantly higher ILBPDS scores of the EG also shows that the participants in the EG have more functional disabilities than their apparently healthy controls since higher scores indicates higher levels of functional disabilities. This supports the fact that low back pain has impact on the functional status of an individuals. The results of this study corroborates with that of Akinpelu et al, 2007   who found a significant difference between IKHOAM (Ibadan Knee /Hip Osteoarthritis Outcome Measure) score of the patients with knee/ hip Osteoarthritis and their apparently healthy controls while validating IKHOAM. The finding that the median ILBPDS score of the control group [21.4 (20 -24.3)] is above 20% showed that the control group have some degree of difficulty in performing the activities of daily living included in ILBPDS.

In addition, the significant correlation between ILBPDS score and pain intensity score on NPRS (r = 0.50) at baseline and (r = 0.35) after a 5 week physiotherapy for the EG further supports the evidence of construct (divergent) validity of ILBPDS and implies that the functional disabilities experienced by the patients is related to low back pain. The correlations between ILBPDS score and NPRS score are not strong as expected because they measure different construct. It has been shown that correlations between 0.20 and 0.60 and very rarely above 0.70 were found when comparing dissimilar constructs (Roos et al, 1999). The results of this study are in agreement with other studies where there was a weak to moderate correlation between pain scores and functional disabilities (Waddell and Main, 1984, Turk and Rudy, 1987, Rainsville et al, 1992). The results of this study is also similar to that of Kopec (1995) who found a moderate correlation of 0.54 between a 7- point scale of pain and Quebec pain Disability scale and to that of Daltroy et al, 1996 who reported a moderate correlation between VAS and ODI (r=0.62, n=94).The results of this study showed the evidence that the ILBPDS meets the criterion for construct validity.

4.2.4 Reliability and Internal Consistency of ILBPDS

A high Intra-class correlation of 0.80 was obtained in this study, indicating that ILBPDS is reliable. Kopec et al (1995) found a higher ICC of 0.92 between repeated measures with almost 4 days interval for QBDS and Roland and Morris, (1983) found a test-retest reliability correlations of 0.91 (same day) for RMDQ . The lower ICC obtained in our study compared to that reported by Roland and Morris could be due to the fact that retest was done on the 3rd day, while their retest was performed on the same day as the initial evaluation therefore, memory effects may have contributed to higher test –retest reliability than observed in the present study. Furthermore, this study reported only Pearson's correlation coefficient, which usually is higher

than the intraclass correlation coefficient. There was a higher ICC 0f 0.92reported by Kopec et al, 1995, despite the fact that there was a longer interval between repeated measures in their study than the present study. The differences in the methods and population used in their study may have contributed to higher ICC than observed in the present study. The internal consistency which indicates the homogeneity among the different domains of the questionnaire is another measure of reliability. A Cronbach's   alpha of ≥ 0.70 is always considered as acceptable (George and Mallary, 2003). The Cronbach's alpha of 0.84 obtained in our study indicates that all items in the scale are correlated with one another and provides assurance that random errors are minimized. This result is comparable to that of Jarvikoski (1995) who reported a Cronbach alpha of 0.84 for RMDQ, and to that of Kopec (1996), who found 0.87 using version 2.0 of ODI. The result of this study is also similar to that of Holt et al (2002) who found a Cronbach alpha of 0.85 for LBOS and to that of Salen et al (1994) who reported a Cronbach alpha of 0.84 for DRI.

4.2.5   Responsiveness of ILBPDS

The finding that ILBPDS score of the EG was significantly lower after five weeks of physiotherapy compared to that at baseline provides evidence of adequate degree of responsiveness of ILBPDS. The significant correlation between change in ILBPDS score and change in NPRS score of the EG further supports the evidence that ILBPDS is a responsive outcome measure. The hypothesis that there would be no significant correlation between changes in ILBPDS and NPRS scores after 5 weeks of physiotherapy was therefore rejected. This shows that ILBPDS is able to detect change in patients' condition / disability level over time. This means that ILBPDS can be used to monitor the progress of patients with back pain participating in treatment or rehabilitation programs.

CHAPTER FIVE

SUMMARY, CONCLUSION AND RECOMMENDATION

5.1 Summary

This study was aimed at developing a Nigerian culture and environmental friendly scale (ILBPDS) for assessing functional disability in patients with LBP in the Nigeria clinical setting and to test for its psychometric properties – validity, reliability, internal consistency and responsiveness. There are many scales that have been developed for patients with LBP for different populations and some have been translated and culturally adapted into different languages to promote their wide use in different populations. None of these scales was developed for Nigerian culture and environment therefore ILBPDS was developed in order to contribute to Nigerian environment biased and culture friendly outcome measures. ILBPDS is a self- report questionnaire containing 18 items which measure the functional disability and assess the outcome/effectiveness of therapeutic interventions in patients with LBP. Literature was reviewed under the sub-headings: epidemiology, aetiology/risk factors, classification, clinical features and management, outcome measure/assessment, steps in developing outcome measures, review of existing standardised outcome measures in LBP.

The sample sizes of participants varied with the investigation. One hundred and forty-two patients (69 male, 73 female) with LBP [Experimental Group (EG)] recruited through a consecutive sampling and one hundred and forty-two (69 male, 73 female) age and sex matched apparently healthy controls without symptoms of LBP [Control Group (CG)] participated in the investigation of the ILBPDS for construct validity.   ILBPDS was found to have evidence of satisfactory construct validity with statistically significant difference in the

ILBPDS scores of patients in the EG and their age and sex matched peers (CG). The pain score correlated moderately (r = 0.50) and significantly (p <0.05) with ILBPDS score of the patients with LBP demonstrating an evidence of divergent construct validity of ILBPDS. One hundred and thirteen patients with LBP in the EG completed ILBPDS twice with two days interval in order to investigate the reliability of ILBPDS. The ICC of 0.80 obtained was satisfactory. Internal consistency was satisfactory with Cronbach's alpha of 0.84. Sixty-four patients in the EG (28 M, 36 F) completed ILBPDS before and after a five week physiotherapy to test ILBPDS for responsiveness.   There was statistically significant difference between ILBPDS scores before and after the 5 – week physiotherapy showing that ILBPDS has evidence of satisfactory responsiveness. It was inferred that the ILBPDS demonstrated adequate evidence of construct validity, reliability, internal consistency and responsiveness.

5.2   Conclusion

The study developed a Nigerian culture and environmental friendly 18 item scale which is short, precise, takes about 5-7 minutes to administer and 2-3 minutes to score. It contains two subscales common ADL and culture specific ADL. Participants with symptoms of LBP have higher functional disability than those without symptoms of LBP. Participants with greater pain level also have higher level of functional disability at baseline and after physiotherapy treatment, therefore pain is a major determinant of functional disability. Ibadan LBP Disability Scale is valid, it measures what it is supposed to measure.

The functional disability experienced by participants with LBP at two different occasions separated by two days in this study is consistent, showing that ILBPDS is reliable. The items

on ILBPDS are related to one another indicating that they are all measuring and assessing different aspects of LBP, indicating that ILBPDS is internally consistent.

Participants with LBP have decreased pain and functional disability after physiotherapy treatment. The reduction in pain level also led to reduced functional disability of the participants with LBP. Ibadan LBP Disability Scale is able to pick the changes in functional disability of the participants with LBP, that is, it is responsive.

Ibadan LBP Disability Scale is a Nigerian culture and environment friendly scale. It is a valid, reliable and responsive scale for measuring disability and treatment outcome in LBP.

