The role of cervical spine Range of Motion

in recovery from

Whiplash Associated Disorders

By

Mark Andrew Williams

A thesis submitted in partial fulfilment of the requirements for the degree of

Doctor of Philosophy in Health Care Research

University of Warwick, Warwick Medical School

April 2011

TABLE OF CONTENTS

Table of Tables

Table of Figures

ACKNOWLEDGEMENTS

Firstly, a thank you to all the patients that became participants and gave their time and effort freely. To Emma Withers and the rest of the MINT administration team for carrying out the unenviable task of sending, chasing and entering postal questionnaires. To the MINT research clinicians for collecting the baseline data that was so critical to my work. A particular thank you goes to Esther Williamson for sharing our PhD journeys in such an open and generous way. To Chris McCarthy and Angeliki Chorti for their contribution as co-reviewers for Chapter Five's systematic literature review. To Warwick Clinical Trials Unit staff for imparting their gems of knowledge (and humour). In particular, thank you to Professor Sallie Lamb for her unceasing support, patience and motivation. To my supervisors Professor Matthew Cooke and Dr Simon Gates for their encouragement, reassurance and sage advice.

To all my family for their support before, during and after this long journey.

This thesis is dedicated to my partner Sam and son Charlie, without whom this wouldn't have been finished

DECLARATION

This thesis is the candidate's own work except where it contains work based on collaborative research. The nature and extent of the author's individual contribution have been indicated where this is applicable.

This thesis has not been submitted for a degree at another university.

The systematic literature reviews contained in Chapters Three and Five of this thesis have been published as manuscripts in peer-reviewed journals and are provided in Appendix Six and Seven.

ABSTRACT

This thesis investigates the role of cervical spine Range of Motion in the recovery from Whiplash Associated Disorders.

In clinical practice, Health Care Professionals attach value to measurements of cervical spine Range of Motion for diagnostic, prognostic and treatment evaluation purposes. A systematic literature review found conflicting evidence as to whether cervical spine Range of Motion was a prognostic factor following a whiplash injury. Greater understanding of prognostic factors such as this may facilitate improvements in patient management.

A second systematic literature review investigated the reliability and validity of methods for measuring cervical spine Range of Motion. The Cervical Range Of Motion (CROM) device was found to be the most rigorously tested and clinimetrically promising method and was subsequently investigated for intra- and inter-observer reliability in a group of whiplash-injured individuals and found to be substantially reliable.

The CROM device was utilised in a longitudinal cohort study of 599 whiplash-injured patients to investigate the prognostic value of cervical spine Range of Motion for neck pain-related disability and patient-reported recovery at short, medium and long-term follow-up. A patient-reported version of cervical spine Range of Motion was also evaluated as a prognostic factor.

Although useful for explaining disability at the time of measurement, active, passive and patient-reported forms of cervical spine Range of Motion were not significant prognostic factors for poor outcome when other physical and psychosocial factors were accounted for. The clinical implication of this research is that if patients are experiencing reduced cervical spine Range of Motion a few weeks after their whiplash injury they will not necessarily have a poor outcome in the longer term as is commonly believed at present.

LIST OF ABBREVIATIONS

ED – Emergency Department

WAD – Whiplash Associated Disorders

QTF – Quebec Task Force

ROM – Range of Motion

cROM – cervical ROM

CROM – Cervical Range of Motion Device

A cROM – Active cervical Range of Motion

P cROM – Passive cervical Range of Motion

PRcROM – patient-rated cervical ROM

F - Flexion

E - Extension

RR – Right rotation

LR – Left rotation

RLF – Right Lateral Flexion

LLF – Left Lateral Flexion

Av Sh Abd ROM – Average Shoulder Abduction Range of Motion

RTW – Return to work

SE – Self Efficacy questionnaire

PCS – Pain catastrophising Scale questionnaire

IES - Impact of Event Scale questionnaire

GHQ-12 - General Health Questionnaire -12

FABQ – PA – Fear Avoidance Beliefs Questionnaire – Physical Activity subscale

ANOVA - analysis of variance

CI - confidence intervals

CV - coefficient of variation

ICC - Intraclass Correlation Coefficient

n - Number of participants/patients in a study

SD - standard deviation

SEM - standard error of measurement

MW- Mark Williams

SG – Simon Gates

CM – Chris McCarthy

AC – Angeliki Chorti

1 CHAPTER ONE – INTRODUCTION

1.1 INTRODUCTION

This chapter will introduce the concept of whiplash injuries, and briefly their epidemiology and management. An overview of the aims and structure of this thesis are also provided.

1.2 BACKGROUND AND AIMS OF THE THESIS

Whiplash is a mechanism of injury applied to the neck that is commonly experienced as a result of a motor vehicle collision and may often lead to pain and disability.

Whiplash injuries and the resulting Whiplash Associated Disorders (WAD) are an increasingly significant healthcare complaint, both globally and in the UK.

Findings from previous research indicate that it is unclear as to the exact factors that influence recovery from WAD. This uncertainty leads to difficulty in managing WAD efficiently and effectively. Should the knowledge of risk factors increase, the efficacy of management for WAD could be improved and benefit individuals and society as a whole.

This thesis is concerned with the assessment of cervical spine range of motion (ROM) in patients with sub-acute WAD. Assessment of ROM is part of the clinical assessment process used by various types of healthcare clinicians. It is believed to assist in the process of diagnosis and prognosis. A number of methods are available to measure ROM. It is unclear which is the most reliable, especially for a WAD population.

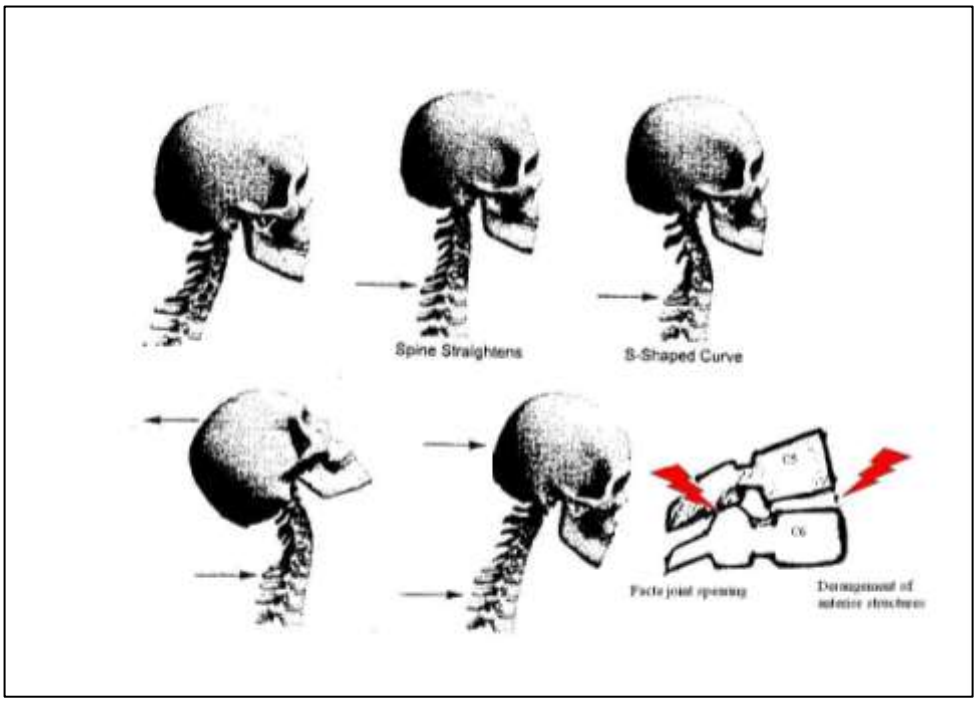
This thesis aims to investigate the value of ROM as a diagnostic and prognostic tool for WAD and to investigate which are the most clinimetrically sound methods for measuring it.

1.3 INTRODUCTION TO WHIPLASH INJURIES AND WAD

1.3.1 THE WHIPLASH MECHANISM OF INJURY AND WAD

The use of the terminology 'whiplash' was first reportedly used by Harold Crowe in the late 1920's and was intended to be used solely to describe the mechanism of injury that is now most commonly associated with Motor Vehicle Collisions (MVC) [1]. The exact direction of the mechanism of injury has gone backwards and forwards, so to speak. Gay and Abbott [2] described the mechanics as a forced flexion then extension movement; however it has since been consistently proven that following a rear end MVC the head remains relatively stationary whilst the body is thrust forward thus creating an initial forced extension of the cervical spine [3, 4]. There is evidence to suggest that an abnormal S-shaped curve is created with lower level hyperextension and upper level flexion in the early phases of the mechanism of injury [4, 5]. This results in abnormal strain being placed on joints and tissues to both the front and rear of the neck (Figure 1). There is also the possibility that a rotation element can be added to the mechanism of injury depending on the direction of the forces involved.

Figure 1- The Whiplash mechanism of injury and associated facet-joint spearing



In the early 1990's a group of experts was assembled to form the Quebec Task Force on Whiplash-Associated Disorders (QTF). The Task Force was charged to improve understanding and make recommendations on " The epidemiology; mechanisms of injury; clinical definitions and syndromes; natural history; evidence of effectiveness of prevention, treatment and rehabilitation; the role of psychosocial factors; and the impact of health services system in general to formulate a rational approach to the problem" of whiplash and its associated disorders [6].

The QTF defined whiplash as "an acceleration-deceleration mechanism of energy transfer to the neck. It may result from rear-end or side-impact motor vehicle collisions, but can also occur during diving or other mishaps." [6]

The term Whiplash Associated Disorders (WAD) was utilised by the QTF to describe the variety of clinical manifestations that arise from a whiplash injury and has become the most widely used definition on an international scale.

The reported incidence of whiplash injuries varies around the world, being dependent on traffic volumes, road conditions and litigation systems. Unsurprisingly, high rates are found in developed countries with high population density and high car ownership. Incidences range from 70 per 100, 000 in Quebec [6] up to 387 per 100,000 in the USA [7]. In the UK, there appears to have been a substantial increase in the incidence of whiplash injuries during the 1980s and 1990s and there were around 250,000 new cases in 2003 [8].

WAD has become a major problem in terms of health-care, personal and medico-legal costs. The annual cost of whiplash injuries to the UK economy has been estimated at circa £3.1 billion per annum, representing a significant amount of the Gross Domestic Product [8]. Health services costs are considerable, with physiotherapy costs representing a substantial component of health care expenditure [6].

1.3.2 DIAGNOSIS AND PROGNOSIS OF WAD

Diagnosis of structural damage is difficult following a whiplash injury when using "objective" measures such as imaging or specific mechanical tests. As a result "subjective" reporting is largely relied on to determine the extent of effect of the whiplash injury on the individual.

The QTF produced a clinical classification of WAD that were not based on cause or source of problems but serve as descriptors of presentation and correspond roughly to severity (see Table 1).

Table 1 - QTF Classification of Whiplash-Associated Disorders [6]

Term	Definition
WAD Grade 0	No neck complaints or signs
WAD Grade I	Complaint of pain, stiffness or tenderness, but no physical signs
WAD Grade II	Complaint of pain, stiffness or tenderness, and musculo-skeletal signs (decreased range of motion, point tenderness etc)
WAD Grade III	Complaint of pain, stiffness or tenderness and neurological signs (decreased or absent deep tendon reflexes, weakness and sensory deficits). Could also have musculo-skeletal signs.
WAD Grade IV	Fracture or dislocation

Prognosis of WAD is still not wholly understood. It is thought to be multi-factorial in nature but the hierarchy of factors continues to be debated. It is believed to involve a combination of physical, psychological and social factors.

Rates of recovery from WAD are variable, with published figures ranging from 16% [9] to 72% [10], however the consensus appears to be that prognosis is favourable and the condition self-limiting. Contrary to this, one study's findings appear to show that general population *beliefs* about prognosis of Whiplash injury are more negative than other conditions [11].

A plethora of outcome measures have been used for researching WAD and failure to recover has yet to reach a standardised definition. According to the QTF, chronic WAD is defined as problems lasting greater than six months. This is consistent with another term found in the literature - Late Whiplash Syndrome (LWS). Balla [12] appears to be the first author to

define this as “the presence of pain, restriction of motion or other symptoms at six months or more following a whiplash injury, sufficient to hinder return to normal activities such as driving, usual occupation and leisure activities.”

1.3.3 MANAGEMENT OF WAD

There is very little good quality evidence for effective conservative treatments for acute WAD and for the prevention of chronic problems. The Cochrane Review by Verhagen et al [13] could not provide a conclusive statement regarding findings of trials evaluating a range of conservative treatments.

Before commencing the Managing Injuries of the Neck Trial (MINT - See Chapter Two) the study team conducted a survey of UK Emergency Department (ED) consultants in order to estimate usual care practice and contents of ED advice sheets [14]. From the responses it was concluded that verbal advice to exercise reinforced with brief written information and pain-relieving medication was used by the majority of departments.

There is no evidence for what treatments are currently used by physiotherapy and other allied health professionals following the initial emergency medical care provided. Treatments used may range from simple exercises, joint mobilisations to use of complex electrotherapeutic agents such as pulsed electromagnetic therapy and acupuncture.

As a result of the uncertainty of the treatment effectiveness and with the knowledge that improvements in treatments are necessary to try and reduce the numbers of patients failing to recover from WAD, the Managing Injuries of the Neck Trial was commissioned on behalf of the UK Department of Health. As well as evaluating treatments, this provided the ideal opportunity to investigate factors that affect recovery following a whiplash injury and hopefully to provide new clinical and research knowledge to benefit patients.

1.4 THESIS OVERVIEW

The studies presented in this thesis are supported, described and analysed in seven further chapters.

Chapter Two describes the Managing Injuries of the Neck Trial, a large randomised controlled trial evaluating conservative treatments for WAD in which the prospective cohort in Chapter Seven was nested.

Chapter Three presents a systematic literature review of physical prognostic factors for poor outcome following a whiplash injury.

Chapter Four provides an introduction to the cervical spine and its assessment and management in the context of WAD.

Chapter Five describes a systematic literature review of reliability and validity studies of methods for measuring cervical spine ROM.

Chapter Six presents studies evaluating the intra- and inter-tester observer for the Cervical Range of Motion Device (CROM) in a WAD population.

Chapter Seven documents cross-sectional and longitudinal analyses of a prospective cohort of individuals with WAD with a focus on factors and mechanisms of recovery after a whiplash injury related to ROM.

Chapter Eight provides a summarising discussion of all the studies presented in preceding chapters, exploring research and clinical implications of the findings.

1.5 SUMMARY

This chapter has introduced the concept of whiplash injuries and their healthcare management. It has sought to provide an overview of what is contained in this thesis (forewarned is forearmed!). The next chapter will describe the Managing Injuries of the Neck Trial which provides the source of participants and resulting data for the work in this thesis.

2 CHAPTER TWO - MANAGING INJURIES OF THE NECK TRIAL (MINT)

2.1 INTRODUCTION

A summary of a large pragmatic randomised controlled trial (RCT) is presented in this chapter in order to provide context for studies described in later chapters of the thesis. A summary of the pertinent methodology and results are described along with the author's contribution to the trial. For a copy of the published manuscript of the study protocol see Appendix 1 [15].

2.2 BACKGROUND

As outlined in the previous chapter, Whiplash Associated Disorders (WAD) are an increasing national and global problem and current conservative treatments are varied and not supported by sound evidence.

As a result of this uncertainty regarding efficacious management, the National Institute for Health Research's (NIHR) Health Technology Assessment (HTA) programme commissioned a study to assess the clinical and cost effectiveness of a conservative, active management approach for acute whiplash injuries to prevent chronic problems. The commissioning brief requested a trial of non-surgical, non-pharmacological treatments for WAD, applied within the first six weeks of injury.

The study team, based at Warwick Clinical Trials Unit at the University of Warwick proposed a multi-centre Randomised Controlled Trial. This was the Managing Injuries of the Neck Trial (MINT).

2.2.1 RATIONALE FOR THE MANAGING INJURIES OF THE NECK TRIAL (MINT)

Clinical pathways for managing acute WAD have been proposed by the Quebec Task Force (QTF) [6] and Scholten-Peeters et al [16] using "expert consensus". These pathways suggest a stepped care approach whereby patients are provided with advice and education initially, then reviewed at approximately three weeks if problems persist and provided with more intensive treatment if appropriate.

Advice and education are considered to be a vital first step in the management of WAD. The QTF concluded that key messages that should be conveyed to patients experiencing acute WAD are:

- Pain is to be expected, is often short-lived and should be manageable
- Early return to normal activities appears to produce a favourable outcome
- The use of soft neck collars may prolong the recovery process

According to a survey conducted by the study team, UK Emergency Department (ED) advice was variable and may have contained conflicting messages [14]. Written information did not appear to mirror the recommendations from the QTF.

Accounting for evidence that suggests that psychological risk factors are important in recovery from WAD [17, 18], a group of "experts" developed a psycho-educational booklet – The Whiplash Book [19]. This publication potentially offers advice and education superior to usual care, providing information about the favourable prognosis of the condition and encouraging active coping strategies to return to usual activity levels assisted by exercises. There is a notable absence of advertisement about pursuing a personal injury claim when

compared to existing advice materials. Although there is evidence that the booklet positively modifies people's thinking about WAD [20], evaluation was required to quantify any healthcare benefit. At the time of commencing the trial only a very small number (<5%) of UK EDs were using The Whiplash Book [14]. Furthermore, most of the literature cited to support the booklet was from the field of low back pain. It is questionable whether this is appropriate due to the difference in the aetiology and course of the two conditions.

Physiotherapy is commonly used by the NHS for patients with WAD who fail to recover. It appeared to be a logical choice as the second component of the stepped care pathway evaluated by MINT. There was and remains no published information that outlines what current UK physiotherapy practice consists of, therefore a number of principles were used to develop the interventions, described later in the methods section. For patients with persistent problems, a package of physiotherapy modalities was compared to an advice session conducted by a physiotherapist, at which the previously-administered ED advice was reinforced.

2.2.2 MINT RESEARCH AIMS

1. To estimate the clinical effectiveness of a stepped care approach for acute whiplash injuries over a 12 month period

Step One: The Whiplash Book and active management approach versus usual care advice in Emergency Departments

Step Two: For patients with persistent symptoms, additional treatment consisting of a package of physiotherapy versus reinforcement of ED advice by a physiotherapist

AND: The combined effect of the differing treatments

2. To estimate the clinical effectiveness in pre-specified sub-groups of patients (pre-injury neck problems, psychological disturbance due to the injury, higher initial injury severity (WAD grade), and those seeking compensation).

3. To estimate the costs of each strategy and to estimate cost effectiveness

4. To gain a qualitative understanding of the patent's perspective of experiencing a whiplash injury and the subsequent NHS treatment within MINT.

2.3 METHODS

2.3.1 STEP ONE: ACTIVE MANAGEMENT APPROACH VS USUAL CARE ADVICE

2.3.1.1 Selection of participants

The first step of the trial was cluster randomised whereby NHS Trusts were allocated to one of the two ED advice interventions.

All patients attending a participating ED following a whiplash injury less than six weeks prior were eligible for the trial.

The following exclusion criteria applied:

- Age less than 18 years.
- Fractures or dislocations of the cervical spine or any other part of the body.
- Head injuries with more than a transient loss of consciousness or with a Glasgow Coma Scale score [21] of 12 or less at any stage of their assessment in hospital.
- Severe psychiatric illness as assessed by the ED staff

- Admission to in-patient services from the ED

Due to Step One of the trial being cluster randomised, patients did not have a choice as to whether to receive the trial advice interventions and therefore consent was not sought individually. This is an accepted procedure for cluster randomised trials [22].

ED clinicians at participating trusts identified eligible participants and recorded a core clinical data set on the ED Proforma. This form contained details on injury severity, pain intensity and WAD grade [6] and was carbonised in order for one copy to be filed in the medical notes and the other copy returned to the study HQ (Warwick Clinical Trials Unit). Departments were provided with information materials (e.g. posters) to advertise the trial to patients and remind clinicians of the selection criteria and trial systems.

Eligible patients were given a trial information pack (letter of introduction and the appropriate advice leaflet) and the trial discussed with them. Patients were not told about the differing advice in the EDs, but that the hospital was taking part in a study of advice given to patients following a whiplash injury. If they were willing to participate they were told to expect a questionnaire in the next few days. Patients were asked for their contact details (address, telephone numbers and email) to assist with follow-up procedures. Patients who did not wish to be contacted had this noted on their ED Proforma. Besides the advice leaflet (The Whiplash Book or usual care advice leaflet), patients were provided with verbal guidance on management of their injury.

Patients were informed about their potential eligibility for Step Two of the study if they continued to have problems after a few weeks following their ED attendance and to contact the study team on a Freephone number if this was the case. The majority of patients who participated in Step One were not expected to have persistent symptoms at three weeks so

detailed information about Step Two was only provided to patients once they contacted the study team.

2.3.1.2 Randomisation

For Step One the unit of randomisation was the NHS trust. Participating Trusts were randomised to Usual Care Advice (UCA) or Whiplash Book Advice (WBA) before the start of recruitment by the project statisticians. Trusts were pair matched on size of the ED (based on number of ED attendances per year, 2004/5 figures). A table of random numbers was used, starting at a random place. The allocation depended on whether the next digit was even or odd.

2.3.1.3 Experimental intervention - Active management approach including the Whiplash Book

ED clinicians (doctors, nurses and allied health professionals) were trained to deliver key messages during the consultation and to highlight the use of The Whiplash Book [19]. The Whiplash Book consists of 26 pages of A5 with illustrations on every page.

The key messages were:

- Reassurance that prognosis following a whiplash injury is good
- Reassurance that pain is normal and analgesia should be used to try and get pain under control
- Encouragement to return to usual activities as soon as possible with the use of exercises to facilitate this
- Advice against using a collar

Existing training slots for rotational inductions or in-service training programmes were used and repeated approximately every four months. Training consisted of a 30 to 40 minute session to educate clinicians about WAD, the trial, recruitment and how to deliver the active management approach. The trial research clinicians were also regularly present in the ED to provide information and support.

2.3.1.4 Control intervention - Usual care

Advice leaflets from all EDs involved in the trial were collected and reviewed prior to randomisation. All of the leaflets were consistent with the most frequently occurring pattern of advice according to the survey carried out by the trial team (Lamb et al [14] - discussed previously). It appeared that all departments potentially randomised to the usual care arm would provide a consistent control arm that would approximately represent usual care in the UK at that time.

Training was developed for the usual care EDs in order to provide a similar length session to the experimental intervention training. Focus was on general information about WAD and how to recruit patients into the trial. No specific instructions were presented about the management of WAD, with emphasis to continue providing advice that was usually given in the department. Frequency of training sessions was the same as the experimental arm of the study.

2.3.1.5 Outcome measures and data collection

The primary outcome measure used for MINT was the Neck Disability Index (NDI) [23]. This is an industry-standard measure of pain-related disability used in a number of previous intervention studies for neck pain/WAD [24, 25]. The 12-Item Short Form Health Survey (SF-12) [26] and EuroQol EQ5D questionnaire (EQ-5D) [27] were administered to assess

generic health-related quality of life alongside health resource questions for both NHS and private treatments.

An in-depth description of the pertinent measures for the cohort study is provided in Chapter Seven (cohort study chapter).

There were six data collection points for both stages of the trial which provided a comprehensive record of patient journeys. Table 2 displays the various measures and the time points at which they were collected.

Table 2 - Measures at the various time points of MINT

Follow-up time point	Data collection	Participants in Step	Measures
Zero	ED Proforma	1&2	Mechanism of injury, pain location and intensity, WAD grade, Medical history
2-week	2 week Questionnaire	1&2	Demographics, Pre-injury neck pain, Symptoms, Pain troublesomeness, SF-12, EQ-5D, ED treatment satisfaction
1 month	Research Clinic Questionnaire and Examination	2 only	Treatment preference, Recovery expectations, Return to work status, Crash details, NDI, Pain ratings, Patient rated cervical ROM, FABQ (P), self-efficacy, PCS, CSOQ, IES, MSPSS, GHQ-12, Treatment expectations, number of physical symptoms, presence of chronic widespread pain, cervical ROM, shoulder abduction ROM
4 months	4 month follow-up questionnaire	1&2	NDI, SF-12, EQ-5D, Health resource, FABQ (Physical), Coping
8 months	8 month follow-up questionnaire	1&2	NDI, SF-12, EQ-5D, Health resource, FABQ (Physical), Coping
12 months	12 month follow-up questionnaire	1&2	NDI, SF-12, EQ-5D, Health resource, FABQ (Physical), Coping

Abbreviations: NDI = Neck Disability Index, FABQ (P) = Fear Avoidance Beliefs Questionnaire (Physical subscale), PCS = Pain Catastrophising Scale, CSOQ = Cervical Spine Questionnaire, IES = Impact of Events Scale, MSPSS = Multidimensional Perceived Social Support, GHQ-12 = General Health Questionnaire – 12 score version.

All eligible patients who were approached whilst attending a participating ED had reference information obtained using the previously described ED Proforma completed by the ED clinicians at the time of assessment. This included the clinician providing a categorisation of the patients' injury severity status in the form of a WAD grade [6].

Participants were then sent a questionnaire as soon as their details had been received and processed at the study HQ (Warwick Clinical Trials Unit). This 2 week questionnaire (Appendix 2) consisted of demographic details, health-related quality of life measure (SF-12), health economic questionnaire (EQ-5D) and an ED treatment satisfaction question.

All patients were then followed up at common time points irrespective of which steps of the trial they participated in, provided they returned either a two week or four month questionnaire. Postal questionnaires were sent out at four, eight and 12 months after their date of ED attendance (Appendix 12). A standardised method of ensuring the maximum response from participants was employed including telephone and postal reminders. If participants had not returned a questionnaire after a pre-defined time, an effort was made to collect core outcome data over the telephone which included the NDI, EQ-5D and health economics questionnaires.

Data were single-entered into a bespoke Microsoft Access database and were cleaned on a weekly basis using a linked computer programme. Administration staff independent of the recruitment, randomisation or intervention processes were responsible for sending and inputting the questionnaire.

2.3.2 STEP TWO: PHYSIOTHERAPY PACKAGE VERSUS ADVICE SESSION

Step Two of MINT is central to the work subsequently presented in this thesis because the participants entering this step of the study simultaneously entered into the prognostic cohort study described in Chapter Seven.

2.3.2.1 Selection of participants

Patients approached in Step One of MINT were asked to contact the study HQ (Warwick Clinical Trials Unit) if they continued to experience symptoms approximately three weeks after their ED attendance. When any patient contacted the study HQ, a research therapist performed an initial screening by telephone and if the patient appeared to be eligible then an appointment was made to attend a research clinic. Information about the second step of the trial was sent to the patients in the days prior to the appointment to ensure they had sufficient time to consider participation and discuss with appropriate parties if necessary. This is in accordance with Good Clinical Practice principles [28]. The research clinics were conducted in the hospital trust where the patient had attended the ED, usually in the ED or in a therapy clinic space. This resulted in assessment of patients in a sub-acute state.

At the research clinic the patients were checked for eligibility for Step Two of MINT and the cohort study according to the following criteria:

- Reporting cervical spine symptoms within the last 24 hours
- Were WAD Grade I-III at time of assessment

- Did not have any contra-indications to physiotherapy treatment. This included central cord compression, upper or lower motor neuron lesion, complete nerve root compression, suspected vascular injury or haemorrhagic event.

If eligible, trial information was discussed and the patient was asked to provide written informed consent prior to randomisation. If the patient did not wish to participate at this point their reasons were requested on a voluntary basis and recorded if given. Once consent and randomisation had occurred, participants were asked to complete a Research Clinic Questionnaire (Appendix 3) and then the research therapist conducted an assessment, completing the Research Clinic Examination form (Appendix 4).

2.3.2.2 Randomisation

Randomisation to the interventions for Step Two of MINT was via a central telephone randomisation service (Birmingham Cancer Trials Unit, University of Birmingham). Randomisation was stratified by centre to ensure balance at each of the sites between the different interventions.

2.3.2.3 Experimental intervention - Physiotherapy package

A number of requirements were identified when considering the design of the physiotherapy package. The experimental intervention needed to be based on high quality evidence and clinical practice guidelines where possible, whilst being feasible to deliver in an NHS setting and not impinging on physiotherapists' autonomy. The intervention was fully documented in a manual to enable consistency and repeatability. The trial team conducted systematic literature reviews of randomised controlled trials and observational studies. There was no published research on the current physiotherapeutic treatments delivered in the UK, despite their widespread use in WAD. As previously described, in the mid 1990's there was a

distinct lack of evidence for conservative treatments for the management of WAD [6] and this had not really changed by 2004 when Verhagen et al [29] conducted a systematic review on behalf of the Cochrane collaboration. A trend for active interventions being effective was apparent alongside some weak evidence suggesting that multimodal approaches consisting of manual therapy, exercise and psychological approaches could be helpful. Numerous observational studies offered some indications of potentially modifiable risk factors for poor outcome from WAD. Two systematic literature reviews were conducted, one for psychological factors and one for physical factors, the latter being described in detail in Chapter Three. Conclusions were that physical and psychological factors may be equally important, with the most important factors highlighted as high initial pain and disability, low self-efficacy and an elevated stress response. To a lesser extent high fear avoidance (fear of re-injury), catastrophising ("excessively negative and unrealistic thoughts or self-statements about pain" [30]), inappropriate coping behaviours, reduced range of motion and joint position and muscle dysfunction were thought to contribute to a poorer prognosis. One set of clinical practice guidelines had been published at the time of the trial intervention development, which endorsed the use of exercise to facilitate a graduated return to activities [16].

Having gleaned as much information from the previous literature, a framework was documented for assessing risk factors for poor outcome and then matching potentially effective treatment strategies to this risk factor profile; a strategy commonly utilised in musculoskeletal physiotherapy. A "treatment planner" was designed to foster consistency in risk factor identification and generating treatment targets between participating physiotherapists. Although treatments were individualised, the physiotherapists were encouraged to consider both physical and psychological factors, resulting in three main

treatment components; manual therapy according to the Maitland approach [31], exercise therapy (including ROM, postural control and proprioception options) and psychological strategies and self-management advice (based on a cognitive behavioural approach). The intervention was to consist of an assessment session then up to six sessions of treatment with the aim to deliver this over an eight week period.

For more details of the Step Two physiotherapy interventions see Williamson et al [32].

2.3.2.4 Control intervention - Re-enforcement of advice session

The control intervention was a single session with a physiotherapist at which advice received in the ED was re-enforced. A brief assessment of symptoms and active cervical Range of Motion was permitted to allow advice to be tailored within the limits of that already prescribed. Physiotherapists were not permitted to provide any 'hands on' treatment or progress exercise beyond that described in the ED advice sheet. Patients were advised to see their General Practitioner if they experienced any further problems. The session lasted approximately 40 minutes.

All physiotherapists were trained to deliver both interventions and were independent of the recruitment and randomisation processes and had received one and a half days training from the trial research physiotherapists. Treatment logs were completed for each participant and returned to the study HQ (Warwick Clinical Trials Unit).

2.3.2.5 Outcome measures and data collection

The outcome measures and data collection methods used in Step Two were identical to those in Step One described above. The only extra data collection point was the research clinic. The development and justification for data collected at this point is documented in more detail in Chapters Six and Seven.

2.3.3 STATISTICAL ANALYSIS

Statistical analyses for MINT were carried out by the trial statisticians using the computer package STATA 10 (StataCorp).

2.3.3.1 Sample size

For the primary outcome (NDI) results from previous studies suggest a minimally clinically important difference lies in the range of three to five absolute points, with a standard deviation of approximately eight [33]. For the purposes of MINT it was decided to aim to be able to detect a three point difference between groups for both steps of the trial. With Step One a cluster randomised comparison, larger numbers were required dependent on the number of clusters. Originally eight centres were planned to participate, however an additional four centres were included in response to slower recruitment than expected. Assuming an intra-cluster correlation of 0.02, 90% power, 1% significance, 0.375 standard deviations and 30% loss to follow up the sample size would be 2004 (167 per cluster). The sample size was inflated to take account of the likelihood that recruitment to clusters would be varied; therefore the target sample size of 3,000 was adopted. For Step Two the same assumptions were used (0.375 Standard Deviations detected between NDI scores, 90% power, 1% significance ICC 0.02 and 30% loss to follow-up) to set a target sample size of 600.

2.3.4 ETHICAL APPROVALS

MINT is registered with ISRCTN, # 3302125. The study was approved by the Trent Multicentre Research Ethics Committee (reference MREC/04/4/003) and by the Local Research Ethics Committee and Research & Development department of each participating NHS trust (See Appendix 5 for the MREC approval letter).

2.3.5 MONITORING

A Trial Steering Committee (TSC) and Data Monitoring and Ethics Committee (DMEC) were convened at regular intervals throughout the trial. The TSC was responsible for ensuring the trial was conducted to rigorous standards to ensure patient safety. The DMEC was responsible for monitoring the ethical and data integrity aspects of the trial.

2.4 RESULTS

A full account of all of the clinical and cost-effectiveness results of MINT is beyond the required explanation for the purposes of this thesis; however presentation of the numbers and characteristics of participants in both steps of the trial and a summary of clinical results will assist in the interpretation of the cohort study described in Chapter Seven. Full results of the trial will be published in an HTA monograph, which is in press at the time of submission of this thesis.

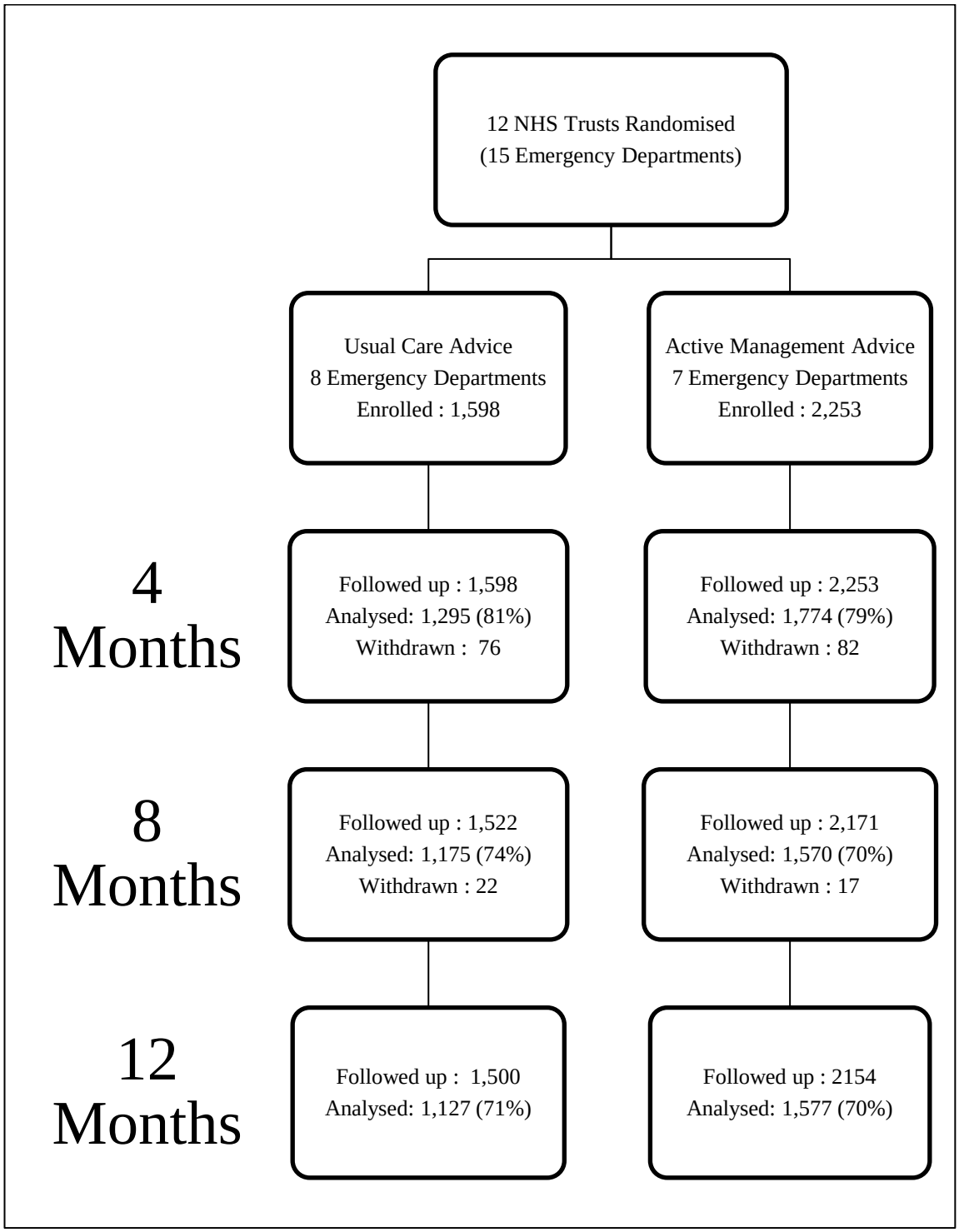
2.4.1 STEP ONE

2.4.1.1 Recruitment

15 Emergency Departments from 12 NHS trusts were involved in the recruitment of patients. Trusts were randomised in a 1:1 ratio to both arms of the trial, resulting in eight EDs (from six trusts) delivering usual care advice (UCA) and seven EDs (from six trusts) delivering whiplash book advice (WBA). Recruitment was set up in a staggered fashion and ran from December 2005 until November 2007. During this period trial proformas were completed for 7,702 patients - 3,034 for UCA arm, 4,668 for WBA arm. Just under 50% of eligible patients attending the EDs had a trial pro-forma completed. There were no major differences in the proportion of patients referred to the trial between the arms. 6952 of the 7,702 patients were eligible for the trial and were sent a two-week questionnaire. 3851 (55%) patients returned

the questionnaire and were thus enrolled to the study. The CONSORT flow chart in Figure 2 summarises the recruitment and the numbers of participants subsequently followed-up at the multiple time points.

Figure 2 - CONSORT Flow diagram for Step One



The two arms of Step One were well matched in terms of the characteristics of participants (see Table 3) with the exception of a small difference in ethnicity, a result of the populations served by EDs in the different arms. The vast majority of participants had suffered their whiplash injury as a result of a motor vehicle collision (94%), were of working age (mean 37 yrs) and diagnosed with a WAD grade of I or II (97%). There were slightly more females recruited to the study (67%), a common occurrence in studies of WAD.

Table 3 – Characteristics of Step One participants by arm*

	UCA	**Missing**	**WBA**	**Missing**
Number enrolled	1,598		2,253	
Gender – Males	666 (42%)	18	995 (44%)	39
Age in years, Mean [SD]	37 [13]	0	37 [13]	0
Ethnic Group		118		224
White	1,336 (84%)		1,586 (70%)	
Mixed	19 (1%)		42 (2%)	
Indian	49 (3%)		95 (4%)	
Pakistani	24 (2%)		179 (8%)	
Bangladeshi	9 (1%)		21 (1%)	
Black or Black British	31 (2%)		69 (3%)	
Chinese or Other	12 (1%)		37 (2%)	
Mechanism of injury		15		14
Road Traffic Accident	1,495 (94%)		2,127 (94%)	
Other	88 (6%)		112 (5%)	
Location of pain		37		73
C-spine only	1,046 (65%)		1,365 (61%)	
C-spine and other spinal area	275 (17%)		400 (18%)	
Other spinal area only	31 (2%)		65 (3%)	
Spinal & other area	141 (9%)		190 (8%)	
Other area only	23 (1%)		56 (2%)	

No pain	45 (3%)		104 (5%)	
Pain intensity (/10), mean [SD]	4.9 [1.9]	349	5.3 [1.9]	574
History				
Previous neck problems	190 (12%)	58	218 (10%)	94
Previous back problems	199 (12%)	308	285 (13%)	396
Neurological symptoms	98 (6%)	57	121 (5%)	86
WAD grades		0		0
I: Complaint of pain, stiffness or tenderness, no physical signs	883 (55%)		1,205 (53%)	
II: Complaint of pain, stiffness or tenderness, musculoskeletal signs	662 (41%)		997 (44%)	
III: Complaint of pain, stiffness or tenderness, neurological signs	53 (3.3%)		51 (2.3%)	
Employment		155		272
Working/earning	1,185 (74%)		1,549 (69%)	
Unpaid work	4 (0.25%)		7 (0.31%)	
Not working	254 (16%)		425 (19%)	

* Mechanism of injury, location of pain, pain intensity, medical history, and WAD grades were collected at ED attendance. Gender, age, and ethnic group were collected on the two week questionnaire.

2.4.1.2 Follow-up

Numbers of participants providing outcome data at the three follow-up time points are presented in Figure 2. Loss to follow-up was 20%, 29% and 30% at 4, 8 and 12 months respectively. Those lost to follow-up as non-responders or withdrawals were well matched between the two arms. There were no unexpected and related serious adverse events reported in Step One of the trial.

2.4.1.3 Clinical results

The majority of participants had recovered at 12 months. 18% of the cohort was classified as having LWS. There were no statistically or clinically significant differences in outcomes for participants in WBA and UCA arms (difference in NDI at 12 months 0.5, 95% CI -1.5-2.5). There was no evidence that the advice interventions were affected by initial injury severity, adverse psychological reactions to injury, pre-existing neck problems, or compensation.

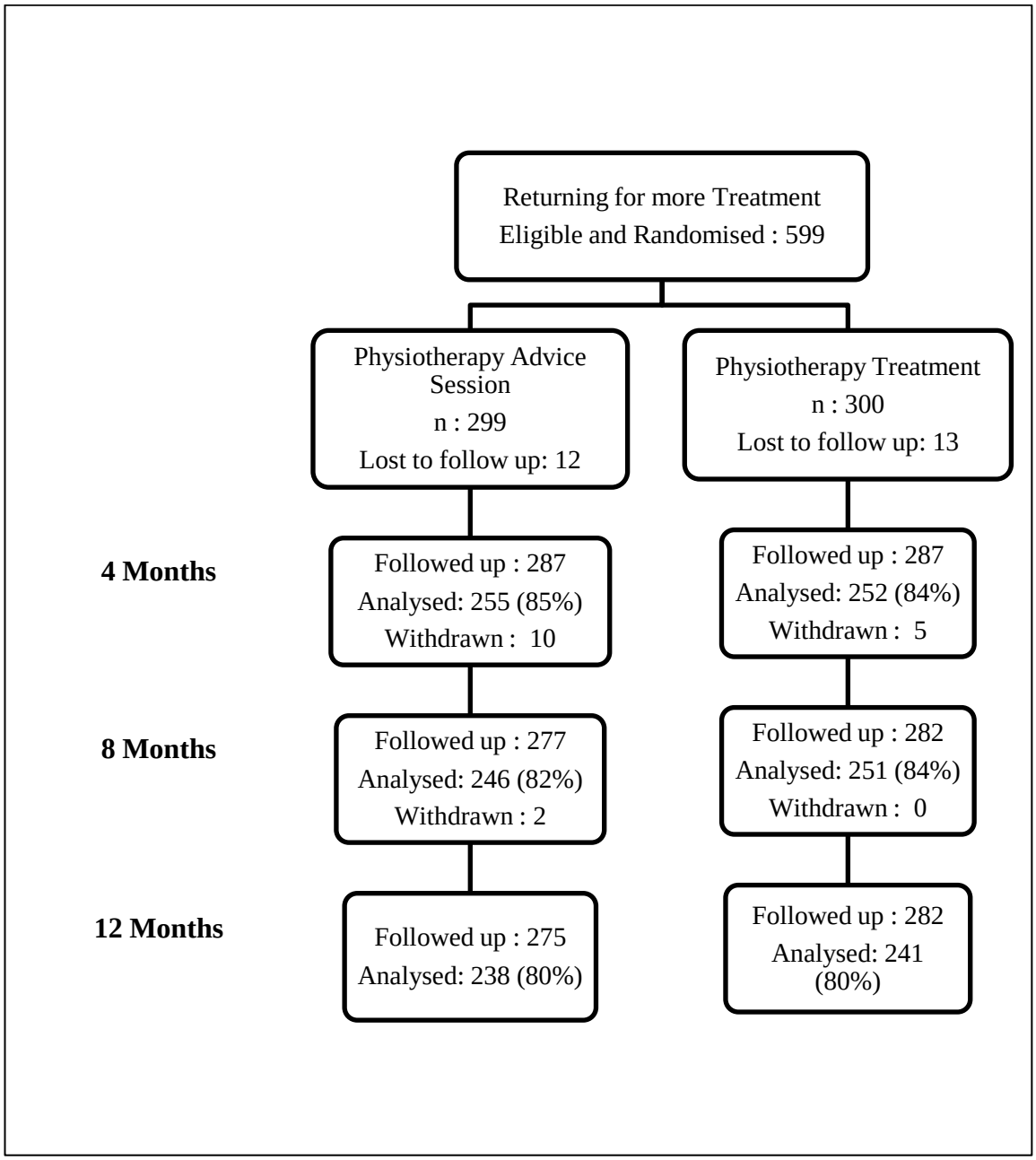
2.4.2 STEP TWO

2.4.2.1 Recruitment

Recruitment to Step Two ran concurrently with recruitment to the first step of MINT (Dec 2005 to Nov 2007). 599 patients gave informed consent and were recruited into Step Two. Figure 3 displays the flow of patients followed-up through the second step of MINT.

949 of the patients recruited to Step One reported on-going problems to the trial HQ, and were considered for the second step of MINT. Of these, 693 were assessed as potentially eligible and were invited to attend a research clinic appointment. 77 patients did not attend or cancelled their appointment leaving 616 patients to be assessed for eligibility at the research clinics. Two patients were ineligible and 15 declined to participate resulting in 599 consenting patients recruited.

Figure 3 – CONSORT flow diagram for Step Two



There were slight differences in characteristics between Step One participants who entered Step Two and those who did not (see Table 4). There were a lesser proportion of males entering Step Two; a greater proportion of participants had experienced neck pain in the month prior to their injury, a greater proportion with neurological signs (WAD grade III) and a lower health-related quality of life (as measured by the SF-12).

Table 4 – Characteristics of those entering and not entering Step Two of MINT

	Randomised	Missing	Not randomised	Missing
Number of patients	599	0	3,277	0
Sex – Males	221 (37%)	0	1,456 (44%)	50
Age in years, Mean [SD]	40 [13]	0	36 [13]	0
Had previous neck pain	77 (13%)	40	334 (10%)	115
WAD grades		0		0
0: No neck complaints or signs	0 (0%)		0 (0%)	
I: Complaints of pain, stiffness or tenderness, no physical signs	275 (46%)		1,823 (56%)	
II: Complaint of pain, stiffness or tenderness, musculoskeletal signs	299 (50%)		1,375 (42%)	
III: Complaint of pain, stiffness or tenderness, neurological signs	25 (4%)		79 (2%)	
IV: Fracture/Dislocation	0 (0%)		0 (0%)	
SF-12v1 scores, Mean [SD]				
Mental component score	36 [12]	108	42 [13]	692
Physical component score	36 [7]	108	41 [9]	692
Received public fund	192 (33%)	37	748 (23%)	332

The two arms of Step Two were also well matched other than a difference in NDI score at point of randomisation. The physiotherapy arm had a mean NDI 5 points greater than the advice session arm representing a greater amount of disability. Table 5 summarises the characteristics of the Step Two population by arm.

Table 5 – Characteristics of Step Two participants

	Advice	Missing	Physiotherapy	Missing
Number randomised	299		300	
Sex – Males	115 (38%)	0	106 (35%)	0
Age in years, Mean [SD]	40 [13]	0	40 [13]	0
Ethnic Group		16		20
White	229 (77%)		226 (75%)	
Mixed	2 (.67%)		3 (1%)	
Indian	18 (6%)		19 (6.3%)	
Pakistani	19 (6.4%)		19 (6.3%)	
Bangladeshi	2 (.67%)		1 (.33%)	
Black or Black British	10 (3.3%)		7 (2.3%)	
Chinese or Other	3 (1%)		5 (1.7%)	
Mechanism of injury		2		1
Road traffic accident	284 (95%)		286 (95%)	
Other	13 (4.3%)		13 (4.3%)	
Location of pain		8		12
C-spine only	178 (60%)		178 (59%)	
C-spine and other spinal area	60 (20%)		62 (21%)	
Other spinal area only	10 (3%)		3 (1%)	
Spinal + other area	27 (9%)		27 (9%)	
Other area only	3 (1%)		6 (2%)	
No pain	13 (4%)		12 (4%)	
Pain intensity (/10), mean [SD]	5.4 [1.9]	69	5.6 [1.9]	91
History				
Previous neck problems	36 (12%)	21	41 (14%)	19

Previous back problems	40 (13%)	56	43 (14%)	65
Neurological symptoms	26 (9%)	19	29 (10%)	15
WAD grades		0		0
0: No neck complaints or signs	0 (0%)		0 (0%)	
I: Complaints of pain, stiffness or tenderness, no physical signs	39 (13%)		45 (15%)	
II: Complaint of pain, stiffness or tenderness, musculoskeletal signs	222 (74%)		220 (73%)	
III: Complaint of pain, stiffness or tenderness, neurological signs	38 (13%)		35 (12%)	
IV: Fracture/Dislocation	0 (0%)		0 (0%)	
SF-12v1 scores, mean [SD]				
Norm-based MCS	37 [12]	53	35 [12]	55
Norm-based PCS	36 [7]	53	36 [6.9]	55
Received any public funds	91 (30%)	17	101 (34%)	20
Neck disability index[2] (%), mean [SD]	39 [16]	3	44 [16]	7
Employment		22		29
Working/Earning	225 (75%)		210 (70%)	
Unpaid work	0 (0%)		2 (1%)	
Not working	52 (17%)		59 (20%)	

2.4.2.2 Follow-up

Numbers of participants providing outcome data at the three follow-up time points are presented in Figure 3. Loss to follow-up was 8%, 13% and 20% at 4, 8 and 12 months respectively. Those lost to follow-up as non-responders or withdrawals were well matched between the two arms. There were no unexpected and related serious adverse events reported in Step Two of the trial.

2.4.2.3 Treatments delivered

Fifty-five senior physiotherapists (median qualification time 6.5 yrs [IQR 4.5-18]) received one to one and a half days training in order to deliver treatments for both arms of Step 2 of the trial. Table 6 displays the treatment attendance rates for both arms. The majority of participants completed treatment as recommended in the protocol.

Table 6 – Step Two treatment attendance rates

	Physiotherapy Package (n=300)	**Advice session (n=299)**
Failed to attend any appointments	34 (11%)	60 (20%)
Attended for assessment only*	26 (9%)	N/A
Partial completion of treatment**	45 (15%)	N/A
Completed treatment*	201 (67%)	239 (80%)

*Six patients attended the assessment session and required no further treatment; therefore these participants are included in both categories.

**Partial completion of treatment was attendance of an assessment session and at least one treatment session but treatment not being deemed completed as intended with mutual agreement on discharge between participant and therapist.

For the 239 participants that attended the advice session, most had their ED exercises reviewed and were given advice on pain control and posture or positioning (see Table 7). Almost all (97%) of the participants had their cervical Range of Motion assessed. About one third of participants had a neurological examination.

Table 7 – Management of advice session participants

	Number of participants receiving advice n=239 (%)
Assessed ROM	232 (97)
Neurological examination	87 (37)
Referred on due to serious complication	2 (1)
Reviewed exercises given in ED	228 (95)
Postural or positioning advice	200(84)
Advice re: collar	56 (22)
Advice re: pain control or medication use	185 (78)
Advised to see GP if have ongoing problems	220 (92)
Other advice	25 (8)

Information on the content of the physiotherapy package treatment sessions was complete for 259 of the 266 participants who attended more than one appointment (see Table 8). The majority received a combination of manual therapy, exercises and psychological strategies (73%). Almost all participants received guidance on cervical range of movement exercises (94%).

Table 8 – Types of treatments and combinations delivered within the physiotherapy package (adapted from Williamson et al[32])

Type of treatment delivered		**Number of patients receiving treatment (%) (n=259)**
Combinations of treatments delivered	Manual therapy, exercises and psychological strategies	190 (73)
	Exercises and psychological strategies	45 (18)
	Manual therapy and psychological strategies	10 (4)
	Manual therapy and exercises	9 (4)
	Manual therapy only	2 (1)
	Exercises only	2 (1)
	Psychological strategies only	1 (1)
Manual therapy techniques	Soft tissue techniques	123 (48)
	Maitland cervical mobilisations	123 (48)
Exercises	Cervical range of movement exercises	244 (94)
	Cervical or scapular stability exercises	118 (46)
Psychological strategies and self-management advice	Advice about posture and positioning	194 (76)
	Reassurance	194 (75)

2.4.2.4 Clinical results

The physiotherapy package resulted in short term improvements in neck disability in comparison to the advice session with a physiotherapist (difference in NDI at 4 months -3.7, 95% CI -6.1, -1.3), but these effects were not maintained at 12 months (difference in NDI at 12 months -2.0, 95% CI -4.6, 0.6). The physiotherapy package was also accompanied by a significant reduction in work days lost (Difference at 12 months -4 days, 95% CI -7.5- to -0.02days).

2.5 DISCUSSION

The clinical and cost effectiveness analyses of MINT suggest that an active management approach applied in the ED is no more effective than usual care. A physiotherapy package provided to individuals with ongoing symptoms was beneficial in the short term when compared to a single session of advice and resulted in a reduction in lost work days.

Assessment and treatment of cervical Range of Motion was a consistent feature in both arms of Step Two of the trial, indicating that therapists felt this was an important factor to address when attempting to facilitate recovery from WAD.

To date, MINT is the largest trial evaluating conservative treatments for acute WAD. This size brings both opportunities and threats. Having recruited from a number of areas across the UK and given the broad selection criteria it is likely that the findings are generalisable to patients using acute NHS services in England and Wales. The challenge of a multi-site study is the variation in pre-existing care processes specific to the individual EDs. The infrastructure assembled for this project enabled, as far as practically possible, rigorous training and delivery procedures including a quality assurance programme to ensure that

recruitment, assessment and intervention tasks were conducted according to the appropriate protocols, which the author co-ordinated. Having a dedicated team of administration staff provided the opportunity for a thorough, standardised follow-up protocol including repeat questionnaires and telephone interviews for core outcomes if necessary. This resulted in a lower than expected attrition of patients at follow-up compared to similar studies of acute injury [34], thus reducing potential bias.

Using standardised early treatments (both Steps One and Two) and multiple outcome measures evaluating a range of constructs at numerous time points provides an excellent opportunity to explore the recovery process of WAD and the factors that influence this.

2.6 AUTHORS CONTRIBUTION TO MINT

As the clinical trial co-ordinator/Research Fellow of MINT the author was responsible for the day-to-day running of the study. Processes that the author was involved with:

- Post-award modification and operationalising of trial design
- Leading design of ED trial proforma
- Design of questionnaire for ED survey of WAD management
- Operationalising and modifying recruitment process between Steps One and Two
- Development of the interventions for Steps One and Two
- Development and delivery of training programmes for Steps One and Two
- Development of the outcome measures and data collection methods for research clinics and postal questionnaires– in particular the decision to measure cervical

ROM and the resulting method and protocol for measurement were solely conducted by the author.

- Visiting EDs to support Step One recruitment
- Conducting Research Clinic assessments (recruitment and randomisation to Step Two for a number of the West Midlands sites)
- Monitoring of recruitment
- Applying for ethical and NHS trust governance approvals
- Monitoring of administrative procedures (follow-up questionnaires)
- Contributing to the publication of the trial protocol (Appendix One)
- Writing and editing the final report to the funders (NCCHTA)

2.7 SUMMARY

This chapter has summarised the background, methods and pertinent results of MINT to provide contextual orientation for subsequent chapters. The large, robust RCT described was not only well powered for pre-specified main and sub-group analyses but provides an opportunity to explore diagnostic and prognostic factors in a large, generalisable group of patients affected by WAD. The subject of the next chapter is to establish what is known about physical risk factors for outcome in a WAD population through a systematic literature review.

3 CHAPTER THREE – A SYSTEMATIC LITERATURE REVIEW OF PHYSICAL PROGNOSTIC FACTORS FOR THE DEVELOPMENT OF LATE WHIPLASH SYNDROME

3.1 INTRODUCTION

The last chapter described the Managing Injuries of the Neck Trial (MINT), a large RCT aiming to evaluate the conservative management of acute WAD. The cohort assembled for Step Two of MINT provides an ideal opportunity to evaluate factors influencing recovery in a prospective way because it was assembled at a common point in the disorder (within 6 weeks of injury) and participants received standardised initial management [35]. A prognostic study can be carried out in a variety of ways, investigating one variable in particular, or a number of variables simultaneously and looking at predicting outcome or response to treatment. These studies may help to increase the understanding of a disease process and groups at greatest risk of developing it and subsequently improving design of clinical trials and treatment selection [36]. This chapter aims to evaluate the literature regarding physical prognostic factors for WAD.

3.2 BACKGROUND

There have been numerous prognostic studies of WAD, with many factors being cited to influence recovery. A review of the literature is necessary to summarise what is and more often what is not known about an area, with a systematic literature review being defined as a "scientific tool which can be used to summarise, appraise, and communicate the results and

implications of otherwise unmanageable quantities of research."[37] It offers several advantages over a traditional narrative review including objectivity, repeatability and quantitative summarisation. In the case of prognosis it may offer a point estimate for the relative importance of the specified factors (using appropriate meta-analysis) and/or insights into what new studies are required. As explained in the previous chapter, treatment regimes have so far proven to be largely ineffective for the management of WAD [13] so identifying factors that influence outcome might allow for development of more effective treatment strategies.

In the mid 1990's the Quebec Task Force on Whiplash Associated Disorders (QTF) published an extensive systematic literature review including studies completed and published as of September 1993 [6]. They reviewed 66 articles pertaining to prognostic factors for WAD and deemed 11 as "acceptable" in terms of clinical relevance. Only five of these studies had "acceptable" design but were still flawed methodologically. The QTF concluded that there was a pressing need for studies of all types of prognostic factors in WAD populations.

Subsequently there was an upsurge in prognostic studies which were summarised in two systematic literature reviews [18, 38]. All three reviews have drawn conflicting conclusions on the relative importance of mechanism of injury, demographic, physical, psychological and social prognostic factors. It is likely that these reviews produced different findings because of their differences in methodology (sourcing studies, extraction of data and assessing quality). There is currently no consensus on how systematic reviews of prognostic studies should be conducted and reported. In the period between publication of the recent review by Scholten-Peeters et al [38] and the conduct of this review, a considerable (>10 studies) amount of literature focussing on prognostic factors was published, highlighting a

need for an up-to-date review to inform the prognostic investigations that were critical to this thesis. Since the conduct of the review described in this chapter there have been a number of additional systematic reviews which are discussed in Chapter Eight.

It was decided to limit the review of literature to purely physical prognostic factors due to the large number of studies involved and the interest and clinical importance of physical factors that are routinely used to make decisions about management of WAD. The emerging importance of psychosocial factors is acknowledged and readers are referred to other contributions by the author regarding psychosocial prognostic factors [39]. The review presented in this chapter has been published as a manuscript in the journal Spine [40] (Appendix 6). The reporting of this review incorporated guidelines from the QUOROM statement [41] which, although aiming to improve quality of reporting of systematic reviews and meta-analyses of RCTs, is generally pertinent to improving the transparency of reporting of systematic reviews of other types of studies. The author is aware that this statement has subsequently been updated to the PRISMA statement [42] since the conduct of this review.

The aim of this review was to identify physical prognostic factors for the development of Late Whiplash Syndrome in adults.

3.3 METHODS

Approaching a systematic review in the same way one would conduct an observational study, appropriate steps were undertaken to minimise sources of bias where possible. Steps included formulating a research question and developing a study protocol a priori, collecting and analysing studies, interpreting the results and reporting in a transparent way [43].

The first step was to assemble a group who had experience in research synthesis and knowledge of the content area. As a result of MINT, an appropriate group of researchers

with experience in both the assessment and treatment of WAD and conducting research synthesis were assembled.

3.3.1 RESEARCH QUESTION

The next step was to define the research question. This was generated partially as a result of the decision to take a risk factor modification approach for the MINT physiotherapy intervention (see Chapter Two) and also because of the group's awareness of the existing literature. Having made initial approaches to the literature, it was evident that the volume of studies investigating prognostic factors was large and increasing. With the awareness of the quantity of resources required to do a thorough systematic literature review, discussions were had between supervisors and other members of the review team to ensure the research question that was proposed would be feasible to answer. It was decided to split the potential literature into two reviews, one investigating studies of psychological factors and the other of physical factors. The author's interest in physical factors was due to the way health care practitioners predominantly assess WAD patients in an acute setting with a focus on physical measures. The author therefore led the review on physical factors which attempted to answer the research question:

What are the 'physical' prognostic factors for poor outcome following a whiplash injury?

3.3.2 STUDY SELECTION

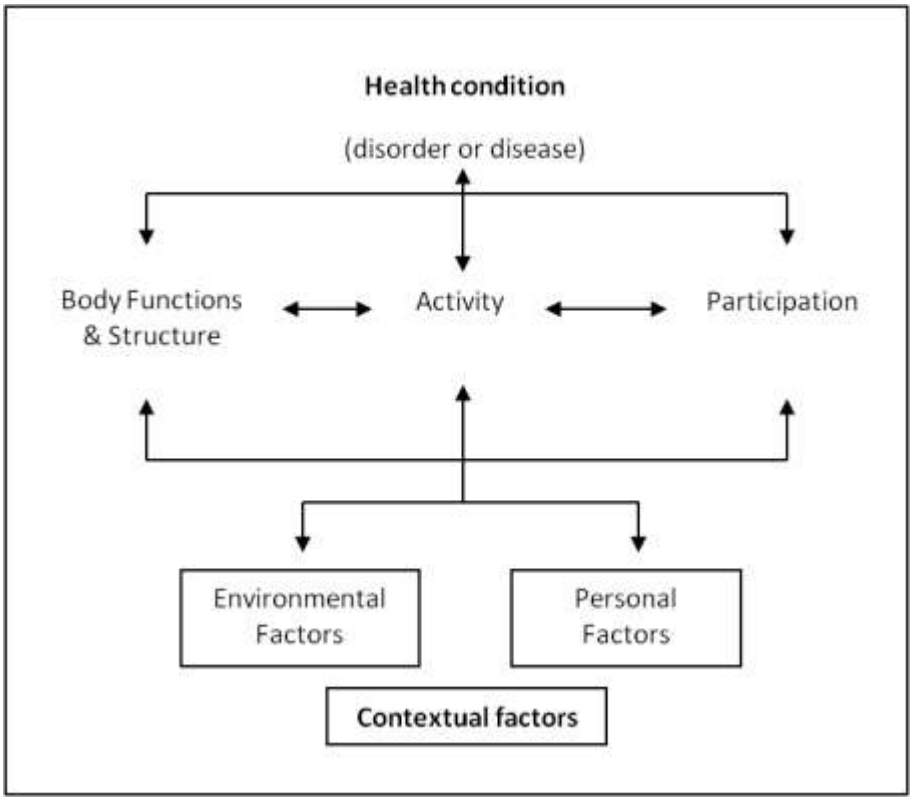
The next step was to define the selection criteria for studies. It was decided that only prospective studies should be included due to the increased likelihood of bias that comes with retrospective analysis of prognostic factors [43]. In this context prospective is taken to mean collection of data on prognostic factors between exposure to a whiplash injury and the

development of chronic or long term problems. Clinical or population-based cohorts or case-control studies were deemed acceptable designs for assessing prognostic factors.

It was important that studies were investigating individuals suffering from acute WAD (of less than 6 weeks duration) as it is expected that most recovery occurs early in the condition [6]. This time point was also taken due to the chosen design for MINT. It was important to ascertain the risk factors for this particular patient group in order to develop an accurate evidence based intervention. Studies that included children were excluded due to the brief for MINT and adults being the main recipients of this type of healthcare management. For the purposes of this review a 'physical' factor was considered to be one that directly involves a body function and/or structure. This definition was informed by the biopsychosocial model of dis/ability that is the basis for the World Health Organisation's International Classification of Functioning, Disability and Health [44]. Measuring body or body part functions and structures informs a healthcare professional of any presenting impairments which, alongside environmental and personal factors, may lead to activity limitations and participation restrictions (see Figure 4).

Examples of functions and structures pertinent to measurement in studies reviewed include sensory functions and pain via the nervous system and also and movement related functions via the neuromusculoskeletal system.

Figure 4 – Representation of model of disability that is the basis for International Classification of Functioning, Disability and Health



Outcome from WAD can be measured using a plethora of tools and constructs. Most commonly, pain and disability are evaluated. In Chapter One the definition of Late Whiplash Syndrome (LWS) was introduced and defined as "ongoing (chronic) problems affecting activities of daily living for at least 6 months following a whiplash injury". Considering that this was the outcome of interest for this thesis and MINT, a minimum follow-up period of 6 months was set for eligible studies.

There are a small number of studies that have used time to closure of compensation claim as an outcome measure [45-47]. It was decided to exclude this method of outcome measurement from our review because its clinical relevance is questionable. Time to

symptomatic or disability recovery may not always coincide with ending of disability compensation [48].The considerations for the selection of studies described above were formulated into inclusion and exclusion criteria as follows:

3.3.2.1 Inclusion criteria

- Prospective clinical or population-based cohort studies or case-controlled studies.
- Studies investigating at least one physical prognostic factor at baseline and the development of LWS.
- Inception cohorts to have been assembled within 6 weeks of whiplash injury.
- Follow up for a minimum of 6 months post whiplash injury.
- Subjects to be 18 years or older.
- Outcome measures to be related to the clinical presentation of LWS (e.g. pain or disability due to neck problems > 6 months post injury).
- English language

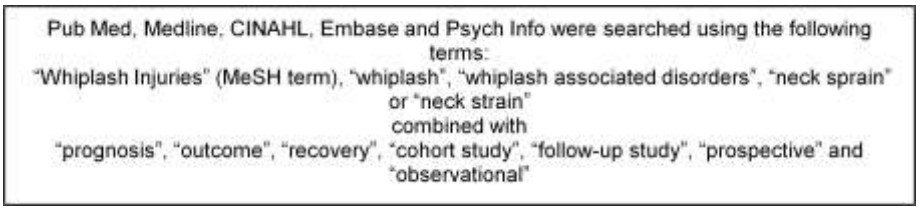
3.3.2.2 Exclusion criteria

- Studies about neck pain other than that arising from a whiplash mechanism of injury
- Studies using outcome measures unrelated to LWS e.g. “time-to-claim closure”

3.3.3 SEARCH STRATEGY

A search strategy was developed and is shown in Figure 5. Five electronic databases were searched for relevant studies. It was deemed important to not only use MEDLINE but also EMBASE to address the bias in geographical areas they cover (MEDLINE has more North American journals listed, whereas EMBASE tends to cover European research better – their overlap is estimated to be about a third [49]). Recommendations from information scientists regarding appropriate words and MeSH terms were considered [50]. Electronic databases were searched from their inception to August 2006. Bibliographies of previous systematic reviews of prognostic cohort studies were searched. Articles were eligible for the review if they fulfilled the selection criteria described above.

Figure 5 – Search strategy



Following the literature search, abstracts were screened independently by two reviewers referring to full papers if necessary. Where consensus could not be achieved by discussion between the two reviewers, a third reviewer was consulted and final decision made through discussion. It was felt that the possible benefits that result from blinding reviewers to publication details could not be justified in light of the limitations on resources (cost and time) for the purposes of this doctoral study.

3.3.4 DATA EXTRACTION AND QUALITY ASSESSMENT

A standardised data extraction form was created and used to document study characteristics, methodology, prognostic factors studied, outcome measures used and results. Two independent reviewers extracted data and assessed quality for each article. Any discrepancies were discussed to achieve consensus. If disagreement persisted then another of the authors was consulted and a final decision made.

A quality assessment tool (See Figure 6) was developed using recommendations of Altman [51] and taking into consideration the study populations. Quality scoring was divided into three sections; patient sampling, measurements used and analysis. Rather than using a total score to decide quality ratings, scores were considered from each of the three sections. Each section was designed to have equal importance. This was to prevent studies that scored very highly in one section but very poorly in others gaining a rating that may exaggerate the

overall methodological quality. Both the data extraction and quality assessment tools were piloted on two papers and slight adjustments were made to the contents regarding recording sample sizes and if multivariate analysis was done, which methods were used. Ideally study analyses would adjust for covariates to control for other variables that may influence outcome; this may additionally allow for investigation of how much prognostic value the factor of interest offers over what has been shown in previous research.

Figure 6 – Quality Assessment tool

Section one: sampling of patients

1.Inclusion criteria are defined
2.Exclusion criteria are defined
3.Source population is defined
4.Adequate description of diagnostic criteria for classifying patients with a whiplash injury
5.Clinical and demographic characteristics are fully described
6.The sample is representative of the majority of patients with a whiplash injury
7.The sample is assembled at a common point (within a 2 week period)
8.The sample is complete
9.Any treatment received is fully described (including no treatment)

Section two: measurements used

1.Standardized or valid outcome measures have been used
2.Unbiased evaluation of outcome (blinded)
3.Prognostic factors are fully defined, including details of method of measurements
4.Measurements used for the prognostic factors are standardized or validated

Section three: analysis

1.Was the sample size adequate for the number of prognostic factors included in the analysis (minimum of 10 per factor)?
2.Loss to follow up is < 20%
3.Was multivariate analysis carried out?
4.Statistical adjustment for important prognostic factors including age, sex, previous neck pain and initial pain severity

Scoring

Fully described = 2
Poorly described = 1
Not described/insufficient information = 0

Individual articles were assessed rather than providing an overall rating for each cohort because data, analyses and reporting often differed between articles of the same cohort. Scholten-Peeters et al [38] combined articles in the same cohorts to provide an overall quality assessment and suggested this may have introduced some bias. This could lead to an overestimation in quality for cohorts with multiple publications. Therefore, each *article* was rated according to the following definitions (similar to those used by Scholten-Peeters et al [38]):

High-quality: study scores 75% or above for all 3 sections

Adequate-quality: study scores at least 50% for all 3 sections

Low-quality: score of less than 50% for any one section

3.3.5 DATA ANALYSIS AND SYNTHESIS

Following quality assessment and rating, results from the studies were tabulated. Meta-analysis was not performed due to the heterogeneity between studies and insufficient data. It follows that it was not appropriate to construct forest plots. This may have been feasible with individual patient data but this additional work was not feasible within the constraints of resources of this project.

Levels of evidence were generated instead by grouping similar findings from cohorts using a "vote counting" procedure. The overall levels of evidence for an association of a prognostic factor with LWS were defined according to the definitions below. These are similar to the definitions used in previous systematic reviews [38].

Strong evidence: Consistent findings in at least 2 high quality studies from different cohorts.

Moderate evidence: Consistent findings in at least 2 adequate quality studies from different cohorts.

Limited evidence: Findings in one adequate quality cohort or at least 2 low quality studies from different cohorts.

Inconclusive evidence: Inconsistent findings or insufficient research (e.g. evidence from one low quality cohort only)

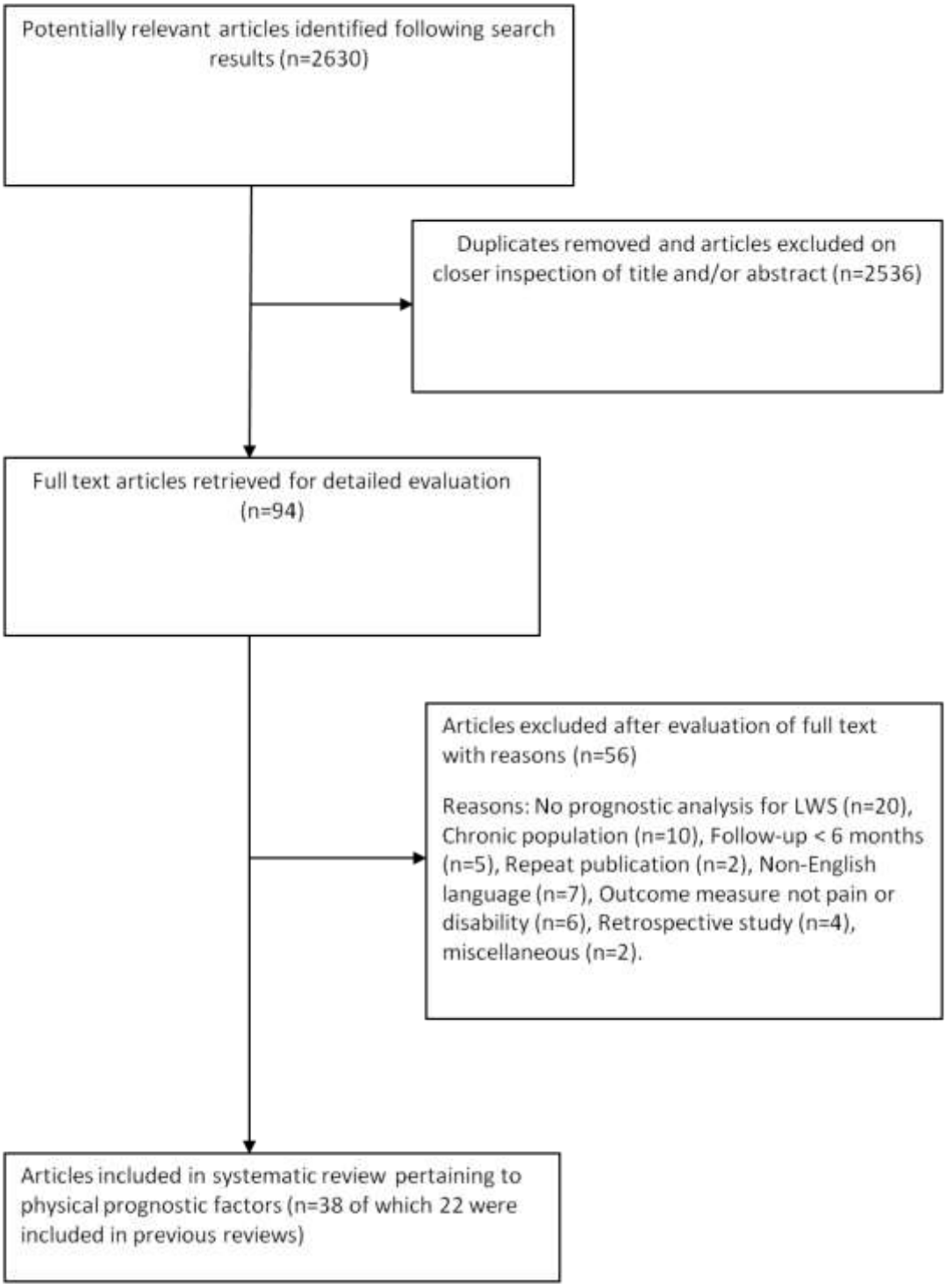
Levels of evidence were defined using findings from cohorts rather than articles. Failing to do this could lead to a situation where several high quality publications from the same cohort could be used to classify a factor as having strong evidence for an association with LWS even if not replicated in other cohorts.

3.4 RESULTS

3.4.1 STUDY SELECTION

Figure 7 shows the results of the search strategy and flow of the study. A large number of articles were discarded following the initial search result (n=2536). Following detailed full text evaluation 38 articles reporting data from 26 cohorts were deemed eligible for this review.

Figure 7 – Study flow diagram for systematic review of physical prognostic factors for Late Whiplash Syndrome



Characteristics of the eligible studies are presented in Table 9. All studies were prospective cohort studies except one case control study [52] and publication year ranged from 1974 to 2006. The majority of cohorts (n=15) recruited solely from emergency departments. The number of participants ranged from 29 [53] to 1030 [54]. Follow-up ranged from 6 months (minimum specified by inclusion criteria) to >5 years [55]. Loss to follow-up varied between 0% [56, 57] and 73% [55]. 32 different physical factor constructs were studied.

Table 9 - Study Characteristics

Cohort	Author	Population	Number of subjects*	Loss to follow-up	Length of follow-up	Physical factors studied	Other prognostic factors studied
1	Atherton et al, 2006[58]	Emergency Department	480/765	37	1 year	Previous neck pain, presence of widespread chronic pain, initial injury severity (VAS), number of WAD symptoms, WAD grade, bony tenderness, neurological signs, limited Range of Motion	General health, number of GP visits in previous 12 months, collision factors, initial disability (NDI),age, gender, Psychosocial work factors (WS), psychological state (GHQ), somatisation (MSPQ)
2	Borchgrevink et al, 1995[59] Borchgrevink et al, 1997a[60]	Emergency Department	50/52 88/99	4, 11	6 months	MRI Results, XR Findings	Personality profile

3	Brison et al, 2000[61] Hartling et al, 2001[62]	Emergency Department	334/380 126/380	12, 67	6 months 2 years	Height BMI Presence of WAD at baseline WAD Grading at baseline	Gender, Age, Crash characteristics
4	Gargan and Bannister 1994[63] Gargan et al, 1997[56]	Emergency Department	50/50 50/50	0	2 years	Cervical ROM, Symptom severity	Psychological state (GHQ-28)
5	Gun et al, 2005[64]	Emergency Department, medical and physiotherapy practices	135/147	8	1 year	Quality of Life (SF-36) (bodily pain score)	Age, Consulting a lawyer, Vehicle damage, Use of Head Rest, Previous claim for MVA, Treatment by physio/chiropractor
6	Hendriks et al, 2005[65]	Emergency Department and General Practice	119/125	5	1 year	Cervical ROM, Neck pain intensity, Number of complaints, Radicular complaints, Diagnostic imaging	Age, Gender, Education, Marital status, Crash Related factors, Pre-existing health factors, Pain medication, Ability to perform ADL, Psychological symptoms (SCL-90), Work activities,Absent

							from work, Use of collar
7	Herrstrom et al 2000[66]	Emergency Department and Primary Care	125/158	21	1 year (mean)	Previous chronic headache/neck pain	Sex, Type of accident, Sick leave
8	Hildingsson and Toolanen, 1990[67]	Orthopaedic Department	93/97	4	25 months (mean)	X-ray findings, History of Neck Pain, Height, Neck pain, Neck stiffness, Headache, Shoulder pain, Arm pain/numbness, Dizziness, Visual symptoms, Auditory symptoms	Crash Factors, Gender

9	Hohl 1974[55]	Private Orthopaedic Clinic	146/534	73	>5 years	Unconsciousness, Radiating pain, Forward head posture, Muscle spasm, Cervical ROM, X-ray findings	Age, Gender, Property damage, Site of initial treatment, Use of collar, Time until claim settlement, Amount of settlement, Surgical treatment
10	Karlsborg et al, 1997[10]	Emergency Department	34/39	13	7 months	WAD Grade, Number of symptoms at baseline, Neuropsychology, MRI results, Motor evoked potentials	Gender, Age, Presence of stress unrelated to the accident
11	Kasch et al, 2001[68] Kasch et al, 2005[69]	Emergency Department	132/141	6	1 year	BMI, Pain severity, Presence of neurological symptoms, Number or symptoms, Active cervical ROM, Work load (cervical muscles), Cold induced pain ratings (cold pressor test), Discomfort following cold pressor test, Pressure induced pain	Gender, Age, Health behaviour, Speed difference between vehicles, claiming compensation within the first month

						threshold.	
12	Kivioja et al, 2005[70]	Emergency Department	91/96	5	1 year	Initial pain severity, Previous neck and shoulder pain (month before RTA).	Age, Gender, Coping strategies
13	Kyhlback et al, 2002[71]	Emergency Department	83/98	15	1 year	Pain intensity (VAS), WAD Grade	Self-efficacy (SES), Pain and disability (PDI), Age, Gender
14	Mayou and Bryant, 1996[72]	Emergency Department	57/63	10	1 year	Initial Physical Symptoms	Sex, Age, Psychological factors, Previous psychological problems, Driver/passenger status

15	Miettinen et al, 2004[73]	Insurance Company Records	144/312	54	3 years	Neck pain, Headache, Symptoms of the upper extremities, Previous symptoms	Depression (BDI), Psychological status (GHQ-12), Neck disability (NDI), Ability to work, Crash characteristics
16	Miles et al 1988[57]	Emergency Department	73/73	0	2 years	X-ray findings	
17	Minton et al 2000[74]	Emergency Department	134/174	23	1 year	Height, Weight	Gender, Impact speed, Impact direction, Head rest type, Head rest distance, Awareness of impending accident, Seating position
18	Nederhand et al 2003, 2004[75, 76]	Emergency Department	141/154 82/90	8, 9	6 months	Muscle EMG, Pain intensity	Gender, Age, Collision direction, Functional status, NDI, Fear of Movement (TSK), Catastrophising (PCL-E)
19	Olsson et al 2002[77]	Emergency Department	123/130	5	1 year	Pain intensity, Condition severity (WAD Grade)	Psychological response to pain (MPI), Age, Gender,
20	Pettersson et al 1997[78]	Orthopaedic Department	39/40	2	2 years	MRI imaging, Neurological examination	Quality of Life (unvalidated Qu)

21	Radanov et al, 1991[79] Radanov et al, 1993a[80] Radanov et al, 1993b[81] Radanov et al, 1994b[82] Radanov et al, 1995[83] Sturzenegger et al, 1995[84] Di Stefano and Radanov, 1995[52]	General Practice	78/92 98/113 117/137 117/137 117/137 117/137 42/42	15,13, 15, 0	6 months 1 year 6 months 1 year 2 years 1 year 2 years	Initial pain intensity, Initial subjective complaints, Neurological examination, Timing of onset of symptoms, Baseline cervical ROM, Radiological examination, History of pre-traumatic headache, Previous head trauma, Previous whiplash injury, Type and frequency of pre-traumatic headache	Gender, Age, Educational attainment, Vocational related variables, Crash related variables, Psychosocial stress, Psychological variables, Sleep disturbance, Wellbeing, Personality traits, Cognitive function
22	Richter et al, 2004[85]	Emergency Department	32/43	25	6 months	Clinical findings, Radiological findings	Pain control, Quality of Life (SF-36, EDLQ), Speed difference between vehicles, Litigation
23	Sterling et al, 2005[86] Sterling et al, 2006[87]	Emergency Department, General Practice, Advertisement	76/80 65/80	5	6 months 2-3 years	Physical measures (ROM, JPE, EMG, PPT's, TPT's, BPPT, Sympathetic function), Pain	Psychological Distress (IES), Fear of movement (TSK), Disability (NDI)

						Intensity (VAS)	
24	Sterner et al, 2003[88]	Emergency Department and General Practice	296/356	17	16 months (mean)	WAD Grade, Pre-injury Neck Pain, Pre-injury headache, Pre-injury back complaint	Educational level, Gender, Age, Accident type,
25	Voyvodic et al, 1997[53]	Physiotherapy and Private Medical Practice	27/29	7	6 months	MRI Findings, Cervical ROM, Pain intensity, Neurological assessment	Crash Factors, Gender, Previous Neck Injury
26	Warren and Warren, 2001[54]	Emergency Department	1027/1030	1	3 year or until recovery	Radiation of pain	Gender, Age, Time to pain onset, Occupation

*Number of participants at final follow up/number recruited

Abbreviations: WAD = Whiplash Associated Disorders, ROM = Range of Motion, BMI = Body Mass Index, MRI = Magnetic Resonance Imaging, PPT's = Pressure Pain Thresholds, TPT's = Thermal Pressure Thresholds, BPPT = Brachial Plexus Provocation Test, EMG = Electromyography, TSK =Tampa Scale of Kinesiophobia, PCL-E = Pain Cognition List – Experimental, NDI = Neck Disability Index, VAS = Visual Analogue Scale, MPI = West Haven-Yale Multidimensional Pain Inventory, PDI = Pain Disability Index, SES = Self Efficacy Scale, SCL-90 = Symptoms Checklist-90, TCI = Temperament and Character Inventory, GHQ-12 or 28 = General Health Questionnaire-12or 28, BDI = Beck Depression Inventory, WS = Karasek's demand-support-control model of workplace strain – 8 items, IES = Impact of Events Scale, MSPQ =Modified Somatic Perceptions Questionnaire.

3.4.2 QUALITY ASSESSMENT

The results of the quality assessment and quality ratings are shown in Table 10. The majority of articles (n=25/38) were rated as "low-quality" with the remaining 13 rated as "adequate-quality". No studies were rated as "high-quality". Only six articles failed to gain an "adequate score" for the "patient sampling" section. The main failing for most studies (30 articles) was lack of any description of treatment received during the follow-up period. Approximately half of the studies (n=18) scored below 50% for the "measures used" section. Only 14 studies (from 9 cohorts) used validated prognostic measures and even fewer (8 studies from 6 cohorts) used validated outcome measures. The majority of studies (n=31) did not report blinding of assessors to baseline data when evaluating outcome. Scores in the "analysis" section were commonly the lowest of the three sections. Only 16 studies (from 13 cohorts) had adequate sample sizes, 17 articles reported using multivariate techniques, and two studies reported having adjusted for the pre-specified prognostic factors.

For each article univariate and multivariate results for association to poor outcome (LWS) are presented in Table 11 and Table 12. Only statistically significant results are presented ($p<0.05$), however all results were considered in the vote counting procedure. Results have been separated into 2 tables depending on the type of outcome measure used. Table 11 presents results of studies that used disability-based outcome measures (e.g. Neck Disability Index). Table 12 presents results of studies that used symptom-based outcome measures (e.g. presence of pain or pain intensity). This separation was used because factors associated with a symptom-based outcome measure may not automatically be associated with a disability-based one [89]. Seven studies did not report any statistically significant results for physical prognostic factors (Cohorts 7, 8, 10, 17, 19, 22, 25) and two studies did not carry out analysis for association between prognostic factors and outcome (cohorts 4 and

20). If adequate data were presented unadjusted odds ratios were calculated by the reviewers and presented in the tables (Cohorts 20 and 25).

Table 10 - Quality Assessment ratings

	Patient Sampling										Measures Used					Analysis					Quality Rating
Article	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**%**	**10**	**11**	**12**	**13**	**%**	**14**	**15**	**16**	**17**	**%**	
Atherton et al, 2006[58]	2	2	2	2	1	1	2	2	0	78	0	0	2	2	50	2	0	2	2	75	**Adequate**
Borchgrevink et al, 1995[59]	2	2	2	1	1	?	2	0	0	56	0	?	2	0	25	2	2	0	0	50	**Low**
Borchgrevink et al, 1997a[60]	2	2	2	2	1	?	2	0	0	61	0	0	2	2	50	2	2	0	0	50	**Adequate**
Brison et al, 2000[61]	2	2	2	2	1	1	?	1	2	72	0	?	2	1	38	2	2	0	0	50	**Low**
Hartling et al, 2001[62]	2	2	2	2	0	1	?	1	0	56	0	2	2	2	75	2	2	2	1	88	**Adequate**
Gargan & Bannister 1994[56]	2	0	1	0	1	1	2	2	1	56	0	2	1	0	38	2	2	0	0	50	**Low**
Gargan et al, 1997[56]	2	0	1	0	1	1	2	2	1	56	0	2	1	1	50	0	2	0	0	25	**Low**
Gun et al, 2005[64]	1	2	2	0	2	0	2	0	0	50	1	?	2	1	50	?	2	?	0	25	**Low**
Hendriks et al, 2005[65]	2	2	1	2	2	0	2	0	2	72	2	2	1	1	75	0	2	2	0	50	**Adequate**
Herrstrom et al 2000[66]	2	2	2	2	1	1	?	2	0	67	0	?	1	0	13	2	0	0	0	25	**Low**
Hildingsson and Toolanen, 1990[67]	2	2	2	0	2	2	2	2	1	83	0	?	0	0	0	0	0	2	0	25	**Low**
Hohl 1974[55]	1	1	2	0	1	0	0	0	1	33	0	?	1	0	13	?	0	0	0	0	**Low**
Karlsborg et al, 1997[10]	2	2	2	2	2	0	2	0	1	72	0	?	1	1	25	0	2	2	1	63	**Low**
Kasch et al, 2001[68]	2	2	2	2	1	0	1	2	2	78	1	0	2	1	50	2	2	2	1	88	**Adequate**
Kasch et al, 2005[69]	2	2	2	2	1	?	2	0	2	72	1	0	2	1	50	2	2	0	0	50	**Adequate**
Kivioja et al, 2005[70]	2	2	2	2	2	1	2	1	0	78	0	0	2	2	50	2	2	2	2	100	**Adequate**
Kyhlback et al, 2002[71]	2	2	1	2	1	0	1	0	0	50	2	2	1	2	88	2	2	2	0	75	**Adequate**

Mayou and Bryant, 1996[72]	1	0	2	1	1	1	0	2	1	50	1	?	1	1	38	2	2	0	1	63	**Low**
Miettinen et al, 2004[73]	1	1	2	2	1	0	?	0	0	39	0	2	2	1	63	?	0	2	0	25	**Low**
Miles et al, 1988[57]	1	2	2	0	1	?	?	0	1	39	0	0	1	0	13	?	2	0	0	25	**Low**
Minton et al, 2000[74]	1	1	1	0	0	0	?	0	0	17	0	?	1	0	13	?	?	0	0	0	**Low**
Nederhand et al, 2003[75]	2	2	2	0	2	0	2	0	1	61	2	?	2	2	75	0	2	0	0	25	**Low**
Nederhand et al, 2004[76]	2	2	2	2	1	0	2	0	0	61	2	?	2	2	75	2	2	0	0	50	**Adequate**
Olsson et al, 2002[77]	2	2	2	2	2	1	1	2	2	89	1	?	2	2	63	2	2	2	0	75	**Adequate**
Pettersson et al, 1997[78]	2	2	2	2	2	0	2	2	2	89	1	2	1	1	63	0	2	0	0	25	**Low**
Radanov et al, 1991[79]	1	1	1	2	1	1	2	0	0	50	0	?	2	1	38	0	2	2	1	63	**Low**
Radanov et al, 1993a[80]	2	2	1	2	2	1	2	0	0	67	0	0	2	2	50	0	2	0	1	38	**Low**
Radanov et al, 1993b[81]	2	1	1	2	2	1	2	0	0	61	2	0	2	2	75	0	2	2	?	50	**Adequate**
Radanov et al, 1994b[82]	2	2	1	2	2	1	2	0	0	67	0	0	2	1	38	0	2	2	?	50	**Low**
Radanov et al, 1995[83]	2	2	1	2	2	1	2	0	0	67	0	0	2	1	38	0	2	0	0	25	**Low**
Sturzenegger et al, 1995[84]	2	2	1	1	2	1	2	0	1	67	0	0	2	1	38	0	2	2	0	50	**Low**
Di Stefano and Radanov, 1995[52]	2	2	1	2	2	0	2	0	0	61	0	0	2	2	50	0	2	0	0	25	**Low**
Richter et al, 2004[85]	2	2	2	2	2	0	2	0	2	78	1	0	1	1	38	0	0	?	0	0	**Low**
Sterling et al, 2005[86]	2	2	2	2	2	1	?	?	2	72	2	?	2	2	75	0	2	2	1	63	**Adequate**
Sterling et al, 2006[87]	2	2	2	2	2	1	?	?	0	61	2	?	2	2	75	0	2	2	1	63	**Adequate**
Sterner et al, 2003[88]	2	2	2	2	1	2	?	2	0	72	0	2	1	0	38	2	2	2	0	75	**Low**
Voyvodic et al, 1997[53]	2	1	1	1	1	0	0	0	0	33	0	0	0	0	0	0	0	0	0	0	**Low**
Warren and Warren, 2001[54]	1	1	1	0	1	0	0	0	0	22	0	0	0	0	0	2	?	0	0	25	**Low**

Table 11 - Results based on disability outcomes

Study	Outcome Measure	Univariate results	Test used	Multivariate results	Final Model Included
Gun et al, 2005[64]	Improvement in Neck Pain Outcome Score Questionnaire			Patients with a higher SF-36 Bodily Pain score (lower degree of bodily pain) had a greater improvement after 1 year (ß=0.18, p<0.01). Patients with a higher SF-36 Role Emotion score had a greater improvement after 1 year (ß=0.07, p<0.05). Patients who consulted a lawyer had less improvement after 1 year (ß=-7.1, p<0.01) Patients who had made a previous claim had less improvement after 1year (ß=-10.5, p<0.01).	Initial NPOS score.
Hendriks et al, 2005[65]	Functionally recovered vs. Non-recovered: VAS <30mm for neck pain or VAS >78mm for activities AND no pain medication use during follow-up.		Multiple logistic regression	Female Gender (OR 4.596 [1.507-14.015]), Low level of education (OR 3.511 [1.054-11.696]), high initial neck pain intensity (OR 1.020 [1.002-1.038]), Greater severe work activity limitation (OR 0.986 [0.975-0.998]), higher levels of somatisation (OR 1.110 [1.030-1.195]).	Seen by PT or GP.

Kasch et al, 2001[68]	Disabliity: Patients completed a 6 point scale to rate work capacity and handicap. Patients were considered handicapped (or non-recovered) if they selected items 3,4,5, or 6.		Cox regression analysis	Reduced total cervical ROM (measured in the first week post injury) (B=2.53 CI 1.26-5.11, p =0.01). ROM was considered to be a risk factor if it was 2 standard deviations below the total cervical ROM of the control group.	Included in model: Cervical muscle workload, pain (VAS), number of symptoms, gender, speed differences >26km/hr, age>31, BMI>30, lawsuit during first month post injury.
Kasch et al, 2005[69]	Disability: Patients completed a 6 point scale to rate work capacity and handicap. Patients were considered handicapped (or non-recovered) if they selected items 3,4,5, or 6.	Cold pressor test: Non-recovered patients had reduced time to peak pain ratings during the cold pressor test compared to recovered ($p=4.5x10^{-8}$) on initial testing. Peak pain ratings were higher in non-recovered than recovered ($p=1.8x10^{-6}$). Greater pain ratings over duration of the test measured by area under the curve ($p<3.5x10^{-7}$). Non-recovered patients reported greater discomfort on completion of the test ($p<0.02$). McGill Pain Questionnaire: Non-recovered patients had higher initial Pain Rating Index Scores than recovered ($p<0.001$)	Mann-Whitney U, Student t-test with Bonferoni correction.		

Kyhlback et al, 2002[71]	Pain related disability (PDI, 0-70)		General Linear Model	Lower self-efficacy (β= -0.56, P<0.001), Male gender (β= 0.27, P<0.05), Older age(β= 0.21, P<0.05).	Age, Sex, WAD Grade, WAD grade/sex.
Mayou and Bryant, 1996[72]	Poor Social Outcome: Determined by interview. The interviewer gave a global rating of all work, leisure and other social changes attributable to the injury.		Logistic regression	Patients with a history of previous psychological problems were 5 times more likely to have a poor social outcome (p<0.05)	Age, sex, driver-passenger status, neuroticism, previous psychological problems, memories of the accident, mood score and neck symptoms immediately following the accident.
Miettinen et al, 2004[73]	Change in Health Status: Self report. Subjects rated the effect of whiplash injury on their health as no change, slightly worse or significantly worse. Slightly worse or significantly worse were both considered to be a change in health status.	Neck Disability Index score (p<0.05, OR 7.4) neck pain (p<0.05, OR????), lower back pain (p<0.05, OR 3.4) reported soon after injury were associated with a change in health status. Initial BDI score>9, total score of GHQ>2, score of NDI >20 and WAI score<27 were all significantly associated with poor outcome.	Binomial linear regression (logistic regression)	NDI score (>20) was significantly related to poor outcome (p<0.05, OR 11.2).	Age, gender, marital status, condition al health before the accident, symptoms after the accident, Scores on Beck's depression inventory, General Health Questionnaire, Neck Disability Index, Work Ability Index.

Nederhand et al, 2003[75]	Recovered vs. disabled using NDI (Categorised - recovered 0-4, mild 5-14, moderate 15-34, severe 25-50)	Patients with moderate or severe disability had significantly less isometric muscle activity at 1 week compared to those who had recovered or had mild disability (p=0.000 for both). Patients with moderate or severe disability had significantly less dynamic muscle activity at 1 week compared to those who had recovered (p= 0.003).	Two-way ANOVA, Post-hoc with a Bonferroni procedure
Nederhand et al, 2004[76]	Recovered vs. disabled using NDI (Dichotomised <15 = recovered and >15 = disabled	Disabled patients had a higher mean BMI (p=0.015), more intense initial neck pain (p=0.000) had higher responses on the TSK (p=0.000) and PCL-E (p= 0.000) and lower isometric muscle activity (p=0.004). Initial NDI score (>15) was predictive of poor outcome at 6 months and was more predictive when combined with TSK score >40. Initial NDI score >15 is predictive of poor outcome with 54% probability. If this is combined with an initial TSK score >40 this increased to 83% probability.	Mann-Whitney U, Student t-test, chi-square test. ROC curves.

Sterling et al, 2005[86]	Neck Disability Index (Categorical: <8=Recovered, 10-28=Mild, >30=Mod/Severe)	Multiple logistic regression	Factors predictive of mod/severe disability (NDI score >30) P<0.05: High Initial NDI score (OR 1.06 [1.007-1.12]), Older age (OR 1.13 [1.03-1.23]), reduced cold pain threshold (OR 1.29 [1.05-1.58]) and elevated IES score (OR 1.11 [1.03-1.2]). Was able to correctly classify 86.7% of patients as to whether they had mod/sev symptoms or not using these variables. Factors predictive of mild disability (NDI score 10-18): Initial NDI score (OR 1.15 [1.03-1.28]), GHQ-28 total (OR 1.15 [1.04-1.28]) and Cervical Extension ROM (OR 1.1 [1.03-1.25]).	Age, sex, initial NDI score, physical measures of motor function, measures of sensory function, sympathetic nervous system function, and psychological questionnaires.
Sterling et al, 2006[87]	Neck Disability Index (Categorical: <8=Recovered, 10-28=Mild, >30=Mod/Severe)	Multiple logistic regression	Factors predicting NDI score: Initial NDI (p=0.001), Age (p=0.008), Cold pain thresholds (p=0.026), Impact of Events Scale scores (p=0.018). Factors predicting mod/sev disability (NDI>30): High Initial NDI (OR=1.05, CI 1.0-1.1) Older age (OR=1.1, CI 1.0-1.13) Reduced cold pain threshold (OR=1.1, CI 1.0-1.13) High Impact of Events Scale score (OR=1.03, CI 1.03-1.20).	Left cervical rotation, sympathetic nervous system function, compensation status, initial NDI, age, cold pain thresholds, Impact of Events Scale.

Sterner et al, 2003[88]	Disability questionnaire (None vs. Minor for analysis)		Logistic Regression and multiple regression.	Female Gender (OR 2.02 [1.13-3.63]). Patients educated below University level (OR 2.08 [1.09-3.98]). Patients with WAD grades 2-3 (OR 2.03 [1.08-3.88]). Patients with prior neck complaints (OR 3.17 [1.37-7.46]).	Age, sex, education, WAD Grade, accident type, previous neck complaint, prior headache, prior back complaint.
Warren and Warren, 2001[54]	Time off work (Self report). Measured in days.	Older age ($p<0.05$). Patients who experienced neck pain within 24 hours of injury ($p<0.05$). Patients whose symptoms were confined to the neck region ($p<0.01$). Those involved in driving occupations took the longest off work and this was significantly different compared to those who did heavy manual work, secretarial work and sedentary work but not light manual work (No p values given). Those who did secretarial work had the least time off work and this was significant compared to all the other occupations (No p values given).	Student t test		

Table 12 - Results based on symptomatic outcomes

Study	Outcome Measure	Univariate results	Test used	Multivariate results	Final model included
Atherton et al, 2006[58]	Report of persistent neck pain. Defined as neck pain at all follow up time points (1,3 and 12 months post injury).		Multiple logistic regression	Factors associated with persistent neck pain: Pre-collision widespread body pain (one month) (OR 1.9, CI 1.1-3.2), Vehicle other than a car (OR 1.8, CI 1.04-3.2), Number of other symptoms reported = 6-10 (OR 2.0, CI 1.2-3.3), Initial NDI >22 (OR 1.9, CI 1.2-2.9)	Gender, age, GHQ, presence of widespread body pain, vehicle type, initial NDI, number of symptoms.
Borchgrevink et al, 1995[59]	Report daily or constant symptoms (neck pain, stiffness and headache)	6/12 MRI Findings: Patients with preexisting spondylosis had more headaches than patients in the other groups ($p<0.01$). 6/12 X-Ray Findings: Patients with spondylosis had more headaches than patients with no findings or postural abnormalities only ($p<0.01$).12/12 X-Ray findings: Patients with postural abnormalities only had significantly less symptoms than the other groups. No p value given.	Kruskal-wallis followed by a Mann Whitney U Test.		