5.3   Recommendations

The following are recommended based on the findings of this study:

1) The ILBPDS could be used for measuring disability and treatment outcomes in LBP in the Nigeria clinical setting.

2) Clinicians involved in management of patients with LBP should integrate the use of ILBPDS into their daily practice.

3) Future studies should further validate the scale.

4) Future studies should translate ILBPDS into the major indigenous Nigerian languages to ensure that patients who could not speak or read English are not excluded from assessment using ILBPDS

5) The clinicians involved with management of LBP patients should promote awareness programmes on how to prevent LBP among the general public.

6) Physiotherapists should be involved in the ergonomics of the work place to eliminate work related risk factors associated with LBP.

REFERENCES

Adegoke, B.O.A., Akodu, A.K. and Oyeyemi, A.L. 2008. Work – related musculoskeletal disorders among Nigerian Physiotherapists.*Biomed Central Musculoskeletal Disorders* 9:112.

Akinpelu, A.O. and Olowe, O.O. 2002. Correlative Study of 3 Pain Rating Scales Among Obstetric Patients. *African Journal of Medical Science* 31:123-126.

Akinpelu, A.O. and Eluchie, N. C. 2006. Familiarity with knowledge and utilisation of standardized outcome measures among physiotherapists in Nigeria. *Physiotherapy Theory and Practice* 22.2: 61 – 72.

Akinpelu, A.O., Maruf, F. and Adegoke, B. 2006. Validation of a Yoruba Translation of the World Health Organization of Quality of Life Scale Short from among Stroke Survivors in Southwest Nigeria. *African Journal of Medicine and Medical Sciences* 35: 417-424.

Akinpelu, O.A, Odetunde, M.O. and Odole, C.A. 2012. Cross cultural adaptation and initial validation of the stroke specific quality of life scale into the Yoruba language. *International Journal of Rehabilitation Research* 35.4:339-344.

Akinpelu, A.O., Odole, A.C., Adegoke, B.O.A. and Adeniyi, A.F. 2007. Development and initial validation of the Ibadan Knee/Hip Osteoarthritis Outcome Measure. *South African Journal of Physiotherapy* 63.2:3-8.

Akinpelu, A.O., Odole, A.C. andOdejide, A.S. 2010. Prevalence and Pattern of Musculoskeletal Pain in a Rural Community in Southwestern Nigeria. *The Internet Journal of Epidemiology*8.2:1-13.

Akinpelu, A.O., Odole, A.C. and Raheem, S. 2011. Minimal Clinically Important Difference of Ibadan Knee Hip Osteo Arthritis Outcome Measure (IKHOAM) - A Cross-Sectional Study. *Journal of Physical Therapy* 3.2:52-60.

Atlas, S.J., Deyo, R.A., van den Ancker, M., Singer, D.E., Keller, R.B. and Patrick D.L. 2003. The Maine-Seattle Back Questionnaire: a 12-item disability questionnaire for evaluating patients with lumbar sciatica or stenosis. Results of a derivation and validation cohort analysis. *Spine* 28:1869–1876.

Basler, H., Jakle, C. and Kroner-Herwig, B. 1997. Incorporation of cognitive behavorial treatment into the medical care of chronic low back patients: a controlled randomized study in German pain treatment centers. *Patient Education and Counselling* 31:113–124.

Beaton, D.E., Boers, M. and Wells, G.A. 2002. Many faces of the minimal clinically important difference (MCID): a literature review and directions for future research. *Current Opinion in Rheumatology* 14:109 -114.

Beaton, D.E., Bombardier, C., Kartz, J.N. and Wright, G.J. 2001. Taxonomy for responsiveness. *Journal of Clinical Epidemiology* 54.12:1204 - 1217.

Beaton, D, Claire, M. and Guillemin, F. 2000. Guidelines for the process of Cross-Cultural Adaptation of Self Report Measures. *Spine*. 24:3186 – 3191.

Beck, A. 1967.*Depression: Clinical, Experimental and Theoretical Aspects*New York, NY: Harper & Row.

Becker, J.A. and Stumbo, J.R. 2013. Back pain in adults. *Primary Care* 40.2:271-288.

Bergner, M., Bobbitt, R., Carter, W.B. and Gilson, B.S. 1981. The Sickness Index Profile: Development and final revision of health status measure. *Medical Care*19:787 – 809.

Beurskens, A.J., de Vet, H.C. and Köke, A.J. 1996. Responsiveness of functional status in low back pain: a comparison of different instruments. *Pain*65.1:71-76.

Bigos, S., Battle, M. and Spengler, D.M. 1991. A prospective study of work perceptions and psychosocial factors affecting the report of back injury. *Spine* 16:1 – 6.

Black, N. 2013. Patient reported outcome measures could help transform healthcare. *The British Medical Journal* 28.346:f167.

Bombardier, C. 2000. Outcome assessments in the evaluation of treatment of spinal disorders: summary and general recommendations. *Spine*25:3100–3103.

Bond, S. and Thomas, L. 1991. Issues in measuring outcomes in Nursing. *Journal of Advanced Nursing* 16:1492–1520.

Bolton, J.E. and Breen, A.C. 1999. The Bournemouth Questionnaire: a short form comprehensive outcome measure: Psychometric properties in back pain patients. *Journal of manipulative physiology and therapeutics* 22:503–510.

Boscainos, P.J., Sapkas, G., Stilianessi, E., Prouskas, K. and Papadakis, S.A. 2003. Greek versions of the Oswestry and Roland–Morris Disability Questionnaires. *Clinical Orthopedics and Related Research*411:40–53.

Boulter, L.M., van Tulder, M.W. and Boes, B.W. 1998. Methodological issues in low back pain in primary health care. *Spine* 23:2014-2020.

Bronfort, G., Haas, M., Evans, R., Kawchuk, G. and Dagenais, S. 2008. Evidence-informed management of chronic low back pain with spinal manipulation and mobilization. *Spine Journal* 8.1:213-225.

Brox, J.I., Storheim, K., Grotle, M., Tveito, T.H., Indahl, A. and Eriksen, H.R. 2008. Evidence-informed management of chronic low back pain with back schools, brief education and fear-avoidance training. *Spine Journal* 8.1:28-39.

Buchanan, H. 2007. The Psychometric properties of the Roland-Morris Disability Questionnaire for patients with Chronic Mechanical Low Back Pain. *South AfricaJournal of Physiotherapy* 63:9-15.

Buchbinder, R., Blyth, F.M., March, L.M., Brooks, P., Woolf, A.D. and Hoy, D.G. 2013.Placing the global burden of low back pain in context.*Best Practice and Research Clinical Rheumatology*27.5:575-589.

Buton, A., Conway, J. and Holstate, S. (2000): Reliability, what is it and how is it measured? *Physiotherapy* 86.2:94–99.

Cailliet, R. 2003.*Low back pain: a medical enigma* Philadelphia: Lippincott Williams and Wilkins.

Carey, T.S., Garret, J., Jackman, A., McLaughlin, C., Fryer, J. and Smucker, D.R. 1995. The outcomes and costs of care for acute low back pain among patients seen by primary care practitioners, chiropractors and orthopaedic surgeons: the North Carolina Back Pain Project. *New England Journal of Medicine* 333.14:913-917.

Carr, J. and Shepherd, R. 1989. Modified motor assessment scale. *Physical Therapy* 69.9:780.

Carragee, E.J. 2005. Persistent low back pain. *New England Journal of Medicine* 352:1891-1898.

Casazza, B.A. 2012. Diagnosis and treatment of acute low back pain. *American Family Physician* 85:343-350.

Chapman–Smith, D. 1992. Measuring Results – The new Importance of patient questionnaires. *The Chiro Report* 7.1:1–6.

Cherkin, D.C. and MacCornack, F.A.1989. Patient evaluations of low back pain are from family physicians and chiropractors. *Western Journal of Medicine* 150:351-355.

Childs, J.D. and Piva, S.R. 2005. Psychometric properties of the functional rating index in patients with low back pain. *European Spine Journal* 14:1008-1012.

Choi, B.K.L., Verbeek, J.H., Tam, W.W.S. and Jiang, J.Y. 2010. Exercises for prevention of recurrences of low-back pain. *Cochrane Database of Systematic Reviews*20.1:CD006555. doi: 10.1002/14651858.CD006555.pub2.

Chou, R., Qaseem, A., Snow, V., Casey, D., Cross, J.T., Shekelle, P. and Owens, D.K. 2007. Diagnosis and treatment of low back pain: a joint clinical practice guideline from the American College of Physicians and the American Pain Society. *Annals of Internal Medicine* 2007.147:478-491

Christensen, T.H., Bliddal, H., Hansen, S.E., Jensen, E.M., Jensen, H., Jensen, R. and Bay, H. 1993. Severe low-back pain. I: Clinical assessment of two weeks conservative therapy. *Scandinavian Journal of Rheumatology* 22:25–29.

Clarke, J.A., van Tulder, M.W., Blomberg, S.E., de Vet, H.C., van der, Heijden G.J., Bronfort, G. and Bouter, L.M. 2007. Traction for low-back pain with or without sciatica. *Cochrane Database of Systematic Reviews* (2): CD003010. http://dx.doi.org/10.1002/14651858. CD003010.pub4.

Coasta, L.O., Maher, C.G. and Latimer, J. 2007a. Self-report outcome measures for low back pain: searching for international cross-cultural adaptations. *Spine* 32:1028-1037.

Cohen, S.P., Bicket, M.C., Jamison, D. and Rathmell, J.P. 2013. Epidural steroids: a comprehensive, evidence-based review. *Regional Anesthesia and Pain Medicine* 38.3:175-200.

Coste, J., Le Parc, J.M., Berge, E., Delecoeuillerie G. and Paolaggi, J.B. 1993. [French validation of a disability rating scale for the evaluation of low back pain (EIFEL questionnaire)]. *Revue du rhumatisme (Ed. Française)* 60:335–341.

Coulter, A. 1994. Assembling the evidence: patient-focused outcomes research. *Health Library Review* 11.4:263-268.

Currier, D.P. 1981.*Elements of Research in Physical Therapy* Baltimore:Lippincott Williams and Wilkins

Dalpiaz, A.S., Lordon, S.P. and Lipman, A.G. 2004. Tropical lidocaine patch therapy for myofascial pain. *Journal of Pain, Palliative Care and Pharmacotherapy*18.3:15-34.

Daltroy, L.H., Cats-Baril, W.L., Katz, J.N., Fossel, A.H. and Liang, M.H. 1996. The North Americanspinesociety lumbarspine outcome assessment instrument: reliability and validity tests. *Spine (Phila Pa 1976)*15; 21.6:741-749.

Davidson, M. and Keating, J.L. 2002. The comparison of five low back disability questionnaires: reliability and responsiveness. *Physical Therapy* 82.1:8-24.

Deshpande, A.,Furlan, A.D., Mallis-Gagnon, A., Atlas, S. and Turk, D. 2007. Opioids for Chronic low-back pain. *Cochrane Database Systematic Review* 3:CD004959. http://dx.doi.org/10.1002/14651858.CD004959.pub3.

Deyo, R., Diehl, P. and Diehl,A. 1991. Reproducibility and Responsiveness of Health Status Measures. *Controlled clinical trials*. 12.4:S142-S158.

Deyo, R.A., Battie, M., Beurskens, A.J., Bombardier, C., Croft, P., Koes, B., Malmivaara, A., Roland, M., Von Korff, M. and Waddell, G. 1998. Outcome measures for low back pain research. A proposal for Standardized use. *Spine*23.18:2003-2013.

Deyo, R.A. 1986. Comparative Validity of the sickness Impact Profile short scales for functional assessment in low back pain. *Spine* 11.9:951 -954.

Deyo, R.A. and Weinstein, J.N. 2011. Low Back Pain. *New England Journal of Medicine*. 1; 344.5:363-370.

Dilke, T.F., Burry, H.C. and Grahame, R. 1973. Epidural corticosteroid in injection in management of lumbar root compression. *British Medical Journal*162:653-657.

Dionne, C.E., Koepsell, T.D., Von Korff, M., Deyo, R.A., Barlow, W.E. and Checkoway, H. 1997. Predicting long-term functional limitations among back pain patients in primary care settings. *Journal of Clinical Epidemiology* 50:31–43.

Dropsy, R. and Marty, M. 1994. Quality-of life indexes for assessment of low back pain. *Revue du rhumatisme (Ed. Française)* 61:S44–S48.

Djiken, C.B., Fjellman-Wikland, A. and Hildingsson, C. 2008. Low Back pain, Lifestyle Factors and Physical Activity: A population-based study. *Journal of Rehabilitation Medicine*40:864-869.

Dubs, L. 1999. [Everything was done – patient still not satisfied: introduction to evidence – based surgery]. *Swiss Surgery* 5.4:160–166.

Duruoz, M.T., Ozcan, E., Ketenci, A. and Karan, A 2013. Development and validation of a functional disability index for chronic low back pain. *Journal of Back and Musculoskeletal Rehabilitation* 26:45–54.

Ducksworth, M. 1999. Outcome Measure Selection and Typology. *Physiotherapy* 85.1:21– 27.

Ehrlich, G.E. 2003. Low back pain. *Bulletin of the World Health Organization* 81:671-676.

Enzmann, D. F. 1997. *Surviving in Healthcare* St. Louis Mo: Mosby – Year Book.

Fairbank, J. 1995. Use of Oswestry Disability Index (ODI). *Spine* 20:1535–1537.

Fairbank, J.C.T., Davies, J.B. and Couper, J. 1980. The Oswestry Low-back pain Disability Questionnaire. *Physiotherapy* 66:271 – 273.

Fairbank, J.C.T. and Pynsent. P.B. 2000. The Oswestry Disability Index. *Spine* 25:2940–2953.

Feise, R.J. and Menke, J.M. 2001. Functional rating index. A new valid and reliable instrument to measure the magnitude of clinical change in spinal conditions. *Spine* 26:78–87.

Feise, R.J. and Menke, J.M. 2010. Functional Rating Index: literature review. *Medical Science Monitor* 16.2:RA25-36.

Fisher, K. and Johnson, M. 1997. Validation of low back Pain disability questionnaire, its sensitivity as a measure of change following treatment and its relationship with other aspects of chronic pain experience. *Physiotherapy Theory Practical* 13:67 – 80.

Freburger, J.K., Holmes, G.M., Agans, R.P., Jackman, A.M., Darter, J.D., Wallace, A.S., Castel, L.D., Kalsbeek, W.D. and Carey, T.S. 2009. The rising prevalence of Chronic Low Back pain. *Archives of Internal Medicine* 169.3:251-258.