Borchgrevink et al, 1997a[60]	Report daily or constant symptoms (neck pain, stiffness and headache)	6/12 MRI Findings: patients with disc pathology or spondylosis reported more headaches than patients with no pathology or postural abnormalities only ($p<0.05$). 6/12 X-Ray findings: Patients with no pathology had less stiffness than patients with spondylosis or postural abnormalities ($p<0.01$). 12/12 X-Ray findings: Patients with no pathology had less stiffness ($p<0.01$) and neck pain ($p<0.05$) than patients with spondylosis or postural abnormalities	

Brison et al, 2000[61]	WAD present (as defined by the authors) - the subject experienced pain in the neck, upper back or shoulders. And they experienced moderate pain regularly or daily or severe pain occasionally, regularly or daily.	<u>Crash characteristics:</u> Reduced risk of ongoing symptoms in a stopped car compared to a moving car RR 0.7 (CI 0.5-1.0). Increased risk of ongoing symptoms if RTA occurred on a Highway compared to a Parking lot RR 2.8 (CI 1.0-7.9). Increased risk of ongoing symptoms if the posted speed limit is 60-80km/hr compared to ≤ 50 km/hr RR 1.4 (CI1.0-1.9) <u>Personal characteristics:</u> Increased risk of ongoing symptoms if BMI is high (25-26.9) compared to low (<20) RR1.8 (CI 1.03-3.3). Increased risk of ongoing symptoms if aged between 31 - 50 years compared to 18-30 years RR 1.5 (CI 1.0-2.1). Increased risk of ongoing symptoms if aged between 51-70 years compared to 18-30 years RR 2.1 (CI 1.4-3.0). Increased risk of ongoing symptoms in patients that had WAD at initial presentation compared to those that did not RR 3.3 (CI 2.2 - 4.7).	Univariate analysis of relative risks.	Unadjusted.

Hartling et al, 2001[62]	WAD present (as defined by the authors) - the subject experienced pain in the neck, upper back or shoulders. And they experienced moderate pain regularly or daily or severe pain occasionally, regularly or daily.		Logistic regression and X^2 test for trend.	Modified WAD grading: The risk of ongoing symptoms compared to WAD grade 0 are: Grade I RR 0.78 (CI 0.78 - 1.88) - not significant. Grade IIa (normal ROM) RR 1.17 (CI 0.49 - 2.77) - not significant. Grade II (undefined) RR 1.87 (CI 0.69-5.07) -not significant. Grade IIb (reduced ROM) RR 3.10 (1.18 - 8.19). X^2 for trend = 12.17 ($p<0.01$) indicating a trend for increasing risk of ongoing symptoms with increasing WAD grade.	Age, sex, presence or absence of prior neck pain, shoulder or upper back pain.
Gargan et al, 1997[56]	Recovered vs. non-recovered. Non-recovered= Symptoms intrusive or disabling	Reduced cervical ROM at 3/12 post injury is associated with non recovery (OR = 13.29 (CI 2.36-85.83). Abnormal GHQ score at 3/12 post injury is associated with non recovery (OR = 7.27 (CI 1.01-64.58).	Student t tests X^2 test with Yates correlation		

Gun et al, 2005[64]	Improvement in neck pain severity (VAS)			Patients with a higher SF-36 Bodily Pain score (lower degree of bodily pain) had a greater improvement after 1 year (ß=0.02, p<0.05). Patients with a higher SF-36 Role Emotion score had a greater improvement after 1 year (ß=0.01, p<0.05). Patients who had made a previous claim had less improvement after 1 year (ß=-1.13, p<0.05). Patients who had treatment by a physiotherapist or chiropractor had less improvement after 1 year (ß=-0.94, p<0.05)	Initial VAPS score.
Hohl 1974[55]	Recovered = patient's opinion there was no residual problems. Non-recovered = patient's opinion there were residual problems.	Age was significantly lower (p<0.05) in recovered individuals. There was significantly higher incidence of recovery in males than in females (p<0.01). Hospitalised patients recovery was significantly poorer (p<0.01). Radiating pain and/or numbness showed a positive correlation with symptomatic individuals (p>0.05).	Not stated.		
Karlsborg et al, 1997[10]	Number of symptoms		Logistic regression	The presence of stress unrelated to the whiplash injury predicted the number of symptoms (p=0.0078)	unadjusted

Kivioja et al, 2005[70]	Recovered = no neck pain now	Significantly more females (22/49) reported neck pain at follow up than males (9/42) (p,0.05). Lower initial pain intensity was associated with recovery (P<0.05)	Mann-Whitney U test. X^2 test, Logistic regression	Report of neck pain or shoulder pain in 4/52 prior to RTA (Exp (B) = 4.5 CI (1.1-8.76), p=0.035)	Sex, age, neck and shoulder pain before accident, initial pain intensity , catastrophising.
Kyhlback et al, 2002[71]	Pain Intensity (VAS, 0-100)		General Linear Model	Males gender (ß=0.43, p=<0.01). High initial self-efficacy scores (ß=0.32, p<0.01), Patients with higher WAD grades (ß=0.23, p=<0.05).	Unadjusted.
Mayou and Bryant, 1996[72]	Presence of Physical symptoms		Logistic regression	A report of neck pain at the time of the accident (p<0.01) Female passengers were at greater risk of non-recovery than a driver of either sex (p<0.01)	Age, gender, driver-passenger status, neuroticism, previous psychological problems, memories of the accident, mood score and neck symptoms immediately following the accident.

Miles et al, 1988[57]	Patients were questioned about nature, onset and durations of symptoms at a clinic review.	More patients with degenerative changes on x-ray at baseline had symptoms ($p<0.025$) and abnormal neurological signs ($p<0.01$) than those without changes. Fewer patients with an angular deformity at baseline on x-ray had symptoms ($p<0.05$) Unadjusted OR's calculated by reviewers: Initial degenerative changes on x-ray for symptoms at 2 years (OR 3.96 [1.04-15.33]). Initial degenerative changes on x-ray for abnormal neurology at 2 years (OR 6.75 [1.26-37.90]). Initial angular deformity for symptoms at 2 years (OR 0.30 [0.07-1.12])	Not stated.		
Olsson et al, 2002[77]	Residual pain at 1 year. Question: "Do you have residual pain which you relate to the accident?"		Cluster analysis followed by regression	Patients who perceived interference caused by pain and preventing or hindering the patient from pursuing a variety of activities were significantly linked with reporting pain b = -2.451 exp (b) = 0.086. Pain severity, life control, affective distress, support and general activity did not significantly predict poor outcome.	Age, Sex, WAD grade and MPI variables.

Pettersson et al, 1997[78]	Standardised Questionnaire and clinical examination (symptoms vs. no symptoms)	A descriptive study. No prognostic results presented. Unadjusted OR's calculated by reviewers: More likely to have symptoms if MRI shows bulging disc (OR= 15 [1.53-359.19]). More likely to have symptoms if moderate disc changes (OR=2.88 [0.16-19.48]). More likely to have symptoms if sever disc changes (OR=3.00 [0.18-93.04])			
Radanov et al, 1991[79]	Recovered vs. Non-recovered. Non-recovered = Presence of symptoms		Stepwise regression	Initial neck pain intensity (p=0.0019), older age (p=0.0036), injury related subjective cognitive impairment on Cognitive Function Questionnaire (p=0.0009).	Age, injury mechanism, lifetime history of psychological or behavioural problems, psychosocial stress (current and lifetime), personality dimensions, Well being, cognitive function, initial neck pain intensity, initial headache intensity, neurotic symptoms in childhood.

Radanov et al, 1993a[80]	Recovered vs. Non-recovered. Non-recovered = Presence of symptoms	Greater neck pain intensity at baseline ($p<0.01$, $U=937.5$), Greater headache intensity at baseline ($p<0.06$, $U=807.5$), Restricted neck ROM at baseline ($p=0.007$, $X^2=7.05$), Report of neck pain earlier post injury (Mean time = 7.0 hours +/-15.1(SD) vs. Mean time=11.0 hours +/-16.9 (SD)) ($p=0.038$).	Mann Whitney U test X^2 test		
Radanov et al, 1993b[81]	Presence of trauma related headache vs. no headache	A greater percentage of those with headaches had a history of pretraumatic headache ($p<0.0001$).	X^2 test. Multivariate analysis.	A history of pretraumatic headache was associated with headache at 6/12 post injury ($p<0.001$)	Presence of neck pain at 6/12 and neck pain intensity at 6/12. Age, gender, mechanism of injury, timing of initial symptoms, Personality traits.

Radanov et al, 1994b[82]: 6 month follow up	Recovered vs. Non-recovered. Non-recovered = Presence of symptoms		Stepwise regression	Greater initial neck pain intensity (t=4.595, p<0.001), sleep disturbances (t=4.381, p<0.0001), older age (t=4.222, p =0.0001), previous history of head trauma (t=3.287, p=0.0014), forgetfulness (t=3.129, p=0.00023), history of pretraumatic headache (t=3.037, p=0.0003), symptoms of radicular irritation (t=2.422, p=0.0172), score on neuroticism scale on Freiburg personality inventory (t=-2.334, p =0.0215), complained of poor concentration at baseline (t=-2.568, p=0.00117).	The following factors were entered into the initial model but the final model is not reported: Age, gender, injury mechanism, a history or head injury or whiplash, the type and frequency of pre-traumatic headaches and all findings from the baseline assessment (neck pain, headache, fatigue, shoulder pain, anxiety, sleep disturbances, back pain, sensitivity to noise, poor concentration, blurred vision, irritability, sensitivity to light, dizziness, forgetfulness, difficulty swallowing).

Radanov et al, 1994b[82]: 1 year follow up	Recovered vs. Non-recovered. Non-recovered = Presence of symptoms	Stepwise regression	Older age (t=3.824, p=0.0002, previous history of head trauma (t=3.333, p=0.0012), sleep disturbances (t=3.097, p=0.0025), intensity of initial neck pain (t=3.068, p=0.0028), pre-traumatic headache (t=2.951,p=0.0039), score on nervousness scale of Freiburg personality inventory (t=2.277, p=0.0249) and score on neuroticism scale on Freiburg personality inventory (t=-3.249, p=0.0016).	As above.

Radanov et al, 1995[83]	Presence of symptoms	Older age ($p<0.03$), Head rotated or inclined ($p<0.008$), History or pretraumatic headache(migraine) ($p<0.0001$).Initial neck pain intensity ($p<0.008$), Initial headache intensity ($p<0.004$), Anxiety $p<0.023$), Sleep disturbance ($p<0.0001$), Blurred vision ($p<0.008$), Forgetfulness ($p<0.006$), Symptoms of radicular deficit ($p<0.043$), Symptoms of cranial nerve or brainstem disturbance ($p=0.004$), Multiple symptom score ($p<0.026$), Radiological findings - sign of degeneration (osteoarthrosis) ($p<0.017$), Score on wellbeing scale ($p<0.033$), Cognitive variables: number connection test ($p<0.0001$), Trail making part A ($p<0.026$), Trail making part B ($p<0.012$), PASAT ($p<0.023$)	X^2 test. Mann-Whitney U test.

Sturzenegger et al, 1995[84]	Presence of symptoms	There were significant differences between symptomatic patients and asymptomatic patients in regard to head position at impact. Symptomatic patients were more likely to have had a rotated head position (x^2= 4.33, p=0.037) or inclined head position (x^2 = 4.48, p=0.034). Recovered patients were more likely to have a straight head position (x^2=7.87, p=0.005). Symptomatic subjects had significantly more neurological symptoms(p=0.008) at baseline than asymptomatic subjects. This included more cranial nerve and brain stem symptoms (p=0.009) and radicular irritation (p=0.015). Initial neck pain intensity was higher in the symptomatic group (p=0.0009). Symptomatic patients complained of more headaches initially (p=0.004) and higher initial headache intensity (p=0.0002). Symptomatic patients had higher multiple symptom scores (total number of symptoms) than asymptomatic patients (p=0.004).	Mann Whitney U test X^2 test

Di Stefano and Radanov, 1995[52]	Presence of symptoms	Greater neck pain intensity at baseline (U= 93.0 p=0.001) and headache intensity (U= 126.0, p=0.01). Worse scores on the Number Connection Test at baseline (p=0.003).	Mann Whitney U test Wilcoxon signed ranks test with Bonferroni correction
Voyvodic et al, 1997[53]	Recovered = no signs or symptoms were evident. Non-recovered = continuing to have signs and symptoms.	No significant findings. Unadjusted OR's calculated by reviewers: Injury rating at baseline (symptom free/mild vs. moderate severe) for recovered and non recovered at 6 months: OR undefined, RR 18.00 (2.68 - 120.92). Presence of spondylosis on MRI for recovered and non recovered at 6 months: OR 0.63, RR 0.75 (0.6-2.16).	X^2 test

3.4.3 DATA SYNTHESIS

Levels of evidence for each physical factor are presented in Table 13. For the majority of the 32 physical factors studied there is inconclusive evidence for an association with development of LWS.

No physical factor was identified as having strong evidence to support an association (negative or positive) with the development of LWS.

Three factors had moderate evidence for an association with LWS. These were high initial neck pain intensity; high initial neck pain related disability and cold hyperalgesia. High initial neck pain intensity was studied in six cohorts (cohorts 6, 11, 12, 14, 18, 21) and measured using a variety of methods. High initial neck pain related disability was measured at baseline in three cohorts (cohorts 15, 18, 23). Cold hyperalgesia was studied in two recent cohorts (cohorts 11 and 23).

Limited evidence was found to support an association of pre-injury chronic widespread pain with LWS (cohort 1). Finally, limited evidence from one cohort was found to support a lack of association between reduced pressure pain thresholds and LWS (cohort 23).

Table 13 - Levels of evidence for physical prognostic factors for the development of Late Whiplash Syndrome

Prognostic Factor	Adequate Quality Cohorts supporting an association with LWS	Low Quality Cohorts supporting an association with LWS	Adequate Quality Cohorts failing to show an association with LWS	Low Quality Cohorts failing to show an association with LWS	Overall level of evidence
Pre-injury Neck Pain	Kivioja et al (S)	Sterner et al (D)		Miettinen et al (D)	**Inconclusive**
Pre-injury Headache	Radanov et al (S)			Miettinen et al (D)	**Inconclusive**
Pre-injury Back pain				Miettinen et al (D)	**Inconclusive**
Pre-injury Widespread Chronic Pain	Atherton (S)				**Limited evidence for an association with development of LWS**
Pre-injury degeneration	Borchgrevink et al (S)			Voyvodic et al (S)	**Inconclusive**
Initial Neck Pain Intensity	Nederhand et al (D); Hendriks et al (D&S); Kasch et al (D)	Radanov et al (s), Mayou and Bryant (D);	Kivioja et al (S)		**Moderate evidence for an association with development of LWS**
Initial Shoulder Pain Intensity				Radanov et al (S)	**Inconclusive**
Initial Back Pain Intensity				Radanov et al (S)	**Inconclusive**

Initial Headache intensity		Radanov et al (S)			**Inconclusive**
High Initial Neck pain related disability	Nederhand et al (D); Sterling et al (D)	Miettinen et al (D)			**Moderate evidence for an association with development of LWS**
Baseline WAD grade (based on QTF definition of signs and symptoms)	Hartling et al (S)	Sterner et al (D)	Olsson et al (S)	Karlsborg et al (S)	**Inconclusive**
Early onset of Symptoms		Brison et al (S); Warren and Warren (D)		Radanov et al (S)	**Inconclusive**
Greater number of symptoms	Atherton (S)	Radanov et al (S)	Hendriks et al (D)	Karlsborg et al (S); Kasch et al (D)	**Inconclusive**
Restricted ROM	Sterling et al (D); Kasch et al (D)	Radanov et al (S); Gargan et al (S)	Hendriks et al (D); Atherton (S)	Hohl (S)	**Inconclusive**
Radicular symptoms		Hohl (S); Warren and Warren (D); Radanov et al (S)	Hendriks et al (D)		**Inconclusive**
Cranial nerve or brainstem disturbance		Radanov et al (S)			**Inconclusive**
Blurred vision		Radanov et al (S)		Radanov et al (S)	**Inconclusive**
Baseline sensitivity to noise or light				Radanov et al (S)	**Inconclusive**
Baseline dizziness				Radanov et al (S)	**Inconclusive**
Difficulty swallowing				Radanov et al (S)	**Inconclusive**

Fatigue at baseline				Radanov et al (S)	**Inconclusive**
Reduced superficial neck muscle EMG activity		Nederhand et al (D)	Sterling et al (D); Kasch et al (D)		**Inconclusive**
Resting neck muscle EMG activity level				Nederhand et al (D)	**Inconclusive**
Motor evoked potentials measurements				Karlsborg et al (S)	**Inconclusive**
Muscle Spasm				Hohl (S)	**Inconclusive**
Cold hyperalgesia	Sterling et al (D); Kasch et al (D)				**Moderate evidence for an association with development of LWS**
Reduced Pressure pain thresholds			Sterling et al (D)		**No association based on limited evidence**
MRI Imaging				Karlsborg et al (S)	**Inconclusive**
Abnormal X-ray findings		Radanov et al (S); Miles et al (S); Borchgrevink et al (S)	Hendriks et al (D)	Richter et al (S); Hohl (S); Voyvodic et al(S); Miles et al (S)	**Inconclusive**
Increased BMI score	Nederhand et al (D)	Brison et al (S);	Kasch et al (D)		**Inconclusive**
Increased Height				Brison et al (S); Minton et al (S)	**Inconclusive**
Increased Weight				Minton et al (S)	**Inconclusive**

3.5 DISCUSSION

This systematic review finds that increased initial neck pain intensity, pain related disability and cold hyperalgesia are linked to the development of LWS (poor outcome following a whiplash injury for at least six months). It also finds there is inconclusive evidence for the majority of physical factors being associated with the development of LWS. There are a number of limitations to consider when drawing conclusions.

3.5.1 LIMITATIONS

It was not possible for meta-analysis to be conducted due to the heterogeneity between studies. The main sources of heterogeneity were the different methods of measuring prognostic factors and the outcome measures used. Data necessary to carry out meta-analysis and thus provide objective estimation of effect sizes were often not present or poorly reported. Instead a "vote counting" method was used. This is not an ideal method as it may result in large and small studies being given equal weighting. Meta analysis also has the benefit in that it may reveal a significant association from the combination of a number of studies each showing a non-significant association.

Studies that show significant associations are more likely to be published and in duplicate [90] leading to possible exaggeration of strength of associations, a common criticism of systematic reviews. The inclusion of duplicated data may not only have the potential to lead to overestimation of effect sizes, but also increase the number of counts when a "vote-counting" methodology is utilised. Attempts have been made to negate this by summarising results by cohorts and not individual studies.

Due to resource limitations, non-English language articles and grey literature were excluded from this review. We were unable to include seven articles published in non-English languages that were potentially eligible. Three of these articles were from the same cohort

(cohort number 21) published in English which we have included in the review. Two articles had no abstract available in English [91, 92] and two articles [93, 94] had English abstracts, but it was unclear whether any physical factors were studied. Attempts were made to contact the authors of these last two articles but were unsuccessful.

3.5.2 METHODOLOGICAL CONSIDERATIONS

According to the quality assessment criteria there were no high quality studies assessing physical prognostic factors for the development of LWS. Reporting of methodology and results was often poor. There was a large variation in outcome measures used (23 different types) and only a small number (7) of studies reporting the use of validated measures of symptoms or disability. As well as outcome measurement heterogeneity, there was great diversity in the methods reported to measure physical prognostic factors. For example, of the seven studies that measured Range of Motion as a prognostic factor, two studies did not report the method used and only two studies used the same tool. Very few studies adequately described treatments that individuals were receiving during the study period. Treatments may influence the prognostic factors measured and so there is potential for bias in most of the previous studies.

In terms of methodological quality, statistical analysis was the most challenging area of the review. There was a very limited and varied use of multivariate analysis, which is necessary to control for bias in observational studies. In this review, merit was not only awarded for the use of multivariate analysis techniques, but adequate sample sizes for this analysis to be carried out on. Some studies received points for multivariate analyses but were not actually informative and papers may not have based their conclusions on these analyses.

There had been three previous systematic reviews on the prognosis of WAD [6, 18, 38]. This review has included 16 additional articles to the most recent of these; 13 of these have been published subsequently, and the remaining three were included due to differences in

selection criteria. Different methodological quality criteria were used, meaning it was more difficult for studies to achieve a high quality rating. It was felt that this more accurately reflected the quality of research carried out. Incorporating new publications and refining methodologies has lead to firmer conclusions regarding the value of physical prognostic factors.

Both of the two latest reviews concur with the findings that initial pain intensity appears to be the most important physical factor for the development of LWS. Evidence as a result of studies published since Scholten-Peeters et al [38] have lead to the conclusions that range of motion has inconclusive evidence and cold hyperalgesia moderate evidence for association with a poor outcome.

3.5.3 CLINICAL IMPLICATIONS

From this review increased initial pain intensity, pain-related disability and cold hyperalgesia may play a role in the development of LWS. These are measures that have clinical implications for identifying individuals who may be at risk and may require further intervention. It is difficult to assess the strength of these roles due to the paucity of effect sizes presented, and therefore their precise clinical impact.

Other physical factors commonly used in the clinical setting to make management decisions, e.g. cervical Range of Motion or radicular signs, showed inconclusive evidence for an association with outcome of LWS. Previous reviews [18, 38] have concluded that they may be of prognostic value therefore it appears these factors warrant further evaluation to affirm or refute conclusions here.

3.5.4 RESEARCH IMPLICATIONS

Scholten-Peeters et al [38] called for consensus regarding methodological criteria for prognostic studies. It is clear from this review that this has yet to be achieved. It is important

to continue to build on previous research using high quality studies to evaluate physical factors alongside psychological and social prognostic factors. This should allow for meta-analysis which will aid clinicians and policy makers alike in the search for more efficacious management of WAD [6] and prevention of LWS.

Cold hyperalgesia - an increased sensitivity to cold temperature suggestive of changes in central nocioceptive pathways [95] – shows promise for prognostic value. However, reduced pressure pain thresholds showed no association to LWS development, even though both measures are thought to represent a measure of central nervous system sensitisation. It has been hypothesised that their sensitivity to indicate central sensitisation may be different [86]. This warrants further investigation.

3.6 CONCLUSIONS

From the results of this review, it is concluded that pain has a central role to play in the development of LWS. Evidence suggests that increased initial pain intensity, pain related disability and cold hyperalgesia are associated with a poorer outcome following a whiplash injury. Pre-injury widespread chronic pain also shows limited evidence that it may affect outcome.

It is clear that there is a need for the consistent use of validated measures of both prognostic factors and outcome in order to provide a clearer picture of the prognosis of WAD. This review found an absence of high quality prognostic studies. Rectifying this for future studies will require strict adherence to appropriate sampling, statistical analysis and reporting methods.

3.7 SUMMARY

This chapter described a systematic literature review of physical prognostic factors for the development of LWS. The limitations in previous prognostic studies highlighted by this systematic review should be noted for future prognostic studies in this area, such as the one described in Chapter Seven.

This review has highlighted that the prognostic value of cervical ROM, a commonly used clinical tool, is uncertain at present and further studies are warranted. It also raises awareness of the necessity of using valid and reliable measures of prognostic factors. These findings resulted in the systematic review of reliability and validity studies of measurement methods for cervical ROM presented in Chapter Five. The following chapter provides an in-depth description of the cervical spine, its assessment and management relative to WAD.

4 CHAPTER FOUR – ASSESSMENT OF THE CERVICAL SPINE IN WAD

4.1 INTRODUCTION

The previous chapters have provided evidence that the cervical spine is affected by a whiplash mechanism of injury, which can lead to long term symptoms and disability (LWS) for a substantial proportion of individuals. The preceding chapter described findings from a systematic review which concluded that the measurement of cervical spine range of motion is uncertain as a prognostic factor for LWS. This chapter describes clinical assessment of the cervical spine in relation to WAD, with a particular focus on measurement of Range of Motion (ROM).

The objectives of this chapter:

-
- To explain which body structures and functions are affected by a whiplash injury that will impact on the cervical ROM and function of the cervical spine
- To describe how cervical ROM is measured in clinical practice and how it is used for diagnosis and monitoring of treatment response
- To describe the rationale behind why clinicians measure ROM
- To describe how cervical ROM has previously been measured in WAD populations and justify why further investigation is necessary

4.2 FUNCTIONAL ANATOMY AND PHYSIOLOGY OF THE CERVICAL SPINE

The focus of this thesis is cervical ROM assessment and therefore the primary concern is with kinematics, the branch of mechanics that describes motion of a body without regard of what produces this motion. It is important to have an understanding of the structures that are potentially affected by a whiplash injury and the subsequent impact on motion and ultimately function. Readers are referred to Bogduk and Mercer [96] for an excellent review article of the normal kinematics of the cervical spine, including a summary of the key anatomy. The cervical spine consists of bones and joints, soft tissues such as ligaments and muscles, neurological tissues (the spinal cord and nerves) and vascular tissues (blood vessels and in particular the vertebral artery).

The cervical spine has three functions. Firstly it forms a stable osteo-ligamentous axis for support of the head. Secondly it provides a variety and range of movements that are essential for human tasks – varying the direction of the senses. The cervical spine allows the sensory organs contained within the head to move and orientate in a three-dimensional space. Lastly it forms a protective conduit for the spinal cord and its nerves and vertebral arteries that supply the brain.

Considering all of these functions it is apparent that there is a compromise between mobility (for the senses- indeed it is the most mobile section of the spine) and stability (for protection of literally vital structures).

4.2.1 SEGMENTAL AND GLOBAL KINEMATICS

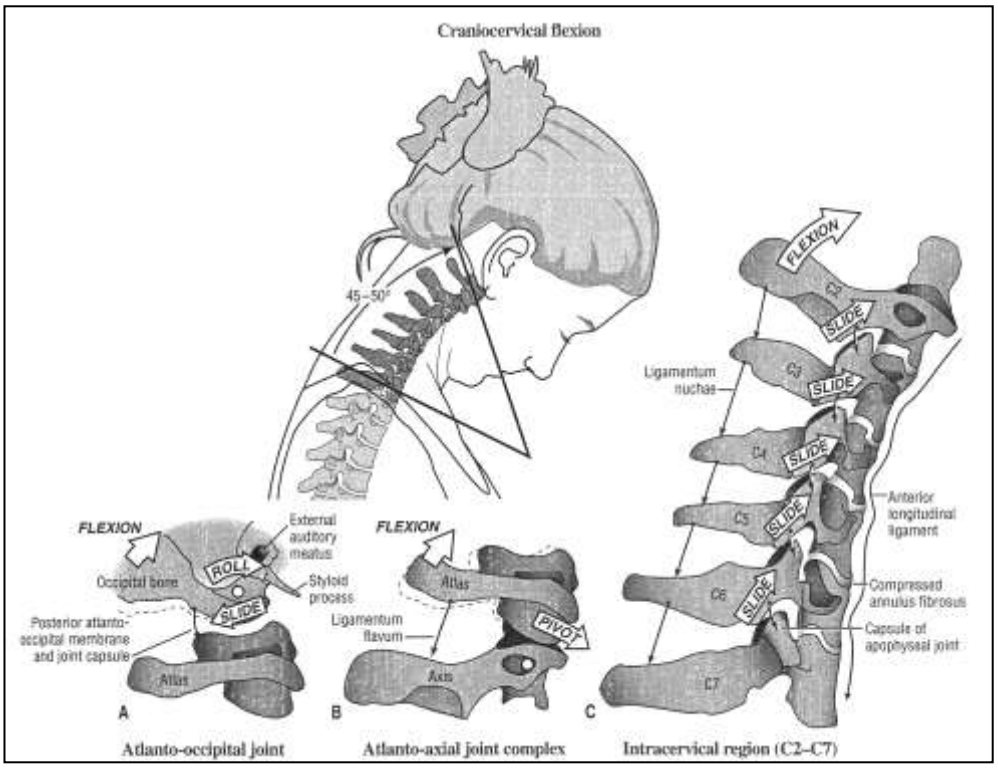
The joints between the bones of the cervical spine all vary in their contribution to the multitude of possible movements of the head. The following sections describe the

movements occurring at the individual joints grouped by movement direction. Kinematics are dependent on the muscles and bony anatomy of the neck.

4.2.1.1 Flexion kinematics

Flexion of the cervical spine involves the head being brought forwards and downwards so that the chin is brought to the chest. Figure 8 illustrates the movements of cranio-cervical flexion demonstrating the movements at the atlanto-occipital, atlanto axial and intra-cervical joints. At the atlanto-occipital joint the occiput slides backwards and rolls forwards simultaneously creating a nodding of the skull. At the atlanto-axial joint the atlas pivots on the axis and in the intra-cervical region the facet joints slide. Flexion is resisted by a number of structures. Anteriorly, the chin obstructed by the chest. Posteriorly, the posterior longitudinal ligament, the ligamentum flavum, the capsules of the zygapophyseal joints, and the inter-spinous ligaments will prevent further movement.

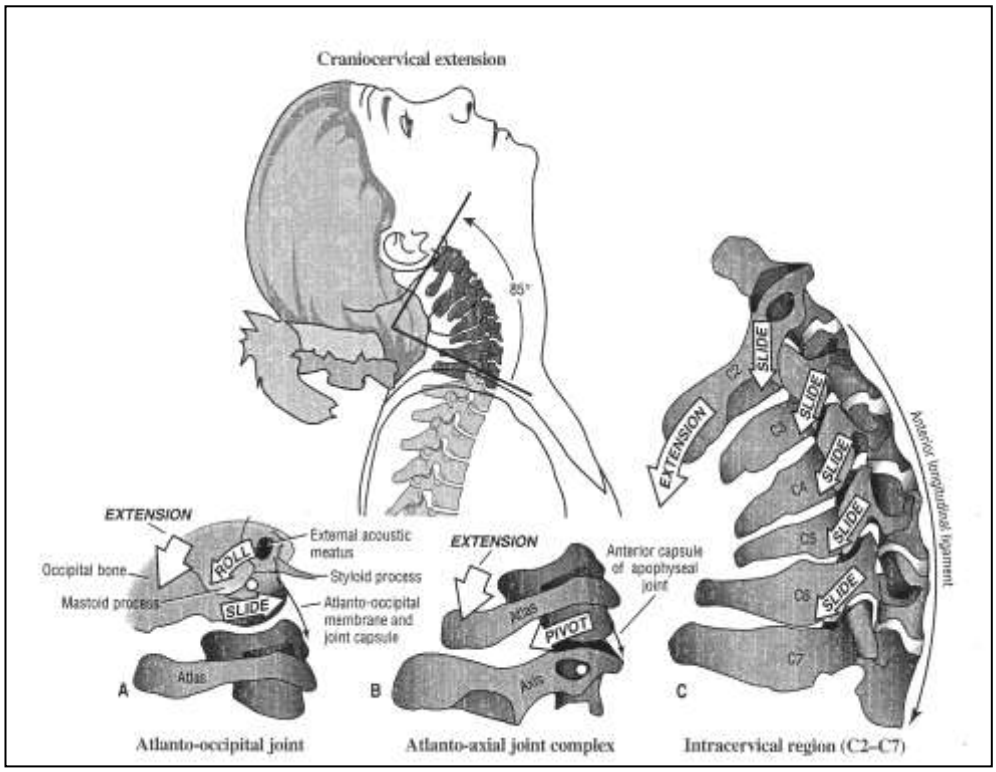
Figure 8 - Illustration of cranio-cervical flexion (From Neumann [97] with permission)



4.2.1.2 Extension kinematics

Extension of the cranio-cervical spine involves the head being taken backwards, resulting in the individual being able to look upwards. Figure 9 illustrates cranio-cervical extension and the movements involved at the different types of joints. The reverse of the rolling, sliding and gliding movements described for cranio-cervical flexion occur. Extension is primarily restricted by the anterior longitudinal ligament (anteriorly), the annulus fibrosis of the intervertebral discs, and ultimately by the spinous processes impacting on one another posteriorly.

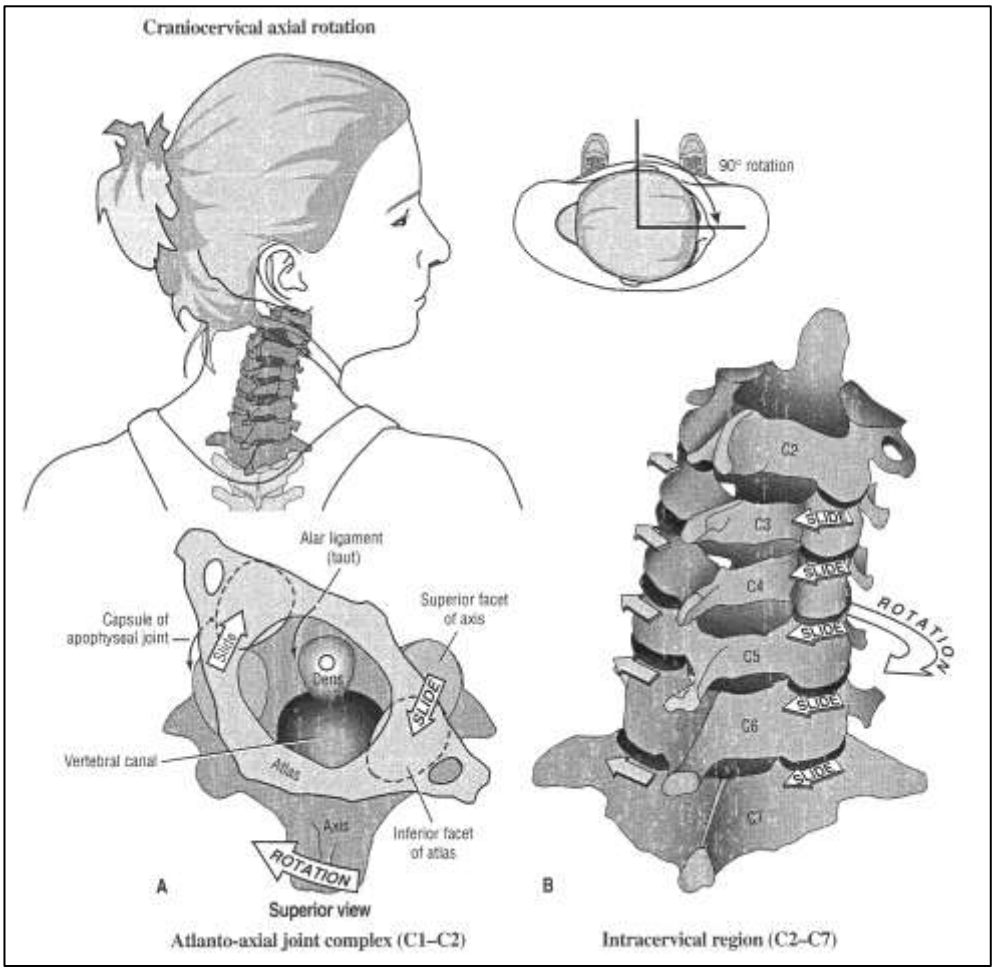
Figure 9 – Illustration of cranio-cervical extension (From Neumann [97] with permission)



4.2.1.3 Rotation kinematics

Rotation of the cranio-cervical spine is the turning motion to either side which results in being able to look over ones shoulder. Figure 10 illustrates the movements occurring at the atlanto-axial and intervertebral joints that produce this. A significant amount of rotation occurs at the one joint between the atlas and the axis, with the dens as a pivot. At the same time, the coupling motion of sliding and tilting of facet joints occurs. Rotation is limited by the alar ligament, capsules of the zygapophyseal joints and the anterior annulus fibrosis of the intervertebral disc.

Figure 10 - Illustration of cranio-cervical rotation (From Neumann [97] with permission)

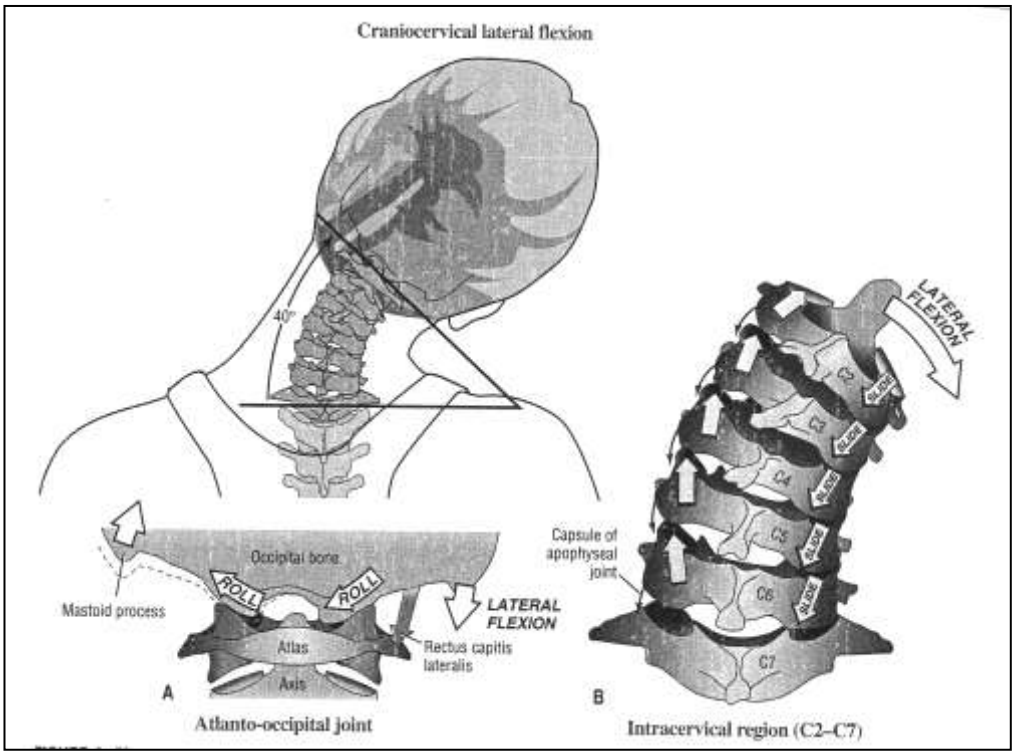


4.2.1.4 Lateral flexion kinematics

Lateral flexion – also known as side flexion – is the movement which results in the ear tilting to the shoulder.

Figure 11 illustrates the movements occurring within the cranio-cervical spine. At the atlanto-occipital joint the occiput rolls to the side at the same time as the ‘coupled’ sliding and tilting motion of zygapophyseal joints of the column. Tissues that limit the range of lateral flexion are the capsule of the apophyseal joints and other more superficial soft tissues including muscles on the contralateral side.

Figure 11 – Illustration of Craniocervical lateral flexion (From Neumann [97] with permission)



4.2.2 RANGES OF CERVICAL SPINE MOTION

Table 14 summarises "textbook" values for approximate ranges of motion for each of the joints in the cervical spine. Normatively, the greatest ranges of motion in the cervical spine are afforded to rotation and extension. This table demonstrates that the majority of movement occurs at the atlanto-axial and intervertebral joints.

Table 14 – Approximate active Range of Motion values for the joints of the cranio-cervical region [100]

Joint or region	Flexion and Extension (Saggital Plane, Degrees)	Axial Rotation (Horizontal Plane, Degrees)	Lateral Flexion (Frontal Plane, Degrees)
Atlanto-occipital joint	Flexion:5 Extension:10 Total:15	Negligible	Approximately 5
Atlanto-axial joint complex	Flexion:5 Extension:10 Total:15	40-45	Negligible
Intra-cervical region (C2-7)	Flexion:35 Extension:70 Total:105	45	35
Total across Cranio-cervical region	Flexion:45-50 Extension:85 Total:130-135	90	Approximately 40

As far as the author is aware, there is only one published meta-analysis of studies reporting cervical spine range of motion values that attempts to provide estimates for normative values for ROM [101]. Table 15 summarises this work, providing overall active and passive cervical ROM values for each half and full-cycle direction of cervical spine movement. The authors of this meta-analysis did warn that the estimates were potentially confounded by the diversity of different methods used to obtain measurements, sometimes with a dramatic variation between them. On average passive motion is greater than active motion by approximately 10%.

Table 15 – Mean values for active and passive cervical Range of Motion summarised from Chen et al [101]

	Mean value for ROM (SD) / º		**Mean Value for ROM (SD) / º**
Active Flexion	52 (7)		
Active Extension	71 (5)		
Total Flexion-Extension	126 (12)	Total Passive Flexion-Extension	**140 (4)**
Active Right Rotation	73 (11)		
Active Left Rotation	71 (11)		
Total Rotation	151 (23)	Total Passive Rotation	**174(18)**
Active Right Lateral Flexion	44 (0)		
Active Left Lateral Flexion	42 (2)		
Total Lateral Flexion	**86 (5)**	**Total Passive Lateral Flexion**	**109 (13)**

4.3 THE EFFECT OF A WHIPLASH INJURY ON THE CERVICAL SPINE

Having described the normal kinematics of the cervical spine, this section discusses which cervical spine structures and functions that could potentially be altered as a result of a whiplash mechanism of injury.

Chapter One described the whiplash mechanism of injury, but to briefly summarise, the result of an acceleration-deceleration force to the head causes the cervical spine to extend at a rapid rate causing an abnormal S-shaped curve, after which, depending on the forces involved, the head may then be moved forwards bringing the chin to the chest. Accurately ascertaining the precise structures that have been damaged as a result of a whiplash injury is very difficult. Experimental and cadaveric studies have failed to show a consistent 'lesion' related to a whiplash injury. Autopsy studies that do show tears to muscles, rim lesions of the cervical disc and injuries to the facet joints are usually of deceased individuals involved in severe accidents [102, 103] and therefore have limited generalisability to the average whiplash-injured patient seeking attention for further management. With an awareness of anatomy and the mechanism of injury described above, there are numerous tissues that could potentially be affected.

Radiographs are primarily used to rule out serious bony and joint injury i.e. fractures and dislocations. The absence of any findings leads to the diagnosis of a soft-tissue injury, which is the case for the vast majority of patients following a whiplash injury. Soft-tissue injuries may include damage to ligaments, muscles, blood vessels, nerves and also articular cartilage within joints which may not be picked up by the relatively insensitive radiographic image and interpretation. With the knowledge that the formation of an abnormal S-shape curve can occur, it is extrapolated that the posterior joints of the cervical spine and their contents could easily be affected. In particular the facet joints of the lower cervical spine (and upper spine if the amount of force is sufficient) could be affected, which are known to have a nerve supply

which can produce symptoms of pain in both the neck and the head. Using local anaesthetic blocks, Barnsley et al [104] found evidence that approximately half of whiplash-injured patients' long standing neck pain originated from these facet joints.

Magnetic Resonance Imaging (MRI) is not routinely used in the clinical setting; however researchers have used its sensitivity to investigate possible soft tissue injuries resulting from whiplash injuries. Most commonly, MRI has been used to research abnormalities of ligaments. Krakenes and Kaale [105] reviewed a selection of the published literature, and along with a study of their own, concluded that whiplash injury is associated with changes to craniovertebral ligaments and these are associated with impairment. Similarly, Johansson and Bengt [106] found that three severely injured patients had visible ligament and joint capsule abnormalities to functional MRI following whiplash injury. Kongsted et al [107] performed MRI scans on 178 WAD patients and found a range of abnormalities, ranging from separation of discs from vertebral end-plates and inter-spinal bleeding (n=1 and 3 respectively) to bulging disc contours (n=36) and pre-existing degeneration (n=56), but despite all these abnormal findings, only seven patients (4%) were deemed to have abnormal findings related to trauma and so they concluded that trauma-related MRI findings are rare. Matsumoto et al [108] conducted a 10 year MRI follow-up on WAD patients (n=133) and controls and found there was no statistically significant correlation between symptoms (neck pain) and MRI findings in either group. Pettersson et al [78] used MRI imaging in a prospective cohort to detect injuries to intervertebral discs. They found that only minor changes had occurred in a minority of patient's discs concluding that MRI may not be useful in diagnosing patient's in the acute phase. MRI can also be used to evaluate muscle tissue; however it is apparent that no studies have directly evaluated muscle damage using this imaging technique. It has been used to evaluate muscle recruitment, but failed to show a difference between WAD patients and controls although this technology is in its infancy[109]. So it appears that despite MRI's ability to detect disc, ligament, muscle and

joint abnormalities following whiplash injury, these do not necessarily explain symptoms that patients describe, suggesting possible involvement of additional non-tissue (psychosocial) factors.

Researchers have explored dysfunction of muscles following a whiplash injury and effects on function. There is an emerging body of evidence linking neck disorders with changes in motor control patterns of cervical muscles [111], and that these changes occur soon after the onset of neck pain . Kasch et al [112] investigated muscle tenderness as an indicator of tissue pathology and concluded that whiplash-injured individuals were sensitive initially and increasingly so for those that failed to recover. Biochemical changes in muscles have been found, although the findings are far from definitive [110].

Measures of nervous system sensitivity have demonstrated relevance for diagnosis and prognosis for WAD disorders. Sterling et al [87, 113] performed a range of tests on a prospective cohort of acute WAD patients (n=80) . Brachial plexus provocation tests [114] demonstrated that those with greater disability had reduced neural mobility and greater responsive pain levels when compared to control subjects. The group of moderately/severely disabled patients also had reduced pressure pain and thermal thresholds indicating sensory hypersensitivity. Proprioception – the awareness of the body in space – of the cervical spine may be assessed by evaluating the ability to reposition the head. This has been found to be less precise in whiplash injured subjects compared to controls[115]. Also, correlations have been observed between both cervical ROM and oculomotor function and head repositioning and oculomotor function. However, contrary findings have more recently been presented by Armstrong et al [116], where no difference in proprioception was found for WAD patients compared to healthy controls. Balance is another function which may be affected by dysfunction of the cervical spine and dizziness is a symptom that has commonly been reported following a whiplash injury [117].

Disturbance to the structures and functions of the cervical spine described above could potentially lead to a complex pattern of signs and symptoms. The following section describes how musculoskeletal clinicians assess the cervical spine to provide a diagnosis and plan for management of WAD.

4.4 CLINICAL ASSESSMENT OF THE CERVICAL SPINE

Clinical assessment is the process by which the health care professional (HCP) seeks to find out and understand the disease or problem that the patient has consulted them for. This is an immediate and continuous process; the assessment begins from the very first moment the patient enters the clinical setting and is continued with each visit.

Clinical assessment is complex and consists of a combination of listening to the patient and observing in structured and unstructured tests. The HCP has to assimilate information from a variety of sources including the subjective (what the patient 'reports') and the objective (what the examiner 'finds'). Both the 'reporting' and 'finding' are mediated through the complexities of communication and may be affected by factors such as age, gender and ethnicity.

The Maitland concept for assessment and management of musculoskeletal problems underpins modern formal teaching for physiotherapists and other HCPs in the UK and beyond. The concept "emphasises careful and comprehensive examination leading to the precise application of treatment by movement and followed in turn by the assessment of the effects of that movement on the patient." [118]. The following section aims to highlight the mechanisms by which HCPs apply this concept.

4.4.1 PATIENT HISTORY

The patient history has numerous functions, including ascertaining the problems faced by the patient and their concerns or fears and expectations of assessment and treatment. By the end

of the discussion the HCP should have an awareness of the severity, irritability and nature of the problem and possible structures involved which may guide the composition and order of the physical examination (e.g. which parts of the examination should be omitted or delayed to prevent aggravation until final stages). Patients should be questioned carefully to identify "red flags" which may indicate serious pathology and the need for urgent medical consultation (e.g. severe unremitting pain, severe night pain, unexplained weight loss [119]).

With particular regard to ROM, the HCP should question whether particular movements cause problems, frequently in the form of pain, and whether this pain is at the end, in a particular section or throughout the whole movement. The HCP may also ask what the patient's opinion is of what is the cause of these movement problems.

Following the patient history, the physical examination will take place.

4.4.2 PHYSICAL EXAMINATION

4.4.2.1 Observation

Although highlighted within the physical examination section, observation should already have commenced when the patient enters the clinical setting and throughout the patient history taking. The HCP will be looking at general posture, ease of movement and any signs of pain behaviour such as guarding, rubbing and grimacing. The HCP will also be looking for body alignment, any deformities, any deviations from normal skin colour and texture, any swelling or redness that may indicate inflammation, and finally the patient's reaction – any apprehension, restlessness etc. Following observation, the HCP may alter the 'working diagnosis' constructed from the patient history.

4.4.2.2 Examination

The examination is performed in a logical order in order to elicit any need for changing the working diagnosis. The HCP is searching for the source of the patient's problems, informed

by the patient history and observation. James Cyriax was the first author to document a truly systematic examination process:

‘Only by sticking to a standard sequence will the physician be sure to leave nothing out and only by leaving nothing out are true findings feasible” [120]

Principles of the examination include comparison of one side of the body to the other and using previous knowledge to understand where the findings of this patient fit into the wide variation of what is considered normal. The unaffected side is tested first as applicable. For tests where the patient is passive, movements or stresses should be applied in a gradual manner; often ‘the less you press, the more you feel’. At the end of the examination patients should be warned of the potential for symptom exacerbation. “Every effort should be made to objectify the patient’s report of pain and discomfort” [121].

Cyriax advocated a scanning examination to direct a further in-depth examination, which should then concentrate on the spinal or peripheral system [120].

The examination routinely begins with palpation, checking of vital signs (pulses) if appropriate and then progresses with the HCP observing movements, conducting a neurological examination and then application of more in-depth, passive tests of movements and joints. As stated in section 4.3, neurodynamic tests have also received recent attention as useful assessment procedures for whiplash-injured patients.

4.4.3 Examination of Range of Motion

When examining movements of the body the HCP will seek to evaluate different types of movements in order to include or, more likely, exclude structures that are contributing to the patient’s symptoms. By performing these movements in different ways the HCP can differentiate further between structures.

Active movements are defined as movements of the body that the patient performs by themselves unassisted and provide information on the patient's willingness to move, coordination, muscle strength, range, state of contractile and non-contractile structures and the functional limits of the patient's condition.

Passive movements are defined as movements of body parts that are performed by the examiner without assistance from the patient [31]. Passive movements are sub-divided into physiological and accessory movements. Physiological movements are defined as those which are able to be performed actively[122]. An accessory movement is the opposite, a movement which cannot be performed actively by the muscles surrounding the joint(s) and therefore must be performed by an external force [31], for example when an HCP applies pressure to a particular spinal segment.