French, S.D., Cameron, M., Walker, B.F., Reggars, J.W. and Esterman, A.J. 2006. A Cochrane review of superficial heat or cold for low back pain. *Spine* 31:998-1006.

Fritz, J.M. and Irrang, J.J. 2001. A Comparison of a Modified Oswestry Low Back Pain Disability Questionnaire and the Quebec Back Pain Disability Scale. *Physical Therapy* 81.2:776-788.

From the Centers for Disease Control and Prevention. (2001): Prevalence of disabilities and associated health conditions among adults –United States, 1999. *Journal of the American Medical Association* 285.12:1571-1572.

Foster, L. 2001.Botulinum Toxin A and chronic low back pain: a randomised double blind study. *Neurology* 56:1290–1293.

Frymoyer, J.W. 1992. Predicting disability from low back pain. *Clinical Orthopaedics* 279:101-119.

Gagnier, J.J., van Tulder, M.W., Berman, B.M. and Bombardier C. 2007. Herbal medicine for low back pain. Cochrane Database Systematic Reviews. 2:CD004504. http://dx.doi.org/10.1002/14651858.CD004504.pub3.

Gammon, G.D. and Starr, I. 1993. Studies on the relief of pain by counter-irritation. *Journal of Physiology* 72:392.

Gay, R.E. and Brault, J.S. 2008. Evidence-informed management of chronic low back pain with traction therapy. *The Spine Journal* 8.1:234-242.

George, D. and Mallary, P. 2003.*SPSS for windows step by step: A simple guide and reference - 11.0 update* 4th Ed. Boston: Alyn and Bacon. 231.

Gibbon, B. 1991. Measuring Stroke recovery. *Nursing Times* 87.44:32–34.

Gilson, B.S., Gilson, J.S., Bergner, M., Bobbit, R.D., Kressel, S., Pollard, W.E. and Vesselagom, M. 1975. The Sickness Impact Profile. Development of an outcome measure of health care. *American Journal of Public Health* 65.12:1304-1311.

Goertz, C.M. 1994. Measuring Functional Health status in the chiropractic office using self – report questionnaires. *Topics in Clinical Chiropathy* 1:51–59.

Granger, C.U. 1984).*A conceptual model for functional assessment in rehabilitation medicine.*Baltimore: William and Wilkins. 14–28.

Granger, C.V. and Wright, B.D. 1993. Looking ahead to the use of functional assessment in ambulatory physiatric practice. *Physical Medicine and Rehabilitation Clinics North America* 4:595–605.

Granger, C.V., Ottenbacher, K.J., Baker, J.G. and Sehgal, A. 1995. Reliability of a brief outpatient functional outcome assessment measure. *American Journal of Physical Medicine Rehabilitation* 74:469–475.

Greenough, C.G. and Fraser, R.D. 1992. Assessment of Outcome in patients with low – back pain. *Spine* 17.1:36–41.

Grenville, J. and Lyne, P. 1995. Patient centered evaluation and rehabilitation care. *Journal of Advanced Nursing* 22:965-972.

Gronblad, M., Jarvinen, E., Hurri, H., Hupli, M. and Karaharju, E.O. 1994. Relationship of the Pain Disability Index (PDI) and the Oswestry Disability Questionnaire (ODQ) with three dynamic physical tests in a group of patients with chronic low-back and leg pain. *Clinical Journal of Pain* 10:197–203.

Grotle, M., Brox, J.I. and Vøllestad, N.K. 2003. Cross-cultural adaptation of the Norwegian versions of the Roland-Morris Disability Questionnaire and the Oswestry Disability Index. *Journal of Rehabilitation Medicine* 35:241–247.

Grotle, M., Brox, J.I. and Vollestad, N.K. 2004. Functional status and disability questionnaires: What do they assess? A systematic review of back-specific outcome questionnaires. *Spine* 30:130–140.

Golob, A.L. and Wipf, J.E. 2014. Low back pain. *Medical Clinics of North America* 98.3:405-428.

Guillemin, F., Constant, F., Collin, J.F. and Boulange, M. 1994. Short and longterm effect of spa therapy in chronic low back pain. *British Journal of Rheumatology* 33:148–151.

Guo, H.R., Tanaka, S., Halperin, W.E. and Cameron, L.L. 1999. Back pain prevalence in US industry and estimates of lost workdays. *American Journal of Public Health* 89.7:1029-1035.

Hagg, O., Fritzell, P., Romberg, K. and Nordwall, A. 2001. The General Function Score: a useful tool for measurement of physical disability. Validity and reliability. *European Spine Journal* 10:203-210.

Haldeman, S., Chapman – Smith, D. and Petersen, D. 1992.*Guidelines for Chiropractic Quality Assurance and practice parameters* Gaithersbury, Md:Aspen.

Hansen, D.J. 1994. Determining how much care to give and reporting patient's progress. *Topics in Clinical Chiropathy* 1.4:1–8.

Hansen, D.T. and Vernon, H. 1999. Applications of quality improvement to the Chiropratic Profession In: Lawrence D, Ed. *Advances in Chiropratic* St. Louis, Mo: Mosby – Year Book.

Haas, M. and Nyiendo, J. 1992.Diagnostic utility of the McGill Pain Questionnaire and the Oswestry Disability Questionnaire for classification of low back pain syndromes. *Journal of Manipulative and Physiological Therapeutics*15.2:90-98.

Hazard, R.G., Haugh, L.D., Reid, S., Preble, J.B. and MacDonald, L. 1996. Early prediction of chronic disability after occupational low back injury. *Spine* 15;21.8:945-951.

Harper, A.C., Harper, D.A., Lambert, L.J., de Klerk, N.H., Andrews, H.B., Ross, F.M., Straker, L.J. and Lo, S.K. 1995.Development and validation of the Curtin Back Screening Questionnaire (CBSQ): a discriminative disability measure. *Pain*60.1:73-81.

Henderson, D., Chapman – Smith, D., Moor, S. and Vernon, H. 1994. Clinical guidelines for Chiropractic practice in Canada. *Journal of Canadian Chiropractic Association (supplementary)* 38.1:1–203.

Hinderer, S. R. and Hinderer, K.A. 1993. Quantitative methods of evaluation. In Delisa J., Ed. *Rehabilitation Medicine: principles and practice* 2nd ed. Philadelphia: JB Lippincott. 96 – 121.

Holm, S. and Nachemson, A. 1988. Nutrition of the intervertebral disc: acute effects of cigarette smoking (an experimental animal study). *Journal of Medical Sciences* 93.1:91-99.

Holt, A.E., Shaw, N.J., ShettyA. and Greenough, C.G. 2002. The reliability of the Low Back Outcome Score for back pain.*Spine (Phila Pa 1976)* 15;27.2:206-210.

Hoy, D., Bain, C., Williams, G., March, L., Brooks, P., Blyth, F., Woolf, A., Vos, T., and Buchbinder, R. 2012. A systemic review of the global prevalence of low back pain. *Arthritis and Rheumatism* 64.6:2028–2037.

Hsieh, C-Y.J., Philips, R.B., Adams, A.H. and Pope, M.H. 1992. Functional outcomes of low back pain. Comparison of four treatment groups in a randomized controlled trial. *Journal of Manipulative Physiology Therapeutics* 15:4–9.

Husted, J.A., Cook, R.J., Farewell, V.T. and Gladmann, D.D. 2000. Methods for assessing responsiveness: a critical review and recommendations. *Journal of Clinical Epidemiology* 53:459–468.

Huan, H.C., Chang, H.J., Lin, K.C., Chin, H.Y., Chung, J.H. and Tsai, H.C. 2014. A closer examination of the interaction among risk factors for low back pain. *American Journal of Health Promotion* 28.6:372–379.

Imamura, M., Furlan, A.D., Dryden, T. and Irvin, E. 2008. Evidence-informed management of chronic low back pain with massage. *Spine Journal* 8.1:121-133.

Itzkovich, M., Katz, A., Tamir, A., Ronen, J., Philo, O., Steinberg, F., Tabacaru, E., Spasser, R. andGepstein, R. 2001. The Spinal Pain Independence Measure - a new scale for assessment of primary ADL dysfunction related to LBP. *Disability and Rehabilitation* 23:186–191.

Jarvikoski, A., Mellin, G., Estlander, A.M., Harkapaa, K., Vanharanta, H., Hupli, M., and Heinonen, R. 1993. Outcome of two multimodal back treatment programs with and without intensive physical training. *Journal of Spinal Disorders* 6.2:93-98.

Johansson, E. and Lindberg, P. 1998.Subacute and chronic low back pain. Reliability and validity of a Swedish version of the Roland and Morris Disability Questionnaire. *Scandinavian Journal of Rehabilitation Medicine* 30:139–143.

Khatri, H., Greenough, C.G. 2002. Minimum clinically important difference in the low back outcome score. *Proceedings of the Low Back Research Society meeting* Leeds, UK.

Katz, J.N. 2006. Lumbar disc disorders and low-back pain: socioeconomic factors and consequences. *Journal of Bone and Joint Surgery America*88.suppl 2:21-24.

Kaplan, R. and Anderson, J. 1982. A general health policy model update and applications. *Health Service Resource* 23:203–235.

Kettenmann, B., Wille, C., Lurie-Luke, E., Walter, D. and Kobal, G. 2007. Impact of continuous low level heat-wrap therapy in acute low back pain patients: subjective and objective measurements. *Clinical Journal of Pain* 23.8:663-668.

Kinkade, S. 2007. Evaluation and treatment of acute low back pain. *American Journal of Family Physician*75:1182-1188

Kirschner, B. and Guyatt, G.H. 1985. A methodological framework for assessing health indices. *Journal of Chronic Disorders* 38:27-36.

Kopec, J.A., Esdaile, J.M., Abrahamowicz, M., Abenhaim, L., Wood, D.S., Lamping, D.L. and Williams, J.I. 1995. The Quebec Back Pain Disability Scale. Measurement properties. *Spine* 20:341–352.

Kopec, J.A., Esdaile, J.M., Abrahamowicz, M., Abenhaim, L., Wood-Dauphinee, S., Lamping, D.L. and Williams, J.I.1996. The Quebec Back Pain Disability Scale: conceptualization and development. *Journal of Clinical Epidemiology* 49.2:151-161.

Korbon, G.A., Schroeder, M.E., Schwartz, D.P. and Shutty, M.S. 1989. The development of a somatic amplification rating scale for Low Back pain. *Spine* 12.8:987-991.

Kovacs, F.M., Llobera, J., Gil Del Real, M.T., Abraira, V., Gestoso, M., Fernández, C. and Primaria Group KA.2002. Validation of the Spanish version of the Roland-Morris questionnaire. *Spine* 27:538–542.

Kressel, S., Polland, W.E. and Vesselago, M. 1975. The sickness impact profile. Development of an outcome measure of health care. *American Journal of public health*65.12:1304–1311.

Kopec, J.A. and Esdaile, J.M. 1998. Occupational role performance in persons with back pain.*Disability Rehabilitation* 20:373–379.

Kucukdeveci, A.A., Tennant, A., Elhan, A.H. and Niyazoglu, H. 2001. Validation of the Turkish version of the Roland–Morris Disability Questionnaire for use in low back pain. *Spine*26:2738–2743.

Kurande, V.H., Waagepetersen, R., Toft, E., Prasad, R. 2013. Reliability studies of diagnostic methods in Indian traditional Ayurveda medicine: An overview.*Journal of Ayurveda Integrated Medicine*2.4:67-76.

Lankhorst, G.J., Van de Stadt, R.J. andVogelaar, T.W. 1982. Objectivity and repeatability of measurements in low back pain. *Scandinavian Journal of Rehabilitation Medicine*14: 21 – 26.

Last, A.R. and Hulbert, K. 2009. Chronic low back pain: evaluation and management. *American Family Physician*79:1067-1074.

Lawlis, G.F., Cuencas, R., Selby, D. and McCoy, C.E. 1989. The development of the Dallas Pain Questionnaire an assessment of the impact of spinal pain on behaviour. *Spine*14: 511–516.

Leboeuf-Yde, C., Lawrisen, M.J. and Lauritsen, T. 1997. Why has the search for causes of low back pain been largely in conclusive.*Spine* 22.8:881-887.

Lee, H., Nicholson, L.L. and Adams, R.D. 2006. Developmental and Psychometric testing of Korean language versions of 4 neck pain and disability questionnaires. *Spine* 31:1841-1845.

Leung, A.S., Lam, T.H., Hedley, A.J. and Twomey, L.T. 1999. Use of a subjective health measure on Chinese low back pain patients in Hong Kong. *Spine* 15;24.10:961-966.

Lequesne, M.C. 1997. The Algo functional induces for hip and knee osteoathritis. *Journal of Rheumatology* 24:779 – 781.

Lewis, P.S. and Latney, C. 2002. Achieve best practice with an evidence-based approach. Create a collaborative environment that improves patient care through consistent outcomes measurement. *Nursing Management* 33.12:24-30.

Liau, A.K, Chow, D., Tan, T.K. and Senf, K. 2011. Development and Validation of the Personal Strengths inventory Using Exploratory and Confirmatory Factor Analyses. *Journal of Psychoeducational Assessment* 29:14–26.

Lings, S. and Leboeuf-Yde, C. 1998. Whole body vibrations and low back pain. *Ugeskr Laeger* 160:4298-4301.

Loisel, P., Poitras, S., Lemaire, J., Durand, P., Southiere, A. and Abenhaim, L. 1998. Is work status of low back pain patients best described by an automated device or by a questionnaire? *Spine (Phila Pa 1976)*15; 23.14:1588-1594.

Long, A., Sheldon, T., Freemantle, N., Ibbotson, S., Matson, J. and Pollock, C. 1992.*Effective Health Care: Stroke Rehabilitation* Nuffield Institute, University of Leeds, and Centre for Health Economics, University of New York.

Louw, Q.A., Morris, L.D. and Grimmer-Somers, K. 2007. The prevalence of low back pain in Africa: a systematic review. *BioMed Central Musculoskeletal disorders* 8:105.

Luo, X., Pietrobon, R., Sun, S.X., Liu, G.G. and Hey, L. 2004. Estimates and patterns of direct health care expenditures among individuals with back pain in the United States. *Spine.* 29:79–86.

Maaroufi, H., Benbouazza, K., Faık, A., Bahiri, R., Lazrak, N., Abouqal, R., Amine, B. and Hajjaj-Hassouni, N.2007. Translation, adaptation, and validation of the Moroccan version of the Roland Morris Disability Questionnaire. *Spine* 32:1461–1465.