Both passive physiological and accessory movements are believed to provide information about the integrity of the articular surfaces and the extensibility of the surrounding soft tissues. Because these passive movements are not performed by the patient, contractile structures are not being tested, allowing the HCP to generate new information than that provided by active movements. For the cervical spine, physiological movements are an accumulation of movements from numerous cervical spine segments and are thought of as rather a crude measure [123]. In order to obtain more specific information the HCP may use palpation to feel movements between the individual vertebral segments. The limitation of assessment of individual segments is that their reliability and validity is more questionable than that of gross physiological measurement of movement, especially between different examiners [124].

It is recommended that active movements should be tested first because the patient will perform these within their pain limits and are therefore safer. Active movements will indicate the severity of the condition and therefore how forcefully passive movements should

be performed. When examining a movement the clinician is collecting information about the quality, quantity, if and when pain is experienced, the behaviour of that pain, and the effect of adding combinations of movement together, changing speed or adding compression.

The quantity of movement is referred to as Range of Motion (ROM). Technically ROM is defined as "the arc of motion that occurs at a joint or series of joints" [125]. Active ROM is therefore the arc or displacement of a joint or series of joints attained during unassisted voluntary joint motion and passive ROM is the arc or displacement of a joint or series of joints attained by the HCP without assistance from the subject. ROM may be classified as normal, reduced or increased and is often quantified using instruments such as a tape measure, universal goniometer or visual estimation by the HCP [122] . Normal movement is defined as pain-free and full range.

It is commonly acknowledged that ROM is influenced by a number of factors such as age, time of day, temperature, emotional status, effort, medication, injury and disease [123], and yet the HCP has expectations about judging whether the ROM is normal for that patient within a small assessment period. These factors are discussed in more detail in a subsequent section.

The physical examination of ROM is part of trying to 'make features fit' [31] – comparing with the information obtained about movement in the patient history. The patient is continuously questioned as to their experience of movement within the assessments, particularly whether they are experiencing pain or stiffness. This interplay between physical 'objective' findings including ROM and the patient's 'subjective' reporting will lead the HCP to a working diagnosis from which a treatment plan is constructed. Treatment techniques may be movements that relieve or provoke symptoms depending on the nature, severity and irritability of the disorder [126]. Assessment is continuous throughout the

treatment, with ROM contributing information about whether treatment techniques are the correct ones or not.

4.4.3.1 Movements measured in the cervical spine

Typically, movements of the cervical spine are measured in three planes; Saggital, Frontal and Transverse. Clinically, these measurements are usually carried out in the form of 'half-cycle' movements i.e. from a 'neutral' or middle position to the end of range of one of the planes. This results in six movements to measure in the cervical spine, flexion (F; moving chin down to chest), extension (E; looking up to the sky), right lateral flexion (RLF; moving right ear down to right shoulder), left lateral flexion (LLF; moving left ear down to left shoulder), right rotation (RR; turning to look over right shoulder) and left rotation (LR; turning to look over left shoulder) [127]. The alternative to measuring half-cycle movements is to use 'full cycle' proportions i.e. the range of movement for the whole plane e.g. maximal flexion to maximal extension (F-E). There is evidence that this is more reliable to measure [101] but one disadvantage is that the method may be less clinically useful when attempting to ascertain particular impairments and the structural dysfunctions that cause them. A movement in one half-cycle direction may test a completely different structure to movement in another direction.

4.4.3.2 Patient's reporting of ROM

As described above, HCPs will often ask *what* is limiting patient's ROM in order to help identify what is the source of the problem. However, it is not routine for patients to be asked to quantify their ROM themselves. This is usually performed by the HCP and described in degrees or fraction of whole normal movement. Patient-rated quantification of ROM may provide us with a truer reflection of a patient's function and therefore may have not only diagnostic but prognostic value, a key concept in this thesis.

Toomingas et al [128] proposed that if aspects of the clinical examination were validly and reliably self-reported by the patient, then a considerable reduction in required clinician time and therefore healthcare costs could be made. They evaluated cervical and shoulder ROM in 350 Swedish individuals using diagrams which asked participants to mark at which point their ROM would end. These marks were spaced at 15 degrees increments for neck rotation and flexion-extension and 30 degrees for shoulder abduction and external rotation. They found that when compared to medical examination findings carried out by a specially trained examiner, there was very low agreement between patient reported and clinician measured ROM. They recommended that the self-administered examination was not suitable for identifying positive signs of musculoskeletal disorder of the neck or shoulder. The validity of patient-reported ROM is therefore unclear.

Researchers investigating WAD populations have more commonly evaluated patients' perceptions of movement in the form of "stiffness", probably due to the concept's widespread clinical use and also the potential ambiguity between stiffness and motion [129]. A few examples; Hildingsson and Toolanen [67] asked a cohort of whiplash-injured patients whether they had experienced symptoms of stiffness since their injury, 69% of whom had. Hohl [55] reported 95% of a cohort of orthopaedic WAD patients complained of neck stiffness. Drottning et al [130] reported that 55% WAD patients examined six weeks post injury reported neck stiffness.

Richter et al [85] not only asked whether stiffness was present or not but also got patients to rate neck stiffness on 0-10 VAS. 46% had stiffness and the mean severity was 4.8. Borchgrevink et al [131] also asked patients to quantify stiffness, this time on a 0-5 scale, where 0=none and 5=maximum.

Nevertheless, stiffness and its severity is only one symptom that may limit ROM and is not a measure of ROM itself. The author is unaware of any studies of a WAD population where participants have been asked to provide their perception of the quantity of their ROM.

4.4.3.3 Quantity of ROM needed for everyday function

There are two studies that have published measurements of cervical spine range of motion required for everyday activities of daily living. Bennett et al [132] studied 28 college students and measured end of range motion measurements using a CROM device for 13 different Activities of Daily Living (ADL's). Four of the 13 ADL's required between 30-50% of maximal active ROM. Bible et al [133] studied 60 asymptomatic individuals across a range of ages and continuously measured ROM using an electrogoniometer during 15 ADL's. They concluded that most everyday ADL's require about 20-30% of maximal ROM, however exceptional circumstances such as reversing a car can require up to 90% of maximal rotation. It would appear from these two studies that should ROM be reduced by at least 50% this could have a significant impacts on important activities of daily living.

4.4.3.4 Measuring ROM

HCP's and researchers have devised numerous methods to measure joint range of motion, What is consistent throughout these methods is that the observer must have knowledge and skills to position and stabilize the body part correctly, move a body part through the appropriate range, palpate the appropriate bony landmarks, if using an instrument, aligning it with landmarks, determine the end feel (if evaluating passive movements), and reading the measuring instrument and recording measurements correctly.

When measuring cervical spine ROM particularly, the observer should be aware that the thoracic spine and shoulders have an influence on cervical movements, especially in a seated position. This position is recommended for consistency, stability and patient comfort [31]. Measurement methods range from visual estimation through to complex three dimensional

motion analysis. A detailed description of available methods is provided in the next chapter in section 5.3.3.

The next section will discuss factors which affect how much range the cervical spine has.

4.4.4 FACTORS AFFECTING CERVICAL RANGE OF MOTION

Although ROM is referred to as an objective measure, it should be made clear that the range recorded will depend on many variables including subjective or psychosocial factors such as discomfort or pain, and motivation [134].

4.4.4.1 Physical factors

From a general neuromusculoskeletal perspective possible cause of limitation to ROM in normal joints may include articular surface contact, limit of soft tissue extensibility (joint capsules, ligaments, muscles) and opposition of soft tissues. In damaged joints, reason for loss of ROM may include destruction of bone/cartilage/fracture, foreign body, tearing/displacement of intra-capsular structures, adhesions/ scar tissue, muscle atrophy/ hypertrophy, muscle tear/rupture/denervation, pain, oedema and neurological impairment.

4.4.4.2 Clinical factors

More specifically for the cervical spine, a number of clinical factors may affect ROM. It has been repeatedly demonstrated that cervical ROM is reduced in various symptomatic populations when compared to matched asymptomatic counterparts (e.g.[135]), and in particular whiplash-injured populations [136, 137]. Bergman et al [138] demonstrated that there is greater variation in symptomatic populations compared to asymptomatic individuals, also finding that passive ROM had a greater variation than active ROM.

Pain is often cited as the most common direct cause of cervical ROM limitation [139], although there is a limited amount of published evidence to support this. One study

concluded that neck pain is inversely related to cervical range of motion during the first 6 months after acute whiplash injury [140].

Psychological factors may affect an individual's ability to actively move their neck through anxiety, fear or lack of motivation. Turk et al [141] found weak to moderate correlations between fear of activity and cervical ROM.

4.4.4.3 Demographic factors

In some studies, a trend for women to have greater cervical ROM is apparent, although differences are not consistent [101]. Generally, research evidence indicates that cervical ROM decreases with age [101]. The only exception is axial rotation (primarily at the atlanto-occipital joint), which has been shown to stay the same or to increase to compensate for increasing hypomobility at the lower cervical spine. Cyriax's theory of a capsular pattern means that ROM is believed to be restricted in certain directions when articular degeneration has occurred [120].

One study [142] has evaluated the effect of neck dimension on cervical ROM and concluded that ROM was influenced by neck circumference. This should be interpreted with caution as this is only one study with a narrow age range (20-40) of asymptomatic individuals. Regarding Body Mass Index, only two studies have investigated this anthropometric factor and found conflicting evidence that it has an effect on cervical ROM [143, 144].

The slumped, "forward head posture" that is commonly adopted due to the pull of gravity affects the kinematics of the cervical spine and it is argued that lack of control of starting posture may be responsible for the variation in normal neck ROM values [145]. Penas et al [146] investigated the effect of forward head posture on cervical ROM in chronic headache patients and found that all cervical ROM was reduced apart from right lateral flexion.

4.4.4.4 Measurement protocol

The following factors that may affect cervical ROM are associated with the application of a measurement protocol. Particularly pertinent to this thesis is the effect of whether movements are performed in an active or passive form. As already highlighted, in Chen et al's meta-analysis [101] previous studies have on average found that passive movements have greater ranges than active, but only by a small amount.

ROM varies with time [138, 147], however there is a lack of studies that directly evaluate the effect of diurnal variation on cervical ROM. Evidence is often translated from the lumbar spine [145], where there is an indication that ROM increases during the day. Reliability and validity studies often measure at the same time of day in order to prevent bias that may arise of this variation [148, 149].

Theoretically, performing warm-up movements should reduce soft tissue stiffness, increase extensibility and therefore lead to an increase in ROM. However, there is no convincing evidence to suggest that this actually occurs. Researchers have investigated whether different positions provided different ROMs and also effects on reliability. This may be due to the alteration of the spinal curves. Lantz et al [150] found that a sitting position produces slightly greater cervical ROM. Strimpakos et al [148] also investigated this and found that both sitting and standing positions had very similar ranges. The majority of publications investigating cervical ROM perform the measurements with eyes open. It is only relatively recently that investigation into whether eyes open or closed may affect the ROM has been considered and it does not appear to affect the ROM and the reliability of the measurement [148]. There is no evidence available that directly evaluates the effect of examiner experience on ROM, however Nilsson conducted two consecutive reliability studies[151, 152] and found that reliability estimates were improved when more experienced examiners were involved. Because other potentially influential variables changed between the two

studies, it is not conclusive that the examiner experience was the causative factor for the improvement in reliability results.

4.4.5 CLINICAL IMPORTANCE OF ROM

"Of all the orthopaedic tests that an examiner can perform, none is more crucial than range-of-motion (ROM) testing of the affected articulation. ROM testing often reveals the origin of the patient's discomfort, because movement may reproduce the pain." [121]

It is widely accepted that assessment of ROM plays an important role in diagnosis, assessment of severity and the assessment of treatment outcome in management of musculoskeletal conditions [151]. A wide range of Health Care Professionals use cervical ROM in their management of both acute and chronic patients. In the acute setting of an ED, clinicians utilising the Canadian C-Spine Rule use reduced neck rotation ROM as a key indicator of serious injury. If rotation is reduced by greater than 45 degrees then this is taken as evidence of serious injury to warrant radiographic investigation [153]. As described in Chapter One, the QTF grading system also rates severity of injury and reduced cervical ROM distinguishes Grade I injuries from Grade II [6]. Hartling et al [62] evaluated the prognostic ability of this grading system and as a result proposed that the Grade II category should be divided into those who did and did not have limited cervical ROM due to its specific prognostic value. Furthermore, in the survey of ED consultants' management practice for whiplash injuries described in Chapter One [14] a considerable proportion of consultants reported using cervical ROM to guide referral on for further intensive treatment (physiotherapy).

Assessment and targeting of cervical ROM is a recommendation of current clinical guidelines for whiplash management [154, 155]. The promotion of exercise to improve active cervical ROM has been shown to be effective for WAD patients [156].

The ability to discriminate between symptomatic patients and asymptomatic patients, aid in evaluation of injury severity, provide information on prognosis and treatment effectiveness leads clinicians to place high clinical importance on the measurement of cervical ROM.

4.5 PREVIOUS STUDIES OF CERVICAL SPINE ROM IN WAD

Studies have previously investigated cervical ROM in WAD patients for two main reasons, firstly for its diagnostic and prognostic value and secondly as a treatment evaluation method or outcome measure. Therefore the review of literature below is divided into diagnostic/ prognostic studies and intervention studies (most commonly in the form of clinical trials).

4.5.1 PREVIOUS DIAGNOSTIC/PROGNOSTIC STUDIES

A summary of previous diagnostic and prognostic studies in a WAD population in which cervical ROM has been measured is provided in Table 16. The table demonstrates that numerous diagnostic and prognostic studies have been conducted for a range of populations, injury severities and stages of chronicity. A variety of measurement tools have been used by a variety of different examiners, however this was not well reported. The majority of studies measured active cervical ROM alone, although there were two studies that measured passive cervical ROM [85, 130]. No studies reported measuring both active and passive cervical ROM in the same cohort and no studies reported measuring patient-rated cervical ROM.

These studies provide consistent evidence that ROM is reduced in patients who have WAD compared to healthy control subjects. One study estimated that ROM was reduced by 25-35% of normal [157]. Extension was the movement most commonly cited as having the greatest reduction. Measurements of cervical ROM were most commonly presented in half-cycle plane ROM's, although a number of studies did use a sum score of all planes of motion [115, 136, 158, 159] with total active cervical ROM values ranging from 243 (±66)to 321(±61).

There is some conflicting evidence regarding the change in cervical ROM following the acute phase, with some studies finding that cervical ROM improves over time [140, 160] and others where cervical ROM remains reduced [86]. Prognostic studies of WAD populations showed mixed evidence for whether cervical ROM is a prognostic factor for poor outcome- this literature was reviewed systematically and presented in greater detail in Chapter Three.

Table 16 – Diagnostic and prognostic studies of WAD populations in which ROM was measured

Author, year	Study type	Population, sample size	Injury severity	Time of measure since injury	Follow-up period (LTF)	Measurement tool (AROM/PROM)	Ax	Findings
Antonaci et al, 2002[136]	LC	Secondary care, 70	WAD Gd 1&2	Mixed - 42 <1yr, 28>1yr	6 and 12 mths (83% @ 12 mths	Elite motion capture system (AROM)	NS	All movements except Ext were significantly reduced compared to controls. ROM more reduced for those with a recent injury (<1yr)
Armstrong et al, 2005[116]	XS	Advertisement, 23	WAD Gd 2&3	Whiplash 3 mths-5 yrs old,	NA	FASTRAK (AROM)	NS	F,E, RR, and LR were significantly reduced compared to healthy controls, E most limited
Atherton et al, 2006[58]	LC	ED, 480	NS	within 24 hrs	1,3 and 12 mths (30% @ 12 mths)	NS	ED Dr	Limitation of ROM not significant prognostic factor
Bono et al, 2000[161]	LC	Secondary care, 70	WAD Gd 2&3		6 and 12 mths	Elite motion capture system (AROM)	NS	Cervical ROM reduced. ROM improved with time
Cagnie et al, 2007[149]	C	Advertisement,16	WAD Gd 2	NS	NA	Zebris US motion analyser (AROM)	PT	All movements significantly reduced compared to healthy controls. F-E and LF significantly reduced compared to idiopathic neck pain.

Dall'Alba et al, 2001[137]	XS	Secondary care PT, 114 WAD, 89 control	WAD Gd 1-3	Mean 10 mths	NA	FASTRAK (AROM)	NS	ROM was significantly reduced compared to control group
Drottning et al, 2002[130]	LC	ED, 222	NS	6 weeks	6 and 12 mths (31% @ 12 mths)	Cybex inclinometer (PROM)	NS	Extension was most limited for WAD that had developed long term headache
Dvir et al, 2006[157]	XS	secondary care, 25	WAD Gd 1&2	> 6mths post injury	NA	Zebris US motion analyser (AROM)	NS	AROM homogeneously reduced by 25-35% for all directions although extension was most reduced.
Gargan et al, 1997[56]	LC	ED, 50	NS	3 mths	2 yrs (0%)	goniometer (AROM)	Ortho Dr	ROM significantly reduced in intrusive/disabled group compared to asymptomatic/nuisance and prediction of groups at 2 yrs with accuracy of 44% and 91% respectively
Heikkila and Wenngren, 1998[115]	LC	ED, 27	WAD Gd 2&3	2 mths	2 yrs (4%)	CROM (AROM)	NS	TAROM correlates with oculomotor function
Hendriks et al, 2005[65]	LC	GP and ED, 125	WAD Gd 1&2	2 weeks	1,3 and 12 mths (5% @ 12 mths)	CROM (AROM)	res PT	TAROM not a significant prognostic factor
Highland et al, 1992[162]	C	NS, 70	NS	NS	NA	MedX Cervical Extension Machine	NS	ROM improved over time

						(AROM F-E only)		
Hildingsson and Toolanen, 1990[67]	LC	Orthopaedic, 97	WAD Gd 1-3	ns, majority within 3 days	Mean 25 mths (4%)	NS	NS	Reduced ROM not associated with poor outcome
Hohl, 1974[55]	C	Orthopaedic, 534	NS	72% within 30 days	5 yrs (73%)	NS	Ortho Dr	Reduced ROM not associated with poor outcome
Kaale et al, 2007[163]	LC	Rehab centre, 47	NS	Mean 64 days	9 mths (NS)	CROM (AROM)	NS	Weak correlation between ROM and neck pain and no correlation between ROM and FAM
Kasch et al, 2001/5[68, 69]	LC	ED, 141	NS	1 week	1,3,6 and 12 mths (5% @ 12 mths)	CROM (AROM)	MD	Reduced ROM predictive of non-recovery at 1 year (B=2.53, CI 1.26-5.11)
Kasch et al , 2008[159]	LC	ED and GP, 688	WAD Gd 1-3	Within 10 days of injury	3,6 and 12 mths (9% ~@ 12 mths)	CROM (AROM)	Nurse	Active CROM was significantly reduced in high risk (242.9±70.8) compared to low risk group (330.5±34.5). Reduced AROM was associated with 4.6 risk increase for handicap, but not for long-term neck pain or headache.
Klein et al, 2001[164]	XS	Rheumatology & neuro secondary care departments, 46	WAD Gd 1-3	Mean(SD) 34 (26) mths	NA	Spine Motion Analyzer CA600 (AROM- rot only)	NS	ROM was significantly reduced compared to control group

Norris and Watt, 1983[45]	LC	ED, 61	WAD Gd 1-3	NS	2 yrs (NS)	NS	NS	All pts complained of stiffness. Reduced ROM was associated with poor recovery
Osterbauer et al, 1996[165]	XS	secondary care , 30	NS	mean 9 days	NS (NS)	CROM (AROM)	NS	ROM differentiated between WAD and controls with sensitivity of 77% and specificity of 84%
Ovadia et al, 2002[166]	LC	secondary care, 866	WAD Gd 1-4	NS	mean 32 mths	goniometer (AROM)	MD	ROM significantly reduced in severely affected pts.
Radanov et al, 1993a[80]	LC	Primary care, 113	NS	mean 7 days	6 mths (13%)	NS	NS	Reduced ROM associated with non-recovery (presence of symptoms) (p.007, X^2=7.05)
Richter et al, 2004[85]	LC	ED, 43	WAD Gd 1&2	NS	6 mths (26%)	NS (PROM)	NS	Reduced ROM not associated with poor outcome
Ryan et al, 1994[167]	LC	PT and GP primary care, 32	Mixed	"shortly" after injury	6 mths (6%)	CROM (AROM)	PT	NS
Sterling et al, 2003[168]	LC	GP, ED and advertisement, 66	WAD Gd 2&3	within 1 month	2 and 3 mths (0% @ 3 mths)	FASTRAK (AROM)	PT	Cervical ROM was significantly reduced 1 mth post injury. ROM was still significantly reduced in those with mod/sev disability at 3 mths.
Sterling et al, 2005[86]	LC	GP, ED and advertisement, 80	WAD Gd 2&3	within 1 month	6 months (5%)	FASTRAK (AROM)	PT	Cervical extension (OR 1.1 [1.03-1.20] predictive of mild disability at 6 mths. Cervical ROM not predictive of mod/severe disability.

Sterling et al, 2006[87]	LC	GP, ED and advertisement, 80	WAD Gd 2&3	within 1 month	2-3 years (19%)	FASTRAK (AROM)	PT	Cervical ROM was reduced in those with mod/sev disability at 2-3 years. Baseline cervical ROM was not predictive of disability at 2-3 years.

List of abbreviations: NS= Not stated, NA = Not applicable, LC = Longitudinal cohort study, XS = Cross-sectional cohort study, C = Cohort study, ED = Emergency department, PT= Physiotherapy/Physical Therapy, GP = General practice, WAD Gd = WAD Grade, mths = month, AROM = Active cervical Range of Motion, PROM = Passive cervical Range of Motion, Dr = Doctor, Ortho = Orthopaedic, Res = research

4.5.2 PREVIOUS INTERVENTION STUDIES

A number of systematic reviews of intervention studies on general neck pain have highlighted cervical ROM as a frequently utilised outcome measure [169-171] and studies specifically looking at interventions for Whiplash Associated Disorders are no different [13].

Table 17 displays intervention studies of WAD populations where cervical ROM was used as an outcome measure. As with the diagnostic/prognostic studies previously summarised, numerous methods of measurement were used and the vast majority measured active cervical ROM. Patient-rated cervical ROM was not reported in any study. Where significant differences existed between comparison groups using other outcome measures (e.g. pain and/or disability) it was quite common for there not to be a significant difference in ROM and yet an improvement in ROM over time was often noted. This may mean that cervical ROM measurement is not as accurate as required or that the relationship between this measure and pain and/or disability is not direct.

Table 17- Intervention studies in WAD populations with cervical ROM as an outcome measure

Author (year)	Population, n=	WAD grade	Intervention	Outcome measures	Follow-up period (LTF)	Measurement tool (AROM/PROM)	Ax (position)	Findings
Aigner et al (2006)[172]	NS, 50	NS	laser acupuncture vs. sham	symptoms, drug use, collar use, duration of condition	3 wks (clinic) and 12 mths (10%)	tape measure and goniometer (AROM)	NS (NS)	No significant difference in ROM between groups. ROM improved with time.
Bonk et al (2000)[173]	ED, 97	NS	active vs. collar therapy	symptoms	12 wks	Tape measure for F & E and goniometer for LF and R (AROM)	NS (NS)	No significant difference in pain and ROM. ROM improved with time.

Borchgrevink et al (1998)[131]	ED, 201	NS	Act as usual vs. immobilisation	global improvement question, symptoms, sick leave	2, 6 and 24 wks	Cybex (AROM)	NS (NS)	Significant difference in subjective symptoms such as pain and neck stiffness but no significant difference in objective ROM. ROM improved with time.
Bunketorp (2006)[174]	Secondary care, 49	NS	supervised training group or HEP	Self-efficacy scale, Tampa Scale for Kinesiophobia, Pain Disability Index, pain VAS, sick leave, medication	3 and 9 mths	CROM (AROM)	NS (NS)	

Crawford et al (2004)[175]	ED, 108	NS	Mobilisation vs. collar	ADL(function), pain	3, 12 and 52 wks	NS - sum of F, E, LLF, RLF,LR, RR (AROM)	NS (NS)	No significant differences in ADL or ROM between groups at 1 year. ROM improved with time.
Fialka [176] (1989)	NS, 60	NS	Therapy, traction, massage vs. US vs. Iontophoresis vs. none	neck pain	NS	goniometer (NS)	NS (NS)	No between group analysis
Foley-Nolan (1992)[177]	ED, 40	NS	PEMT collar vs. placebo collar	pain, global rating of progress	2,4 and 12 wks	visual estimation (PROM)	NS (NS)	Significant improvement in pain and global rating for PEMT group at 4 wks. No significant difference in pain at 12 weeks. Significant improvement in ROM in active compared with control group at 12 wks.

Hendriks and Horgan (1996)[178]	ED, 16	NS	physio vs. physio plus ulta-reiz current	McGill Pain Qu and VAS pain	6 wks	Myrin (AROM)	PT (NS)	Pain and rotation ROM significantly improved in ultra-reiz group but F-E and LF ROM not significantly different at 6 wks.
McKinney (1989)[179]	ED, 247	NS	Physiotherapy vs. Advice vs. Rest	Pain VAS	2 mths	goniometer - mean LF used as summary (AROM)	Dr (NS)	Physio and Advice groups significantly improved in pain and ROM compared to Rest group. No significant difference between physio and advice groups at 2 mths. ROM improved with time.
Mealy (1986)[180]	ED, 61	NS	active vs. collar	Pain VAS	8 wks	CROM-like (AROM)	Dr (sitting)	Significant improvement in pain and AROM for active group at 8 wks. ROM improved with time.

Pennie and Agambar (1990)[9]	ED, 135	NS	collar vs. traction	Pain VAS	Max 5 mths (5%)	goniometer (AROM)	NS (NS)	no significant difference between two groups
Provinciali (1996)[181]	NS, 60	NS	multimodal vs. electrotherapy	Pain VAS, global improvement rating, return to work	6 mths	tape measure (AROM)	Dr (NS)	Significant improvement in pain, global improvement and RTW for experimental group at 6 mths. No significant difference between two groups for ROM.
Rosenfeld (2000)[182]	ED, Primary care and private clinics, 97	NS	active vs. standard (early and delayed)	Pain VAS	6 mths	CMS (AROM)	Lab tech/nurse (NS)	Significant improvement in pain VAS but not for ROM for active group at 6 mths.
Rosenfeld (2003)[24]	ED, Primary care and private clinics, 97	NS	Active vs. standard	Pain VAS, sick leave	3 yrs	CMS (AROM)	Lab tech/nurse (NS)	Significant improvement in pain VAS and sick leave for active group at 3 yrs. No significant difference but trend for improved ROM for active

								group at 3 yrs.
Soderlund (2000)[183]	ED, 66	14% Gd I, 83% Gd II, 3% Gd III	Physio vs. physio plus kinaesthetic and co-ordination exercises	Pain Disability Index, Self-efficacy scale, Coping strategies Qu, pain VAS, posture, kinaesthetic sensibility	3 and 6 mths	CROM-like (AROM)	PT (sitting)	No significant differences between groups for PDI, SES, Pain VAS or ROM. ROM improved over time.

Soderlund (2001)[184]	Orthopaedic, 33	Gd I-III	physio vs. physio plus CBT	Pain Disability Index, Self-efficacy scale, Coping strategies Qu, pain VAS, posture, kinaesthetic sensibility	3 and 6 mths	CROM-like (AROM)	PT (sitting)	No significant differences between groups for PDI, SES, Pain VAS or ROM. ROM improved over time.
Thuile and Walz (2002)[185]	NS, 92	NS	magnetic field treatment vs. Control	Pain VAS	NS	goniometer (AROM)	NS (NS)	Pain and ROM was significantly improved with magnetic field treatment

List of abbreviations: NS= Not stated, NA = Not applicable, LTF = Loss to Follow-up, ED = Emergency department, PT= Physiotherapy/Physical Therapy, GP = General practice, WAD Gd = WAD Grade, CBT = Cognitive Behavioural Therapy, VAS = Visual Analogue Scale, mths = month, wks = weeks, yrs = years, AROM = Active cervical Range of Motion, PROM = Passive cervical Range of Motion, Dr = Doctor, Lab Tech = Laboratory Technician

4.6 SUMMARY

This chapter has introduced assessment of the cervical spine with a particular focus on the assessment of Range of Motion. It has explained how cervical spine motion is essential for everyday function and how this motion is the result of complex interplay between muscles, joints, nerves and central nervous system control.

It has been asserted that assessment of motion and any associated pain response is a keystone of the concept of modern musculoskeletal management. Health care professionals assess motion to help locate the source of dysfunctions. Evaluation of Range of Motion (ROM) is one important aspect of this assessment. It has been noted that a multitude of factors may affect ROM assessment findings. Patient-rated cervical ROM has so far been neglected as a source of information on ROM.

This chapter has presented evidence to show that whiplash injuries result in a loss of cervical spine motion and argues that the assessment of ROM is used by clinicians and researchers to inform diagnosis, prognosis and evaluate treatment response.

Previous studies of WAD populations have used a wide variety of different methods to measure cervical ROM. It is still unclear which the best method to use is. Key concepts to help inform the selection of a method of measurement are the reliability and validity of that method. With this in mind, the next chapter describes a systematic literature review of the reliability and validity studies of methods for measuring cervical ROM.

5 CHAPTER FIVE – A SYSTEMATIC LITERATURE REVIEW OF RELIABILITY AND VALIDITY STUDIES OF METHODS FOR MEASURING ACTIVE AND PASSIVE CERVICAL RANGE OF MOTION

5.1 INTRODUCTION

In the previous chapters the condition of Whiplash Associated Disorders (WAD) has been introduced as a significant healthcare concern. The author's involvement in a large RCT investigating the management of acute WAD has been discussed, including the development of clinical assessment procedures for this population. Assessment of cervical Range of Motion (ROM) has been highlighted as the focus of this thesis. Chapter Three concluded that cervical ROM remains an uncertain prognostic tool, in part due to the inconsistency of methodologies used to assess this. In particular there have been a variety of methods used to measure ROM in prognostic cohort studies.

The preceding chapter described how cervical ROM is measured in various different directions in both active and passive ways. Health care professionals (HCP) and researchers use a variety of methods to operationalise these various measurements, ranging from simple visual estimation to complex 3D motion analysis [186, 187]. Clinicians in their search for practical and clinically relevant instruments, commonly use visual estimation, inclinometers, and goniometers, whereas researchers, in their quest for optimum accuracy use methods such as complex 3-dimensional electromagnetic or audiovisual technologies. Practicality for clinicians means ease of use, reasonable cost, portability, unprohibitive amount of training

required to operate and patient comfort. Some of these factors are not such important considerations for researchers. In terms of accuracy, clinicians are most frequently monitoring *change* in ROM and therefore are concerned with whether these changes are 'true'. The American Medical Association suggest that measurement tools may allow variation of ± 10% of the measured movement to be acceptable as a clinical assessment tool for cervical spine impairment [127]. An example; if a patient's active right rotation is measured to be 30 degrees at time point one and then 35 degrees at time point one then if the tool measures with error of less than 10% (3 degrees in this case), then a clinician should be satisfied that a change in cervical ROM has occurred in this patient.

Cervical spine ROM is particularly challenging to assess accurately because of the complex anatomical structure and resulting 'coupled' movements [96], as described in the last chapter.

This chapter will investigate the validity and reliability of the range of methods available to measure cervical ROM through a systematic review of the literature. This chapter aims to provide the reader with a clear insight into which are the most clinically useful devices for measurement of cervical ROM and also the strengths and limitations of methodologies used up to this point. Implications for further research will be discussed, some of which will be described in subsequent chapters.

Before describing the methodology of this systematic review, a discussion of the topics of reliability and validity will be presented.

Following description of the methods and results of the systematic review, a discussion of findings related to previous work and the impact this will have on the selection of a tool for the prognostic cohort study described in Chapter Seven.

The review presented in this chapter has been published as a manuscript in the Journal of Manipulative and Physiological Therapeutics (Williams et al [188]) (Appendix 7).

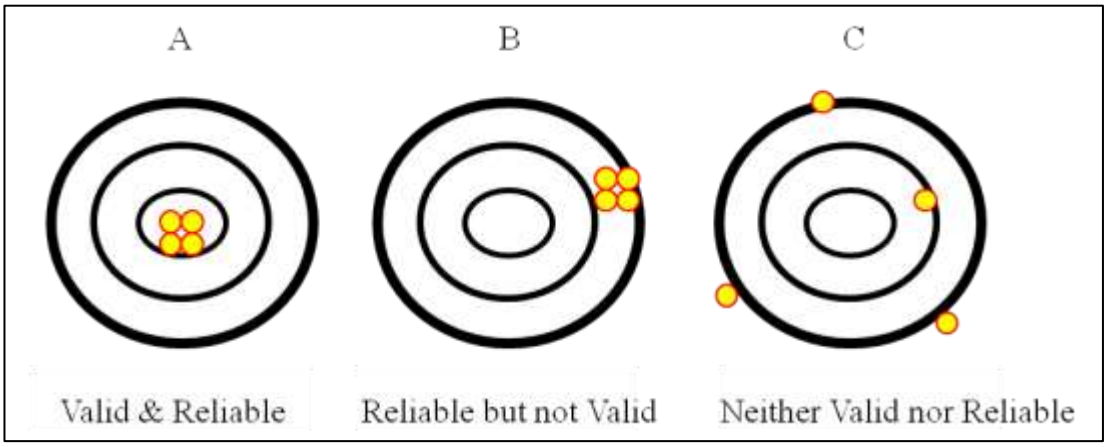
5.1.1 RELIABILITY AND VALIDITY

As previously stated, clinicians' measure cervical ROM in order to assess whether there is a limitation of range or impairment, indicate possible structures that could be causing this limitation/impairment and also to 'objectively' measure treatment progress. In order for a measure of ROM to perform all of the above, i.e. be clinically useful, it must be consistently accurate. In other words the measure needs to be valid and reliable [189].

Reliability and validity can be confusing concepts due to the various synonyms that are used, often interchangeably. For the purpose of this thesis, reliability is defined as consistency of a measurement across time, patients or observers[190]. Validity is defined as the extent to which the method/tool measures what it is intended to measure [189, 190]. More recently this definition of validity has been widened to focus on the degree of confidence we have about making inferences about the population the measurement method/tool was used on; a shift of focus from the method/test to the population it is utilised on.

Several authors have used an analogy of shooting at a target to explain the concepts of reliability and validity as presented in Figure 12 [189, 191]. In order to be defined as a 'good shot' one needs to be accurate and consistent when shooting at a target (A). There is no use in being consistently off-target (B) or inconsistently on-target (C).

Figure 12 – Target analogy for reliability and validity



5.1.1.1 Reliability

The theory of reliability is derived from the discipline of psychology and in particular *Classical Test Theory* [192]. This theory states that any observed measurement consists of a true value and an error value. It is very rare to find a truly consistent clinical measurement method; all methods have some error within them. Only random errors are considered in reliability theory (systematic errors -predictable errors occurring in one direction only- are normally dealt with under the construct of Validity).

Classical test theory provides us with the formula where reliability = $\sigma^2 t / (\sigma^2 t + \sigma^2 e)$ where $\sigma^2 t$ is equal to the true score variance and $\sigma^2 e$ is equal to the error score variance. This results in a unitless number that ranges from zero (all variance due to measurement error or zero reliability) to one (all variance due to true score or perfect reliability).

Two categories of reliability have traditionally been constructed and tested for methods of ROM measurement; Intra-observer reliability – the reliability within a single tester and Inter-observer reliability – the reliability between at least two examiners/ populations/ settings.

One helpful distinction that has been made, particularly with reference to the correct use of statistical techniques, is that of relative reliability and absolute reliability [193]. Relative reliability informs us of whether the differences in one set of measurements are ranked in the same order as a second set of measurements (also known as association). The limitation of this type of reliability is that readings don't necessarily have to agree to result in 'perfect reliability' – therefore this can lead to an exaggeration in degree of reliability.

Absolute reliability is a more recent concept and this is concerned with the degree with which repeated measurements vary for individuals (also known as agreement). It is expressed statistically using the Standard Error of Measurement (SEM) or Limits of Agreement tests (LoA) [194].

The Standard Error of Measurement (SEM) is the standard deviation of measurement errors and provides an estimate of error around a 'true score' of a repeated test on an individual for interval data [195]. When the standard deviation and reliability co-efficient are known it can be calculated as follows:

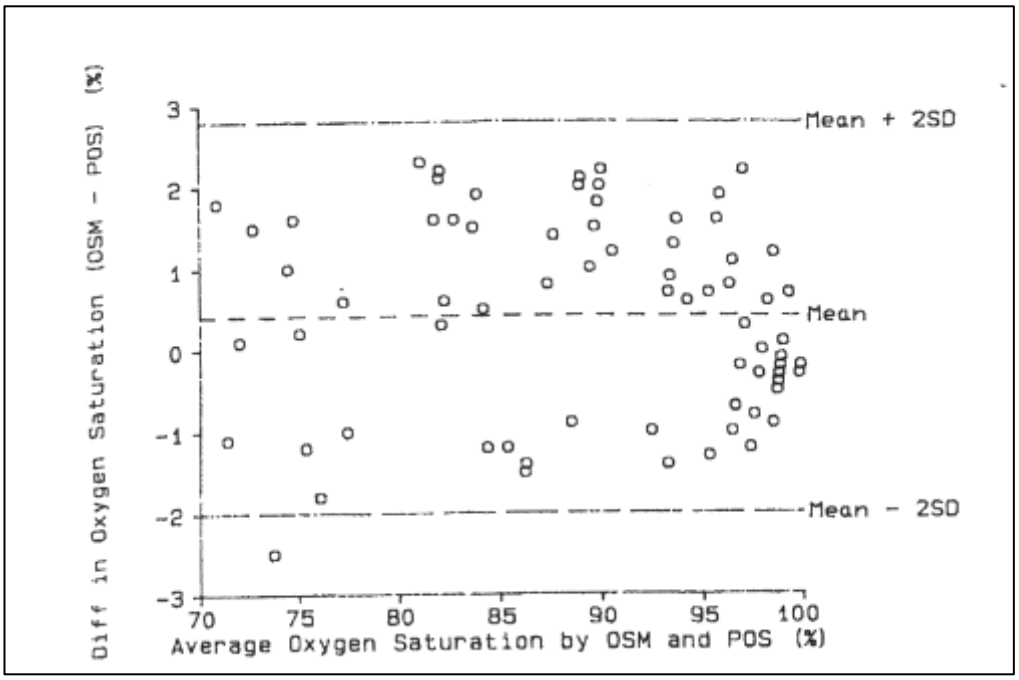
$$SD \sqrt{1 - r}$$

Where SD = Standard Deviation and r = reliability co-efficient.

Limits of Agreement tests are graphical techniques and basic calculations that allow observation of outliers and bias relatively quickly and easily. Differences in results are plotted against the mean value of the two measurements, then mean and SD of the differences between the measures are calculated and then finally 95% limits of agreement with confidence intervals are calculated [194, 195].

Figure 4 below displays an example of good agreement[194]. The mean difference is 0.42 % points (95% CI 0.13 – 0.70). Limits of agreement are -2.0 and 2.8.

Figure 13 - Limit of Agreement plot example[194]



The Intra-class Correlation Coefficient (ICC) reflects both measurement error and degree of consistency and is currently the most commonly used statistical technique to interpret reliability. It expresses the ratio of variance between subjects to total variance in scores. ICC has several versions [196], the use of which depends on assumptions made about the observers and population observed.

5.1.1.2 Validity

Traditionally there have been three categories of validity discussed; content, criterion and construct. Various sub-categories have been proposed, often leading to confusion of this

fundamental area. For the purposes of this thesis a brief outline of the three distinct categories is provided and readers are referred to specific texts for further reading of the fluctuations in definitions [189, 190].

Content validity, the scope of a method/tool, is seldom measured formally; instead the 'face-validity' or clinical credibility of a method/tool is determined from expert opinion. Range of motion is widely accepted to be face-valid i.e. a value of 20 degrees is less ROM than a value of 40 degrees, although there is no statistical evidence that can be provided for this.

Criterion validity is concerned with comparing a method with a definitive 'gold or criterion standard'. A 'gold or criterion standard' is a method/tool or test that hypothetically has a sensitivity and specificity of 100% (no false positives and no false negatives). In practice there are no gold or criterion standards. Therefore there is the potential for a gold standard to change if a more specific and/or sensitive method is found. With regards to cervical ROM, no gold standard exists and it is unlikely there will ever be one confirmed. Radiographs have been considered the closest method to a gold standard; however the method has not undergone sufficient reliability and validity experimentation to be truly classed as this [101]. Criterion validity is divided into two types, concurrent and predictive, depending on when the comparison with the method/tool is compared. Concurrent criterion validity is established when the comparison is made at the same time. For example when ROM is measured using visual estimation and then using a goniometer immediately after. Predictive criterion validity is established when the new method is applied at baseline and compared to subsequent outcomes at a later date. Because of the time delay and the resulting potential for bias, predictive criterion validity studies are rarely conducted.

Construct validity is concerned with the accuracy with which a method represents an attribute that cannot be directly observed. It is determined through deductive reasoning and assessment of convergence to similar methods/tools and divergence to different methods/tools. An example of this is 'neck stiffness'. This is a construct – we cannot definitively prove that an individual has a 'stiff' neck. However we might hypothesise that an individual who complains of a 'stiff' neck might be observed to have difficulty turning their head by a certain amount.

In order to optimise the reporting of this systematic review, the following sections are structured according to the PRISMA statement with modifications appropriate to the nature of the studies within the review [42].

5.1.2 RATIONALE - PREVIOUS REVIEWS

Four literature reviews have previously been published regarding reliability and/or validity studies for the measurement of cervical spine ROM [101, 160, 197, 198].

The primary objective of the first of these reviews by Chen et al [101] was to carry out a meta-analysis of normative cervical motion but also incorporated a narrative review of the reliability and clinical validity of the studies. A search performed solely using Medline (1966 to 1998) was not specific to reliability and/or validity studies, merely using the key words "range of motion" and "cervical". Data extracted for reliability and variability of methods were averaged within each study and organised by technology. 45 papers were retrieved for the meta-analysis of normative motion; 17 of these papers reported reliability studies and seven papers reported validity studies. The authors concluded that reliability was inappropriately and inconsistently analysed and suggested that future studies should include a comprehensive quantitative analysis using Intra-class Correlation Coefficients (ICC) and

Limits of Agreement (LoA) statistical techniques in parallel. They argued that true validation of tools to measure cervical ROM is not possible as there is no true gold standard of measurement and that often intra-technology variability of measurements are as large as or larger than inter-technology variability. From this they deemed measurement protocols and examiner training to be as important as the technology itself. From the papers reviewed it was unclear whether passive ROM is more reliable than active ROM. It was deemed clearer that full cycle movements are more reliable than half-cycle ones.

The second review by Antonaci [160] was purely narrative with no details of how or what types of literature were obtained. 15 reliability studies were referenced however results were only selectively reported. The papers were discussed in groups according to the technology, and tables usefully provide details of advantages and disadvantages of each method. No conclusions were offered as to preferable methods of measurement, although the abstract stated that "Cybex and 3D kinematic analysis by means of opto-electrical scanners (Elite system) seemed to be the most reliable and reproducible methods."[160]

The third review, conducted by Jordan and published in 2000 [197], assessed 21 reliability studies systematically, although meta-analysis was deemed inappropriate due to heterogeneity of the included studies. A search strategy was used on 11 different electronic databases, followed by data extraction and a qualitative assessment of included papers. A thorough discussion of methodological considerations for reliability studies was offered including use of appropriate statistical techniques, sample size calculations, standardised measurement protocol along with analysis of the relative reliability of the various tools. He concluded that the Cervical Range of Motion device (CROM; a combination of gravity and compass goniometers) is the most promising method although further, more rigorous investigation of all tools is warranted.

The final review of studies (de Koning et al[198]) was published in April 2008 as the author's systematic review was being written-up. This clearly justifies there was a need for an up-to-date systematic review to be conducted at this time. The research question was very similar in that it tried to establish the most appropriate method for assessing cervical ROM; however de Koning et al's review was limited to *active* ROM in patients with *non-specific neck pain*. It could be argued that the research objective of the review described below is more ambitious, incorporating passive ROM and all types of sample populations. It was deemed important to consider reliability and validity of methods measuring passive ROM as this is a commonly used impairment measure which is thought to provide its own unique information regarding the state of tissues and in conjunction with active ROM measurement will add to the often complex biopsychosocial picture that is the clinical examination. It is also noted that although de Koning et al [198]stated they were trying to establish the most appropriate method of assessing ROM in non-specific neck pain patients, only nine of the articles included involved a non-specific pain population. Absence of discussion of this point in the article is notable.

The review by de Koning et al [198] included 33 papers and evaluated not only reliability and validity of methods but also responsiveness and interpretability. 23 of the 33 articles are included in the systematic review described below, with the other 10 having been excluded for statistical reasons described in the methods section below. The review of de Koning excluded devices that were not portable, affordable (max 1,000 Euros) and easy to use (time to test max 5 min) by Allied Health Professionals in daily practice. Similarly to Jordan [197], de Koning et al [198] did not attempt meta-analysis, although a progression was made in that a quality assessment tool was devised and utilised to provide some standardised interpretations to be made. Discussion about how this criterion was developed is limited

(probably due to the brevity required in this published format) and this would undoubtedly have been useful for future investigations. De Koning et al's [198]discussions and analyses incorporate both research and clinical issues and conclude that "both the CROM device and single inclinometer can be considered appropriate instruments."

Similar conclusions drawn from all four of the literature reviews suggest that improvements to methodology and reporting can be made to future reliability and validity studies although more recent studies suggest that researchers are heading in the right direction [198]. In terms of the methodology for literature reviews in this area, the main challenge is the lack of agreed quality assessment criteria for these types of studies. It is clear that until methodology and reporting of these studies is more homogeneous the advantages that result from meta-analysis will continue to be absent.

5.1.3 OBJECTIVE

No review exists that includes an evaluation of both reliability and validity studies for methods for measuring both active and passive cervical ROM. The research question that this systematic review is attempting to answer is:

What is the reliability and validity of the various methods for assessing active and passive ROM in the cervical spine?

The objective of this systematic review is therefore to evaluate studies of reliability and validity of methods for measuring active and passive cervical range of motion.

5.2 METHODS

5.2.1 PROTOCOL

A review protocol was developed in accordance with guidelines from NHS Centre for Reviews & Dissemination ensuring research questions, search strategies and data sources were defined *a priori* [37].

A group of researchers who had experience in research synthesis, conducting reliability and validity studies and knowledge of cervical spine assessment were assembled to work up the protocol and conduct the research synthesis itself.

5.2.2 ELIGIBILITY CRITERIA

Selection of papers was according to the following selection criteria, which was developed following discussions with a clinical specialist and researchers with expertise in systematic reviews. Following a pilot study a supplementary exclusion criteria was inserted. It was felt that studies using what have only more recently been deemed inappropriate statistical tests (particularly reliability studies that presented Pearson's r statistics) would cloud the synthesis process which was already challenging enough due to the lack of objective criteria.

5.2.2.1 Inclusion criteria

- Studies aiming to assess a cervical spine Range Of Motion measurement method for intra-observer and/or inter-observer reliability and/or validity.

- Studies investigating methods of measuring global cervical spine Range Of Motion (i.e. angular displacement of the head away from the thorax).

- Studies could involve symptomatic and asymptomatic subjects (no disease group excluded).
- Studies investigating methods that evaluate movement in saggital (flexion/extension), transverse (rotation) or frontal (lateral flexion) planes.
- Studies with participants that were adult subjects >18years

5.2.2.2 Exclusion criteria

- Studies investigating methods that measure Range Of Motion during a whiplash mechanism of injury
- Studies investigating methods that measure individual vertebral segmental motion
- Studies investigating methods that measure of static postures of the cervical spine.

Supplementary exclusion criterion:

- Studies that used inappropriate statistical analysis techniques e.g. Pearson's r for reliability.

5.2.3 INFORMATION SOURCES

A search was conducted in the following electronic databases: PubMed (from 1950 - October 2007), MEDLINE (from 1966 – January 2008), CINAHL (from 1982 - January 2008), EMBASE (from 1980 - January 2008) and AMED (via OVID) (from 1985 – January 2008).

References from retrieved, eligible articles, systematic reviews and theses were searched for supplementary studies. In addition, Physiotherapy and the Journal of Manipulative and Physiological Therapeutics were hand searched for potential studies in the last 5 years.

Google Scholar was also used in a limited capacity with the name of any retrieved tools and the term “cervical spine” to locate any further studies.

5.2.4 SEARCH

The following search terms were used: (Neck OR cervical OR spine OR cervical spine) AND (movement OR motion OR range of motion) AND (Validity OR reliability OR repeatability OR reproducibility) and also MeSH Terms were used in PubMed: “Range of Motion, Articular” AND “Neck” AND “Reproducibility of Results” AND “Validation Studies [Publication Type]”.

5.2.5 STUDY SELECTION

Non-English articles were not excluded from results of the searches, however abstracts of conference presentations were.

Papers were initially screened by the author and deemed potentially relevant based on their abstract or their title (if an abstract was not available). Full text versions were then obtained to ensure studies fulfilled the selection criteria. Study inclusion was assessed by two independent reviewers (MW and CM). Disagreements were discussed and if necessary a third reviewer (SG) would facilitate consensus.

5.2.6 DATA COLLECTION

An electronic format was constructed following discussions with expert reviewers and clinical researchers to facilitate assimilation and interpretation of data. Descriptive data regarding publication details, type of study, movements and device evaluated, subject and observer characteristics, measurement protocol including blinding and statistical analysis methods were recorded. This was carried out independently by the author and then appraised

with a second reviewer (CM). Any discrepancies would be discussed to achieve consensus, using a third person if disagreement persisted. Ideally the data abstraction should have been conducted by two independent reviewers and then agreed on a consensus; however limited resources meant that this was not achievable.

5.2.7 QUALITY ASSESSMENT

There were no established or validated criteria for assessing quality of reliability or validity studies at the time of commencing this review. The most recent systematic review of clinimetric properties of methods to measure cervical ROM [198] composed a checklist for assessing the studies, but as this was published after the design and conduct of this systematic review a comparison will be made in the discussion section.

Authors of systematic reviews of reliability and validity studies for measuring ROM for other related anatomical areas have developed their own criteria or adapted previous works. Separate quality criteria were developed for reliability and validity studies because a significant difference in the methodology and reporting of these studies is present, for example appropriate statistical methods are different. Criteria were designed to assess internal validity, external validity and statistical methods. The quality criteria presented below were developed using previous tools used by Van der Wurff et al [199], Stochkendahl et al [200], and Van Trijffel et al [201].

For the checklist for assessing reliability studies, items 1, 2, 4, 7, 9, 11, 12, and 13 were taken directly from the criteria list of Van der Wurff et al [199]. Items 5, 6 and10 used wording from items used by Stochkendahl et al [200]et al to adapt Van der Wurff et al [199] items. Items 3 and 8 were adapted from items used by Van der Wurff et al [202] and Van

Trijffel et al [201] incorporating important considerations highlighted in previous systematic reviews [101, 197].

For the checklist for assessing validity studies, items 1, 2, 4, 7, 8, 10, 12, 13, 14 were taken directly from the items taken from Van der Wurff et al [202]. Items 5, 6 and 11 used wording from items used by Stochkendahl et al [200] to adapt Van der Wurff et al [199] items. Items 3 and 8 were adapted from items used by Van der Wurff et al [202] and Van Trijffel et al [201] incorporating important considerations highlighted in previous systematic reviews [101, 197].

5.2.7.1 Quality Criteria for Reliability studies

Section A – Sample population

1. Adequate description of study population – symptomatic/asymptomatic, gender, age
2. Description of selection criteria
3. Justification of appropriate sample size (through calculation or guidelines)
4. Withdrawals / Drop-outs described

Section B – Test Procedure

5. Order of observers conducting the test(s) randomised (inter-observer studies)
6. Observers blind to clinical presentation of participants/previous findings
7. Description of standardised measurement protocol incorporating standardisation of positions, movement directions, warm-ups etc. in order that the procedure could be reproduced

8. Description of examiner's experience (clinical and with device and procedure)

9. Consensus procedure / pilot study reported

Section C – Test Results & Analysis

10. Observers blind to other observer's findings

11. Test re-test procedure, description of time interval (participants' characteristics stable during study period?)

12. Appropriate descriptive statistics presented (frequencies and agreements)

13. Appropriate inferential statistics presented (ICC or Kappa with confidence intervals)

5.2.7.2 Quality Criteria for Validity studies

Section A – Sample population

1. Adequate description of study population – symptomatic/asymptomatic, severity, gender, age

2. Description of selection criteria

3. Justification of appropriate sample size (through calculation or guidelines)

4. Withdrawals / Drop-outs described

Section B – Test Procedure

5. Order of tests randomised

6. Observers blind to clinical presentation of participants/previous findings

7. Description of standardised measurement protocol incorporating standardisation of positions, movement directions, warm-ups etc. in order that the procedure could be reproduced
8. Description of Gold or Reference standard
9. Description of examiners experience (clinical and with device and procedure)
10. Consensus procedure / pilot study reported

Section C – Test Results & Analysis

11. Observers blind to previous index/reference device results
12. Description of time interval (participants' characteristics stable during study period?)
13. Appropriate descriptive statistics presented (frequencies and agreements)
14. Appropriate inferential statistics presented (Correlation coefficient or agreement stats with confidence intervals)

Each criterion was rated as Yes/No/Unclear. If there was any difference between reviewers for these 3 categories this was defined as a discrepancy.

A subset of included studies was reviewed to pilot the quality assessment tool. Two reliability articles [135, 203] and two validity articles[204, 205] were assessed to establish consistency in the procedure and further modify the assessment tool[206]. Minor modifications are described in the results.

For the main review, quality assessment was conducted by three independent reviewers (MW, CM and AC), with each article being assessed by two of the three reviewers.

Disagreements were discussed in order to reach consensus. If consensus could not be reached the third reviewer could be consulted to decide on the final rating.

Kappa coefficients were calculated in order to provide an estimation of the strength of agreement between reviewers for both the reliability and validity studies included. Additionally, confidence intervals around the coefficients were calculated along with prevalence and bias indexes to enable informed interpretation [191].

5.2.8 DATA SYNTHESIS

It was the intention as part of this systematic review to carry out meta-analysis if appropriate. According to Deeks [207] this "should only be considered when the studies have recruited from clinically similar populations, used comparable experimental and reference tests, and are unlikely to be biased". Conducting meta-analysis also relies on the data being sufficiently homogeneous in order to make meaningful synthesis and analysis.

If meta-analysis was not possible or advisable a descriptive 'best evidence synthesis' would be provided. Instead, an estimate of the level of reliability and validity would be calculated for each study using the mean of the reliability statistics for each of the ROM's in the three cardinal planes. Half-cycle statistics were used where available. Using this mean value, device reliability was categorised as good, moderate or poor, depending on the type of study (adapted from Swinkels et al, [208]; see Table 18 below). This method was used by de Koning et al [198] in the most recent systematic review of this kind. The ratings are displayed in the last columns of tables three and four.

Table 18 - Categories for levels of reliability and validity (Swinkels et al [208])

Study Type	Good	Moderate	Poor
IaO Reliability	>0.85	0.65-0.85	<0.65
IeO Reliability	>0.80	0.60-0.80	<0.60
Validity	>0.65	0.50-0.65	<0.50

An overall rating for device reliability and validity was calculated according to the following rules:

Good: At least 75% of studies had a rating of good.

Moderate: At least 75% of studies had a rating of moderate (and good if not rated overall as good)

Poor: At least 75% of studies had a rating of poor.

5.3 RESULTS

5.3.1 PILOT STUDY

Two reliability ([135, 203]) and two validity ([204, 205]) articles were piloted for consistency between reviewers to ascertain whether clarification or modification was necessary for the data abstraction and quality assessment procedures. Generally the data abstraction forms were found to be appropriate for both reliability and validity sections and the data produced was satisfactory to the reviewers appraising the system.

The reliability articles that were reviewed produced scores on the quality assessment tool of five and 11 positive items after all discrepancies were discussed and resolved. Neither of the articles used a sample size calculation and there were inconsistencies in how the re-test interval was reported and what descriptive and inferential statistics were presented.

Table 19 shows the number of items rated positively and negatively for each reviewer. From this data the Kappa coefficient for quality assessment of reliability studies was calculated using an online application[209].

Table 19 - Pilot Quality Assessment results table – Reliability studies

		Reviewer B		**Total**
		Item Yes	Item No / Unsure	
Reviewer A	Item Yes	13	5	18
	Item No / Unsure	3	4	7
	Total	16	9	25

The resulting Kappa coefficient (0.27 (95% CI -0.12-0.66)) and agreement (68%, SE 0.2) indicates there was a ‘fair’ strength of agreement between reviewers for use of the quality assessment tool for reliability studies [210]. The prevalence index was calculated as 0.36 indicating a moderate index in favour of obtaining a ‘Yes’ decision. The bias index was calculated as 0.08 indicating a negligible bias between reviewers.

Despite all disagreements being resolved through discussion the reviewers felt that the Kappa coefficient and agreement were unsatisfactory so considerable clarification was made on items that scored poorly. Items that required clarification included what constituted an

appropriate description of drop-outs, blinding, a consensus procedure, a suitable interval period and finally, what appropriate inferential statistics were. This final point lead to the creation of an extra exclusion criterion described above in the study selection section.

For the validity section of the pilot study, the two studies scored five and nine positive items. Again neither of the studies justified their sample size with a calculation. There was some inconsistency with the reporting of blinding and description of experience and training of observers. Table 20 below shows the number of items rated positively and negatively for each reviewer. From this data the Kappa coefficient for quality assessment of validity studies was calculated using the aforementioned 'Kappa calculator'.

Table 20 - Pilot Quality Assessment results table – Validity studies

		Reviewer B		**Total**
		Item Yes	Item No / Unsure	
Reviewer A	Item Yes	10	4	14
	Item No / Unsure	2	12	14
	Total	12	16	28

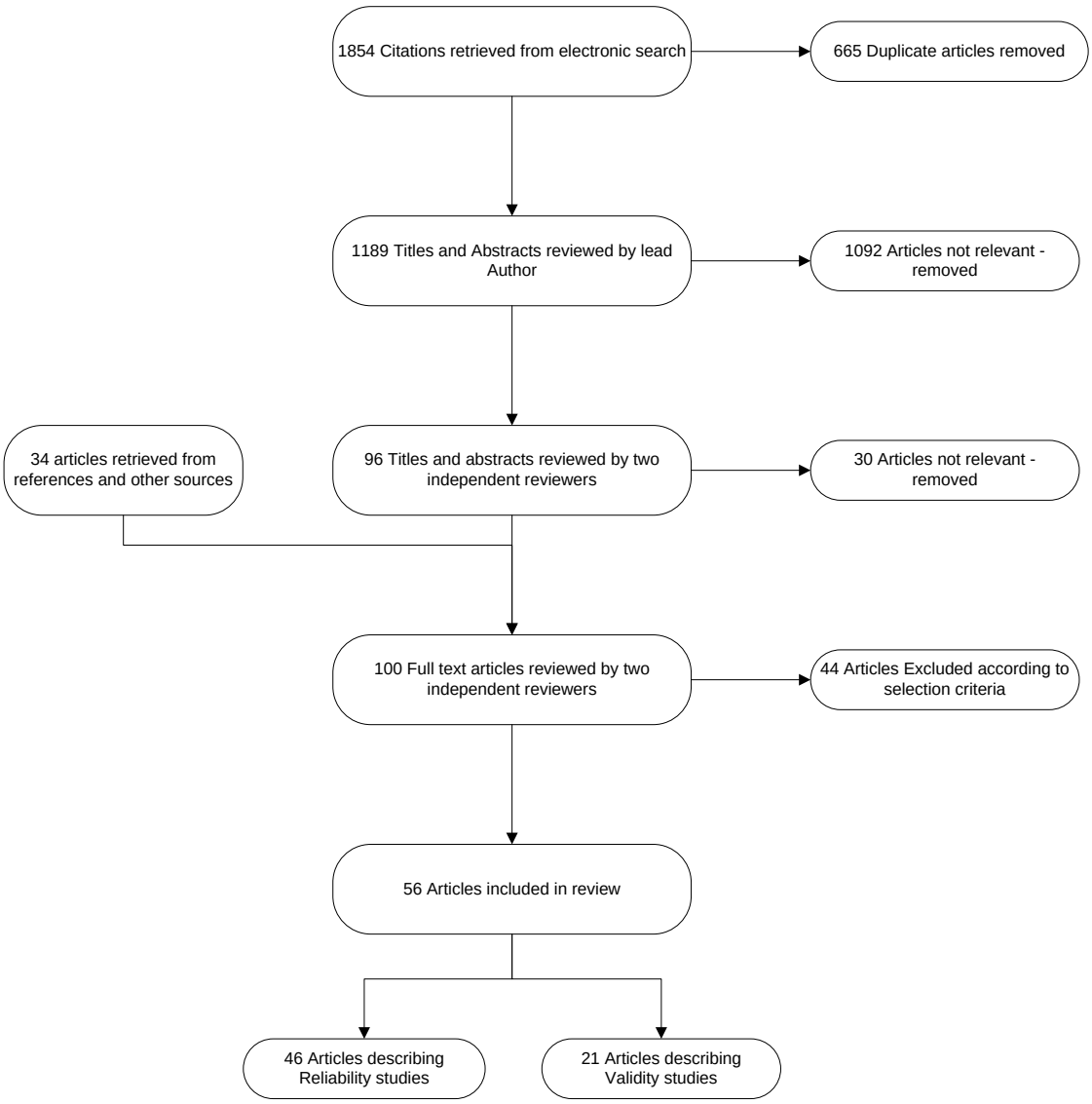
The resulting Kappa coefficient (0.57 (95% CI 0.27-0.87)) and agreement (79%, SE 0.15) indicates there was a 'moderate' strength of agreement between reviewers for use of the quality assessment tool for reliability studies [210]. The prevalence index was calculated as -0.07 indicating a very low index in favour of obtaining a 'No' decision. The bias index was calculated as 0.07 indicating a negligible bias between reviewers. With far fewer discrepancies than the reliability review despite being conducted simultaneously, and all

disagreements being resolved through discussion, there were only a small number of minor clarifications to make with the quality assessment tool for validity studies.

5.3.2 STUDY SELECTION

Figure 14 displays the flow of articles through the reviewing stages. A list of the excluded papers and reasons for exclusion is available.17 articles fulfilled the main selection criteria but did not use appropriate statistical techniques. These were excluded from the main review as not fulfilling the supplementary exclusion criterion.

Figure 14 – Flow of study



A total of 56 articles fulfilled the selection criteria and were included in the review. 46 articles described reliability studies and 21 articles described concurrent validity studies (11 articles described reliability and validity studies within the same paper).

34 articles were retrieved only from sources outside of the electronic search strategy (15 of these were included in the final review).

46 articles reporting 68 reliability studies were included in the review (31 intra-observer and 37 inter-observer studies were described). Where the word ‘study/studies’ is used below this specifically pertains to the 68 individual studies reported within the 46 articles/papers.

5.3.3 STUDY CHARACTERISTICS

5.3.3.1 Reliability studies

Of the 46 articles describing reliability studies, 19 articles reported both intra- and inter-observer studies within the same paper, whereas 10 articles reported intra-observer (IaO) and 17 articles reported inter-observer (IeO) studies solely. Articles were published from 17 different countries; most frequently from the USA (12 articles). The majority of the articles were published after the year 2000 (n=32). Only two eligible articles were published before 1990. Table 21 displays the articles that have been categorised according to eleven different methods of measuring cervical ROM; grouping 15 different instruments. A brief description of the various instruments and their strengths and limitations is offered below.

Digital inclinometry in the form of the Cybex EDI-320 device was investigated and described in four papers [211-214]. The EDI-320 device consists of a hand-held, gravity dependent unit and a portable display unit which is able to record 360° of gross movement. It calculates the differential ROM between initial position reading and final position reading. Although clinically easy to use, this instrument requires accurate location of anatomical landmarks increasing the possibility of inter-observer differences.

Electromagnetic motion analysis was utilised in the form of two devices, the FASTRAK and Flock-of-Birds that work by tracking position of sensors electromagnetically relative to a source transmitter. Each sensor can measure data in three planes of joint motion collecting range of motion and speed over time. The disadvantage of these systems is the relative

expense and lack of portability and the need for substantial calibration procedures. Four articles reported investigation of the FASTRAK system [187, 215-217] and two of the Flock-of-Birds system [218, 219].

Goniometric methodology was described in seven articles. A universal goniometer (a 360° protractor with two arms) or modified version was in five papers [186, 220-223]. The lack of available landmarks and inability to measure the changing axis of rotation is a significant flaw to the use of these devices. The Spin-T goniometer attempts to negate this fact by using a wall as the reference point and consists of "a spectacle-type aluminium frame… with three 360° dials lying in orthogonal planes reflecting the cardinal movement planes of the cervical spine" [224]. Its reliability has been evaluated in two articles [224, 225].

Gravity-plus-compass goniometry was tested using three different devices. The CROM device has been investigated most frequently for reliability (nine articles) [152, 186, 212, 226-231]. The CROM is a spectacle-type plastic frame with two gravity goniometers and a compass goniometer. A magnetic yoke is placed over the shoulders to minimise effects of thoracic rotation as a substitution. The CMS system is a similar device with the addition of two spirit levels to assist with neutral placement of the head, but without the magnetic yoke. This was described in one article [226].

Investigation of the Myrin goniometer was also described in one article [232] and is almost identical to the CROM device, minus the magnetic yoke. The limitations of these devices are that there may be some effect on readings of rotation when lateral tilt occurs and vice versa and also that it is not possible to measure any other joints or parts of the body.

Inclinometers or gravity-dependent goniometers use the effect of gravity on pointers or fluid levels to measure joint position and motion. Usually the devices have a rotating dial so that

the scale can be zeroed with the pointer or bubble in the starting position. Because of the dependence on gravity, participants are usually in a seated position for measurements of Flexion-Extension and Lateral Flexion and lying supine (on back) for rotation, which could be seen as a disadvantage. Reliability studies of inclinometry were reported in four articles [223, 231, 233, 234].

The Elite optical motion analysis system, a TV image processing system, supplies three-dimensional co-ordinates of markers stuck to specified landmarks on the participant's body. In both papers that described this technology [136, 235] cameras were placed behind and above the seated participants, with six markers placed on head and trunk. Opto-electronic systems such as this are limited in that they require a dedicated space and often complex calibration and analysis procedures.

The OSI CA-6000 Spine Motion Analyser is a linkage device with six potentiometers that are connected by a series of bars, a headpiece and shoulder straps. Five IaO and three IeO studies are described in five articles [135, 147, 150, 236, 237]. There is some question over the fixation of the series of bars, with investigators finding that they bind on themselves during extreme saggital movements.

The Multi Cervical Rehabilitation Unit was evaluated in one IaO study [238] and consists of a fixed armchair with lumbar support, armrests and a shoulder restraint system. A head brace is suspended from above which contains a potentiometer which is connected to a PC. This machine requires a large space and is clearly a more expensive option than most devices.

Tape measure was appraised in four studies (two IaO and two IeO) documented in two papers [221, 239]. The tape is used to measure the distance between two landmarks e.g. tip of nose to acromio-clavicular joint or suprasternal notch to tragus for rotation). Only rotation

was measured in these studies. Some authors have postulated that the validity of measurements with a tape measure may be affected by individual's biometric characteristics (e.g. neck size [240].

Zebris Ultrasound system consists of a helmet and shoulder cap each fitted with three ultrasound microphones which receive signals from a transmitter located in a measuring unit on a stand approximately one meter to the right of the participant. The transmitter sends continuous pulses which are interpreted according to the timing of the interval between their emission and reception, providing three-dimensional co-ordinates. This system requires accurate calibration and as a result is not portable to different sites. This technology was investigated in five articles [148, 149, 157, 232, 237].

Visual estimation to judge whether a movement is limited or normal was appraised in six papers (all IeO studies) [186, 203, 241-245]. No reliability study described visual estimation of degrees of ROM. The obvious limitation of this method is there are no reference points and as such can be subject to expectation bias.

One study describes and evaluates the reliability of a miscellaneous method using a combination of protractor, goniometer and tape measure [246].

32 studies involved asymptomatic individuals, 16 involved symptomatic, nine involved both, seven involved a mixed population and two did not report the type of subjects involved. There were nine categories of sources of participants with the most commonly reported being university staff and students (17 studies) followed by secondary care patients (14). A significant number of studies (16) did not have the source of their participants reported. The mean sample size was 30 subjects with a considerable range (3-100) and variance (Standard Deviation 20).

Types of observers were varied; the most commonly reported were allied health professionals (31 studies). Other types were mixed (8 studies), university research staff (1 study), students of health professions (7 studies) and medical doctors (1 study). 18 studies did not have the type of observers documented.

51 studies investigated active ROM solely, whilst only eight investigated passive ROM solely.

Seven studies investigated both active and passive ROM. The interval between tests/ers varied considerably from immediately consecutive measurements to greater than two weeks. The most common interval was consecutive measurement (15 studies) although a number of studies (7) did not report the duration of interval.

The majority of studies (53/66; 80%) were deemed to have used a standardised measurement protocol in sufficient detail to allow replication. 13 studies did not have a standardised protocol reported and two studies had no record at all of how measurements were conducted. A seated measuring position was by far the most commonly reported (47 studies), seven studies used a mixture of supine (lying on back) and seated and ten studies did not report the position used. 25 studies reported using a warm-up procedure with three, four or five repetitions of movements. 41 studies did not report using a warm-up procedure. A small number of studies (6/66; 9%) were reported to have had a sample size calculation conducted. 59 studies used the ICC statistic and six used a Kappa statistic. 23 studies reported a confidence interval with the chosen statistical test.

Table 21 - Study characteristics for reliability studies

Method	Device	First Author (ref)	Study Type	Epoch	Stats method	Results by direction/plane									Level of reliability	No. of +ve QA items /13
						F	E	F-E	RR	LR	R	RLF	LLF	LF		
Digital inclinometry	Cybex EDI-320	Hoving[211]	IaO	A	ICC (95%CI)			0.96 (0.93-0.98)			0.96 (0.91 - 0.98)			0.93 (0.86 - 0.97)	Good	11
		Hoving[211]	IaO	B	ICC (95%CI)			0.97 (0..93 -0.86)			0.93 (0.86 - 0.96)			0.96 (0.92 - 0.98)	Good	11
		Love[212]	IaO		ICC (95% CI range)	0.92 (0.11)	0.91 (0.12)								Good	7
		Tousignant[213]	IaO	t1	ICC (95%CI)	0.77 (0.62-0.87)	0.79 (0.65-0.88)								Moderate	9
		Tousignant[213]	IaO	t2	ICC (95%CI)	0.77 (0.58-0.87)	0.83 (0.63-0.92								Moderate	9
		Zwart[214]	IaO		ICC			0.78			0.94			0.69	Moderate	1
		Hoving[211]	IeO		ICC (95%CI)			0.95 (0.90-0.98)			0.95 (0.90 - 0.98)			0.89 (0.77 - 0.94)	Good	11
		Love[212]	IeO		ICC	0.89 (0.13)	0.80 (0.18)								Moderate	7

		Tousignant,b	IeO	d1	ICC (95%CI)	F 0.66 (0.24-0.84)	0.66 (0.46-0.81)								Good	9
		Tousignant,b	IeO	d2	ICC (95%CI)	0.73 (0.53-0.85)	0.80 (0.64-0.89)								Moderate	9
Electro-magnetic motion analysis	FASTRAK	Amiri	IaO	d1	ICC				0.93	0.92					Good	5
		Amiri	IaO	d2	ICC				0.92	0.9					Good	5
		Jordan , a	IaO		ICC (1-sided CI)	0.64 (0.48)	0.70 (0.55)	0.82 (0.71	0.63 (0.47)	0.54 (0.37)	0.79 (0.68)	0.76 (0.62)	0.61 (0.42)	0.76 (0.6)	Moderate	13
		Jordan , b	IaO		ICC	0.91	0.84	0.95	0.95	0.94	0.98	0.95	0.94	0.97	Good	8
		Sterling	IaO	asymp	ICC	0.67	0.81		0.89	0.93		0.73	0.88		Moderate	6
		Sterling	IaO	symp	ICC	0.64	0.83		0.79	0.66		0.88	0.65		Moderate	6
		Jordan , a	IeO		ICC (1-sided CI)	0.74 (0.59)	0.78 (0.66)	0.89 (0.82)	0.70 (0.54)	0.80 (0.68)	0.85 (0.76	0.64 (0.45)	0.61 (0.42)	0.81 (0.7)	Moderate	13
	Flock of Birds	Morphett	IaO	t1	ICC (95%CI)			0.96 (0.89-0.99)			0.97 (0.63-0.99)			0.94 (0.84-0.98)	Good	6
		Morphett	IaO	t2	ICC (95%CI)			0.96 (0.88-0.99)			0.96 (0.88-0.99			0.95 (0.84-0.99)	Good	6
		Assink	IeO	asymp arom	ICC (95%CI)			0.77 (0.57-0.89)			0.85 (0.71-0.93)			0.79 (0.61-0.89)	Moderate	9
		Assink	IeO	asymp prom	ICC (95%CI)			0.75 (0.53-0.88)			0.77 (0.57-0.88)			0.73 (0.51-0.86)	Moderate	9

		Assink	IeO	symp arom	ICC (95%CI)			0.70 (0.56-0.88)			0.91 (0.83-0.96)			0.77 (0.58-0.88)	Moderate	9
		Assink	IeO	symp prom	ICC (95%CI)			0.72 (0.49-0.85)			0.36 (0.02-0.63)			0.82 (0.66-0.91)	Poor	9
		Morphett	IeO		ICC (95%CI)			0.78 (0.44-0.92)			0.94 (0.75-0.98)			0.80 (0.0-0.95)	Good	6
Goniometry	Gravity goniometer and universal goniometer combination	Cleland	IeO		ICC (95%CI)	0.75 (.50-.89)	0.74 (.48-.88)		0.78 (.55-.90)	0.77 (.52-.90)		0.66 (.33-.84)	0.69 (.40-.86)		Moderate	9
	Modified universal goniometer	Pellecchia	IaO		ICC							0.94	0.91		Good	6
		Pellecchia	IeO		ICC							0.86	0.65		Moderate	6
	Spin-T goiniometer	Agarwal	IaO		ICC (95%CI)	0.98	0.98		0.98	0.98		0.96	0.96		Good	4
		Haynes	IaO	t1	ICC	0.91 (0.8-0.96)	0.96 (0.91-0.98)		0.94 (0.86-0.97)	0.97 (0.91-0.98		0.87 (0.72-0.94)	0.87 (0.74-0.95)		Good	10
		Haynes	IaO	t2	ICC	0.95 (0.89-0.98)	0.91 (0.8-0.96)		0.96 (0.91-0.98)	0.97 (0.83-0.99)		0.98 (0.95-0.99)	0.98 (0.93-0.99),		Good	10
		Haynes	IeO		ICC	0.95 (0.89-0.98)	0.91 (0.8-0.96)		0.96 (0.91-0.98)	0.96 (0.89-0.98)		0.98 (0.95-0.99)	0.82 (0.62-0.92)		Good	10
	Universal goniometer	Maksymowych	IaO	t1	ICC						0.98				Good	8
		Maksymowych	IaO	t2	ICC						0.97				Good	8

		Youdas , a	IaO		ICC	0.83	0.86		0.9	0.78		0.85	0.84		Moderate	6
		Maksymowych	IeO								0.95				Good	8
		Tucci	IeO		ICC	0.232	0.822		0.522	0.604		0.211	0.337		Poor	4
		Youdas , a	IeO		ICC	0.57	0.79		0.62	0.54		0.72	0.79		Moderate	6
Gravity-plus-compass goniometry	CMS	Peolsson	IaO	t1	ICC			0.9			0.76			0.64	Moderate	7
		Peolsson	IaO	t2	ICC			0.88			0.71			0.64	Moderate	7
		Peolsson	IeO		ICC			0.89			0.82			0.61	Moderate	7
	CROM	Hole	IaO		ICC			0.96	0.92	0.92		0.96	0.96		Good	7
		Love	IaO		ICC (95%CI range)	0.97 (0.07)	0.98 (0.06)								Good	7
		Olson	IaO		ICC	0.58	0.97		0.96	0.98		0.96	0.94		Good	3
		Peolsson	IaO	t1	ICC			0.89			0.93			0.88	Good	7
		Peolsson	IaO	t2	ICC			0.91			0.87			0.9	Good	7
		Youdas , b	IaO		ICC	0.83	0.9		0.82	0.66		0.87	0.89		Moderate	5
		Youdas , a	IaO		ICC	0.95	0.9		0.93	0.9		0.92	0.84		Good	6
		Hole	IeO		ICC			0.88	0.94	0.9		0.82	0.86		Good	7
		Lee	IeO		ICC	0.84 (0.72-0.91)	0.81 (0.67-0.89)		0.74 (0.56-0.85)	0.76 (0.59-0.86)		0.81 (0.66-0.89)	0.81 (0.68-0.9)		Moderate	5
		Love	IeO		ICC	0.96 (0.08)	0.97 (0.07)								Good	7
		Nilsson	IeO		ICC	0.65	0.54	0.6	0.41	0.64	0.88	0.64	0.38	0.69	Moderate	4
		Olson	IeO		ICC	0.88	0.99		0.99	0.97		0.98	0.98		Good	3

		Peolsson	IeO		ICC			0.9			0.75			0.9	Good	7
		Rheault	IeO		ICC	0.76	0.98		0.81	0.82		0.87	0.86		Good	3
		Youdas , b	IeO		ICC	0.76	0.94		0.8	0.84		0.85	0.86		Good	5
		Youdas , a	IeO		ICC	0.86	0.86		0.92	0.82		0.88	0.73		Good	6
	Myrin	Malmstrom	IaO		ICC	0.83	0.89	0.95	0.85	0.69	0.93	0.77	0.85	0.9	Good	4
Inclinometry	double inclinometry	Bush	IeO		ICC	0.89	0.93		nr	nr		0.92	0.92		Good	3
	single inclinometer	Hole	IaO		ICC			0.94	0.93	0.84		0.94	0.88		Good	7
		Bush	IeO		ICC	0.92	0.91		0.91	0.91		0.93	0.92		Good	3
		Hole	IeO		ICC			0.84	0.76	0.86		0.82	0.81		Good	7
		Tucci	IeO		ICC	0.839	0.862		0.8	0.911		0.867	0.824		Good	4
	stabilisation inclinometry	Bush	IeO		ICC	0.93	0.89		nr	nr		0.93	0.94		Good	3
Misc	Protractor, goniometer and tape measure combination	Pile	IeO		Coeff of rel	0.21	0.59		0.9	0.84		0.74	0.68		Moderate	4
Optical Motion Analysis	Elite system	Bulgheroni	IaO		ICC	0.92	0.74		0.92	0.95		0.83	0.92		Good	3
		Antonaci	IeO		ICC				0.77	0.78	0.68	0.73	0.47	0.68	Moderate	6
Potentiometry	CA-6000 SMA	Christensen	IaO	t1 arom	ICC	0.93	0.89	0.94	0.94	0.81	0.88	0.9	0.92	0.91	Good	11
		Christensen	IaO	t2 arom	ICC	0.97	0.94	0.94	0.96	0.88	0.95	0.9	0.92	0.9	Good	11
		Christensen	IaO	t1 prom	ICC	0.93	0.92	0.95	0.93	0.95	0.97	0.97	0.96	0.98	Good	11
		Christensen	IaO	t2 prom	ICC	0.9	0.85	0.95	0.97	0.92	0.97	0.94	0.95	0.97	Good	11
		Lantz	IaO	t1 arom	ICC			0.76			0.85			0.89	Moderate	3
		Lantz	IaO	t1	ICC			0.59			0.64			0.87	Moderate	3

				prom												
		Lantz	IaO	t2 arom	ICC			0.9			0.97			0.81	Good	3
		Lantz	IaO	t2 prom	ICC			0.83			0.95			0.73	Moderate	3
		Mannion	IaO		ICC	0.85	0.75	0.82			0.92			0.89	Moderate	5
		Petersen,a	IaO	t1 asymp	ICC	0.78	0.89		0.92	0.85		0.93	0.86		Good	8
		Petersen ,a	IaO	t2 asymp	ICC	0.89	0.82		0.94	0.91		0.81	0.94		Good	8
		Petersen ,a	IaO	symp	ICC	0.68	0.87		0.94	0.88		0.92	0.96		Good	8
		Petersen , b	IaO		ICC	0.995	0.995					0.984	0.987		Good	5
		Christensen	IeO	d1 arom	ICC	0.98	0.88	0.96	0.88	0.76	0.83	0.78	0.9	0.83	Good	11
		Christensen	IeO	d2 arom	ICC	0.97	0.98	0.97	0.93	0.93	0.97	0.94	0.97	0.97	Good	11
		Christensen	IeO	d1 prom	ICC	0.94	0.78	0.91	0.88	0.91	0.91	0.8	0.65	0.72	Good	11
		Christensen	IeO	d2 prom	ICC	0.84	0.8	0.92	0.9	0.94	0.95	0.84	0.69	0.72	Good	11
		Lantz	IeO	arom	ICC			0.89			0.91			0.84	Good	3
		Lantz	IeO	prom	ICC			0.86			0.74			0.8	Good	3
		Petersen	IeO	asymp	ICC	0.89	0.88		0.94	0.92		0.91	0.93		Good	8
	Multi Cervical Rehabilitation Unit	Chiu	IaO	asymp	ICC (95%CI)	0.81 (0.59,0.93)	0.94 (0.7,0.99)		0.85 (0.71.0.92)	0.82 (0.65,0.92)		0.93 (0.88,0.95)	0.96 (0.93,0.97)		Good	7
		Chiu	IaO	symp	ICC (95%CI)	0.96 (0.89,0.98)	0.95 (0.88,0.98)		0.87 (0.76,0.95)	0.90 (0.82,0.95).		0.91 (0.85,0.95)	0.82 (0.66,0.92),		Good	7
Tape measure	Tape measure	Haywood	IaO		ICC (95%CI)				0.88 (0.75-.94)	0.79 (0.58-.90)					Moderate	7
		Maksymowych	IaO	t1	ICC						0.8				Moderate	8
		Maksymowych	IaO	t2	ICC						0.89				Good	8

		Haywood	IeO		ICC (95%CI)				0.68 (0.50-.80)	0.65 (0.25-0.82)					Moderate	7
		Maksymowych	IeO								0.82				Good	8
Ultrasound motion analysis	Zebris system	Cagnie	IaO		ICC	0.77	0.81	0.87	0.7	0.62	0.8	0.73	0.85	0.84	Moderate	5
		Dvir	IaO	degen	ICC	0.82	0.8		0.89	0.8		0.83	0.8		Moderate	6
		Dvir	IaO	WAD	ICC	0.82	0.82		0.86	0.85		0.86	0.81		Moderate	6
		Malmstrom	IaO		ICC	0.82	0.91	0.96	0.76	0.8	0.94	0.83	0.83	0.93	Good	4
		Mannion	IaO		ICC	0.88	0.78	0.86			0.93			0.92	Good	5
		Strimpakos	IaO	Oe sit arom	ICC			0.9			0.76			0.87	Moderate	9
		Strimpakos	IaO	Oe St arom	ICC			0.86			0.73			0.83	Moderate	9
		Strimpakos	IaO	Ce sit arom	ICC			0.86			0.77			0.87	Moderate	9
		Strimpakos	IaO	Ce st arom	ICC			0.87			0.75			0.87	Moderate	9
		Strimpakos	IaO	Oe sit prom	ICC			0.93			0.83			0.89	Good	9
		Strimpakos	IaO	Oe st prom	ICC			0.95			0.84			0.9	Good	9
		Cagnie	IeO		ICC	0.84	0.92	0.94	0.5	0.58	0.92	0.9	0.79	0.92	Good	5
		Strimpakos	IeO		ICC			0.43			0.57			0.68	Poor	9
Visual Estimation	Judgement of normal vs abnormal	Bertilson	IeO	with K	Kappa (SD)	nr	0.42 (0.16)		0.16 (0.12)	0.39 (0.11)		0.16 (0.16)	0.31 (0.14)		Poor	12
		Bertilson	IeO	Without K	Kappa (SD)	nr	0.15 (0.15)		0.2 (0.14)	0.18 (0.14)		0.45 (0.15)	0.37 (0.14)		Poor	12

Judgement of normal vs limited vs markedly limited	Viikari-Juntura	IeO		W Kappa	0.43	0.56	0.56	0.4	0.51	0.41	Poor	5
Judgement of reduced vs normal vs increased	Fjellner	IeO		W Kappa (95%CI)	0.26 (-.03;0. 55)	0.58 (0.34;0.82)	0.6 (0.38;0.82)	0.66 (0.42;0.9)	0.6 (0.36;0.84)	0.52 (0.27;0.77)	Poor	8
Judgement of normal vs restricted movement	Hoppenbrouwers	IeO	arom	Kappa	0.57	0.88	0.54	0.43	0.36	0.33	Poor	12
	Hoppenbrouwers	IeO	prom	Kappa	0.77	0.85	0.54	0.47	0.43	0.23	Poor	12
	Pool	IeO		Kappa	0.19	0.39	0.25	0.61	0.38	0.05	Poor	8
Judgement of restricted vs non-restricted	Van Suijlekom	IeO		Kappa	0.27	0.28	0.44	0.46			Poor	6
Visual Estimation	Youdas , a	IeO		ICC	0.42	0.42	0.82	0.69	0.7	0.63	Moderate	6

Abbreviations: nr = not reported, **Study type:** IaO = Intra-observer study, IeO = Inter-observer study, **Population:** Asymp = Asymptomatic, Symp = Symptomatic, Mix = Mixed population of asymp and symp, Both = groups of asymp and symp evaluated, **Source of subjects**: pcp = primary care patients, pcs = primary care staff, scp = secondary care patients, scs = secondary care staff, uss = higher education/university staff/students, mix = mixed sources, pub = public, rct = randomised controlled trial or other research study, **Type of observer:** ahp = allied health professional, doc = medical doctor, stu = student, mix = mix of types of observers, **Type of Movement:** AROM = Active Range of Cervical Movement, PROM = Passive Range of Cervical Movement, **Interval:** consec = consecutively, **Movements:** F = flexion, E = extension, RR = right rotation, LR = left rotation, RLF = right lateral flexion, LLF = left lateral flexion, F-E: Flexion-Extension, R = rotation, LF = lateral flexion, F-RR = flexion with right rotation, F-LR = flexion with left rotation ER = extension with rotation RF = flexion with rotation, **Statistical methods:** ICC = intra-class correlation coefficient, ICC (CI) = intra-class correlation coefficient with 95% confidence interval, Kappa = Kappa, Kappa (CI) = Kappa with 95% confidence interval, wKappa = weighted Kappa, LoA = Limits of Agreement technique, SEM = Standard Error of Measurement, MDC = Minimal Detectable Change, SDD = Smallest Detectable Difference.

5.3.3.2 Validity studies

21 articles reported concurrent validity studies and the data abstracted from them are presented in Table 22 below. Eight methods were assessed within these studies incorporating 13 different instruments. Gravity-plus-compass goniometry was the most frequently investigated (5 studies); most commonly in the form of the CROM device (4 papers). The Zebris device was also investigated in four studies for validity with a number of different reference devices.

Radiographic imaging was most often used as the reference or index device (9 studies). Nine other devices were used as references: gravity-plus-compass goniometer, optical motion capture system, electrogoniometer, digital inclinometer, electromagnetic motion capture system, single inclinometer, tape measure and CT imaging.

Similarly to the findings for reliability studies, the most frequent country of origin was the USA (8 papers) and year of publication ranged from 1986 up to 2007 with the majority of articles being published post 2000 (15/21; 71%). 12 studies used an asymptomatic population compared to just two that used symptomatic participants. One study performed experiments on separate asymptomatic and symptomatic groups compared to three studies that used a mixed population. The type of population was unknown in three studies. Sample sizes ranged from three to 105 participants (mean 28, SD 25.8). 20 of the 21 papers did not report using a sample size calculation. The one study that did report conducting a calculation did this in order to be able to detect a difference in ROM between symptomatic and asymptomatic subjects but also stated this would be "sufficient to …compare the Fastrak data to the other assessment tools used." [214]

The most common observers in these studies were Allied Health Professionals (11 studies). Eight papers did not report who the observers were. 20 studies investigated the measurement of active ROM compared to just one study that assessed passive ROM. Reporting of the interval between studies ranged from simultaneous (eight studies) to “within 10 days”. The majority of papers reported a standardised measurement protocol (17), with 14 of them describing a warm-up procedure and the same number using a seated position to conduct the measurements. Three studies used a standing position and one study conducting inclinometry reported the use of a combination of seated and supine measures for different planes of movement. A variety of statistical methods were used to estimate the validity of the measurement devices, with some studies utilising multiple methods.

Table 22 - Study characteristics for Validity studies

Experimental Method	Experimental Device	Index/Ref Device	First Author [ref]	Stats method	Epoch	Results by direction/plane									Level of validity	No. of +ve QA items /13
						F	E	F-E	RR	LR	R	RLF	LLF	LF		
Digital inclinometry	Cybex EDI-320	X-ray	Mayer[247]	Pearson's r		**0.99**									Good	2
	Digital and bubble dual inclinometry	X-ray	Wolfenberger[240]	nr												6
	WASP system	Optical Motion System	Syed[248]	Pearson's r (SD)				**0.98 (0.03)**			**0.93 (0.10)**			**0.92 (0.19)**	Good	4
Electro-goniometry	Electrogoniometer	Gravity and compass goniometers	Alund[204]	Pearson's r	d1			**0.99**			**0.63**			**0.92**	Good	5
		Gravity and compass goniometers	Alund[204]	Pearson's r	d2 (1 wk later)			**0.93**			**0.05**			**0.69**	Good	5
		X-ray	Alund[204]	Pearson's r				**0.85**						**0.76**	Good	5
Electro-magnetic motion analysis	FASTRAK	Tape Measure	Jordan[216]	Association					**0.9**	**0.89**					Good	9
	Flock-of-Birds	CROM	Morphett[219]	ICC (95%) CI)				**0.94 (0.76-0.99)**			**0.91 (0.62-0.98)**			**0.78 (0.33-0.94)**	Good	5
Goniometry	Pendulum goniometer	X-ray	Hermann[249]	Pearson's r (ICC)				**0.975 (0.98)**							Good	7
	Spin-T goniometer	MotionStar	Agarwal[250]	Pearson's r				**0.999**			**1**			**0.998**	Good	3

Gravity-plus-compass goniometry	CMS	CROM	Peolsson	Pearson's r (ICC)				**0.92-0.95 (0.92-0.94)**			**0.70-0.85 (0.46-0.69)**			**0.73-0.82 (0.50-0.58)**	Good	8
	CROM	X-ray	Tousignant, a	Pearson's r		**0.97**	**0.98**								Good	8
		X-ray	Tousignant, b	Pearson's r (95% CI)								**0.84 (0.66-0.93)**	**0.82 (0.62-0.92)**		Good	9
		Double Inclinometer	Hole	ICC				**0.8**	**-0.23**	**-0.12**		**0.8**	**0.78**		Moderate	7
		OPTO-TRAK	Tousignant, c	Pearson's r (95%CI)		**0.98 (0.97-0.99)**	**0.99 (0.98-0.99)**		**0.89 (0.81-0.94)**	**0.94 (0.90-0.97)**		**0.91 (0.85-0.95)**	**0.89 (0.82-0.93)**		Good	11
Inclinometry	Gravity inclinometer	Universal goniometer	Tucci	ICC		**0.673,**	**0.907**		**0.49**	**0.378**		**0.8**	**0.784**		Good	6
	Single, double and stabilis'n inclinom'rs	X-ray and CT	Bush	nr												3
Potentiometry	CA-6000 SMA	Dualer & Protractor	Lantz	ICC				**0.965**			**0.999**			**0.937.**	Good	4
		Dualer & Protractor	Lantz	ICC	d2 (1 wk later)			**0.972**						**0.955**	Good	4

		X-ray	Petersen	B&A plots	**36° b'tw'n U & L LoA.**	**37° b'tw'n U & L LoA.**								Poor	5
Ultrasound motion analysis	Zebris system	Myrin	Malmstrom	ICC	**0.9**	**0.92**	**0.96**	**0.78**	**0.81**	**0.94**	**0.82**	**0.85**	**0.93**	Good	6
		CROM device	Wang	Adjusted R^2	**0.66**	**0.88**					**0.86**	**0.87**		Good	0
		X-ray	Strimpakos	ICC	**0.88**	**0.95**								Good	7
		CA-6000 SMA	Mannion	Pearson's r	**0.99**	**0.92**	**0.97**			**0.97**			**0.95**	Good	8

Abbreviations: Nr = not reported, Study type: IaO = Intra-observer study, IeO = Inter-observer study, **Population:** Asymp = Asymptomatic, Symp = Symptomatic, Mix = Mixed population of asymp and symp, Both = groups of asymp and symp evaluated, **Source of subjects**: pcp = primary care patients, pcs = primary care staff, scp = secondary care patients, scs = secondary care staff, uss = higher education/university staff/students, mix = mixed sources, pub = public, rct = randomised controlled trial or other research study, **Type of observer:** ahp = allied health professional, doc = medical doctor, stu = student, mix = mix of types of observers, **Type of Movement:** AROM = Active Range of Cervical Movement, PROM = Passive Range of Cervical Movement, **Interval:** consec = consecutively, **Movements:** F = flexion, E = extension, RR = right rotation, LR = left rotation, RLF = right lateral flexion, LLF = left lateral flexion, F-E: Flexion-Extension, R = rotation, LF = lateral flexion, F-RR = flexion with right rotation, F-LR = flexion with left rotation ER = extension with rotation RF = flexion with rotation, **Statistical methods:** Pearson's r = Pearson's correlation coefficient r, ICC = intra-class correlation coefficient, ICC (CI) = intra-class correlation coefficient with 95% confidence interval, Kappa = Kappa, Kappa (CI) = Kappa with 95% confidence interval, wKappa = weighted Kappa, LoA = Limits of Agreement technique, SEM = Standard Error of Measurement, MDC = Minimal Detectable Change, SDD = Smallest Detectable Difference.

5.3.4 QUALITY ASSESSMENT

5.3.4.1 Reliability studies

The Kappa coefficient for quality assessment of reliability studies was calculated using the data presented in Table 23 and an online application [209].

Table 23 - Agreement on quality assessment scoring for reliability studies

		Reviewer B		**Total**
		Item Yes	Item No / Unsure	
Reviewer A	Item Yes	244	44	288
	Item No / Unsure	80	230	310
	Total	324	274	598

The resulting Kappa coefficient was 0.59 (95% CI 0.52-0.65) and Agreement 79%, SE 0.04. This indicates there was a moderate strength of agreement between reviewers for use of the quality assessment tool for reliability studies [210]. The prevalence index was calculated as 0.02 indicating a very low index in favour of obtaining a 'Yes' decision. The bias index was calculated as -0.06 indicating a negligible bias between reviewers. All disagreements were resolved through discussion.

Appendix 8 displays the individual scores for the Quality assessment for all reliability studies included in the review. The overall rating is presented in the last column of Table 21. The number of positively scored items for all studies ranged from zero to 13 (mean 6.6, SD 2.7). The mean number of positively scored items for the individual methods ranged from 3.7 (inclinometry) to 8.7 (visual estimation). The most common failings were lack of reporting of a sample size calculation (only 5 studies scored positively) and failure to

describe withdrawals for the study (only 8 studies scored positively). It was common for studies to adequately describe the study population (36 studies scored positively), describe a standardised measurement protocol that would enable replication (41 studies scored positively) and also describe the interval between measurements satisfactorily (36 studies scored positively).

5.3.4.2 Validity studies

The Kappa coefficient for assessment of validity studies was calculated using the data presented in Table 24 and the previously referenced 'Kappa calculator'.

Table 24 - Agreement on quality assessment scoring for validity studies

		Reviewer B		**Total**
		Item Yes	Item No / Unsure	
Reviewer A	Item Yes	94	49	143
	Item No / Unsure	31	162	193
	Total	125	211	336

The resulting Kappa coefficient was 0.51 (95% CI 0.41-0.6) and Agreement 76%, SE 0.05. This indicates there was also a moderate strength of agreement between reviewers for quality assessment of validity studies (Landis and Koch, 1977[210]).

The prevalence index was calculated as -0.20 indicating a low index in favour of obtaining a 'No' decision. The bias index was calculated as 0.05 indicating a very low bias between reviewers. Again all disagreements were resolved through discussion.

Appendix 9 presents the individual scores for the Quality assessment for all validity studies included in the review. Table 22 presents the total rating as a summary. The range of positively rated items ranged from zero to 11 (mean 5.9, SD 2.6). For individual methods, the mean number of positive items scored ranged from 4.3 (digital inclinometry) to 8.6 (gravity-plus-compass goniometry). The most common failings were a lack of sample size calculation (only one study scored positively), reporting a randomised order of testing and reporting whether observers were blind to the previous tests findings (both had only two studies score positively on this). Both reporting of a standardised measurement protocol and description of the reference/index device commonly scored positively (both had 18 out of 21 studies score positively).

5.3.5 SYNTHESIS OF RESULTS

It was judged that the studies under review were too heterogeneous to undertake appropriate meta-analysis. The variation in sample populations, types of observers and measurement protocols was wide, even within devices. Table 25 and Table 26 display these ratings by device.

Table 25 - Overall ratings of reliability for each device

	No. of studies with reliability rating:			
Device	**Good**	**Moderate**	**Poor**	**Overall Rating**
SpinT goniometer	3	0	0	Good
CROM	12	2	1	Good
Single Inclinometer	2	0	0	Good
CA-6000 SMA	7	0	0	Good
Multi Cervical Rehabilitation Unit	1	0	0	Good
Cybex EDI-320	4	3	0	Moderate
FASTRAK	2	3	0	Moderate
Flock of Birds	2	1	0	Moderate
Universal goniometer	2	3	1	Moderate
Modified universal goniometer	1	1	0	Moderate
CMS	0	2	0	Moderate
Myrin	0	1	0	Moderate
Double and stabilisation inclinometry	0	1	0	Moderate
Elite system	1	1	0	Moderate
Tape measure	2	2		Moderate
Zebris system	1	5	1	Moderate
Protractor, goniometer and tape measure combination	0	1	0	Moderate

Visual Estimation of normal vs. abnormal	0	1	6	Poor

Table 26 - Overall ratings of Validity for each device

	No. of studies with Validity rating:			
Experimental Device	**Good**	**Moderate**	**Poor**	**Overall Rating**
CMS	1	0	0	Good
CROM	3	1	0	Good
Cybex EDI-320	3	0	0	Good
Electrogoniometer	1	0	0	Good
FASTRAK	1	0	0	Good
Flock of Birds	1	0	0	Good
Inclinometer	1	0	0	Good
Pendulum goniometer	1	0	0	Good
SpinT goniometer	1	0	0	Good
WASP	1	0	0	Good
Zebris system	4	0	0	Good
CA-6000 SMA	1	0	1	Moderate

5.4 DISCUSSION

5.4.1 SUMMARY OF EVIDENCE

This review finds that a large number of devices are available to measure cervical spine ROM; some of which may be too expensive for everyday clinical use and may be more appropriate for purely research purposes. Despite identifying a great number of articles from the initial search, only a small proportion of these were included in this review.

This systematic review finds that devices deemed to have "good" reliability and validity were the Cervical Range of Motion Device (CROM), the Spin-T goniometer and the single inclinometer. The CROM device has been investigated most frequently and in both symptomatic and asymptomatic groups. One study investigating reliability of the CROM device for measuring passive ROM found a poor level of reliability, although methodological quality of this study was rated as low. Other studies investigating passive ROM did find that the CROM device showed 'good' reliability so further investigation of this appears to be warranted. The advantages of the CROM device are its portability, ease of fitting and relative ease of use. Its disadvantages include that it can only measure cervical spine ROM and no other joints which make it limited in a clinical environment and also it is more expensive than other inclinometer systems.

Both the Spin-T goniometer and the single inclinometer had a smaller number of reliability and validity studies conducted. The Spin-T goniometer's reliability and validity has not been assessed in a symptomatic population so generalisability is limited. The Spin-T goniometer has similar advantages to the CROM, although it does require proximity to a wall which may limit some use in some clinical spaces (e.g. curtained cubicles). It also requires two hands for

operation which may not always be convenient, especially if the clinician wishes to correct or guide the subject.