Main, C.J., Wood, P.L. and Hollis, S. 1992. The Distress and Risk assessment method. A simple patient classification to identify distress and evaluate the risk of poor outcome *Spine*17.1:42-52.

Manchikanti, L. and Fellows, B. 2006. Psychological aspects of chronic low back pain. In Manchikanti L., Slipman CW and Fellows B. *Low Back Pain Diagnosis and Treatment* ASIPP: USA. 97-118.

Manchikanti, L., Singh, V., Falco, F.J., Benjamin, R.M. and Hirsch, J.A. 2014. Epidemiology of low back pain in adults. *Neuromodulation* 17.suppl. 2:3-10.

Manek, N.J. and MacGregor, A.J. 2005. Epidemiology of back disorders: prevalence, risk factors and prognosis. *Current Opinion in Rheumatology* 17:134-140.

Mancuso, C.A., Altchek, D.W., Craig, E.V., Jones, E.C., Robbins, L., Warren R.F. and Williams-Russo, P. 2002. Patient's expectations of shoulder surgery. *Journal of Shoulder Elbow Surgery* 11:541–549.

Maniadakis, N. and Gray, A. 2000. The economic burden of low back pain in the United Kingdom. *Pain* 84:95-103.

Manniche, C., Asmussen, K., Lauritsen, B., Vinterberg, H., Kreiner, S. and Jordan, A. 1994. Low back pain rating scale: validation of a tool for assessment of low back pain. *Pain*57:317– 326.

Martini, C.J. and McDowell, I. 1976. Health status: patient and physician judgements. *Health Service Research* 11.4:508-515.

Mawson, J.S. 1993. Measuring Physiotherapy Outcome in Stroke Rehabilitation. *Physiotherapy* 79.11:762–765.

May, S. and Donelson, R. 2008. Evidence-informed management of chronic low back pain with the McKenzie method. *Spine Journal* 8.1:134-141.

Mayer, J., Mooney, V., Dagenais, S. 2008. Evidence-informed management of chronic low back pain with lumbar extensor strengthening exercises. *Spine Journal* 8.1:96-113.

McCormick, J.D., Werner, B.C. and Shimer, A.L. 2013. Patient reported outcome measure in Spine surgery. *Journal of American Academy of Orthopaedic Surgery* 21:99-107.

McDowell I. and Newell C. 2006.*Measuring Health: A guide in rating scales and questionnaire* 3rd Ed. New York, NY: Oxford University Press.

McElheran, L.J. and Sollectio, P.C. 1994. Delivering quality Chiropractic Care in a managed Care Setting. *Topics in Clinical Chiropathy* 1.4:30–39.

McKay – Lyons, M. 1998. Outcome Measures, *Proceedings of Workshop XI Kuwait – Galhousie Physiotherapy project* 60–61.

McKenzie, R.A. 1981.*The lumbar spine: mechanical diagnosis and therapy*Waikanae, New Zealand: Spinal Publications.

Meenan, R.F., Gentman, P.M. and Manson, J.H. 1980. Measuring Health Status in Arthritis; The Arthritis Impact Measurement scale. *Arthritis and Rheumatism* 23.2:146–152.

Melon, M., Aebli, N., Elfering, A., Roder, C., Zweing, T., Barz, T., Herbison, P., Hendrick, P., Bajracharya, S., Stout, K. and Theis, J.C. 2008. Development of a screening tool predicting the transition from acute to chronic low back pian for patients in a GP setting: protocol of a multinational prospective cohort study. *BioMed Central Musculoskeletal Disorders* 19.9:167.

Melzack, R. 1975. The McGill pain Questionnaire: Major properties and scoring methods. *Pain* 1:277–279.

Melzack, R. 1987. The short-form McGill Pain Questionnaire. *Pain* 30:191–197.

Michael, V. 2008. Psychometric properties and clinical usefulness of the Oswestry Disability Index. *Journal of Chiropractic Medicine* 7:161–163.

Millard, R.W. 1989. The Functional Assessment Screening Questionnaire: application for evaluating pain-related disability. *Archives of Physical Medicine Rehabilitation* 70 .4:303-307.

Million, R., Hall, W., Nilsen, K.H., Baker, R.D. and Jayson, M.I. 1981. Assessment of the progress of the back – pain patient 1981 Volvo Award in Clinical Science. *Spine* 7:204 – 212.

Million, R., Nilsen, K.H., Jayson, M.I. and Baker, R.D. 1981. Evaluation of low back pain and assessment of lumbar corsets with and without back supports. *Annals of Rheumatic Diseases* 40.5:449-454.

Mokkink, L.B., Terwee, C.B., Knol, D.L., Stratford, P.W., Alonso, J.,Patrick, D.L., Bouter, L.M. and de Vet, H.C. 2010. The COSMIN checklist for evaluating the methodological quality of studies on measurement properties: A clarification of it's content.*BioMed Central Medical Research Methodology*10.22:doi:10. 1186 / 1471-2288.

Mootz, R.D. 1994. Maximising the effectiveness of Clinical documentation. *Topics in Clinical Chiropathy* 1.1:60–65.

Mousavi, S.J., Parnianpour, M., Mehdian, H., Montazeri, A. and Mobini, B. 2006. The Oswestry Disability Index, the Roland-Morris disability questionnaire, and the Quebec back pain dis- ability scale: translation and validation studies of the Iranian versions. *Spine* 31:E454–E459.

Muller, U., Duetz, M.S., Roeder, C. and Greenough, C.G. 2004. Condition – specific outcome measures for low back pain. Part 1: Validation. *European Spine Journal* 13:301–313.

Muller, U., Roeder, C., Dubs, L., Duetz, M.S. and Greenough, C.G. 2004. Condition – specific outcome measures for low back pain. Part II: Scale Construction. *European Spine Journal* 13:314–324.

Nunally, J.C. and Bernstein, L.H. 1994. *Psychometric Theory* 3rd Edition New York: McGraw-Hill.

Nork, S.E., Hu, S.S., Workman, K.L., Glazer, P.A. and Bradford, D.S. 1999. Patient outcomes after decompression and instrumented posterior spinal fusion for degenerative spondylolisthesis.*Spine* 15;24.6:561-569.

Nusbaum, L., Natour, J., Ferraz, M.B. and Goldenberg, J. 2001. Translation, adaptation and validation of the Roland– Morris questionnaire – Brazil. *Brazilian Journal of Medical and Biological Research* 34:203–210.

Nwuga, V.C.B. 1993. Old age and low back pain among Nigerians. *Health Team* 1:6-8.

Odole, A.C. and Akinpelu, A.O. 2009. Translation and alternate forms reliability of the visual analogue scale in the three major Nigerian languages. *The Internet Journal of Allied Health Sciences and Practice*. 7.3:1540-1580.

Onel, D., Tuzlaci, M., Sari, H. and Demir, K. 1989. Computed tomographic investigation of the effect of traction on lumbar disc herniations. *Spine* 14:82-90.

Omokhodion, F.O. 2002. Low back pain in a rural community in South West Nigeria. *West African Journal of Medicine* 21.2:87-90.

Omokhodion, F.O. 2004. Low back pain in a rural community in South-West Nigeria. *Tropical Doctor* 34:17-20.

Omokhodion, F.O. and Sanya, A.O. 2003. Risk factor for low back pain among office workers in Ibadan, Southwest Nigeria. *Occupational Medicine* 53:287-289.

Omokhodion, F.O.and Osungbade, O.O. 1996. Health problems of automobile mechanics in Nigeria. *Tropical Doctor*26.3:102-104.