The inclinometer's reliability was tested in a mixture of populations, although the single validity study that resulted in its categorisation as 'good' did not report the type of population or observers used, again providing some questions as to how externally validity the study's findings are. It is clear that despite its practical strengths, visual estimation is the least reliable and valid method for measuring cervical ROM according to the studies assessed in this review, findings that are in concordance with previous reviews' [197, 198].

Methodological quality scores were varied for both reliability (range 0-13) and validity studies (range 0-11). The most common failing for both constructs was lack of consideration of sample size. For a thorough discussion of this topic see Jordan [197].

Contrary to previous reviews findings [198], methodological quality did not appear to have significantly improved as time has progressed.

Generalisability to clinical settings should be an important aim for studies in this review. A minority of reliability and validity studies achieved this aim in terms of reporting populations and observers investigated. 32 reliability studies reported investigating asymptomatic individuals compared to just 16 studies with symptomatic individuals. Even fewer validity studies investigated devices measuring symptomatic populations (12 used an asymptomatic population; 2 used a symptomatic population). The source of these populations may also affect generalisability and in the case of 16 reliability and 10 validity studies this was unknown. Similar poor reporting meant that the types of observers using the devices were not known in a significant number of reliability and validity studies (18 and 8 articles respectively).

Internal validity of the studies was brought into question in a number of ways. A very limited number of articles overtly stated whether or not there were any withdrawals or missing data (8 reliability and 6 validity articles) and no paper provided a flow chart of recruitment/testing which would provide useful information when attempting to assess sources of bias. 25 reliability articles and seven validity articles reported observers being blinded to previous test results leaving a considerable number of studies potentially being flawed. This is despite the awareness that, in some cases (especially with validity studies), blinding is not always necessary/appropriate if computers are recording readings simultaneously. Blinding is a fundamental method of negating bias in studies of this nature and this highlights the importance of sound reporting and there is still considerable room for improvement in this area of research. On a more positive note, the majority of studies did document a detailed description of the measurement protocol enabling replication and interpretation or at least provided a reference to an adequate source of this information. The effect of variation in protocols will be discussed in further research suggestions.

5.4.2 LIMITATIONS

The conclusions of this review are to be treated with caution due to a number of limitations that are common in systematic reviews of studies of this type. Significant sources of heterogeneity meant that meta-analysis was not appropriate. Considerable differences in studies were variations in the sample populations studied, significant variations in measurement protocols and the use of different statistical analysis techniques. The resulting technique of providing an ‘average’ level of reliability and validity uses an arbitrary (albeit previously used[198]) categorisation into ‘good’, ‘moderate’ and ‘poor’. These categories are independent of the judgement on study quality. By providing a mean score for the quality assessment of articles for each device the authors attempt to give the reader further

information to interpret the categories of reliability and validity, however it is still challenging to know which results warrant more weight than others.

There is no consistently used method for evaluating quality of studies of clinimetric properties, let alone a 'Gold standard'. Previous reviews had not included a quality assessment checklist until De Koning et al in 2008[198], which was published after this review was conducted. Their use of a trichotomous outcome (Adequate design, method and result vs. doubtful method used vs. no information) has advantages of simplicity in interpretation although what actually constitutes 'adequate' for design and method is considerably subjective. It was decided that categorising a study as adequate or inadequate or high or low quality based on the numerical value of criteria should be avoided as it was felt that it was not feasible to determine the relative weight of each quality criterion.

Studies that show significant reliability and/or validity of a method are more likely to be published leading to the possibility of publication bias. One could argue however, that the advantage of assessing published studies is that the peer-review process ensures a basic level of quality. Non-English language articles and grey literature were not included due to resource limitations. A small number of non-English language articles were retrieved from the initial search (three German, one Dutch, one Spanish and one Polish) and it is unclear from the abstracts how many would actually have been eligible for inclusion in the final review. It is possible that some Grey literature could have been retrieved in this review due to the limited use of Google Scholar, although in actuality this did not happen. Reviewer bias is also another possible limitation of this review as reviewers were un-blinded to authors of the studies.

5.4.3 RESEARCH IMPLICATIONS

Poor reporting of reliability and validity studies has been highlighted in previous sections as a significant factor in increasing the difficulty of conducting a review such as this, but also of interpretation by researchers and clinicians attempting to select an appropriate measurement tool. Numerous previous systematic reviews in related areas have called for studies of this nature to use an adaptation of the STARD checklist [201, 251, 252] and the results from this review require continued echoing of this sentiment. Reliability and validity studies are often inadequately indexed in electronic databases making it difficult to retrieve all published evidence. This appears to be supported by the result that 34 articles were located from sources other than the electronic search of the selected databases. It is recommended that MeSH headings be used to facilitate searches by researchers and clinicians. Two headings appear in the PubMed MeSH database; 'Reproducibility of results' and 'Validation Studies', the latter having been introduced only very recently in 2008.

Although radiographic measurement was most frequently used as the reference device for concurrent validation it should be noted that this should not be unconsciously accepted as the Gold Standard. This is because sufficient reliability and validity studies are still required, although this is unlikely to occur due to the risk associated with using x-rays. It may be more appropriate to conduct concurrent validation with multiple methods to give us a greater understanding of the validity of the device under examination.

There is a significant omission in the validation of the single inclinometer (which is recommended for use by the American Medical Association [127]) which should be rectified to provide further information alongside the satisfactory evidence of its reliability.

5.4.4 PREVIOUS RESEARCH

45 different articles were reviewed in the four previous reviews discussed in the introduction section. This study included an additional 28 studies. Three of these were published after the commencement of the last review. This leaves a considerable number of studies that had been missed by previous reviews or did not fulfil their alternative selection criteria. 18 studies from the previous reviews did not fulfil the selection criteria for inclusion in this review. The findings of this review concur with the most recent review by De Koning et al [198] in that the CROM device and inclinometer had the most favourable results for reliability and validity studies. In the only other 'systematic' review of reliability studies Jordan also concluded that the CROM appeared to be the most promising device. It is noted that although different methodology has been utilised in each of these reviews the conclusions have broadly been the same.

5.4.5 CLINICAL IMPLICATIONS

In an age when we are striving to practise Evidence-Based Medicine and required to justify effectiveness by a variety of healthcare stakeholders, it is vital that the fundamental process of assessment and reassessment of motion should be clinimetrically robust. Findings from this review indicate that visual estimation is not reliable enough as an assessment method and the use of the CROM device or single inclinometer is preferable. The advantages of the CROM device is that it allows measurement of both active and passive cervical spine movements in a sitting position with the clinician's hands free to assist unlike the inclinometer. The fixed nature of the position of the CROM may facilitate its apparent greater reliability over the inclinometer. The advantage of the inclinometer is it is considerably more portable (pocket sized) and affordable. In terms of how cervical ROM

should be assessed, it is unclear whether there is a difference in reliability and validity terms between active and passive ROM.

5.4.6 FURTHER RESEARCH

This review has posed a great deal more research questions than answers and a number of potential research areas have been highlighted. The studies in this review conducted the research on a limited number of symptomatic populations. Assessment of cervical ROM is conducted on a wide variety of patients and so it would be beneficial to assess reliability and validity in different conditions especially as clinimetric properties are only applicable to a certain population using a certain measurement protocol. Work on the effect of different measurement protocols on reliability of cervical ROM measurement has been limited so far (e.g. sitting vs. standing position [148, 150], different repetitions and time intervals [253] and eyes open vs. closed [254]) and continuation of this would be invaluable if a consensus is to be reached on what an optimised protocol should consist of. An example of variation in aspects of protocols is provided in tables 3 and 4 on the subject of whether a warm-up was used (46% and 33% respectively did in reliability and validity studies*)* – a feature that requires further examination. The Spin-T goniometer shows promising reliability and validity but requires work in symptomatic populations to really ascertain its value as an assessment tool.

5.5 CONCLUSIONS

This systematic review found a number of reliability and concurrent validity studies have been published on the subject of cervical spine ROM measurement. The CROM device has undergone significantly more investigation and has been shown to be clinimetrically sound.

No studies were found assessing the reliability instruments in a specific population of individuals with WAD.

5.6 CHAPTER SUMMARY

This chapter has described a systematic literature review of reliability and validity studies of methods for measuring cervical ROM.

This review has highlighted that the CROM device appears to be the most clinimetrically sound method for measuring cervical ROM at this time. However, a well conducted and reported study to assess the reliability of the device in a WAD population is absent. The following chapter will describe both intra- and inter-observer reliability studies for measuring cervical ROM in individuals with WAD using the CROM device.

6 CHAPTER SIX – INTRA- AND INTER-OBSERVER RELIABILITY STUDIES OF THE CROM DEVICE IN A WAD POPULATION

6.1 INTRODUCTION

In the preceding five chapters the challenge of managing Whiplash Associated Disorders (WAD) has been introduced and discussed. The importance of sound clinical assessment and in particular assessment of cervical spine Range of motion (ROM) has been highlighted as the focus of this thesis. Current evidence of cervical spine ROM's role as a prognostic factor has been reviewed and the conclusion drawn that further investigation of this measure is required. A fundamental aspect of the conduct of such a prognostic study and subsequent clinical use would require the use of a reliable and valid measurement tool. There is a belief among clinicians and researchers that cervical spine ROM is hard to measure reliably due to the nature of the structures and movements involved [101]. In the preceding chapter current evidence for reliability and validity of cervical spine ROM measurement tools was synthesised and it was concluded that, although there was no definitive tool, the Cervical Range of Motion (CROM) device appears to be the most promising. However, no adequate studies have been conducted on individuals with WAD. Clearly there is a great need for a well conducted population-specific reliability study to assess intra- and inter-observer reliability for measuring cervical spine ROM in a WAD population because clinimetric findings are highly population dependent [101].

In this chapter two studies are presented that evaluate the reliability of the CROM device in a WAD population. Initially a justification for the selection of this instrument is provided. As

far as the author is aware this is the first time the reliability of this device has been investigated for both active and passive ROM in individuals with WAD.

This doctoral work does not include a concurrent validity study for a number of reasons. Firstly it was not feasible within the constraints of the time allocated for a PhD; secondly because the study would need to use some form of medical imaging which would be either too expensive (Magnetic Resonance Imaging) or potentially harmful (X-ray) to administer for all planes of movement required within this project.

Methodological considerations are discussed. Results are presented in the form of Intra-class Correlation Coefficients (ICC), Standard Error of Measurement (SEM), and Limits of Agreement (LOA). The discussion focuses on the strengths and limitations of the two studies and the implications for the cohort study described and discussed in Chapter Seven.

6.2 OBJECTIVE

The objective of the two studies described was to determine the intra- and inter-observer reliability of the CROM device for measuring active and passive cervical ROM for individuals with sub-acute WAD.

This was to answer the research question: "How reliable is the CROM device for measuring active and passive cervical ROM in a sub-acute WAD population?"

6.3 JUSTIFICATION FOR SELECTION OF THE CERVICAL RANGE OF MOTION (CROM) DEVICE

The selection of the CROM device for the cohort study was a result of a decision based on a balance of numerous factors. These included validity, reliability, accuracy, cost effectiveness and appropriateness for the setting in which it was to be used. The device was to be used in a

number of research clinics by solitary observers and needed to be portable as financial and logistical constraints made it impossible to have one device stationary in every clinic, a scenario not uncommon in the UK NHS clinical setting.

There were multiple sites of assessment across the UK so it was necessary to purchase multiple units of the device, rendering cost as a secondary selection factor. The tool was required to be able to measure both active and passive cervical spine ROM with a single observer. Additionally, clinical inter-observer reliability was an important factor as multiple observers would collect data for the main cohort study.

Figure 15 below represents the cognitive "funnelling" processes which lead to the selection of the CROM device. Of the tools that fulfilled all four selection points the CROM device has been investigated most frequently for its reliability and validity and in a variety of populations, as discussed in Chapter Five. In the systematic review it was concluded that the CROM was deemed to have "good" reliability and validity.

Figure 15 - Selection process for cervical ROM measurement device

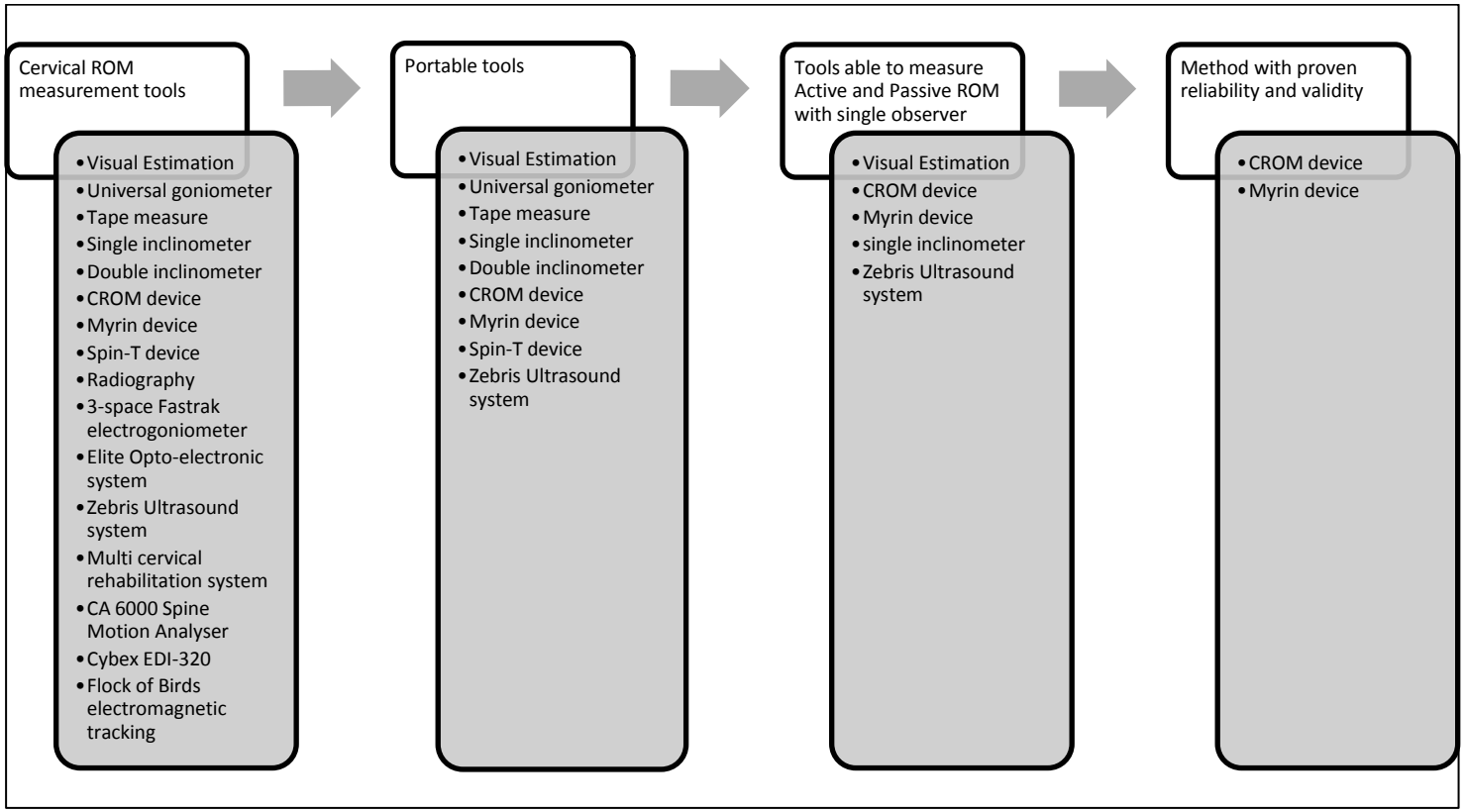


Table 27 and Table 28 display the studies from which the conclusion was drawn that the CROM device has "good" reliability. Studies that were excluded due to inappropriate statistical techniques are included in the table for completeness.

There have been 11 published reliability studies evaluating the CROM device, none evaluating solely intra-observer reliability, four evaluating solely inter-observer reliability and seven evaluating both intra and inter-observer reliability in the same cohort. The majority (73%) of these studies have used active rather than passive ROM (8 active: 3 passive) and no studies utilised both methods in the same cohort.

The most frequently studied type of subjects was asymptomatic (six studies) with a minority of studies investigating symptomatic participants (three studies). The reliability studies conducted thus far have generally been small scale with sample sizes ranging from 12 to 40 participants. Four concurrent validity studies have been conducted for the CROM device, all of which evaluated active ROM. One compared the CROM to a single inclinometer [231], two studies compared to radiography [205, 255] and the final study compared it to an optical motion analysis system [256]. Sample sizes ranged from 31 to 55 for a range of symptom populations.

Reviews by Jordan [197] and de Koning [198] agreed that the CROM device is the most promising in terms of reliability but that "further studies need to be performed …on subjects with specific neck pathologies" [198]. The CROM device is available commercially unlike other potential devices that could be evaluated in this study i.e. the Myrin device (See Chapter Five for description).

Table 27 - CROM device reproducibility studies

Author	Year	Population	Sample size	AROM/ PROM	Results
Intra-tester studies					
Capuano-Pucci	1991	Asymptomatic	20	AROM	Pearson's r = 0.63-0.91
Hole	1995	Asymptomatic	30	AROM	ICC 0.92-0.96
Nilsson	1995	Asymptomatic	14	PROM	Pearson's r = 0.61 - 0.85
Olson	2000	Symptomatic	12	AROM	ICC 0.88-0.99
Peolsson	2000	Asymptomatic	30	AROM	ICC 0.87-0.93
Youdas	1991	Symptomatic	20	AROM	ICC 0.84-0.95
Youdas	1992	Asymptomatic	30	AROM	Median ICC 0.76-0.94
Inter-tester studies					
Capuano-Pucci	1991	Asymptomatic	20	AROM	Pearson's r =0.74-0.87
Hole	1995	Asymptomatic	30	AROM	ICC 0.82 - 0.94
Lee	2003	Both	40	AROM	ICC 0.74 - 0.84
Love	1998	Both	27	PROM	ICC 0.96 - 0.99
Nilsson	1995	Asymptomatic	14	PROM	Pearson's r = 0.29 - 0.66
Nilsson	1996	Asymptomatic	35	PROM	ICC = 0.60 - 0.88
Olson	2000	Symptomatic	12	AROM	ICC 0.58 – 0.98
Peolsson	2000	Asymptomatic	30	AROM	ICC 0.75 - 0.90
Rheault	1992	Symptomatic	22	AROM	ICC 0.76 - 0.98

Youdas	1991	Symptomatic	20	AROM	ICC 0.73-0.92
Youdas	1992	Asymptomatic	20	AROM	ICC .66 - .90

Table 28 - CROM device validity studies

Author	**Year**	**Population**	**Sample size**	**AROM/PROM**	**Results**
Hole	1995	Asymptomatic	30	AROM	ICC= 0.12 – 0.80
Tousignant	2000	Asymptomatic	31	AROM (F-E)	Pearson's r = 0.97-0.98
Tousignant	2002	Symptomatic	24	AROM (LF)	Pearson's r = 0.82-0.84
Tousignant	2006	Mixed	55	AROM (F-E and R)	Pearson's r = 0.89-0.99

The CROM device consists of two gravity dependent goniometers and one compass dial on a head-mounted frame, allowing measurement of cervical spine ROM in three planes (See Figure 16). A magnetic yoke is supplied, which is rested over the front and back of the chest, to reduce the influence of thorax rotation. ROM is measured in two degree increments.

Figure 16 - CROM Device with magnetic yoke

6.4 METHODS

6.4.1 RELIABILITY STUDY DESIGN

In order to assess reliability, measures need to be repeated at least once, utilising a test-retest design. Often inter-observer studies contain multiple sources of error and therefore an argument can be made that it is unnecessary to carry out intra-observer studies if the inter-observer reliability is high. However if the reliability is found to be poor one cannot be sure of the source of the variation, this could be between or within (or both) observers. Another

justification for carrying out an intra-observer study first was to observe the stability of measuring a symptomatic patient group repeatedly.

In order to estimate observer effect variance it is logical to try and involve as many observers as possible, however very often in reliability studies, practicality limits the number of observers [257]. In the circumstances of this study the issue of patient tolerance was important.

Two studies were conducted to evaluate reliability for measurement of both active and passive ROM, firstly to assess intra-observer reliability (Study One) and secondly to assess inter-observer reliability (Study Two).

6.4.2 PARTICIPANTS

Participants were recruited as part of the multi-centre Randomised Controlled Trial 'Managing Injuries of the Neck Trial' (MINT) and were screened according to the selection criteria described below.

Potential participants were identified by the MINT telephone screening service that the patients were made aware of at their initial Emergency Department visit. If they were continuing to have problems a few weeks after their whiplash injury, they were encouraged to call a freephone number as they may be eligible to receive physiotherapy as part of step two of MINT. In order to be randomised into step two of MINT, participants had to attend a hospital research clinic. Patients were screened over the phone to ensure they were within six weeks of their ED visit and they were experiencing cervical spine problems. If they appeared eligible, they were provided with verbal information regarding the study and a research clinic appointment was arranged. Two information sheets were sent to the patients at least 24 hours prior to the appointment. One provided information regarding step two of the main

trial and the other regarding the reliability study. Before patients were enrolled in either of the reliability studies they had the procedures explained and were given the opportunity to ask any questions. Written informed consent was then obtained for those who agreed to participate.

6.4.3 SELECTION CRITERIA

- Aged 18 years or over
- Experienced a whiplash injury less than six weeks before initial recruitment in the Emergency Department
- Able to provided written informed consent
- WAD grade I – III reported in the last 24 hours
- No fractures of spine or other bones.

6.4.4 OBSERVERS

In Study One the measurements were carried out by the author - a physiotherapist, with approximately 6 years of clinical experience at that time. Prior to the conduct of the study the author undertook two hours training in the use of the CROM device with practice on asymptomatic subjects (colleagues). Subsequently the device and protocol were used in weekly research clinics for approximately 8 months (circa 30 patients) prior to the commencement of Study One.

In Study Two, the observers were the author and another research physiotherapist who had at least 10 years clinical experience at that time. This observer had also taken part in a 2 hour

training session and had experience of using the CROM device and the measurement protocol in research clinics prior to commencing Study Two.

6.4.5 DEVISING THE MEASUREMENT PROTOCOL

The procedure used when cervical movement is measured will have a significant effect on the results obtained [253] and therefore it was vital to develop a rigorous measurement protocol to test and, if acceptable, utilise in the cohort study described in Chapter Seven.

Using a combination of knowledge acquired from performing the systematic review of reliability studies described in the previous chapter, reading other systematic reviews of the subject, reading the instruction manual of the CROM device (Performance Attainments Associates ^{tm}) and clinical expertise, a measurement protocol was devised to pilot. Particular attention was paid to deficiencies identified in systematic reviews to ensure sources of error/bias in previous studies were not repeated or minimised for this study.

The protocol was piloted on a number of asymptomatic colleagues who were also experienced clinicians and minor amendments were made, mainly regarding patient instruction and positioning with the chair.

The full finalised measurement protocol is presented in Appendix 10. The following is a justification of the important elements of the protocol.

The participant was asked to sit with hips and knees at 90 degrees, feet flat on the floor, arms resting in lap and as far as possible a neutral pelvic position. The cervical spine is most frequently functioning in an upright position where its configuration and conjunct movements are at the greatest advantage. Ideally an assessment process should mimic the functional position to provide greatest information on impairment and resulting potential

disability. The CROM device has the advantage of measuring movement in all three planes in an upright position, unlike some instruments – e.g. standard inclinometer – which require rotation to be measured in supine. Furthermore, it has been demonstrated that measurement of rotation in supine is not concurrently valid to measurements taken in an upright position [231].

Next the CROM device was fitted – instructions were provided on the details for fitting the magnetic yoke which needs to be in a certain orientation to magnetic north.

The measurement process was explained to the participant in lay terms prior to any movements being conducted – ability to understand instructions was thought critical to the measurement process. Although research into the affect of the nature of instructions is absent [254], patients have indicated that explanation of the process of assessment is key to putting them at ease [258] and this will have an effect on motivation for movement [259]. Clinicians were provided with a script to ensure a consistent message was provided to the patient and in a consistent manner – for example volume of voice has been demonstrated to affect impairment assessment results [260].

For active movements participants were asked to move their heads as far as they felt able whilst keeping back and shoulders as still as possible. Researchers demonstrated the movements to the participant and recorded ROM in degrees and limiting factors after each movement. For passive movements researchers moved the participant's head as far as the participant would allow. For both active and passive movements, the order of the single movements was Flexion, Extension, Right rotation, Left rotation, Right lateral flexion and Left lateral flexion as recommended in the Guides to the evaluation of permanent impairment [127].

A single repetition was made for each direction of movement. Lea and Gerhardt [261] give the opinion that if there is an acute condition present it is preferable to take a single measurement and there is some evidence that a single measurement is as reliable as taking an average of multiple measurements [262]. With a WAD population, not only may individuals suffer from an exacerbation of pain but also dizziness is quite common and therefore single repetitions may be less provocative. Short rest periods were utilised to attempt to negate this provocation of symptoms [253].

Very few studies have evaluated both active and passive range of movement within the same study and none have evaluated whether the order of this affects measurement stability. On average more ROM is usually achieved passively than actively. This is thought to be due to the examiner having the ability to generate more force than the patient's muscles and also because the contractile tissues should be in a relaxed state. Bearing this in mind, it was believed that the active ROM would act as some kind of 'warm-up' to ready the periarticular structures for a greater motion when examined passively. Measuring active before passive ROM is also the most commonly used sequence in a clinical setting, probably for the reason stated above.

The protocol did not include warm-up repetitions of the movements. Approximately half of previous reliability studies used a warm-up procedure, usually consisting of three repetitions. Warm-up is thought to benefit reliability of a measurement protocol by minimising creep during tests which is associated with multiple movements. However, Solinger et al[263] found very little warm-up effect on magnitude of cervical motion. This uncertainty coupled with potentially sensitisation lead to the decision that no warm-up would be included.

The effect of participants having their eyes open or closed is unclear. From the only study that evaluated the effect of this variable on reliability [148], results suggested that when examined in a seated position, it is more reliable to request participants keep their eyes open.

It was decided that the history and subjective examination would be conducted first and unblinded as this has been shown not to affect/influence reliability [203, 241]. The advantage of this is that it mirrors clinical practice and the observer can find out any contra-indications to testing e.g. VBI symptoms which may in turn provide confidence to fully test the participants ROM.

A short test-retest interval for both studies was adopted to minimise any opportunity for clinical change but long enough for participants symptoms not to be aggravated continuously.

Two observers were used for the inter-observer study for practical reasons. Firstly this made it logistically easier to conduct the study (both observers worked in the same department and had permission to conduct assessments in the same NHS trusts) and also this minimised the potential for symptom aggravation as described above for numbers of repetitions.

6.4.6 MEASUREMENTS

Measurements for both studies were completed in a single session lasting approximately 15 minutes.

Data on demographics (age, sex), 'today's' neck pain severity (Visual Analogue Score 0 'no pain' -10 'as bad as a pain could be'), clinical factors (chronic widespread pain [264], number of complaints) and disability (Neck Disability Index[23]) were collected prior to the cervical ROM measurements in both study one and two.

In Study One a single reading for each maximal range motion for active then passive cervical spine motion were taken by the author for the following movements:

Flexion, extension, right rotation, left rotation, right lateral flexion and left lateral flexion.

The subject would return to their neutral with a pause for 5-10 seconds for the observer to note the measurement.

After the sequence was complete the subject then had a rest period of approximately two minutes and then the movements were repeated in the same order. Results of the first sequence of measurements were blinded to the author by covering them with a taped piece of paper.

In Study Two, the order of the observers was randomised and results were blinded from both the subjects and the other observer by using a separate recording sheet for each observer.

6.4.7 ETHICAL APPROVAL

Ethical approval was granted for study one and two by Trent Multi-Centre Research Ethics Committee and appropriate Local Research Ethics Committees as a substantial amendment to the approval for the MINT study. The confirmation letter is presented in Appendix 11. Participants attended research clinics at a number of hospitals across the West Midlands.

6.5 STATISTICAL ANALYSIS

All statistical analysis was performed using SPSS (Version 15). Before data were analysed they were checked using double data entry by the author, screening for implausible values and range of values and appropriately labelling missing items [265].

Descriptive statistics (frequency counts for categorical variables and mean (SD) for continuous variables) were collated to summarise the demographics of the population studies. This is important as interpretation of the reliability of any device will always need to be population specific [189].

The next stage was to screen the distributions of each of the variables of range of motion measurements. The use of reliability statistics discussed below rely on normal distribution of continuous data i.e. are parametric.

The normality of the data distribution was assessed qualitatively and quantitatively. Firstly frequency distribution histograms and cumulative frequency plots were produced and studied for deviations from normal curves/lines and secondly the data was subjected to the Shapiro-Wilk *W* test for normality [265]. If the data was not normally distributed the utilisation of log transformation may be necessary. It was decided a-priori that reliability statistics would be run in both raw and transformed states to see if the transformation added benefit.

Choice of statistical techniques to assess reliability has evolved relatively recently with an increased understanding of the subtleties of differences between agreement, association and consistency and the strengths and limitations of techniques available.

The reliability coefficient is the ratio of variance between the subjects to error variance. We can interpret this to give us a percentage of the variance that results from 'true' variation among the patients. This is called the Intra-class Correlation Coefficient or ICC. There are different versions of this depending on the assumptions we make. In this study the observers were considered to be a sample of all possible observers (we want the findings to be generalisable to all clinicians who would use the CROM device) and therefore were treated as a random factor and as such calculated absolute agreement. Therefore the version ICC

(2,1) was used. The 2 reflects class or model 2 ICC which means all subjects are evaluated by all observers and the 1 reflects the form which in this case is the reliability of a single measurement. 95% Confidence intervals were calculated for each ICC value.

Other reliability coefficients exist, however they were inappropriate for use in these studies (see previous chapter for further explanation).

One other method of analysing measurement error has become popular in the medical statistics sphere, that of calculating Limits of Agreement (LoA) proposed by Bland and Altman [194]. The method is closely linked to that of calculating the ICC but it has the advantage of producing graphical displays of differences in observations or observers, thus allowing assessment of whether systematic bias is present. It has been suggested that the ICC and LoA plots be presented in parallel [266].

Therefore it was decided that LoA plots would be presented alongside the ICC calculations for each of the measurements (these consist of a plot of the difference between two observations against the mean of the pair).

With the ICC being a dimensionless ratio it is hard to elucidate what the reliability means from a clinical perspective in the units of interest (in this case degrees). The Standard Error of Measurement (SEM) can be calculated from the ICC and Standard Deviation (SD) and allows the provision of a 95% confidence interval around a measurement. The SEM can also be used to calculate the Minimal Detectable Change (MDC) which may be used as a threshold for judging the value of a change in the measurement with repeated measures over time. This however is superfluous to the current studies as they are concerned with the value of a one-off measure, although this would be an obvious next step for further investigation.

6.5.1 SAMPLE SIZE CALCULATION

As discussed in the previous chapter, sample size calculations for reliability studies have been omitted from the majority of reporting of previous studies.

In order to demonstrate a 'substantial' level of reliability, the ICC value required was chosen as 0.80 [210]. Therefore H0 (null hypothesis): ICC is less than or equal to 0.60 versus H1 (experimental hypothesis): ICC is greater than 0.80

Po = 0.6, P = 0.8

Theta-0 = 0.6/1-0.6 = 1.5

Theta = 0.8/1-0.8 = 4

Co = (1+[2x1.5]) / (1+[2x4]) = 4/9 = 0.444

$K = 1+ [2(1.6449 + 0.8416)^2 x2 / (\ln 0.444)^2 (2-1)]$

$= 1+ 24.730 / 0.659$

$= 38.526$

The sample size calculated for a 90% power of testing a 5% significance level required 39 subjects [267].

6.6 RESULTS

6.6.1 STUDY ONE - INTRA-OBSERVER STUDY

39 patients were recruited to the study, however one subject was unable to continue with the assessment after consenting to participate and one other subject had incomplete data as symptoms were too severe to complete the entire range of motion assessment.

Demographics for the 38 participants in the reliability study are shown in Table 29. The vast majority of participants had a WAD grade of II meaning they had objective signs of cervical spine dysfunction. Only two subjects had neurological signs (WAD grade III). Pain related neck disability in the form of Neck Disability Index scores categorised into none (0-4), mild (5-14), moderate (15-24), severe (25-34) and complete (>34) [23]. The majority (>70%) were either moderately or severely disabled.

Table 29 - Population demographic summary data

	IaO Study (n=38) **Mean (SD) unless stated**
Sex (F:M)	19:19
Age	38 (11.3)
WAD Grade – n (%)	I=2 (5), II=34(90), III=2(5)
Injury due to MVC - n (%)	37 (97)
Days between injury and Ax	27 (8.4)
Pain VAS 0-10	6 (2.3)
NDI score	22 (9.1)

Mean ROM, ICC (95% CI), SEM and LoA ranges are presented in Table 30. Both active and passive ROMs were approximately 25% less than expected normal cervical ROM values[101].

One set of measurements for passive RLF was not normally distributed (significant S-W test) and so its data was converted using a natural log transformation. Results (ICC = 0.98 (95% CI 0.95-0.99) were consistent with raw data results and the measurement maintained its high level of reliability.

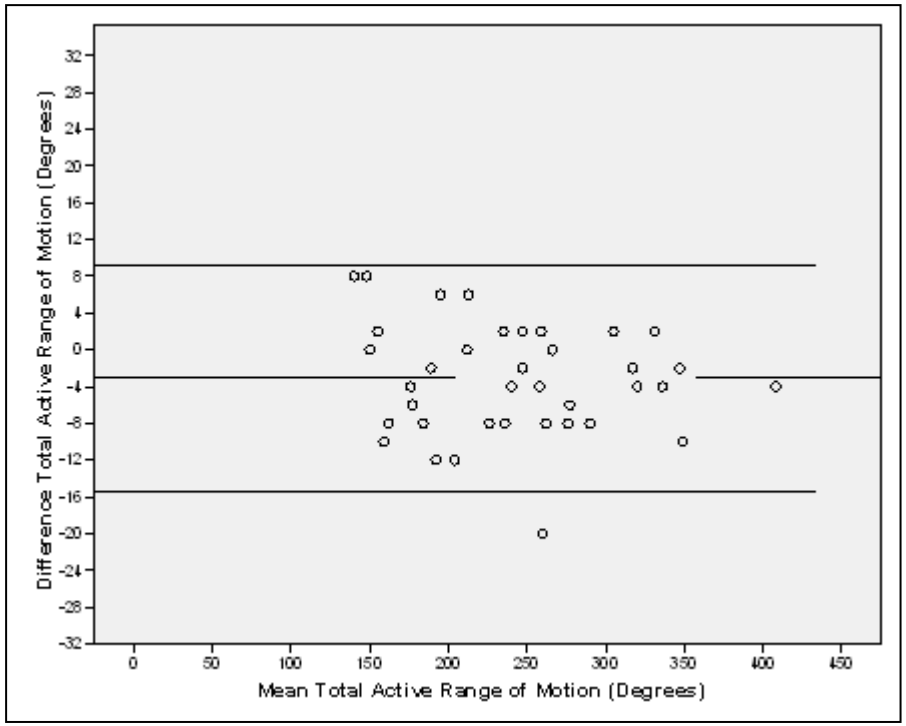
Intra Class Correlations ranged from 0.98 to 0.99 for both active and passive movements indicating a high level of reliability. The Standard Error of Measurement ranged from 1.3 to 2.1 degrees and Limits of Agreement ranged from -6.4 to 5.0 and -6.9 to 6.3 degrees for half-cycle active and passive movements respectively. Figure 18 displays Limit of Agreement plots for total active and passive ROM.

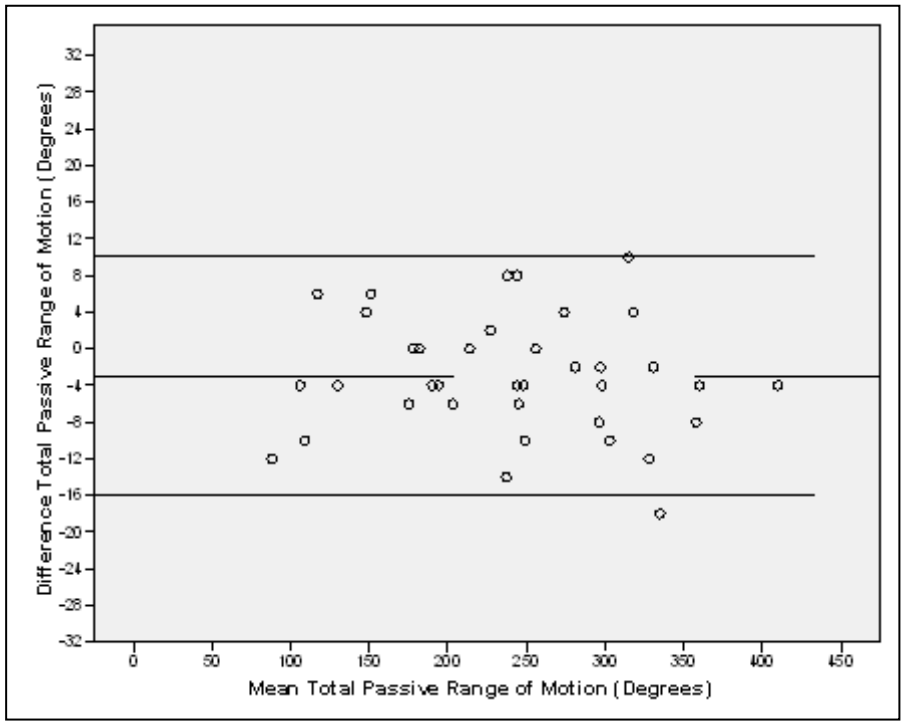
Table 30– ROM Summary and Reliability Statistics –ROM, ICC (95% CI) and Standard Error of Measurement (SEM)

Movement direction	Mean (SD) ROM / degrees	ICC (95%CI)	SEM / degrees	LoA/ degrees
Active Flexion	38 (14.3)	0.99 (0.98-0.99)	1.4	-4.3 to 2.9
Active Extension	41 (16.4)	0.99 (0.98-0.99)	1.6	-5.0 to5.0
Active Right Rotation	51 (14.1)	0.98 (0.96-0.99)	2.0	-6.4 to 4.5
Active Left Rotation	51 (15.0)	0.99 (0.98-0.99)	1.5	-4.6 to 3.4
Active Right Lateral Flexion	26 (9.4)	0.98 (0.96-0.99)	1.3	-4.2 to 3.0
Active Left Lateral Flexion	34 (9.8)	0.98 (0.96-0.99)	1.4	-4.3 to 3.3
Total Active ROM	241 (66.3)	0.99 (0.99-0.99)	6.6	-15.5 to 9.3
Passive Flexion	35 (14.8)	0.98 (0.96- 0.99)	2.1	-6.5 to 4.6
Passive Extension	42 (17.8)	0.99 (0.99-1.00)	1.8	-4.3 to 4.4
Passive Right Rotation	50 (18.8)	0.99 (0.97-0.99)	1.9	-6.9 to 5.1
Passive Left Rotation	55 (20.3)	0.99 (0.98-0.99)	2.0	-6.6 to 6.3
Passive Right Lateral Flexion	26 (11.0)*	0.98 (0.97-0.99)	1.6	-4.4 to4.0
Passive Left Lateral Flexion	32 (9.3)	0.98 (0.95-0.99)	1.3	-4.3 to 2.6
Total Passive ROM	240 (80.1)	0.99 (0.99-0.99)	8.0	-16.0 to 10.1

*denotes significant for Shapiro Wilk test $p<.05$

Figure 17 - Intra-observer Limits of Agreement Plots for active (left) and passive (right) cervical ROM





6.6.2 STUDY TWO – INTER-OBSERVER RELIABILITY

19 patients were recruited to the study; however one subject was unable to continue with the assessment after consenting to participate. The pre-specified sample size of 39 participants was not achieved due to completion of the main MINT study recruitment before the completion of this study and therefore no availability of further similar patients. The implications for this are considered in the discussion section.

Demographics for the 19 participants that participated in the reliability study are shown in Table 31. The vast majority of participants (95%) had a WAD grade of II meaning they had objective signs of cervical spine dysfunction. As with Study One, pain related neck disability in the form of Neck Disability Index scores were categorised and the majority of participants (68%) were either moderately or severely disabled. Demographically the samples of subjects in Study One and Two were broadly similar.

Table 31– Inter-observer population demographic summary data

	IeO Study (n=19) **Mean (SD) unless stated**
Sex (F:M)	13:6
Age	41 (14.8)
WAD Grade – n (%)	I= 0, II= 18 (95), III= 1(5)
Injury due to MVC - n (%)	17 (90)
Days between injury and Ax	35 (9.2)
Pain VAS 0-10	5 (2.4)
NDI score	21 (9.7)

Mean ROM, ICC (95% CI) and SEM are presented in Table 32. As with the intra-observer study, ROM was consistently reduced in all planes of motion when compared to asymptomatic normative values. Intra Class Correlation coefficients ranged from 0.82 to 0.95 and 0.77 to 0.96 for active and passive half cycle ROM measurements. Standard Error of measurement ranged from 3.6 to 8.5 degrees for half cycle ROM measurements. Limits of agreement ranged from -21.6 to21.5 and -16.9 to 29.4 for active and passive half-cycle measurements respectively (Figure 18).

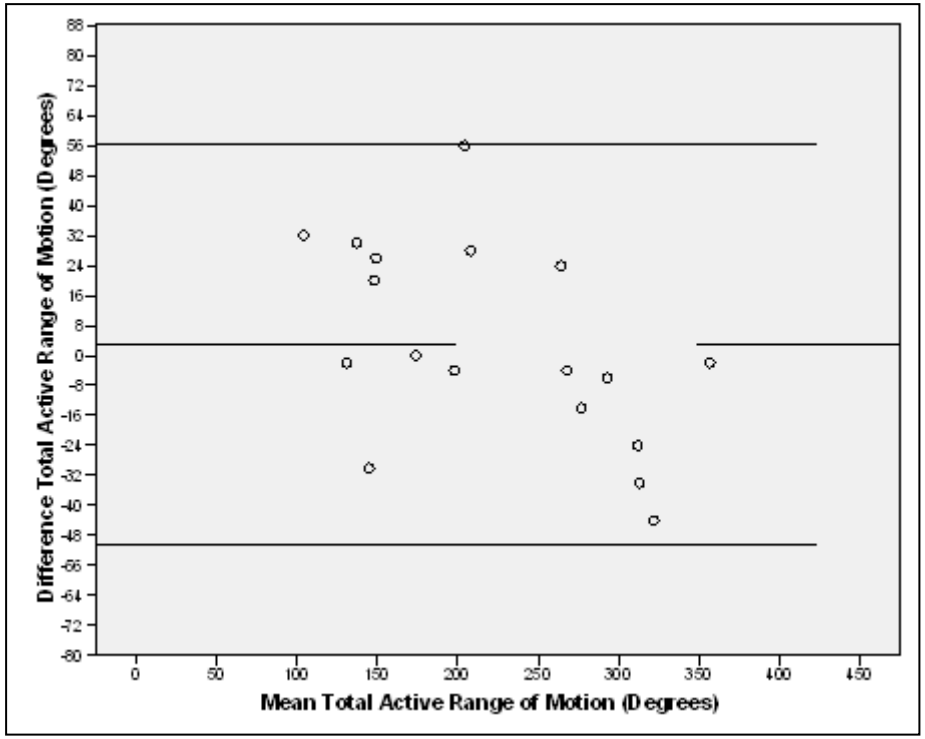
A number of the movements (active LR, passive E, LR, LLF and total passive ROM) were not normally distributed (significant S-W test) and so data were converted using a natural log transformation. Results were consistent with raw data results and the measurements maintained their categories of reliability.

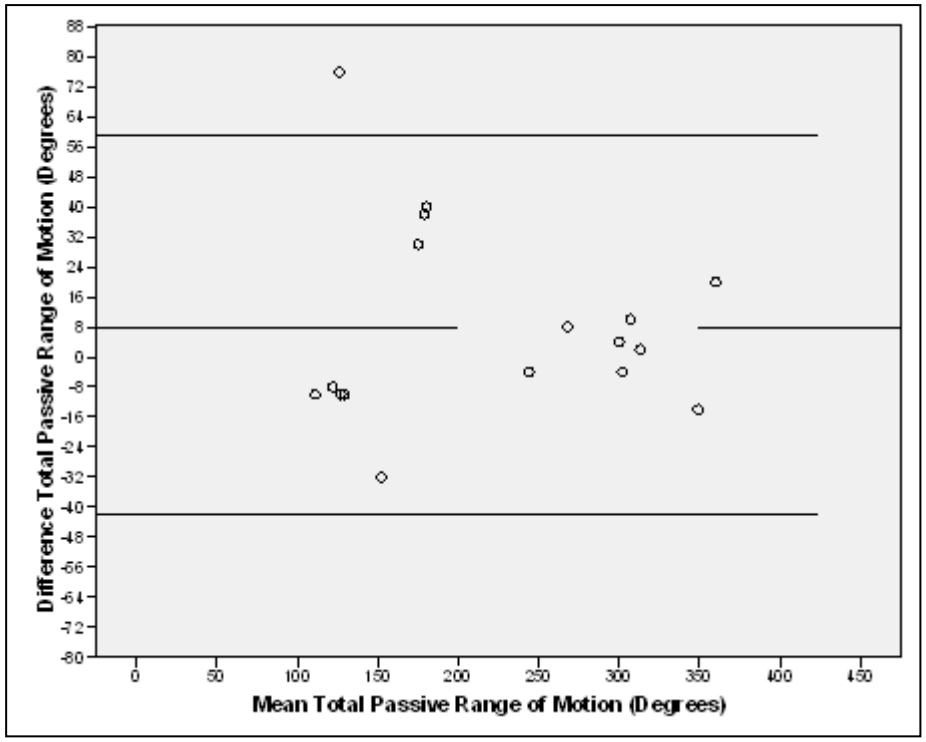
Table 32– Inter-observer ROM Summary and Reliability Statistics –ROM, ICC (95% CI) and Standard Error of Measurement (SEM)

Movement direction	Mean (SD) ROM / degrees	ICC (95%CI)	SEM / degrees	LoA/ degrees
Active Flexion	33 (15.8)	0.83 (0.61-0.93)	6.5	-21.6 to 16.1
Active Extension	41 (18.1)	0.88 (0.72-0.96)	6.3	-18.9 to 18.0
Active Right Rotation	45 (17.5)	0.92 (0.80-0.97)	4.9	-13.5 to 15.5
Active Left Rotation	45 (17.9)*	0.87 (0.68-0.95)	6.5	-13.8 to 21.5
Active Right Lateral Flexion	25 (8.7)	0.82 (0.59-0.92)	3.7	-12.0 to 10.0
Active Left Lateral Flexion	32 (10.6)	0.88 (0.70-0.95)	3.7	-7.9 to 12.4
Total Active ROM	222 (79.0)	0.95 (0.86-0.98)	17.7	-50.7 to 56.4
Passive Flexion	32 (17.0)	0.90 (0.76- 0.96)	5.4	-16.9 to 14.0
Passive Extension	40 (19.4)*	0.96 (0.89-0.98)	3.9	-9.5 to 12.8
Passive Right Rotation	44 (18.4)	0.89 (0.67-0.96)	6.1	-10.3 to 19.9
Passive Left Rotation	45 (22.0)*	0.85 (0.60-0.94)	8.5	-17.4 to 29.4
Passive Right Lateral Flexion	23 (10.1)	0.77 (0.36-0.92)	4.8	-16.6 to 7.9
Passive Left Lateral Flexion	30 (10.5)*	0.88 (0.70-0.95)	3.6	-8.4 to 12.2
Total Passive ROM	213 (90.4)*	0.96 (0.895-0.985)	18.1	-42.1 to 59.3

*denotes significant for Shapiro Wilk test $p<.05$

Figure 18 - Inter-observer Limits of Agreement Plots for active (left) and passive (right) cervical ROM





6.7 DISCUSSION

6.7.1 RESULTS AND COMPARISON TO PREVIOUS RESEARCH

Both researchers and clinicians search for convenient, accurate and reliable methods for characterising patients through ROM and monitoring changes. The CROM device promised good reliability and validity from findings of previous studies.

Reliability is represented by a number of measures – absolute reliability and measurement error. The primary aim of this study was to assess the reliability as measured by the ICC as the use of the CROM in the cohort study in the next chapter was to differentiate/categorise patients using a single measurement in time.

The results of this study indicate that measurement of both active and passive cervical ROM can be performed with substantial within and between-observer reproducibility in a WAD population using the CROM device.

Findings indicate that intra-observer reliability is greater than inter-observer for both active and passive measurement methods (ICC ranges 0.98-0.99 vs. 0.77-0.96 respectively). This is to be expected given conclusions from previous studies [77, 186]. The confidence intervals (CI) for ICC results of the inter-observer study are wider than those of the intra-observer study. One direction of passive movement in particular – right lateral flexion – has a particularly wide CI which means less confidence should be afforded to the reliability of this particular movement. The variation in reliability between active and passive cervical ROM for the Inter-Observer study was small.

The ICC values for this study were comparable to previous studies investigating intra-observer reliability of active cervical ROM measurements with the CROM device in

symptomatic individuals [77, 186]. The only previous study that evaluated Intra-observer reliability for passive cervical ROM [151] did so for asymptomatic participants and used the Pearson's R statistic so the current study's findings cannot be validly compared.

Previous inter-observer studies found equivalent results when evaluating reliability of the CROM for active cervical ROM in symptomatic participants [186, 227, 230]. When comparing these findings to previous studies involving asymptomatic populations, reliability appears to be slightly greater [226, 229].

The ICC values for passive cervical ROM in this inter-observer study fell between results of the two previous studies [152, 212]. Unlike Nilsson et al [152] measurement of half cycle passive cervical ROM with the CROM was found to be substantially reliable. The difference in findings may be due to the differences in measurement protocol. Nilsson et al [151]also blinded the examiner to the readings which doesn't appear to be clinically relevant.

An analysis of full cycle measurements (not presented) showed these measurements were more reliable than half cycle ones. Previous studies have shown that full cycle measurements are more reliable [101, 232]. This may be as a result of eliminating the problem of the 'neutral' head position required for consistent half-cycle evaluation. The dilemma the clinician is faced with is that if full-cycle measurements are used, measurements are unable to elucidate unilateral dysfunctions of the cervical spine. Previous research has demonstrated there can be unilateral differences in ROM in a WAD population. For example Sterling et al [168] found a consistent difference between left and right rotation for patients with sub-acute WAD.