Owolabi, M.O. and Platz, T. 2008. Proposing the stroke levity scale: a valid, reliable, simple, and time-saving measure of stroke severity. *European Journal of Neurology* 15:627-633.

131

Padua, R., Padua, L.,Ceccarelli, E., Romanini, E., Bondi, R., Zanoli, G. and Campi, A.2001. Cross-cultural adaptation of the lumbar North American Spine Society questionnaire for Italian-speaking patients with lumbar spinal disease. *Spine* 1;26.15:E344–E347.

Patrick, D.L., Deyo, R.A., Atlas, S.J., Singer, D.E., Chapin, A. and Keller, R.B. 1995. Assessing health-related quality of life in patients with sciatica. *Spine* 1;20.17:1899-1908.

Patrick, N., Emanski, E. and Knaub, M.A. 2014. Acute and Chronic Low Back Pain. *Medical Clinics of North America* 98.4:777-789.

Petering, R.C. and Webb, C. 2011. Treatment options for low back pain in athletes. *Sports Health* 3.6:550-555.

Price, D.D. 1994. Psychophysical measurement of normal and abnormal pain processing In: Boivie J, Hanson P, Lind blom U, eds. *Touch, Temperature, and Pain in Health and Disease: Mechanism and Assessments. Progress in Pain Research and Management* Vol. 3. Seattle, Wash: IASP Press. 3–25.

Price, D.D., Bush, F.M., Long, S. and Harkins, S.W. 1994. A comparison of pain measurement characteristics of mechanical visual analogue and simple numerical pain rating scales. *Pain* 56:217-226.

Poitras, S. and Brosseau, L.2008. Evidence-informed management of chronic low back pain with transcutaneous electrical nerve stimulation, interferential current, electrical muscle stimulation, ultrasound and thermotherapy. *The Spine Journal* 8.1:226-233.

Pose, B., Sangha, O., Peters, A. and Wildner, M. 1999. Validation of the North American Spine Society Instrument for assessment of health status in patients with chronic backache. *Z OrthopIhreGrenzgeb* 137:437–441.

Post, M.W.M., Boosman, H., Van Zandroot, M., Passier, P.E., Rinkel, G.J. and Visser-Meily, J.M. 2011. Development and validation of a short version of the 'Stroke-Specific Quality of Life Scale'.*Journal of Neurology, Neurosurgery and Psychiatry*3.82:283-286.

Rainville, J., Ahern, D.R., Phalen, L., Childs, L.A. and Sutherland, R. 1992. The association of pain with physical activities in chronic low back pain. *Spine* 17.9:1060-1064.

Rebbeck, T.J., Refshauge, K.M., Maher, C.G. and Stewart, M. 2007. Evaluation of the core outcome measure in whiplash. *Spine* 32:696-702.

Resnik, I. and Dobrykowski, E. 2005. Outcomes measurement for patients with low back pain. *Orthopaedic Nurses*. 24:14–24.

Refshauge, K.M. and Maher, C.G. 2006. Low back pain investigations and prognosis; a review. *British Journal of Sports Medicine* 40:494-498.

Ricci, J.A., Stewart, W.F., Chee, E., Leotta, C., Foley, K. and Hochberg, M.C. 2006. Back pain exacerbations among and lost productive time costs in United States workers. *Spine* 31.26:3052-3060.

Riddle, D.L., Stratford, P.W. and Binkley, J.M. 1998. Sensitivity to change of the Roland-Morris Back Pain Questionnaire: part 2. *Physical Therapy* 78.11:1197-1207.

Roach, K.E., Brown, M.D., Albin, R.D., Delaney, K.G., Lipprandi, H.M. and Rangelli, D. 1997. The sensitivity and specificity of pain response to activity and position in categorizing patients with low back pain. *Physical Therapeutics* 77:730–738.

Roach, K.E., Leithanonakus, Y., Songsindg, N. and Budman – Mak, E. 1995. Development of a shoulder pain Disability Index.*Orthopaedic Division Review* 5:17–19.

Robertson, V.J., Ward, A., Low, J. and Reed, A. 2006. *Electrotherapy Explained: principles and Practice* 4th Ed. Massachusetts, USA: Elsevier, Butterworth-Heinemann.

Roelofs, P.D., Deyo, R.A., Koes, B.W., Scholten, R.J. and van Tulder, M.W. 2008. Nonsteroidal anti-inflammatory drugs for low back pain: an updated Cochrane review. *Spine* 15;33.16:1766-1774.

Roland, M. and Morris, R. 1983. A study of the natural history of back pain. Part I: development of a reliable and sensitive measure of disability in low – back pain. *Spine* 8:141–144.

Roos, E.M., Roos, K.P. and Lohmander, L.S. 1999. WOMAC Osteoarthritis Index – additional dimension for use in subjects with posttraumatic Osteoarthritis of Knee. *Osteoarthritis and Cartilage* 7.2:216–221.

Rothstein, J.M. 1985.*Measurement in Physical Therapy* New York, NY: Churchill – Livingstone, 1–46.

Rubin, D.I. 2007. Epidemiology and risk factors for spine pain. *Neurology Clinic*. 25(2):353-371.

Russell, D.J., Avery, L.M., Rosenbaum, P.L., Ralni, P.S., Walter, S.D. and Palisamo, R.J 2000. Improved scaling of the Gross Motor function measure for children with cerebral palsy. *Physical Therapy* 80.9:873–885.

Ruta, D.A., Garratt, A.M., Wardlaw, D., and Russell, I.T. 1994. Developing a valid and reliable measure of health outcome for patients with low back pain. *Spine* 19:1887–1896.

Rytokoski, U., Puukka, P. and Talo, S. 1997.Biodisabilities in relation to other disease consequences in the functional assessment of patients with chronic low back pain. *International Journal of Rehabilitation Research* 20:225–244.

Salen, B.A., Spangfort. E.V., Nygren. A.L. and Nordemar, R. 1994. The Disability Rating Index: an instrument for the assessment of disability in clinical settings. *Journal of Clinical Epidemiology* 47.12:1423-1435.

Schoppink, L.E., van Tulder, M.W., Koes, B.W., Beurskens, S.A. and de Bie, R.A. 1996. Reliability and validity of the Dutch adaptation of the Quebec Back Pain Disability Scale. *Physical Therapy* 76:268–275.

Scuderi G.R., Bourne R.B., Noble P.C., Benjamin J.B., Lonner J.H. and Scott W.N. 2012. The new knee society scoring system. *Clinical Orthopaedics and Related Research* 470:3–19.

Seltzer, G.B., Granger, C.V. and Wineberg, D.E. 1982. Functional assessment: bridge between family and rehabilitation medicine within an ambulatory practice. *Archives of Physical Medicine and Rehabilitation* 63:453–457.

Shen, F.H., Samartzis, D. and Anderson, G.B. 2006. Nonsurgical management of acute and chronic low back pain. *Journal of the American Academy of Orthopedic Surgery*14:477-487.

Solomon, L., Warwick. D. and Nayagam, S. (Eds.) 2010. *Apley's System of Orthopaedics and Fracture* 9thEdition. Florida: Taylor and Francis Group. 476-488.

Soyanwo, O.A., Amanor– Boadu, S.D., Sanya, A.O. and Gureje, O. 2001. Pain assessment in Nigerians – Visual Analogue Scale and Verbal Rating Scale compared. *West African Journal of Medicine* 19:242–245.