Besides half and full cycle measurements, this study also calculated a measurement of total cervical ROM by summing readings of all six half-cycle movements. This has been

conducted in previous reliability and cohort studies [68, 136, 157]. The reliability of measuring this total cervical ROM is greater than individual half-cycle measurements as would be expected. The next chapter will investigate the validity of using this total cervical ROM as a summary measure by comparing the association with the individual half-cycle measurements.

With ICC values being unitless, this makes them difficult to interpret into clinically meaningful information and therefore Standard Error of Measurement (SEM) values and Limit of Agreement (LoA) results have been provided alongside them which use the units of measurement (degrees). The SEM values obtained in this study allow clinicians in particular to determine whether a clinical difference is observable when using the CROM device. For active and passive cervical ROM measurements this equated to approximately two degrees if the same examiner was performing the measurements or approximately six degrees if it was a different examiner.

LoA results indicate that there were no systematic errors in these particular studies. The graphical plots provide evidence to conclude that there was no effect of warm-up or examiner bias (consistent relative over or under measuring). The LoA statistics also provide confidence that 95% of measures will be a true measure within about 7 degrees for active cervical ROM and about 9 degrees for passive cervical ROM for a half cycle movement when different observers are used.

6.7.2 STRENGTHS

As a result of conducting both intra- and inter-observer studies, there can be some confidence that the majority of error observed in the inter-observer study was due to the difference between observers. Both studies benefitted from a robust and standardised

measurement protocol that had been manualised and revised to ensure as much consistency as possible. Both observers involved in these studies had considerable experience of using the CROM device in the patient population studied. Following the pilot work for the development of the measurement protocol the observers did not practice or measure study participants together. This was important so as not to introduce artificial reliability- for clinical purposes the reliability should only be dependent on realistic familiarity with the measurement protocol and the instrument itself.

The standardisation and considerable training can be seen as a strength of the design and conduct of the studies, however the associated limitation is that the findings may not be generalisable to inexperienced clinicians. Experience has previously been found to affect reliability study findings [151, 152]. A key message proposed therefore is that clinicians planning to use the CROM device should be adequately trained and perhaps practice on asymptomatic volunteers prior to using on their patients in order to obtain the most reliable measurements.

From a methodological perspective, these studies sought to use what are currently considered the most appropriate statistical methods for evaluating intra- and inter-observer reliability. Previous studies have been criticised for use of incorrect and misleading techniques (see previous chapter). By reporting findings of a number of different techniques, hopefully the interpretation of these studies will be clear and worthwhile for future researchers and clinicians.

6.7.3 LIMITATIONS

The studies presented do have a number of limitations which should be considered when interpreting them. The most obvious shortcoming is the sub-optimal sample size for the

inter-observer study (Study two), with 19 participants recruited instead of the target of 39. Having a reduced sample size reduces the confidence in the results of the study and may lead to a greater chance of a type I or more commonly type II errors (i.e. rejection of the null hypothesis when it is true or failing to reject the null hypothesis when it is in fact false). However, the 95% confidence intervals around the ICC estimates are relatively narrow, indicating the variability wasn't excessive.

During the assessment of participants for both studies, the re-test period was short in order to minimise burden to the participant. Consequently, these results cannot necessarily be generalised to re-testing over longer, potentially more unstable, periods e.g. one week which clinicians typically use for reassessment and quantification of change/treatment effect. This is an obvious area for further research and has been highlighted as such in the later section.

With the afore-mentioned short re-testing period there is the potential for introduction for recall bias for the intra-observer study. The assessor (and author) was aware of this possibility and sought to minimise this by using separate recording sheets and the distraction of performing all movements in one cycle and then repeating after the rest period.

With the assessment of cervical ROM in sitting, other areas of the spine were mobile and there is the possibility that participants may have contributed to the cervical spine movements with these other areas. Attempts were made to minimise these supplementary movements through instructions by the assessors. Other studies have used restraints of the thorax to minimise the potential confounding movement, although the effect of this has not been formally investigated as yet. Other potential measurement inconsistency due to actions of the participants may have come in the form of apprehension or 'guarding'. Again attempts were made to reduce this through education, reassurance and monitoring by the assessors.

The ‘neutral’ or central position of the cervical spine was not standardised, but was based on the participants feeling of where their neutral was. It is argued that this would not affect the maximal ROM measurements that this study recorded, however posture has been shown to affect cervical ROM in previous studies [146]. For this reason, assessors were instructed to encourage patients to sit as upright as possible to try and negate a forward head posture.

6.7.4 CLINICAL IMPLICATIONS

Besides the previously discussed reliability of the CROM device already demonstrated in a clinical setting there are other clinical relevancies of findings from these studies. Firstly, the CROM device was found to be easy to use. There are no requirements to locate anatomical landmarks as with other tools, thus the device is very quick to fit and take recordings, a distinct advantage in clinical settings where time is often short.

The disadvantage from an NHS clinical perspective is the cost and limited applicability of the device – unfortunately the CROM device cannot be used for other parts of the body, unlike other commonly used devices such as the Universal Goniometer.

For clinicians assessing WAD patients, the CROM device is a reliable tool to use, however, clinicians will often be managing patients with other insidious causes of neck dysfunction. Previous studies indicate reliability may be as substantial as this study demonstrated, however this would have to be formally investigated for both active and passive cervical ROM to be able to make a judgement for assessing these other clinical groups.

6.7.5 RESEARCH IMPLICATIONS

First and foremost, from the reliability studies performed, the CROM device and the associated measurement protocol could be confidently utilised for the cohort study presented

in the proceeding chapter. This includes use by a number of research clinicians across the multiple centres of the MINT study.

These studies have also generated some questions that could be answered with further research. It would be beneficial to investigate intra-observer reliability for more than one observer to estimate the variability of this reliability between different observers. This could also involve the quantification of the effect of the amount of training. The previously highlighted short coming of the sub-optimal sample size of the inter-observer study could be addressed. It would be most efficient if all of the above could be organised into one research project using a single group of participants.

Evaluation of reliability for longer re-test intervals would be advantageous and highly clinically relevant, as cervical spine ROM measurement is often used for monitoring change and response to treatment over periods of weeks and sometimes months within a therapeutic setting.

The findings of these studies are obviously applicable to sub-acute WAD population; however it is not certain whether they are generalisable to a more chronic population. This chronic population is frequently encountered in an NHS therapy setting and therefore investigation of the reliability of this device is warranted.

Finally, musculoskeletal clinicians commonly measure passive cervical ROM in supine, despite the previously discussed recommendations in Chapter Four. It would be interesting to see whether a development of a version of the CROM for use in supine would demonstrate any differences in reliability.

6.8 SUMMARY

This chapter has described two reliability studies of the CROM device, performed with symptomatic individuals that had experienced a whiplash injury. The results demonstrate that the CROM device is substantially reliable for both within and between observers.

The following chapter describes the use of this device as part of a large prospective cohort study to investigate the diagnostic and prognostic role of cervical ROM in individuals with sub-acute WAD.

7 CHAPTER SEVEN –EVALUATING CERVICAL ROM IN A PROGNOSTIC COHORT OF INDIVIDUALS WITH WAD

7.1 INTRODUCTION

So far this thesis has presented two systematic literature reviews that have concluded that the prognostic value of cervical ROM is uncertain in a WAD population and that the CROM device has shown promise as a valid and reliable measurement tool for cervical ROM. Further studies as part of this doctoral work have led to the conclusion that the CROM device is substantially reliable both within and between observers for a sub-acute WAD population. This chapter assimilates this knowledge to investigate cervical ROM in a large cohort of sub-acute WAD patients using the CROM device.

This chapter will begin with a justification of this cohort study, followed by a description of the methods used. A description of baseline characteristics of the cohort will then precede an investigation into the cross-sectional relationships of various physical and psychosocial measures. Lastly, analysis of the longitudinal prognostic value of the cervical ROM measurements will be presented with a thorough discussion of the implications for research and clinical settings. The chapter will utilise the structure and content of reporting recommended in the STROBE guidelines [268], whilst acknowledging that not all items are appropriate as they would be for a journal article, as the guidelines were originally designed.

7.1.1 JUSTIFICATION FOR THIS COHORT STUDY

There is still a significant amount of knowledge to gain regarding the relative importance of factors that affect recovery following a whiplash injury. This is despite the large number of studies synthesised in the literature review presented in Chapter Three. Generally, previous studies have used sub-standard methodology and reporting (insufficient sample sizes, retrospective, unstandardised outcome measurement and inappropriate statistical analysis techniques). Thus, conclusions so far are tentative. A greater understanding should allow for development of improved treatment strategies, which to date, have proven to be largely ineffective for the majority of patients who are struggling to recover from WAD (as demonstrated by the MINT study described in Chapter Two). Evaluation of trials would also be improved through greater understanding of prognostic factors, with trialists stratifying a-priori or subsequently adjusting analyses for these factors and therefore providing more accurate treatment estimates.

Cervical ROM is one such factor that has previously been studied but its prognostic value is still inconclusive, as concluded by the systematic review presented in chapter Three. There was considerable variation in quality of the seven studies that this conclusion was based on.

The majority of the studies measured active cervical ROM with the remainder not stating what type of movement was assessed. A variety of measurement methods were used, most frequently these were not stated, however it is of note that two studies used the CROM device. Regarding statistical analyses, four out of the seven studies used multivariate techniques but only one provided an appropriate sample size according to the

recommendations of Simon and Altman [35] (that there should be at least 10 cases per predictor in the model).

The measurement of the different forms of cervical ROM and its clinical importance is discussed in Chapter Four, but it is restated here that it is integral to the clinical reasoning process and thus healthcare management for patients with WAD. An investigation into the differences between active, passive and patient-rated cervical ROM and their relative prognostic ability is warranted because clinicians place emphasis on the different types of ROM and use these measures to categorise patients. As far as the author is aware, active, passive and patient-rated cervical ROM have not been previously studied together in the same cohort of sub-acute WAD patients, either for cross-sectional or longitudinal purposes.

In sum a high quality cohort study investigating cervical ROM in WAD patients is warranted that utilises and documents sound methodology and analyses.

7.1.2 OBJECTIVES, RESEARCH QUESTIONS AND HYPOTHESES

The objectives of this chapter are twofold:

1. To describe the cross-sectional baseline characteristics of a sub-acute WAD population, particularly with reference to different types of cervical ROM
2. To investigate the prognostic value of these different types of cervical ROM cross-sectionally (observation of the population at baseline) and longitudinally (observation of the population over time) for neck disability and response to treatment

Along with cervical ROM measurements, other potential physical and psychological prognostic factors identified in the systematic literature review documented in Chapter Three were collected in order to answer the following questions:

- Is cervical spine ROM a prognostic factor for poor outcome in WAD?
- What is the prognostic value for measures of:
  - Active cervical ROM
  - Passive cervical ROM
  - Patient-rated cervical ROM
- Do patterns of loss of active or passive cervical ROM predict poor outcome?
- What is the relationship between cervical ROM and other prognostic factors?

- Are there predictors of amount of cervical ROM? (e.g. pain, injury mechanisms, regional pain, psychological factors)

- Does treatment type interact with cervical ROM's prognostic value?

7.2 METHODS

7.2.1 RECRUITMENT AND SELECTION

Patients in this cohort study were recruited as part of the multi-centre RCT Managing Injuries of the Neck Trial (MINT) which was described in detail in Chapter Two. Participants were recruited between December 2005 and November 2007. A brief summary of the methods important to the cohort study are presented here in bullet point form.

- Patients attending participating Emergency Departments (ED) with an acute whiplash injury of WAD grade I to III were eligible for Step One of the trial. Brief clinical details were recorded on the ED proforma.

- All patients attending for treatment who did not ask to be excluded were sent a Two week questionnaire (Appendix 3) within approximately two weeks of their ED attendance. This questionnaire included demographic information (participant's age, sex, ethnicity, employment status) presence of neck pain in the month before their injury, Short Form 12 Health Survey (SF-12), and EuroQol (EQ-5D).

- Patients were asked to contact the co-ordinating centre (Warwick CTU) if they continued to experience symptoms approximately three weeks after their ED attendance. If the patients did contact the co-ordinating centre, a research therapist performed an initial screening by telephone and if the patient appeared to be eligible then an appointment was made to attend a research clinic.

- Research clinics were conducted in the hospital where the patient had attended the ED. This resulted in assessment of patients in a sub-acute state [269].

- At the research clinic the patients were checked for eligibility for Step Two of MINT and the cohort study according to the following criteria:

1. Reporting cervical spine symptoms within the last 24 hours
2. Were WAD Grade I-III at time of assessment
3. Did not have any contra-indications to physiotherapy treatment. This included central cord compression, upper or lower motor neuron lesion, complete nerve root compression, suspected vascular injury or haemorrhagic event.

- If eligible, trial information was reinforced and the patient was asked to provide written informed consent prior to randomisation. Randomisation was via a central telephone randomisation service.

- Once consent and randomisation had occurred, baseline data collection was carried out using methods detailed in the next section.

7.2.2 BASELINE DATA COLLECTION

Participants were first asked to complete a Research Clinic Questionnaire booklet (Appendix 3). This booklet included demographic questions and validated outcome measures evaluating physical, psychological, functional and social aspects of the participants' experience following a whiplash injury. Once this had been completed then the research clinician conducted a clinical assessment, completing the Research Clinic Assessment form accordingly (Appendix 4). Research clinicians received three hours training on how to perform the research clinic appointment and therefore collecting of baseline questionnaire

and assessment measures. This included time spent practising the cervical Range of Motion measures with a CROM device and shoulder abduction measures with a universal goniometer.

Some data was extracted from the questionnaires completed prior to the participant entering the cohort study which were completed as part of Step One of MINT. These were the ED proforma for ED WAD grade and the two week questionnaire (Appendix 2) for demographics and pre-injury neck pain.

7.2.2.1 Selection of baseline measures

Baseline measures for this cohort study were made following the conduct of the systematic review presented in Chapter Three, which included literature published up until August 2006. The aim was to conduct a comprehensive assessment to allow for a detailed description of the cohort and to capture all potentially influential factors to be incorporated into multivariate models. However, choices were also made with the awareness of participant burden in mind, in some cases shorter versions of measures were taken to reduce the time and effort of performing an assessment.

Where possible, measures that had published evidence of validity, reliability and responsiveness were used. For definitions and discussions of these important concepts please see Chapter Five.

A detailed description of all the baseline measures follows.

7.2.2.2 Demographics and pre-injury neck pain

Age, sex, ethnicity and whether the participant had experienced neck pain in the month before their injury were extracted from the MINT two-week questionnaire. The participant

was also asked at the research clinic whether they had returned to work (RTW - if they worked prior to their injury) and details of the motor vehicle accident (if appropriate).

7.2.2.3 Pre-injury Chronic Widespread Pain (CWP)

Pre-injury pain problems were investigated using the Manchester definition of chronic widespread pain (CWP). In order for subjects to be labelled as having pre-injury CWP, pain must have been reported in at least two sections of two contra-lateral limbs and in the spine, and have been present for at least three months prior to their whiplash injury [264]. This validated definition has been used in a previous cohort study of a WAD population [58].

7.2.2.4 WAD Grade

This was recorded by the research clinician at the time of assessment according to the definitions stated by the QTF [6] and described in Chapter One. WAD grade as assigned by the ED clinician and documented on the ED proforma was also extracted. A higher WAD grade is perceived to mean a greater severity of injury.

7.2.2.5 Modified Von Korff Neck Pain Intensity Rating (MVK Pain)

Initial neck pain intensity was measured using 11-point numerical pain rating scales graded from 0-10 where 0 is ‘no pain’ and 10 is ‘as bad as a pain could be’ [270]. Participants were asked to rate their worst pain in the last week, their pain as an average in the last week and their pain at the time of the research clinic assessment [271]. The mean of the three pain scales is multiplied by 10 to give an overall pain score out of 100. A higher score indicates greater the pain intensity.

7.2.2.6 Number and Location of symptoms (No. Of Sx) and Neurological symptoms (Neuro. Sx)

Physical symptoms were evaluated using elements from the Cervical Spine Outcomes Questionnaire (CSOQ) [272, 273]. The questionnaire originally consisted of six subscales and reportedly has good test-retest reliability [272]. Location and number of symptoms have been used in previous cohorts of whiplash-injured patients, however there has been no standardisation of this measure. The physical symptom scale of the CSOQ appeared to be the most valid measure of this construct [273]. Participants were asked whether symptoms were present in 10 body areas and whether they had experienced 5 other types of symptoms (difficulty swallowing, headaches, neurological symptoms in arms, problems with upper limb function and finally neurological symptoms in legs). This resulted in a maximum score of 15, with a higher score indicating a greater number of physical symptoms. Presence of neurological symptoms in the arm was extracted from this measure to be used as a prognostic factor on its own. This was used as a dichotomous outcome (present or not present).

7.2.2.7 Cervical Range of motion (cROM)

Cervical spine ROM was measured in degrees with a CROM device. Active c ROM was measured first and then passive cROM. Main reasons for limitation of range were also noted. These reasons were divided into pain, stiffness and spasm. The measurement protocol and justification is described in detail in Chapter Six.

7.2.2.8 Patient Rated cervical Range of Motion(PRcROM)

As previously discussed in Chapter Four, no previous cohort study of a WAD population has described using a subjective measure of movement or lack thereof and the only published work closely related to this is that of Borchgrevink et al[131] who used a rating of neck stiffness from 0-5 where 0=none and 5=maximum. This work was used as the basis for constructing this exploratory measure. Participants were asked how much they felt they could move their neck in two directions; firstly turning side to side and secondly looking up or down. A numerical rating scale was used ranging from 1 'unable to move' to 5 'able to move normally'. Therefore a higher score indicated the participant reported more normal movement for that plane.

7.2.2.9 Shoulder abduction range of motion (ShAbd ROM)

Shoulder abduction ROM was measured immediately after cervical ROM. This was conducted with the patient sitting. The patient was asked to lift their arm out to the side and up as far as they could take it. The research clinician provided a demonstration of full range prior to the patient performing the movement. Measurements were made in degrees using a universal goniometer. The centre the fulcrum of the goniometer was placed close to the anterior aspect of the acromial process with the proximal arm was aligned so that it was parallel to the midline of the anterior aspect of the sternum. The distal arm was then aligned with the anterior midline of the humerus. Research clinicians were instructed to ensure the participant maintained the same thoracic and lumbar spinal position throughout the ROM assessment. Participants were asked whether pain or stiffness was the predominant limiting factor as appropriate. The range of movement in degrees and limitation (if appropriate) was recorded in the Research Clinic Examination Form after each movement. As with clinician-

measured cervical ROM, a greater number of degrees indicate greater ROM. Shoulder ROM measurements have been shown to be reliable within and between observers using a universal goniometer [274].

7.2.2.10 Fear Avoidance Beliefs Questionnaire (FABQ-PA)

The Fear Avoidance Model attempts to explain why individuals develop chronic pain. Fear of pain leads to avoidance of social and physical activities that are expected to cause pain and/or re-injury [275]. Fear avoidance beliefs were measured using the FABQ Physical Activity subscale [276]. The FABQ consists of two subscales to ascertain fear-avoidance beliefs about physical activity (four items) and work (seven items) and was originally developed for use in patients with low back pain. It has subsequently been validated and used in neck pain populations[277] and the subscales have been demonstrated to have substantial test-retest reliability[278]. The FABQ-PA was used in isolation in order to minimise questionnaire burden to the participants and also because not all participants were working and therefore the work sub-scale had potential to be redundant. Scores for the FABQ-PA subscale range from 0-24 with a higher score representing an increase in fear of movement.

7.2.2.11 Self-efficacy measure (SE)

Self-efficacy is "a personal belief of how successfully one can cope with difficult situations" [279]. Self-efficacy was measured using a single item 7 point numerical scale question; how much do you agree with the statement 'I feel I am able to cope with my neck problem even when it is painful' where 0 is completely disagree and 6 is completely agree. A higher score represented greater self-efficacy or ability to cope. This question was constructed for this cohort study as there was no suitable existing measure for this population of sufficient brevity [280].

7.2.2.12 Pain catastrophising scale (PCS)

Catastrophising in relation to pain is defined as "excessively negative and unrealistic thoughts or self-statements about pain" [30]. Catastrophising can be described as an exaggerated catastrophic interpretation of one's pain and disability, leading to a more intense pain experience and greater emotional distress [281].

Catastrophising was measured using the Pain Catastrophising Scale [282]. This is a 13 item questionnaire which asks the respondent to indicate the degree to which they have the thoughts and feelings listed when they are in pain. It assesses three different dimensions of pain-related catastrophic thinking – rumination, magnification and helplessness. Items have a five point scale to assess frequency of catastrophic thoughts, from 0 (not at all) to 4 (all the time). A higher score pertains to greater catastrophic thinking. This measure has been validated and used with patients with WAD [282-284].

7.2.2.13 Impact of Events Scale (IES)

Psychological distress was measured using the Impact of Events Scale (IES). This measures psychological distress related to a specific life event[285]. It consists of a 15 item questionnaire with higher scores representing a more distressed state. The participant is asked about frequency of psychological distress symptoms in the past seven days with four potential responses (not at all, rarely, sometimes and often). Responses were scored 0,1, 3 or 5 respectively, resulting in a maximum score of 75. This measure has been previously used in a WAD population[86].

7.2.2.14 General Health Questionnaire 12 (GHQ-12)

The General Health Questionnaire 12 (GHQ-12) is a measure of current mental health. The questionnaire was originally developed as a 60-item instrument but shortened versions of

several lengths are currently available [286]. The GHQ-12 was selected for its brevity, ease of completion and its proven application in research settings as a screening tool [287]. The scale asks whether the respondent has experienced a particular symptom or behaviour recently. Each item is rated on a four-point scale (less than usual, no more than usual, rather more than usual, or much more than usual) resulting in a total score of 12 when using a bi modal scoring method (0-0-1-1). A higher score indicates a greater degree of general psychological distress.

7.2.3 FOLLOW-UP DATA COLLECTION AND OUTCOME MEASURES

7.2.3.1 4, 8 and 12 month follow-up Questionnaires

Postal questionnaires were dispatched at four, eight and 12 months after the participant's initial visit to the ED. The format for all three follow-up questionnaires was identical other than the time point label and at the 12 month follow-up an additional question asked whether the participant had pursued and settled a compensation claim. The questionnaires included NDI, SF-12, EQ-5D and health resource use questionnaires (Appendix 12).

All participants were followed up using a standardised procedure. If questionnaires were not received after one week a phone call reminder was made (where possible). If there was no response after two weeks a second copy was dispatched. If there was no response after three weeks then another phone call was attempted and the participant was asked for a core set of data over the phone. If we were unable to contact the participant after three attempts at different times of the day to obtain core outcomes the participant was classified as a non-responder at that time point.

A research assistant was responsible for managing the follow-up questionnaires and data entry. Therefore the blinding of the author was maintained until recruitment and follow-up were complete in order to avoid potential biases.

7.2.3.2 Selection of outcome measures

Choosing the outcome measure for recovery for this cohort study was complex. As documented in a recent review [288], there are numerous different methods to assess recovery in prognostic studies for WAD. It was important to choose a measure which adhered to a biopsychosocial model of recovery and not one where the presence or absence of symptoms was determinant. From a clinical and a research perspective the latter is not a realistic or useful way to determine outcome. Ideally a validated outcome measure that draws on each part of the biopsychosocial model was to be used. The Neck Disability Index was chosen as the primary outcome measure for the MINT study and also for this study.

7.2.3.3 Neck Disability Index (NDI)

The Neck Disability Index is a frequently used condition-specific measure of pain-related disability. It consists of ten-items purported to measure self-report functional status. Seven items assess functional activities (personal care, lifting, reading, work, driving, sleeping and recreation) while the other three address symptoms of concentration, headache, and pain intensity. Each item is scored on a six-point scale from zero (no disability) to five (full disability). The individual scores are summed resulting in a possible total score ranging from zero to 50. Higher scores represent increased disability. Some researchers choose to convert this into a percentage score, giving the added advantage of being able to deal with missing data and inapplicable questions (primarily driving) [289].

The NDI is the most widely validated measure of neck pain related disability for use in this group of patients [290]. The NDI has been shown to be valid and reliable in a whiplash injured population [23]. It has been proposed that both the minimal detectable difference and the clinically important difference lie in the range of five to seven NDI points [33, 291, 292]. The NDI has been used in numerous prospective cohort studies of WAD populations [73, 75, 87, 293] and it was selected for use within this cohort study to allow for some form of comparison to previous studies. The limitations of the NDI are that there is a potential ceiling effect for a small minority of patients (patients who are very disabled may reach the maximum score leading to inability to detect any subsequent deterioration [294]) and that it may not capture psychosocial aspects of the disease. This final point may be contested as Riddle and Stratford [294] concluded that the NDI appears to measure both mental and physical health-related factors when compared to the SF-36 (a generic health related Quality of Life measure with physical and mental components). Moreover it has recently been proposed that the NDI does tap into all elements of a biopsychosocial model of recovery using the ICF model as a framework [295].

Besides using the NDI as a continuous scale, some authors have converted scores to categories. Vernon and Mior [23], originally proposed 5 categories for the NDI; No disability (score <4), mild disability (5-14), moderate disability (15-24), severe disability (25-34) and complete disability (>35). This was used by Crouch et al [293] in a UK cohort of ED patients with WAD. Subsequently, Vernon [296] proposed three categories; recovered (score <8/100), milder pain and disability (10-28) and moderate/severe pain and disability (>30). The latter was used by Sterling et al [86, 87, 168] in their cohort studies. They performed a cluster analysis (K-means algorithm) to validate the groupings. Certainly the

mean NDI scores and variance are separate (Mean NDI score (SD) at 6 month follow-up: recovered 2.9(2.9), Mild 16.5(5.6), Moderate/severe 42.8(12.2) [86].

Nederhand et al (1999) used the NDI as a dichotomised dependent variable [76]. Recovered was classified as <15/50. Miettinen et al [73] dichotomised the NDI for its use as a prognostic tool into score of 0-19 and >20. Neither of the studies had reported sound methodological reasoning for the chosen cut point.

Participants categorised as recovered and not recovered according to categories derived from NDI scores as defined by Vernon [168, 296] and Sterling et al [168] were used as a secondary outcome measure.

7.2.3.4 Patient Reported Recovery question (PRR)

Patients were asked to answer a question on all follow-up questionnaires to ascertain if they perceived a change in the condition in their cervical spine.

At the 4 month follow-up time-point they were asked:

"Is your neck better, just the same or worse after the treatment you received 4 months ago?" Possible responses were Much Better, Better, Same, Worse and Much Worse.

At the 8 and 12 month follow-up time-points they were asked:

"Is your neck better, just the same or worse since your last questionnaire?" Possible responses were identical to the 4 month ones.

7.3 STATISTICAL ANALYSIS

All statistical analyses were carried out in SPSS Version 17 (SPSS Inc. Chicago).

7.3.1 SAMPLE SIZE

As described in Chapter Two, a target sample size of 600 patients was calculated for Step Two of the trial (to detect 0.375 Standard Deviations between group NDI scores, 90% power, 1% significance ICC 0.02 assuming 30% loss to follow-up). All cases that were recruited for Step One and reported ongoing problems to the trial team were potentially recruited. This would determine the sample size for the cohort study. A general recommendation for multivariate analyses that there should be least 10 cases of data for each predictor variable in a model was made by Simon and Altman [35]. More recently, Field [297] summarised work by Miles and Shevlin, who concluded that if we are looking to detect a medium effect then a sample size of 200 will always suffice (for up to 20 predictors). This study would comfortably fulfil these recommendations provided no more than 20 predictors were entered into the final model.

7.3.2 BASELINE AND FOLLOW-UP DATA

Descriptive statistics for demographics, other baseline measurements and outcome measures at the three time points were tabulated. For follow-up/outcome data a flow chart displaying the numbers of participants at each time point was produced. For normally distributed continuous data means and standard deviations are presented and for non-normally distributed and categorical data median and inter-quartile range are supplied. Continuous data was checked for normal distribution using observation of histograms, Kolmogorov-Smirnov and Shaprio-Wilk tests.

7.3.2.1 Presentation of cervical ROM data

Cervical ROM data are presented in Means (Standard Deviation) and frequency counts for the types of limitations and number of limited directions. Cervical ROM data and correlations between them were analysed to decide on what was the best summary measure to use for further analyses. It was decided that summary variables for cervical ROM should be chosen to go forward into the cross-sectional (and longitudinal) multivariate models. If all cervical ROM measurement variables were entered into the models this could provide a potentially unstable model which is over-fitted. Summary variables were selected on the following basis:

- If the variable significantly and strongly associated with the outcome variables of interest
- If the variable significantly and strongly associated with the other ROM measurement variables
- If the variable was believed to have clinical importance not provided by any of the other variables.

Following this, other baseline measures were analysed using univariate correlations and independent t-tests/Mann Whitney U tests to make up a picture of the cohort particularly with reference to the relationship between ROM and other factors.

7.3.3 MULTIVARIATE ANALYSES

One of the criticisms of previous prognostic studies for recovery from WAD is that multivariate analyses have not been performed or performed inappropriately. It is clear that recovery from WAD is multi-factorial in nature with inter-relationships between different prognostic factors and therefore analyses should account for this.

There is still ongoing debate about how to select risk factors for multivariate analyses. It has already been highlighted that there should be a limitation on the number of predictor variables entered into a multivariate model. Therefore previous literature was used to limit the number of factors to be recorded and analysed. There have currently been seven systematic reviews that have investigated prognostic factors for poor outcome following a whiplash injury (5 of which have been published since 2007). Table 33 displays their results. This table was used to decide which factors should be included. All factors that were found to be probable or possible factors in the systematic reviews were investigated for univariate association with the outcome measure *unless* there were a greater number of systematic reviews finding it was not a prognostic factor. No variables were excluded on this latter criterion.

Table 33 - Prognostic Systematic Review Findings

Review (number of studies)	Probable PF	Possible PF	Not PF	Inconclusive PF
Cote et al, 2001 (n=13)[18]	Older age, female sex, baseline neck pain intensity, baseline headache intensity, radicular signs and symptoms	Initial health care, compensation	marital status, no. of dependents, income, work activities, education, crash-related factors, past headaches, past neck pain	
Scholten-Peeters et al, 2003 (n=50)[38]		high initial pain intensity, restricted cervical ROM, low muscle workload, high number of complaints, driving occupation, previous psychological problems		Older age, female sex,
Williams et al, 2007 (n=38)[40]	baseline neck pain intensity, initial disability score, cold hyperalgesia	pre-injury CWP	Reduced pressure pain thresholds	pre-injury neck pain, pre-injury headache, pre-injury back pain, pre-injury degeneration, initial shoulder pain, initial back pain, initial headache, WAD grade, early onset of symptoms, no, of symptoms, restricted cervical ROM, radicular

				symptoms, neurological symptoms, low muscle workload, muscle spasm, imaging findings, BMI score
Williamson et al, 2008 (n=38)[39]		self-efficacy, post-traumatic stress	personality trait, general psychological distress, psychosocial work factors, wellbeing, life control, social support	psychosocial stress not related to injury, previous psych problems, blame and anger, perceived threat, cognitive function, anxiety, depression, irritability, familiarity with symptoms of whiplash, fear-avoidance, catastrophising, coping strategies, somatisation
Kamper et al, 2008 (n=67)[298]	High initial pain and disability, psychological distress		female sex, older age, collision factors	
Carroll et al, 2008 (n=47)[299]	WAD grade, baseline neck pain intensity, baseline disability, no. of symptoms, self-efficacy, fear of movement, catastrophising, initial post-injury anxiety	compensation	collision factors	sex, older age, education level, pre-injury pain

Walton et al, 2009 (n=13)[295]	High baseline neck pain intensity, headache, WAD grade, No postsecondary education	catastrophising, presence of neck pain, no seat belt, history of neck pain, female sex	crash factors	disturbed sleep, older age

Factors found to be probable or possible prognostic factors were:

Older age, sex, baseline neck pain intensity, baseline headache intensity, radicular signs and symptoms, initial health care, compensation, low muscle workload, number of and location of symptoms, driving occupation, previous psychological problems, baseline disability score, cold hyperalgesia, pre-injury CWP, self-efficacy, post-traumatic stress reaction, WAD grade, fear of movement, catastrophising, education level, no seat belt, history of neck pain.

Some of the variables listed above (low muscle workload, driving occupation, education level, no seat belt, cold hyperalgesia) were not measured as part of the cohort study due to lack of evidence at the time of planning the MINT trial and this cohort study (2003-4). For some factors, this lack of adequate evidence at the time of the study commencement meant that the purchase of specialised equipment could not be justified at the time of funding (e.g. for cold hyperalgesia).

7.3.3.1 Primary analysis

The primary analysis was performed using multivariate forward stepwise linear regression. The forward stepwise method was chosen as this analysis was exploratory in nature. The SPSS programme used always added another variable at each step and did not seek to remove any redundant predictors. Backwards stepwise selection methods were used as sensitivity analyses in some cases to see if different selection methods would lead to

substantially different models. The dependent variable for the primary analysis was Neck Disability Index score (0-100).

Two different models were constructed with different groups of independent variables. A 'research clinic' model where all factors that had a significant univariate relationship ($p<.05$) with the dependent variable and a 'typical clinical assessment' model where only a limited number of these significantly associated factors that would be available to clinicians in the current typical clinical setting were entered into the model.

Multiple regression results were tabulated displaying standardised betas (with their standard errors), ANOVA significance, R^2, adjusted R^2 and constant. The model accuracy, fit and assumptions were interpreted in a number of ways. Firstly, variables were checked for multi-collinearity using a correlation matrix and checking values of variance inflation factor (VIF). Correlations between the predictors with a value of $r<.9$ and/or VIF values of 10 or more were treated as cause for concern [300]. The fit of the model was interpreted through the value of R^2 and the significance value of the ANOVA. Standardised beta values of predictors were interpreted as to the importance of each factor. The Durbin-Watson statistic was used to check that the assumption of independent errors was met (a value between 1 and 3 was deemed acceptable). Residuals were assessed in a number of ways (scatter plot of *ZRESID vs. *ZPRED, histogram and a normal p-p plot) to determine whether the assumptions of random errors and homoscedasticity had been met [301]. If they did not have a normal distribution this would question the ability to generalise findings beyond this cohort.

7.3.3.2 Secondary analyses

Two secondary analyses were conducted evaluating the effect of using a different outcome on prognostic value of variables and also the effect of treatment allocation group on prognostic value of ROM measurements.

Participant rated recovery

Participant rated recovery (PRR) was utilised as an outcome measure in comparison to the Neck Disability Index score (primary analysis dependent variable) at 4, 8 and 12 month follow-up points. The categories were dichotomised into those participants that reported improvement in their neck symptoms (better or much better) and those participants that reported no improvement or non-recovery (stayed the same or got worse). Multiple logistic regression was used for this analysis to establish predictors for the non-recovery category. It was planned to construct 'research clinic' and 'typical clinical assessment' models similar to the primary analysis.

Effect of treatment group on ROM prognostic value (moderator) analysis

Information on moderators and mediators of treatment outcomes is important to inform clinical applications and may help to identify for whom treatment has most benefit and/or possible mechanisms by which treatment might be effective. Treatment moderation is distinct from prediction of outcome (baseline factors that predict outcome irrespective of treatment arm are termed non-specific predictors).

Cervical ROM has been previously highlighted as a factor that may help to identify individuals who are at risk of a poor outcome following a whiplash injury [68, 159] however there has been no investigation into identifying whether cervical ROM may help to identify

which patients may be most responsive to a physiotherapy package as opposed to an advice session. Anecdotally clinicians often justify referral to physiotherapy for "hands-on" or more intensive treatment with assessment of reduced cervical ROM [14]. Treatment moderators are investigated by adding an interaction term between the treatment group and potential moderator of interest [302]. If there is no significant interaction effect, but the baseline factor predicts outcome then the factor is a non-specific predictor. Interaction tests are more appropriate than looking at p-values for treatment difference in each subgroup due to reducing findings by chance and effect of small subgroups [303].

An interaction term (treatment group X ROM variable) was added to the model predictors when evaluating prediction of primary outcome at the three follow-up time points (4, 8 and 12 month NDI score).

7.3.4 MISSING DATA

All data was cleaned by checking ranges, identifying outliers and coding missing data. The next step was to compare the characteristics of cases that did and did not have missing date to explore what type of bias might be introduced.

Multiple imputation of data was considered out of the scope of this doctoral study and therefore if participants had missing data, they were excluded from the analyses. The Neck Disability Index score was not imputed if completely missing. However, it has previously been published that if one of the ten item responses is missing (usually the driving item) then a percentage conversion of the remaining 9 items score is valid [304] and so all scores at all time points were converted to this format for ease of interpretation.

7.4 ETHICS COMMITTEE APPROVAL

MINT was approved by the Trent Multicentre Research Ethics Committee (ref MREC/04/4/003) and by the Local Research Ethics Committee and Research and Development Committee of each participating centre. This cohort study was included within the main approved protocol [15] (For approval letter see Appendix 6).

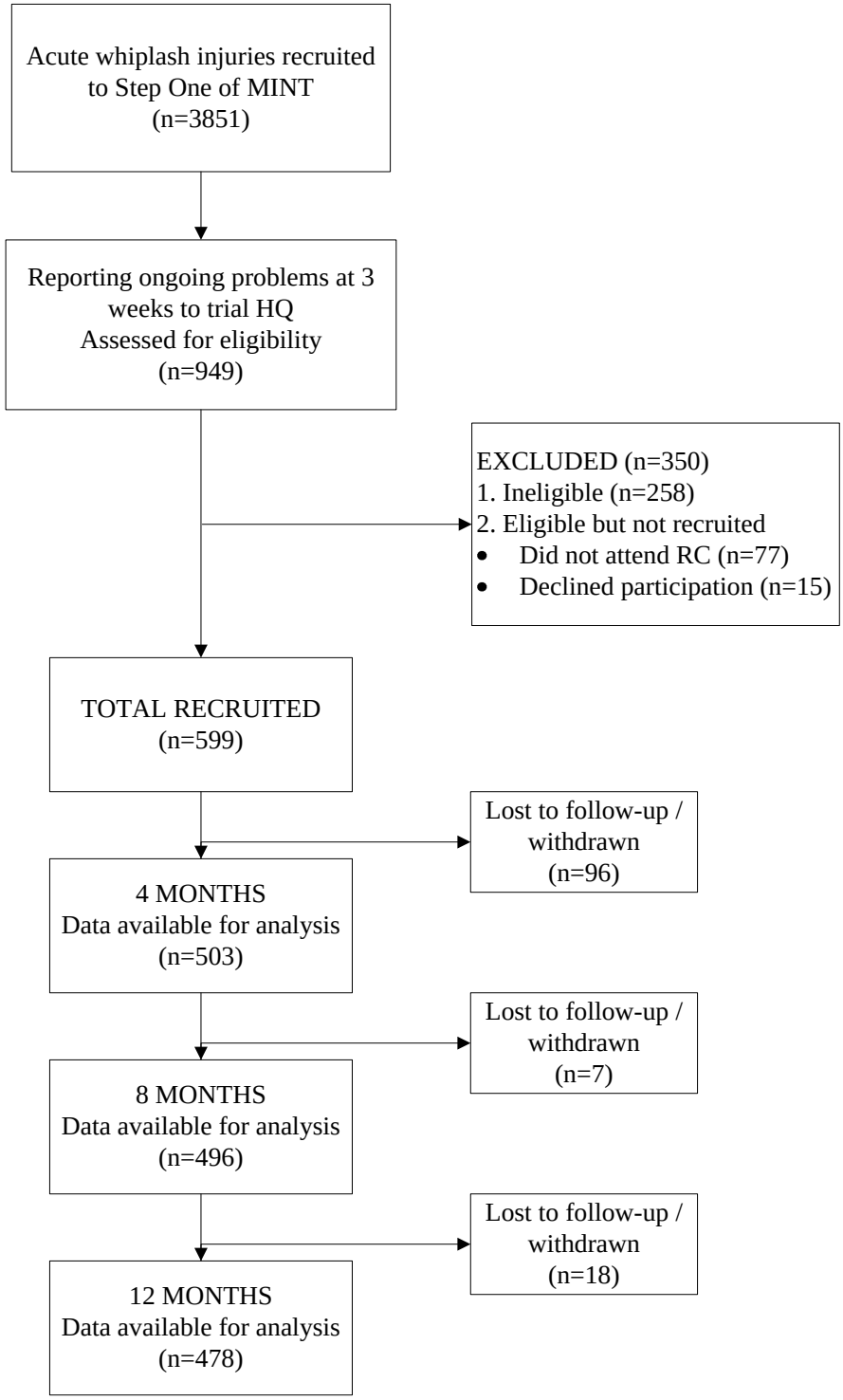
7.5 RESULTS

The results section is structured to describe characteristics of the cohort followed by the results of cross-sectional and then longitudinal multivariate analyses.

7.5.1 RECRUITMENT AND SELECTION

599 patients gave informed consent and were recruited into the cohort study between December 2005 and November 2007 as displayed in Figure 19. 949 of the patients recruited to Step 1 of MINT reported on-going problems to the trial office, and were considered for the second step of MINT. Of these, 693 were assessed as potentially eligible and were invited to attend a research clinic appointment. 77 patients did not attend or cancelled their appointment leaving 616 patients to be assessed for eligibility at the research clinics. Two patients were ineligible and 15 declined to participate resulting in 599 consenting patients recruited.

Figure 19 - Cohort study flowchart



Although different numbers of patients were recruited at each of the sites, the proportion of Step One participants being recruited into Step Two (and this cohort study) was approximately similar (see Table 34). Table 34 shows the number of participants recruited at each site and by treatment allocation. 272 participants were provided with Usual Care advice and 327 with Whiplash Book advice at their previous ED visit.

Table 35 presents characteristics of patients randomised and not randomised into this cohort study from Step One of MINT. The population entered into this cohort included a slightly greater proportion of females, a slightly higher mean age and a higher proportion of patients categorised with WAD grade II and III injury severity compared to participants of Step One that were not recruited into this further piece of research.

Table 34 – Recruitment by site

Cluster #		Advice session	Physiotherapy
		N	**N**
		(% of step 1)	**(% of step 1)**
	Usual Care Advice (n=272)		
1	University Hospitals Coventry and Warwickshire NHS Trust	33 (9%)	32 (9%)
2	North Bristol NHS Trust	19 (8%)	17 (7%)
3	Gloucestershire Hospitals NHS Trust	41 (10%)	42 (10%)
4	Worcestershire Acute Hospitals NHS Trust	24 (8%)	26 (9%)
5	Kettering General Hospital NHS Trust	16 (8%)	14 (7%)
6	Buckinghamshire Hospitals NHS Trust	3 (5%)	5 (9%)
	Whiplash Book Advice (n=327)		
7	Heart of England NHS Foundation	51 (7%)	50 (7%)

	Trust		
8	University Hospitals Birmingham NHS Trust	29 (6%)	31 (6%)
9	Oxford Radcliffe Hospitals NHS Trust	23 (10%)	24 (11%)
10	South Warwickshire General Hospitals NHS Trust	25 (10%)	25 (10%)
11	Gwent Healthcare NHS Trust	14 (11%)	15 (12%)
12	Countess of Chester Hospital NHS Foundation Trust	21 (5%)	19 (5%)
	Total	299	300

Table 35 – Characteristics of participants randomised and not randomised to the cohort study

	Randomised, n (%)	Not randomised, n (%)
Number of patients	599	3,277
Sex – Males	221 (37%)	1,456 (44%)
Age in years, Mean [SD]	40 [13]	36 [13]
Had previous neck pain	77 (13%)	334 (10%)
WAD grades		
0: No neck complaints or signs[†]	0 (0%)	0 (0%)
I: Complaints of pain, stiffness or tenderness, no physical signs	275 (46%)	1,823 (56%)
II: Complaint of pain, stiffness or tenderness, musculoskeletal signs	299 (50%)	1,375 (42%)
III: Complaint of pain, stiffness or tenderness, neurological signs	25 (4%)	79 (2%)
IV: Fracture/Dislocation[†]	0 (0%)	0 (0%)
SF-12v1 scores, Mean [SD]		
Mental component score	36 [12]	42 [13]
Physical component score	36 [7]	41 [9]

Although it was expected that all participants in Step Two would already be taking part in Step One, in fact 25 patients were randomised into Step Two before returning their two week questionnaire, and subsequently failed to return either this or their follow-up questionnaires. Hence some data is only available for 574 participants.

All 599 participants who consented for Step Two and the cohort study started their baseline research clinic questionnaire and assessment, however three participant assessments were not fully completed due to severe increase in their symptoms (n=2) and unavailability of appropriate environment to perform the ROM assessment (n=1).

22 different research clinicians performed the consenting, randomising and assessing at the 15 different hospital sites. The majority of research clinicians were experienced physiotherapists (n=20/91%) with the remainder research nurses with experience in the ED. 300 patients were randomised to receive the MINT physiotherapy package and 299 patients were randomised to receive the reinforcement of ED advice session with a physiotherapist.

7.5.2 BASELINE CHARACTERISTICS

Table 36 displays baseline demographic, pre-injury and accident characteristics. The cohort consisted of a majority of females (n= 376/63%) with a mean age of approximately 40 yrs old. The majority of participants were white (n=456/76%). There was a difference compared to Step One participant demographics (57% females and mean age of 37 years). A small minority of patients (7%) had experienced neck pain in the month prior to their injury and even smaller proportions were defined as having pre-injury chronic widespread pain (3%). 568 (95%) participants experienced a whiplash injury as a result of a Motor Vehicle Collision (MVC), the majority from an impact from the rear (n=350/62%). 80% (n=455) of them were driving. Half of the participants rated the severity of their accident as moderate, whilst over a third rate it as high or very high (n=213/37%).

Table 36- Baseline demographic, accident and pre-injury characteristics

Variable	n available for analysis	n (%) unless otherwise indicated
Age	599	Mean (SD) 39.9 (13.05)
Sex	599	M 222(37%): F 377(63%)
Ethnic group	563	White 456(81%), Mixed 5(1%), Indian 36(6%), Pakistani 37(7%), Bangladeshi 3(1%), Black or Black British 18(3%), Chinese or other 8(1%)
Previous neck pain	599	Yes 44(7%)
CWP	591	Yes 15(3%)
Injured as a result of an MVC	598	Yes 568(95%)
Location of collision	566	Rear 350(62%), drivers side 60(11%), passenger side 50(9%), front 106(19%)
Position of participant in vehicle	567	Driving 455(80%), Front seat passenger 88 (16%), Rear seat passenger 17(3%), Other 7(1%)
Rating of accident severity	568	Very low/low 68(12%), Moderate 287(51%), High/very high 213(37%)

7.5.2.1 Injury severity and neurological assessment

Three quarters of the cohort were classified as having musculoskeletal signs *and* symptoms at the research clinic (WAD grade II, n=442/74%) with the remaining participants being approximately split between the other two grades (WAD I and III).

There was a contrast to this research clinic categorisation of the participants with the grading at their original ED visit, where a greater proportion of them were classified as not having musculoskeletal signs (WAD grade I, n=265/44%).Table 37 displays WAD grade proportions at ED and Research clinic attendance. 4% of the cohort was assessed as having

neurological signs at their ED visit, whereas when they re-attended at the research clinic a few weeks later, 12 % had neurological signs and were therefore categorised as WAD grade III.

Table 37 – WAD Grade proportions at ED and RC attendances

Time point	WAD Grade I n(%)	WAD Grade II n(%)	WAD Grade III n(%)
ED attendance	265 (44)	284 (47)	25 (4)
RC attendance	84 (14)	442(74)	73 (12)

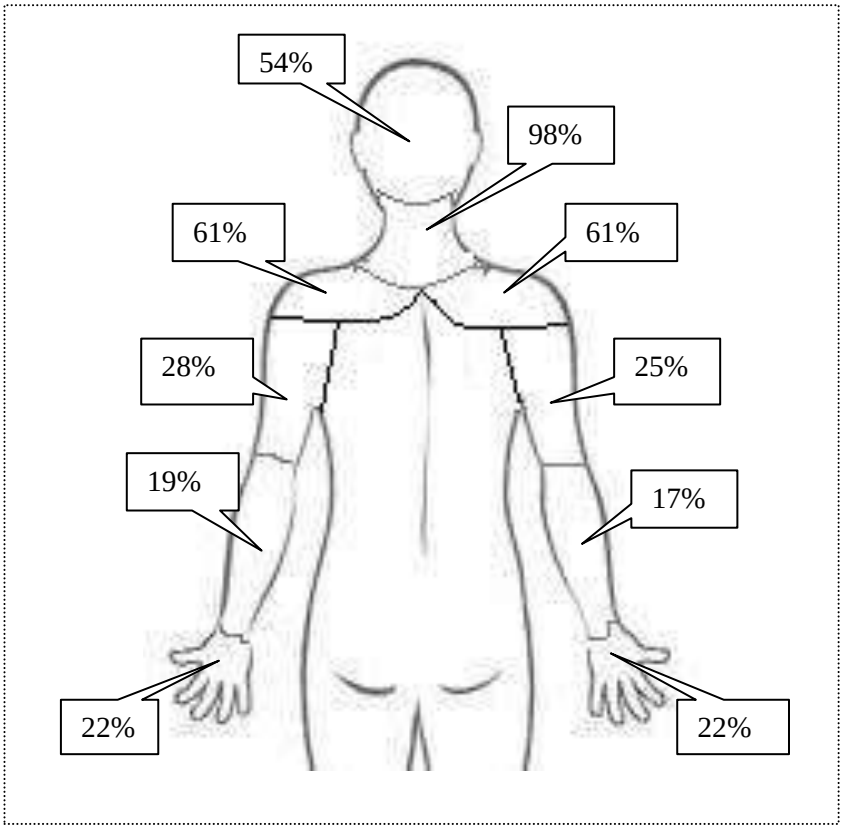
Disability, physical and psychological characteristics

Table 38 presents baseline disability, physical and psychological examination characteristics. The mean Neck Disability Index score for the cohort at baseline was 42 (on 0-100 scale) with a standard deviation of 16.2. This corresponds to the majority of participants (n=454/77%) being categorised as moderately/severely disabled when using the validated categories used by Vernon [296] and Sterling et al [168]. The majority of participants had returned to work by the time they attended the research clinic (n=361 /60%).