Standaert, C.J., Weinstein, S.M. and Rumpeltes, J.2008. Evidence-informed management of chronic low back pain with lumbar stabilization exercises. *Spine Journal* 8.1:114-120.

Stewart, W.F., Ricci, J.A., Chee, E., Morganstein, D. and Lipton, R. 2003. Lost productive time and cost due to common pain conditions in the US workforce. *Journal of the American Medical Association* 290.18:2443-2454.

Stratford, P.W. and Binkley, J.M. 1997. Measurement properties of the RM-18. A modified version of the Roland– Morris Disability Scale. *Spine* 22: 2416–2421.

Stratford, P.W. and Binkley, J.M. 2000. A comparison study of the back pain functional scale and Roand Morris Questionnaire. North American Orthopaedic Rehabilitation Research Network. *Journal of Rheumatology* 27.8:1928-1936.

Stratford, P.W., Binkley, J.M. and Riddle, D.L. 2000. Development and initial validation of the back pain functional scale. *Spine* 15;25.16:2095-2102.

Stratford, P.W., Binkley J., Solomon, P., Gill, C. and Finch, E. 1994. Assessing change over time in patients with low back pain. *Physical Therapy*74:528–533.

Stratford, P.W., Gill, C., Westway, M. and Binkley, J. 1995. Assessing disability and change on individual patients: a report of a patient specific measure. *Physiotherapy Canada* 47: 258-263.

Streiner, L.D. and Norman R. G. 2008.*Health Measurement Scales: A practical guide to their development and use*Oxford, New York: University Press.

Terwee, C.B.,Bota, D.M., de Boera, M.R., Van der Windta, D.A.W.M., Kuola, D.L., Dekkera, J., Boutera, L.M. and de Veta, H.C.W. 2007. Ouality Criteria proposed for Measurement properties of Health Status Questionnaire. *Journal of Clinical Epidemiology* 60:34-42.

Tesio, L., Granger, C.V. and Fiedler, R.C. 1997. A unidimensional pain/disability measure for low back pain syndromes. *Pain*69:269–278.

Turk, D.C. and Rudy, T.E. 1987. Towards a comprehensive assessment of chronic pain patients. *Behaviour Research Therapy* 25:237–249.

Uematsu, T. and Matuzaki, H. 2001. Effects of Nicotine on the intervertebral disc: an experimental study in rabbits. *Journal of Orthopaedic Sciences* 6.2:177-182.

Underwood, M.R., Barnett, A.G. and Vickers, M.R. 1999. Evaluation of two time-specific back pain outcome measures. *Spine* 24:1104–1112.

Valderas, J.M, Alonso, J. and Guayatt, G.H. 2008.Measuring Patient − reported outcomes moving from clinical trials into clinical practice.*Medical Journal of Australia* 189.2:93-94.

Van Tulder, M.W., Malmivaara, A., Esmail, R. and Koes, B. 2000. Exercise therapy for low back pain: a systematic review within the framework of the Cochrane collaboration back review group. *Spine* 25:2784-2796.

Van Tulder, M., Koes, B. and Bombardier, C. 2002. Low back pain. *Best Practice Resource Clinical Rheumatology* 16.5:761-775.

Vianin, M. 2008. Psychometric properties and clinical usefulness of the Oswestry Disability Index. *Journal of Chiropractic Medicine* 7.4:161-163.

Vingard, E., Lindberg, P., Josephson, M., Voss, M., Heijbel, B., Alfredsson, L., Stark, S. and Nygren, A. 2005.Long-term sick-listing among women in the public sector and its associations with age, social situation, lifestyle, and work factors: a three-year follow-up study. *Scandinavian Journal of Public Health* 33.5:370-375.

Volinn, E. 1997. The epidemiology of low back painin the rest of the world. A review of surveys in low- and middle-income countries. *Spine* 22:1747-1754.

Von Korff, M., Ormel, J., Keefe, F. D. and Workin, S.F. 1992. Grading the severity of Chronic pain. *Pain* 50:133–149.

Von Korff, M., Deyo, R.A., Cherkin, D. and Barlow, W. 1993.Back pain in primary care. Outcomes at 1 year. *Spine (Phila Pa 1976)* 1;18.7:855-862.

Waddell, G. 1987. A new clinical model for the treatment of low back pain. *Spine*12:632-644.

Waddell, G. and Main, C.J. 1984. Assessment of severity in low – back disorders. *Spine* 9:204–208.

Waddell, G., Newton, M., Henderson, I., Somerville, D. and Main, C.J. 1993. A Fear-Avoidance Beliefs Questionnaire (FABQ) and the role of fear-avoidance beliefs in chronic low back pain and disability. *Pain* 52.2:157-168.

Waddell, G., Somerville, D., Henderson, I. and Newton, M. 1992. Objective clinical evaluation of physical impairment in chronic low back pain. *Spine* 17:617–628.

Wade, D. 1992. Measurement in Neurological rehabilitation. *Current Opinion Neurology and Neurosurgery* 5:682–686.

Walker, B.F. 2000. The prevalence of low back pain: a systematic review of the literature from 1966 to 1998. *Journal of Spinal Disorders* 13:205-217.

Wallenstein, S.L., Heidrich III G., Kaiko R. and Houde R.W. 1980. Clinical evaluation of mild analgesics – the measurement of clinical pain. *British Journal of Clinical Pharmacology* 10.Suppl 2:319S-327S.

Walsh, D.A. and Radcliffe, J.C. 2002. Pain beliefs and perceived physical disability of patients with chronic low back pain. *Pain* 97:23–31.

Ware, J. and Davies, A.R. 1983. Defining and measuring patient satisfaction with medical care. *Evaluation program plan* 6:247.

Ware, J.E. and Sherbourne, C.D. 1992. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Medical Care* 30.6:473-483.

Wiesinger, G., Nuhur., M, Quittan, M., Ebenbiehler, G., Wolfl, G. and Fialka-Moser, V. 1999. Cross-cultural adaptation of the Roland-Morris questionnaire for German speaking patients with low back pain. *Spine* 24.11:1099–1103.

Wilhelm, F., Fayolle-Minon, F., Phaner, V., Le-Quang, B., Rimaud, D., Bethoux, F. and Calmels, P. 2010. Sensitivity to change of the Quebec Back Pain Disability Scale and the Dallas Pain Questionnaire. *Annals of Physical and Rehabilitation Medicine*53:15–23.

Willke, R.J., Burke, L.B. and Erickson, P. 2004. Measuring treatment impact: a review of patient reported outcomes and other efficacy endpoints in approved product labels. *Controlled Clinical Trials*25:535-552.

Williams, D. and Feuerstein, M. 2000. Measuring psychosocial Outcomes in clinical Practice in Yeomans S 2000.*The clinical Application of Outcomes Assessment* Stanford: Appleton and Lange. 93–106.

W. H. O. Programme on Mental health, Geneva: World Health Organisation 1996.

Woolf, A.D. and Pfleger, B. 2003. Burden of major musculoskeletal conditions. *Bulletin World Health Organisation*81.9:646-656.

Yoemans, S.G. 2000.*The Clinical Application of Outcomes Assessment* Stanford: Appleton and Lange. 601.

Zoega, B., Karrholm, J. and Lind, B. 2000. Outcome scores in degenerative cervical disc surgery. *European Spine Journal*9:137–143.

Zung, W. W. K. 1965. A self – rating depression Scale. *Archives of General Psychiatry* 32: 63–70.