The Modified Von Korff pain score (average: last week's worst, average and current pain) was moderate (57.3/100), as was the rating of pain *at the time of assessment* in the RC (5.3/10).The mean number of symptoms experienced was five, most frequently in the cervical spine (n=584/98%). Figure 20 displays the proportion of participants with symptoms in each of the locations for the cervical spine and upper limbs. Similar proportions of patients had symptoms on the right side as the left, indeed there were identical numbers with symptoms in their right and left hands. Approximately one third (n=218/36%) of participants

reported experiencing neurological symptoms since their injury and three quarters reported having headache symptoms (n=438/73%).

Figure 20 - Percentage of participants experiencing symptoms in each location



The mean fear avoidance (FABQ-PA) score (mean (SD) 15 (5.6)) indicated that, on average, the cohort had fear-avoidant beliefs, with 45% scoring above the previously published cut-off of 15/24 points.

The median self-efficacy score was 4 (IQR 2) indicating a spread in beliefs about how well participants felt they could cope with their neck injury. On average, catastrophic thoughts about pain were at a low level within the cohort and impact of event scale (IES) and general

health questionnaire (GHQ-12) scores indicate moderate levels of distress and depression following their whiplash injury.

Table 38- Baseline disability, physical and psychological assessment characteristics

Variable	n available for analysis	Mean (SD) unless otherwise indicated
RTW	599	n=361 (60%)
NDI score (0-50)	548	20.8 (8.18)
NDI score (0-100)	589	41.8 (16.21)
NDI Category*	587	Recovered n=3(1%); Mild n=130(22%); Moderate/severe n=454(77%)
Pain intensity (MVK, 0-100)	593	57.3 (17.51)
Pain intensity (VAS at assessment, 0-10)	598	5.3 (2.07)
No. of Sx (0-15)	594	5.4 (2.92)
Headache symptoms	598	n=438(73%)
Neuro. Sx	599	n=218(36%)
FABQ-PA score (0-24)	585	14.7 (5.56)
SE score (0-6)	597	3.7 (1.69)
PCS (0-65)	568	17.9 (12.80)
IES score (0-75)	572	27.6 (18.16)
GHQ-12 score (0-12)	593	6.3 (3.87)

*NDI categories published by Sterling et al (2005)

7.5.2.2 Range of Motion characteristics

Table 39 presents the mean ranges of active and passive cervical ROM. Movements to right and left in coronal (rotation) and frontal (lateral flexion) planes were approximately equal for both active and passive forms of cROM. Extension cROM was greater than Flexion as would be expected for normal cROM. The Standard Deviation values for all of the cROM directions indicate that there was a considerable amount of variation in the cROM scores for the cohort, with a greater amount for the passive cROM measurements.

Table 39- Baseline Mean Ranges of Motion

Direction	Mean (SD) active cROM / º	Mean (SD) passive cROM / º
Flexion	36 (13.9)	37 (16.3)
Extension	42 (16.0)	42 (17.3)
Right Rotation	47 (15.9)	49 (18.6)
Left Rotation	47 (15.9)	49 (19.3)
Right Lateral Flexion	28 (10.4)	29 (11.9)
Left Lateral Flexion	32 (10.7)	32 (12.4)
Total cROM	233 (67.1)	238 (83.1)

Total active ROM (the sum score of all individual planes) is presented. This has been used in previous cohorts of WAD patients and shown to be a valid representation [136, 157].

There were very small differences between the active and passive forms of cervical ROM for all the different movements. On average, participants had statistically significantly greater total passive ROM (M=238, SE=3.41) than total active ROM (M=233, SE=2.74), [t (592) =-

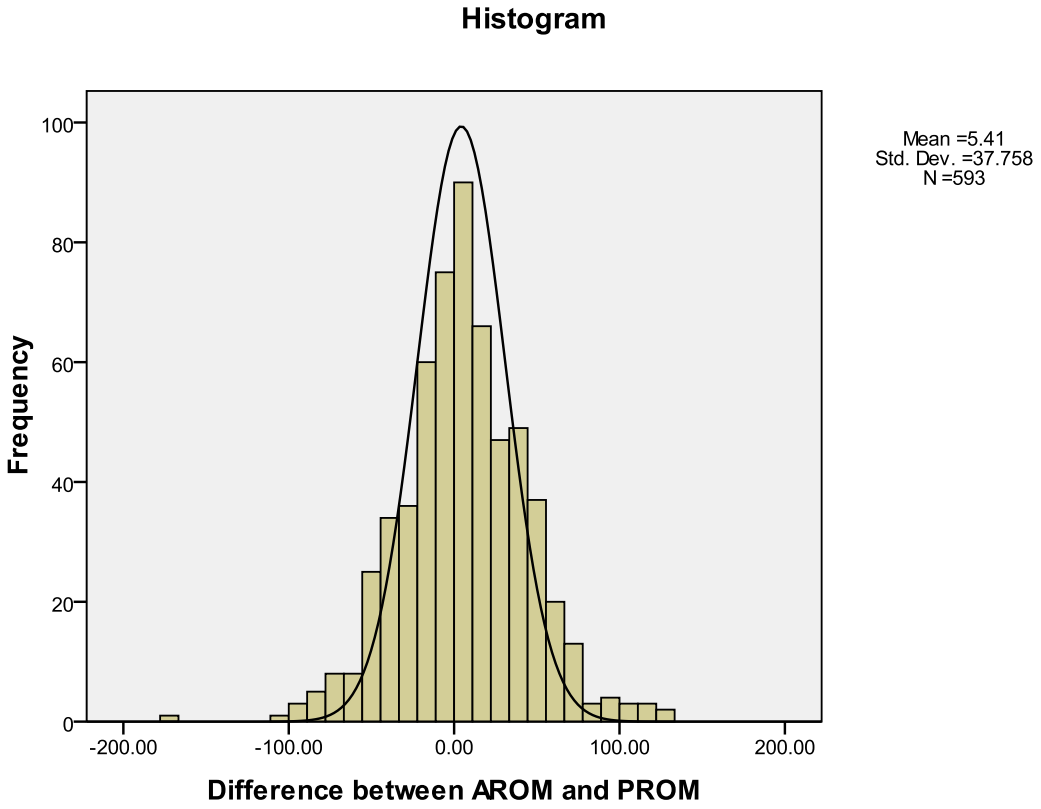
3.49, p<.01, r = 0.14], although clinically, this mean difference of 5 º is regarded as small. Table 40 presents the mean differences for each of the movements.

Figure 21 displays the histogram of distribution for the difference between the total active and passive ROM scores. A normal distribution around a mean of five degrees is shown, with approximately equal numbers of positive (more passive than active cervical ROM) and negative (more active than passive cervical ROM) values.

Table 40 - Differences in values between active and passive cervical ROM

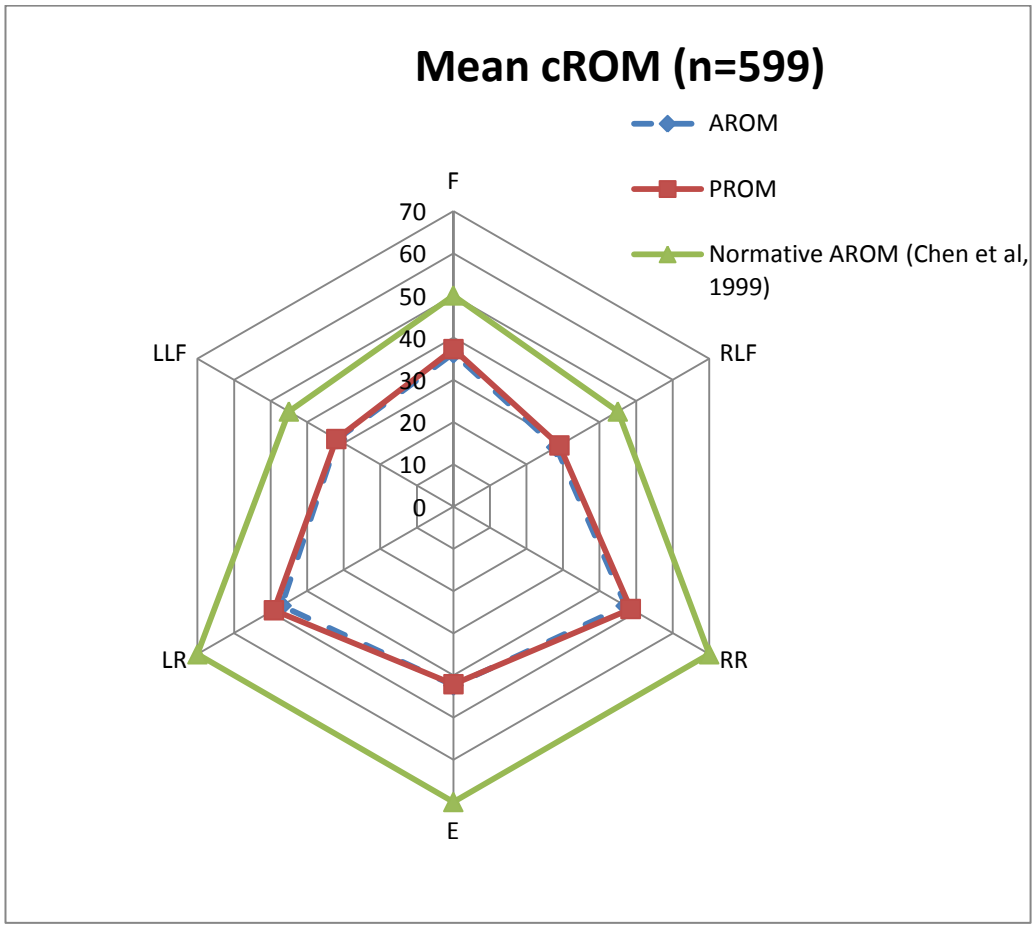
	Mean/degrees (SD)
Difference between AF and PF	1 (9.8)
Difference between AEPE	-0.2 (8.8)
Difference between ARR and PRR	1 (10.5)
Difference between ALR and PLR	2 (10.5)
Difference between ARLF and PRLF	1 (7.4)
Difference between ALLF and PLLF	0.2 (7.9)
Difference between Total AcROM and Total PcROM	5 (37.8)

Figure 21 – Histogram of difference between Total active ROM and Total passive cervical ROM scores



In comparison to average values of total cROM for normal individuals, the cohort had reduced active and passive cROM (233 vs. 360 and 238 vs. 420 degrees respectively – normative values extracted from Chen et al[101]). Mean values for each of the active and passive movements alongside these normative active cROM values are graphically presented in Figure 22. On average, the reduction in cROM is uniform for all of the planes (60-70% reduction). The greatest difference between the whiplash-injured population and the normative data is in the half-cycle measurement of extension (~30 degrees/60%).

Figure 22 - Mean cervical Range of Motion (Active cROM, Passive cROM and normative Active cROM)



Significantly reduced active cROM was defined as less than or equal to 50% of the range of normative motion; F (25), E(35), RLF (25) and LLF (25) RR (35) and LR (35) [132, 133]. Significantly reduced passive cROM was defined as less than or equal to 50% of normative cervical ROM; F (30), E(40), RR(45), LR(45), RSF(30), LSF(30). Table 41 displays the numbers of participants who had less than 50% of normative active and passive cROM. The proportion of participants with limited cROM ranged from 23-39% for active and 39 and 60% for passive cROM directions. In particular extension and right lateral flexion had the greatest proportions of participants with movement limited by at least 50%.

Table 41 – Frequency of participants with significantly reduced active and passive cervical ROM (<50%)

Direction	Pts with ≤50% normative active cROM (n[%])	Pts with ≤50% normative passive cROM (n[%])
Flexion	142(24)	234(39)
Extension	214(36)	297(50)
Right Rotation	136(23)	260(43)
Left Rotation	145(24)	260(43)
Right Lateral Flexion	233(39)	357(60)
Left Lateral Flexion	159(27)	292(49)
Total cROM	135(23)	290(48)

Table 42 presents proportions of participants by how many directions of movement were limited by at least 50%. 40% of participants had no directions of active cROM limited by more than 50%. A much greater proportion of participants had all six directions of passive movement limited when compared to active movement (22% vs. 6%).

Table 42 – Frequency of number of limitations for active and passive cervical ROM.

No. Of directions limited ≥50%	**No. of participants (%) Active cROM**	**No. Of participants (%) Passive cROM**
0	242(40)	147(25)
1	104(17)	83(14)
2	72(12)	61(10)
3	55(9)	52(9)
4	48(8)	54(9)
5	44(7)	63(11)
6	34(6)	134(22)

Participants were asked for what was their limiting their cROM immediately after each direction of movement was performed (pain, stiffness or no limitation). Results in the form of frequency counts are presented for each active and passive movement in Table 43. The most commonly reported reason for limitation of active and passive cROM was pain.

Table 43 – Reasons reported for limitation of cervical ROM by direction

	Number (%) of pts with unlimited movement		**Number (%) of pts limited by pain**		**Number (%) of pts limited by stiffness**	
Direction	**active**	**passive**	**active**	**passive**	**active**	**passive**
F	70(12)	81(14)	310(52)	379(63)	203(34)	122(20)
E	90(15)	78(13)	370(62)	420(70)	123(21)	79(13)
RR	64(11)	62(10)	347(58)	394(66)	177(30)	129(22)
LR	79(13)	69(12)	333(56)	397(66)	168(28)	116(19)
RLF	39(7)	36(6)	404(67)	450(75)	138(23)	101(17)
LL F	65(11)	59(10)	379(63)	411(69)	136(23)	109(18)
Total cROM	53(9)*	65(11)*	373(66)*	403(72)*	147(25)*	94(17)*

* Extrapolated from half cycle movements

There were differences in participants measured active and passive cROM values dependent on their reported reason for limitation of movement (Figure 23 and Figure 24). Participants reporting pain as the limiting factor had the least c ROM on average. Those participants reporting no limitation had lower cROM on average than normative values already quoted above, although these normative values are not age matched. For passive cROM there was a greater difference in mean cROM between those reporting no limitation and those reporting stiffness or pain as the limiting factor when compared to active cROM values.

Figure 23 - Mean active cervical ROM by limitation group

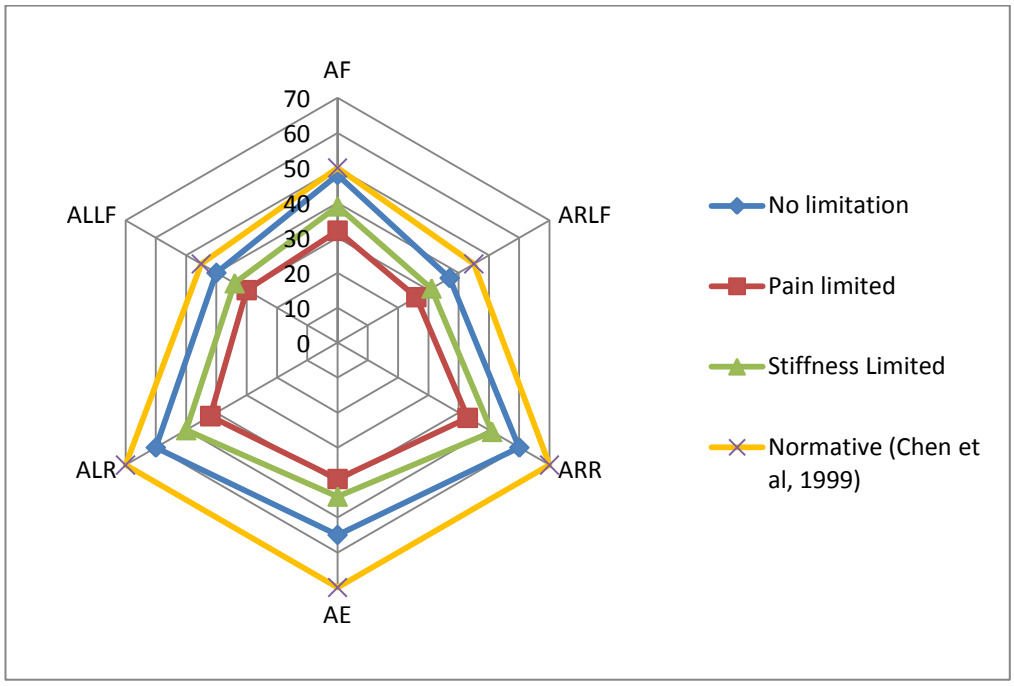
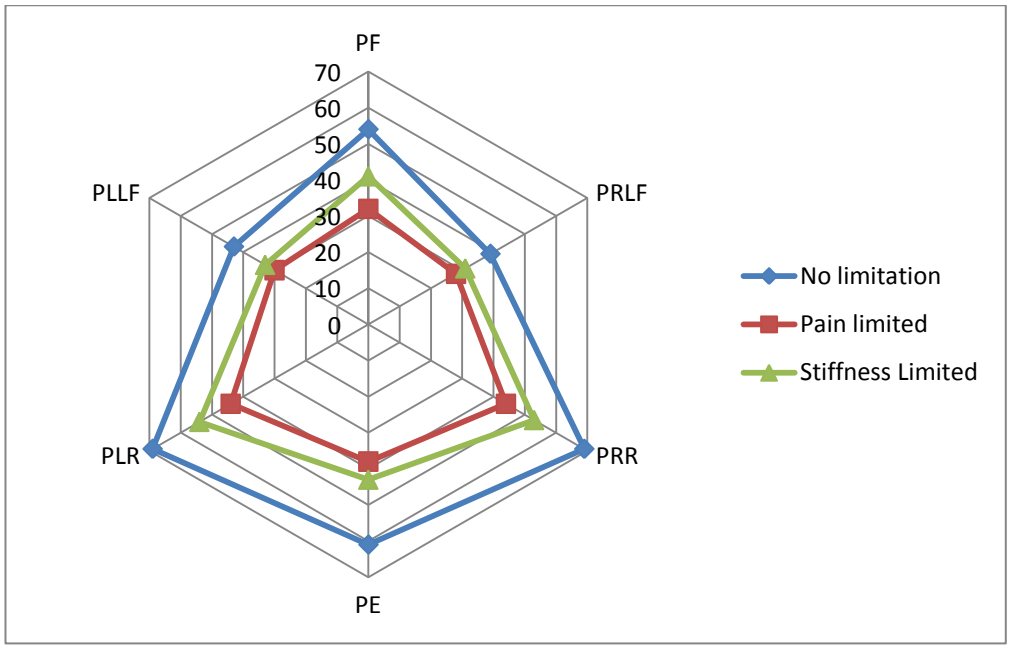


Figure 24 - Mean passive cervical ROM by limitation group



There was a significant difference in both active and passive cROM between the limitation groups according to the findings of a one-way ANOVA (Welch $F=70.6$ $p<.05$ and Brown-Forsythe $F=68.3$ $p<.05$) and results of Post-Hoc tests (Hochberg's GT2 $p<.05$) showed that participants reporting pain as the limiting factor had the significantly smallest cROM (Mean difference [95%CI] for TAROM -51.0[-71—31] and -21.8[-39.5—4.1] for comparison against no limitation and stiffness respectively).

Table 44 displays the frequencies for participant's patient-rated cROM for two planes – coronal (turning head side to side) and saggital (looking up and down). The most common rating was three (the midpoint) for both planes of movement. A very small proportion of patients rated their movements as normal (9% and 12% for Rotation and Flexion-Extension respectively).

Table 44 – Frequencies for patient-rated cervical ROM scale points (PRcROM - Likert scale from 1- unable to move as normal- to 5 – normal movement.

	1 (unable)	**2**	**3**	**4**	**5 (normal)**	**Total**
PR cROM - (rotation)	5	54	269	216	53	597
PRcROM - (flex-Ex)	8	75	239	201	74	597

Table 45 displays descriptive statistics for shoulder abduction ROM (ShAbdROM) measurements. The mean, median and variance values are very similar between right and left sides. ShAbdROM was reduced by approximately 20% in this whiplash-injured cohort compared to expected normative values [274]. This comparison is an approximation as the normative data is not age matched.

Table 46 presents frequencies of participants' reasons for ShAbdROM limitation. The most common reason for limited ShAbdROM was pain and proportions were almost identical for right and left shoulders (46 and 47% respectively). About one third of participants reported no feeling of limitation of ROM for their ShAbdROM.

Table 45 – Descriptive statistics for Shoulder Abduction ROM measurements

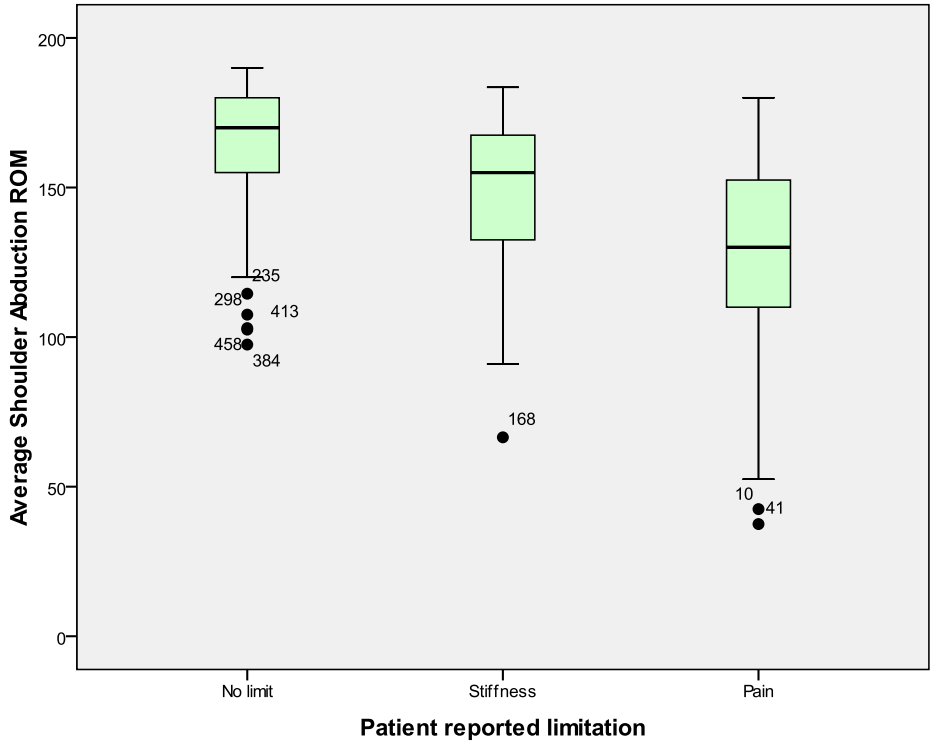
	Right Sh Abd ROM/ Degrees	Left Sh Abd ROM / Degrees
Mean	145	144
Median	155	155
Std. Deviation	32	34
25th Percentile	124	120
75th Percentile	170	170

Table 46 – Frequencies of different reasons for limitations for right and left shoulder abduction ROM

Reason for limitation	No. Of participants - Left Shoulder (%)	No. Of participants - Right Shoulder (%)
No limit	178 (30)	184 (31)
Pain	281 (47)	278 (46)
Stiffness	116 (19)	116 (19)

Sh Abd ROM varied depending on the limitation reported by the participant in the same way as clinician-measured cROM. Figure 25 displays the mean Sh Abd ROM and 95% CI for the different categories.

Figure 25 – Mean Shoulder abduction ROM and 95% CI markers for the different limitation categories.



7.6 RELATIONSHIPS FOR ROM (CROSS-SECTIONAL UNIVARIATE ANALYSIS)

7.6.1 RELATIONSHIP BETWEEN ACTIVE AND PASSIVE CROM

Table 47 presents correlations between each of the planes and their sums of movement for active and passive cROM. All active and passive cROM measurements are significantly and highly correlated with one another with the r value varying from .484 to .895. The corresponding planes correlate highly e.g. active flexion with passive flexion r=.79. The sum scores (Total Active cROM and Total Passive cROM) correlate most highly with one another (r=.893) and with their respective individual movements.

Table 47 – Correlations between half cycle cervical ROM measurements (Spearman's rho due to non-normal distribution)

Plane	AF	AE	ARR	ALR	ARLF	ALLF	PF	PE	PRR	PLR	PRLF	PLLF	TAcROM	TPcROM
AF		.532**	.543**	.558**	.519**	.516**	.790**	.494**	.497**	.484**	.505**	.495**	.752**	.624**
AE	.532**		.585**	.612**	.593**	.595**	.597**	.858**	.610**	.619**	.622**	.644**	.816**	.755**
ARR	.543**	.585**		.719**	.600**	.529**	.570**	.596**	.817**	.660**	.586**	.550**	.833**	.736**
ALR	.558**	.612**	.719**		.536**	.610**	.616**	.652**	.679**	.839**	.567**	.646**	.848**	.778**
ARLF	.519**	.593**	.600**	.536**		.671**	.580**	.591**	.609**	.556**	.785**	.604**	.772**	.695**
ALLF	.516**	.595**	.529**	.610**	.671**		.549**	.601**	.530**	.601**	.595**	.787**	.782**	.688**
PF	.790**	.597**	.570**	.616**	.580**	.549**		.656**	.663**	.656**	.683**	.653**	.765**	.827**
PE	.494**	.858**	.596**	.652**	.591**	.601**	.656**		.707**	.736**	.692**	.722**	.792**	.870**
PRR	.497**	.610**	.817**	.679**	.609**	.530**	.663**	.707**		.797**	.728**	.660**	.784**	.885**
PLR	.484**	.619**	.660**	.839**	.556**	.601**	.656**	.736**	.797**		.682**	.730**	.788**	.895**
PRLF	.505**	.622**	.586**	.567**	.785**	.595**	.683**	.692**	.728**	.682**		.736**	.741**	.846**
PLLF	.495**	.644**	.550**	.646**	.604**	.787**	.653**	.722**	.660**	.730**	.736**		.760**	.846**
TAcROM	.752**	.816**	.833**	.848**	.772**	.782**	.765**	.792**	.784**	.788**	.741**	.760**		.893**

TPcROM	.624**	.755**	.736**	.778**	.695**	.688**	.827**	.870**	.885**	.895**	.846**	.846**	.893**

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

7.6.2 RELATIONSHIP BETWEEN CLINICIAN-MEASURED AND PATIENT-RATED CERVICAL ROM

Table 48 presents correlations between clinician-measured and patient-rated cROM. Both patient-rated flexion-extension and rotation are statistically significantly associated with the active and passive clinician-measured equivalent. These correlations are small to moderate.

Table 48 – Correlation between clinician-measured and patient-rated cervical ROM (Spearman's Rho) – columns are patient-rated cervical ROM cross-tabulated with clinician-measured cervical ROM rows (apart from first 2 rows)

	PRcROM (rot)	**PRcROM (F-E)**
PRcROM (Rot)		.666**
PRcROM (F-E)	.666**	
AF	.315**	.301**
AE	.316**	.299**
ARR	.325**	.289**
ALR	.270**	.274**
ARLF	.307**	.248**
ALLF	.293**	.246**
Total AcROM	.368**	.339**
PF	.330**	.328**
PE	.345**	.331**
PRR	.309**	.281**
PLR	.274**	.285**
PRLF	.322**	.265**
PLLF	.301**	.284**
Total PcROM	.350**	.334**
AF-E	.355**	.341**
AROT	.316**	.299**
PF-E	.357**	.354**
PROT	.304**	.297**

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

Figure 26 and Figure 27 display the associations between clinician- measured cROM and patient-rated cROM.

Figure 26 – Box plot of clinician-measured cervical flexion-extension ROM for each level of patient-rated flexion-extension cervical ROM

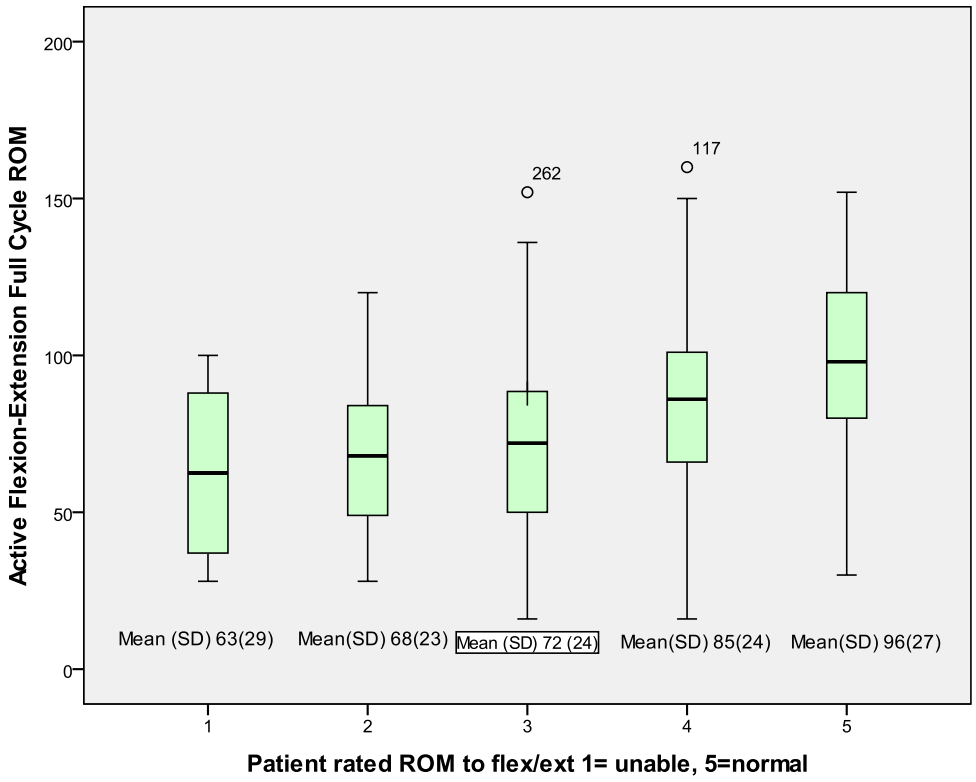
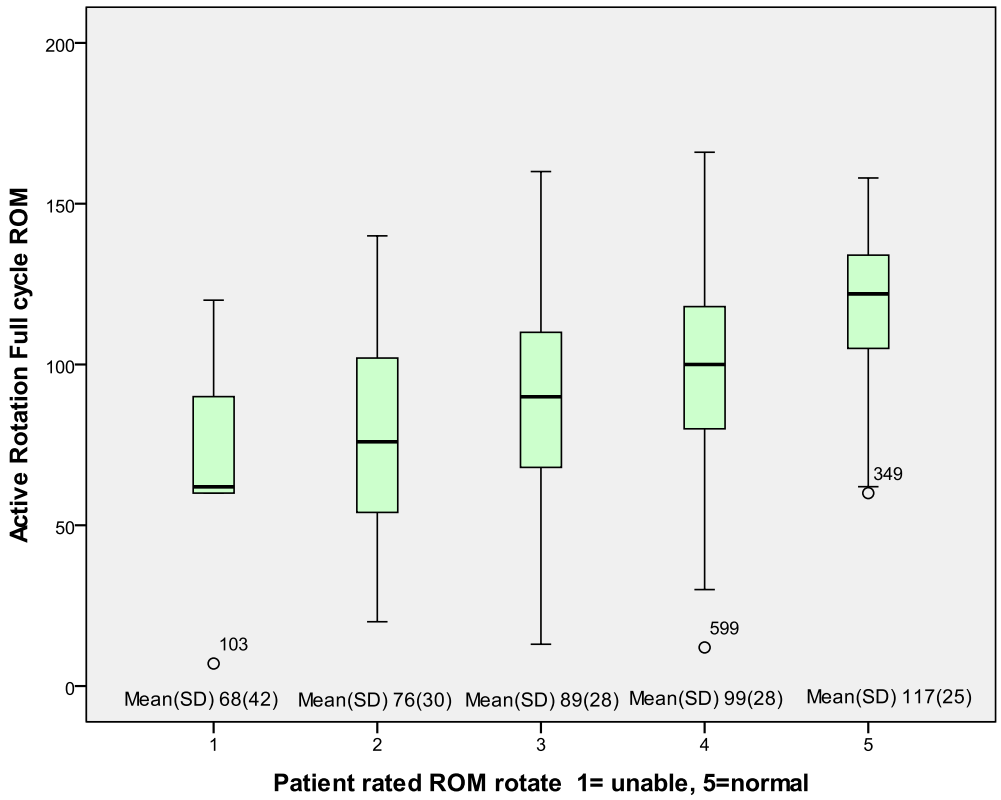


Figure 27 - Box plot of clinician-measured cervical rotation ROM for each level of patient-rated cervical ROM



Those participants reporting that their cROM was unlimited for the clinician-measured rotation and flexion-extension also reported their subjective cROM as normal (Chi-squared statistic significant at $p<.01$ level for ARR limitation vs. PRcROM (ROT) and AF limitation vs. PRcROM (F-E)).

7.6.3 RELATIONSHIP BETWEEN CERVICAL ROM AND SHOULDER ABDUCTION ROM

Table 49 displays the correlation (Spearman's rho) between the various measures of cROM and average Sh AbdROM. Sh Abd ROM correlated at a statistically significant level with all of the cROM measures (clinician-measured and patient-rated), most strongly with total active and passive cROM.

Table 49 – Spearman's Rho correlations between the various measures of cervical and shoulder ROM

Direction	Right Sh Abd ROM	Left Sh Abd ROM
AF	.332**	.398**
AE	.374**	.415**
ARR	.360**	.362**
ALR	.329**	.426**
ARLF	.366**	.390**
ALLF	.319**	.371**
Total AcROM	.420**	.481**
PF	.388**	.420**
PE	.380**	.426**
PRR	.426**	.410**
PLR	.380**	.442**
PRLF	.386**	.387**
PLLF	.375**	.414**
Total PcROM	.444**	.475**

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

7.6.4 RELATIONSHIP BETWEEN ROM AND OTHER BASELINE FACTORS

Results of independent t-tests and Mann-Whitney U tests are presented in Table 50 and correlations in Table 51 for the baseline cross-sectional, univariate analyses. On average there were no significant differences in active or passive cROM between the different sexes, ethnicities, ED advice groups (MINT Step 1) the various crash factors, and previous neck pain.

Table 50 - Comparisons of means/medians for the various ROM measures with significance values for independent t-tests/Mann-Whitney U tests

Group comparison	**Categories (n)**	**mean (SD) Total AcROM**	**mean (SD) Total PcROM**	**Av ShAbdROM**	**median(IQR) PRcROM (rot)**	**median(IQR) PRcROM (F-E)**
Demographics						
Sex	Male(219) vs. Female(376)	226(65.3)vs237(67.9)	231(81.1)vs242(84.0)	145(25.8)vs144(31.2)	3(1)vs3(1)	3(1)vs3(1)
Ethnicity	White(455) vs. remaining(106)	235(66.9)vs226(69.2)	243(82.3)vs222(86.4)	147(28.6)vs137(30.2)**	3(1)vs3(1)**	3(1)vs3(1)**
Trial Treatments						
Step 1 treatment	Usual care advice(271) vs. Whiplash book advice(326)	228(65.0)vs237(68.6)	233(80.3)vs243(85.0)	149(28.8)vs141(29.2)**	3(1)vs3(1)	3(1)vs3(1)
Step 2 treatment	Advice session(299) vs. Physiotherapy package	244(60.5)vs221(71.4)*	252(75.4)vs224(88.1)*	148(27.0)vs141(31.1)**	3(1)vs3(1)	4(1)vs3(1)

Pre-injury factors						
Previous neck pain	Yes(44) vs. No(511)	243(71.1)vs233(67.2)	260(76.1)vs238(84.2)	144(32.1)vs145(2.90)	3(1)vs3(1)	3(1)vs3(1)
CWP	Yes(15) vs. No(573)	175(75.2)vs234(66.6)**	194(74.9)vs239(83.4)*	122(40.1)vs145(28.9)*	3(1)vs3(1)	3(2)vs3(1)
Accident						
MVC	Yes(565) vs. No(30)	232(67.0)vs244(69.0)	238(83.1)vs251(81.5)	144(29.5)vs144(25.5)	3(1)vs4(1)	3(1)vs3(1)
Direction of impact	Rear(349) vs. other(247)	232(65.4)vs233(69.5)	239(82.0)vs238(84.7)	145(29.4)vs144(29.1)	3(1)vs3(1)	3(1)vs3(1)
Position in car	Driving(454) vs. other(142)	233(67.1)vs231(67.1)	239(83.0)vs235(83.4)	145(29.1)vs142(29.7)**	3(1)vs3(1)	3(1)vs3(1)
Injury Severity						
RC WAD Grade	WAD I(84) vs. II/III(512)	272(61.2)vs226(65.9)**	294(71.2)vs229(81.3)**	156(24.7)vs142(29.5)**	4(1)vs3(1)**	4(2)vs3(1)**
Physical Measures						

Headache	No(159) s Yes(436)	241(65.1)vs229(67.6)	250(81.6)vs234(83.3)*	151(24.6)vs142(30.5)**	4(1)vs3(1)*	4(1)vs3(1)*
Neuro. Sx.	No(381) vs. Yes(215)	239(64.7)vs220(69.5)**	245(82.6)vs227(82.9)*	148(28.2)vs137(30.0)**	3(1)vs3(1)**	4(1)vs3(1)**
Disability						
RTW	No(123) vs. Yes(359)	213(70.0)vs244(62.5)**	214(82.3)vs252(80.1)**	137(33.5)vs150(26.5)**	3(1)vs3(1)**	4(1)vs3(1)**

*. T-test/Mann Whitney U test is significant at the 0.05 level (2-tailed).
**. T-test/Mann Whitney U test is significant at the 0.01 level (2-tailed).

Table 51 – Correlations between the various baseline ROM measurements and baseline factors

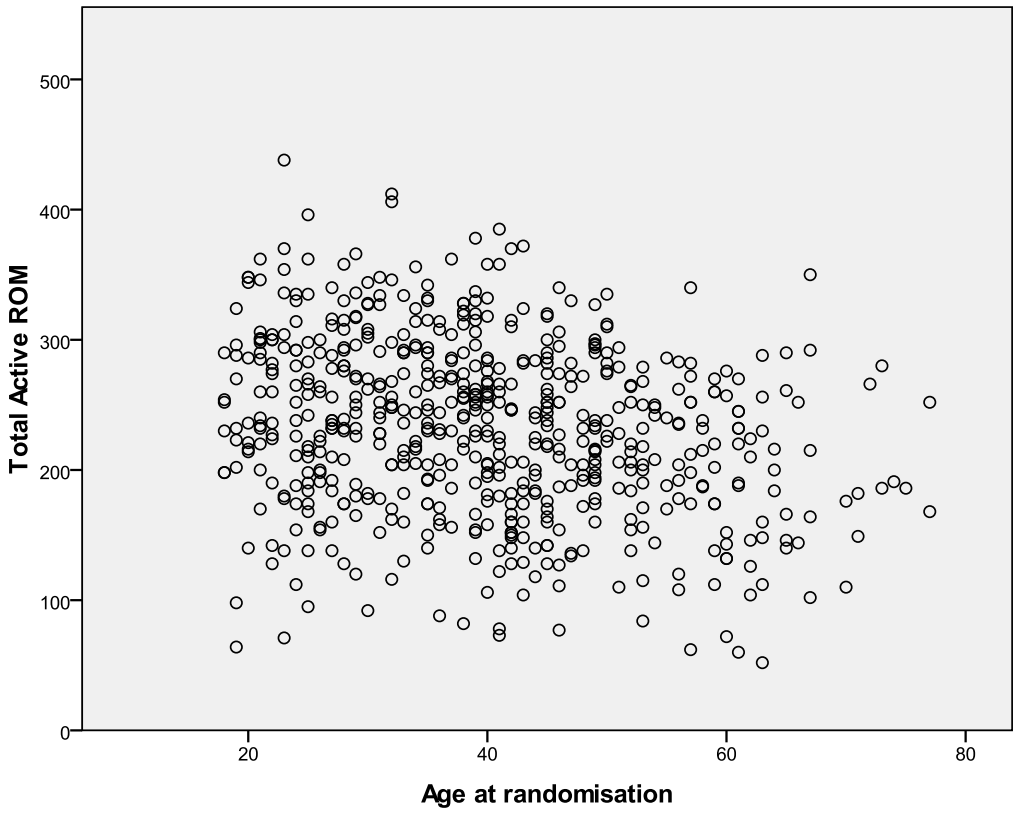
Baseline factor	Total AcROM	Total PcROM	Av Sh Abd ROM	PRcROM (rot)	PRcROM (F-E)
Pain intensity today (0-100)	-.344**	-.361**	-.296**	-.317**	-.296**
Pain intensity (MVK, 0-100)	-.358**	-.388**	-.324**	-.343**	-.311**
SE (0-6)	.189**	.223**	.119**	.258**	.251**
PCS (0-52)	-.321**	-.353**	-.302**	-.316**	-.277**
FABQ-PA (0-24)	-.211**	-.255**	-.182**	-.220**	-.190**
IES (0-75)	-.236**	-.251**	-.280**	-.202**	-.171**
GHQ-12 (0-12)	-.275**	-.289**	-.224**	-.253**	-.215**
Right ShAbdROM	.420**	.444**	.891**	.243**	.215**
Left ShAbdROM	.481**	.475**	.906**	.209**	.210**
No.of Sx. (0- 15)	-.191**	-.188**	-.305**	-.193**	-.171**
NDI Score (0-5)	-.462**	-.479**	-.435**	-.381**	-.359**
NDI Score (0-100)	-.454**	-.473**	-.434	-.369**	-.355**

Age at RC	-.242**	-.165**	-.142	-.041	-.042

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

There was a statistically significant correlation between age and clinician-measured cROM, with the strongest correlation for total active cROM (r=-.242). There was no significant correlation between patient-rated cROM and age. This was the only variable where clinician-measured cROM did not have the same findings as the patient-rated cROM versions. Figure 28 displays total active cROM plotted against age to illustrate the correlation.

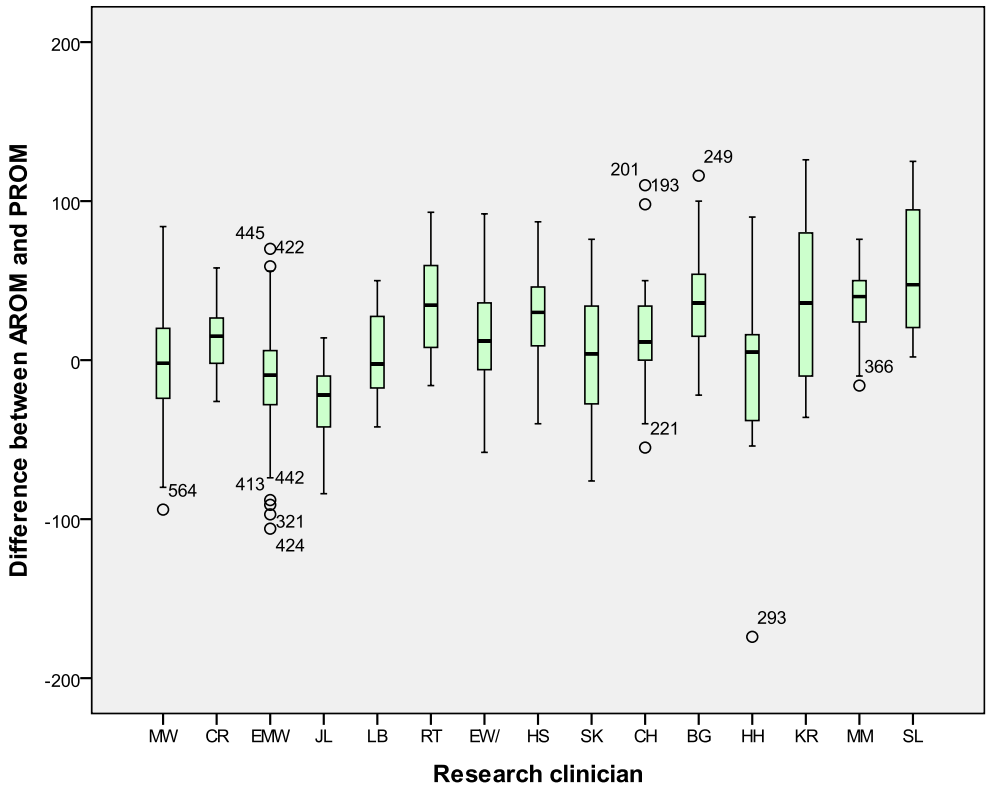
Figure 28 – Scatter plot for total active cervical ROM against age



On average, participants randomised to receive the physiotherapy arm had lower cROM than those randomised to the advice session.

Some sites had more than one research clinician performing clinic assessments and therefore cROM is presented by research clinician. There were differences in the mean cROM measured by the research clinicians. The differences were broadly similar for both active and passive cROM. Figure 29 displays the difference between mean total active and passive cROM values for each research clinician.

Figure 29 – Box plot of difference between total active and passive cervical ROM (passive minus active) values by research clinician



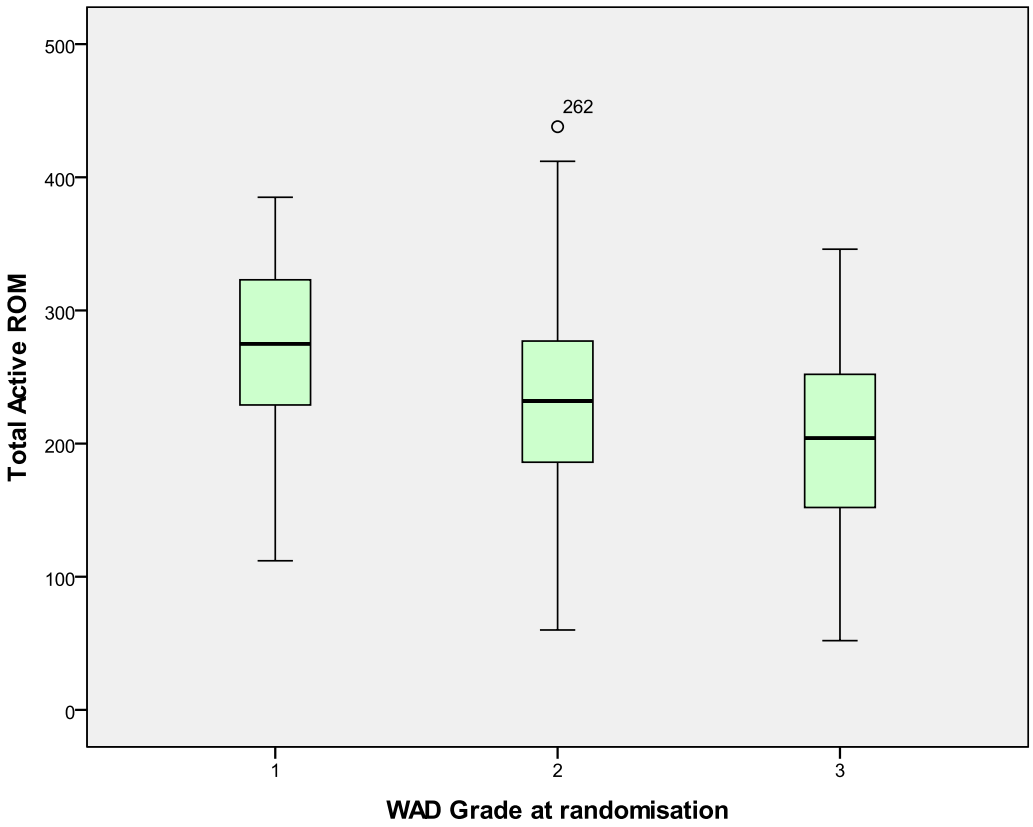
There was, on average, a statistically significant difference in total active cROM and total passive cROM for those that were classified as having chronic widespread pain prior to their whiplash injury compared to those that did not.

There was a trend for reduced cROM as rating of injury severity increased (represented by WAD grade) as Figure 30 illustrates.

Even WAD grade I had reduced cROM - patients classed as not having musculoskeletal or neurological signs - when compared to normative values (Mean TAcROM = 272 degrees, mean normative TAcROM from Chen[101] = 360 degrees).

On average, participants with musculoskeletal signs and symptoms (WAD grade II/III) had lower total active and passive cROM than participants who had just symptoms (WAD grade I).

Figure 30 – Box plot of total active cervical ROM by WAD grade



All physical assessment findings significantly correlated with cROM measurements. Rating of pain had a greater strength of association than the number of symptoms present. All psychosocial assessment findings correlated with cROM measurements. Factors in descending order of strength of association were catastrophising (PCS), depression (GHQ-12), distress (IES), Fear avoidance beliefs (FABQ-PA) and self-efficacy (SE) (see Table 51).

7.6.5 RELATIONSHIP BETWEEN ROM AND BASELINE DISABILITY

All types of ROM measurements (clinician-measured and patient-rated cROM and ShAbdROM) significantly correlated with baseline NDI scores and with moderate strength of association (ranging from r=-.355 to -.473). This negative correlation shows lower ROM correlates with a more disabled state. Figure 31 provides a graphical example of these relationships (r=-.454 p<.01). For active and passive cROM measurements the number of limited directions significantly correlates (r=.438, p<.01) with the disability score as illustrated in Figure 32.

Figure 31 – Scatter plot of total active cROM plotted against baseline NDI score

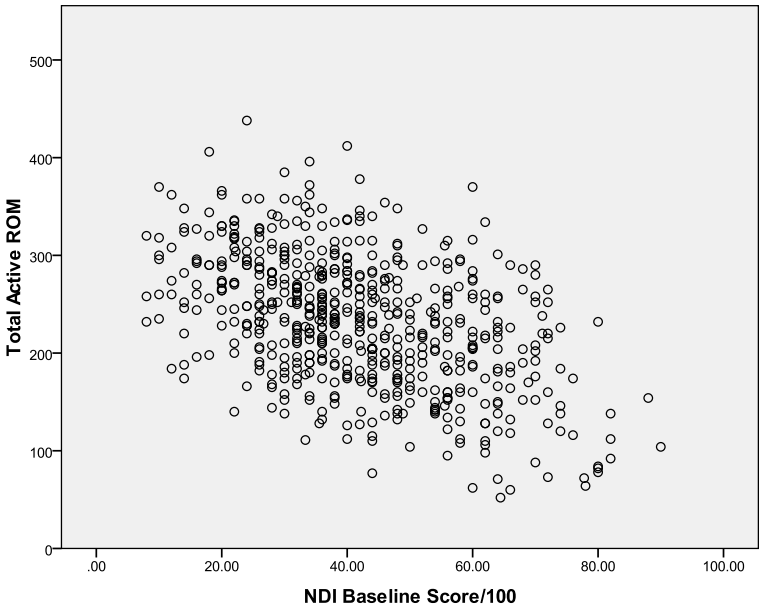
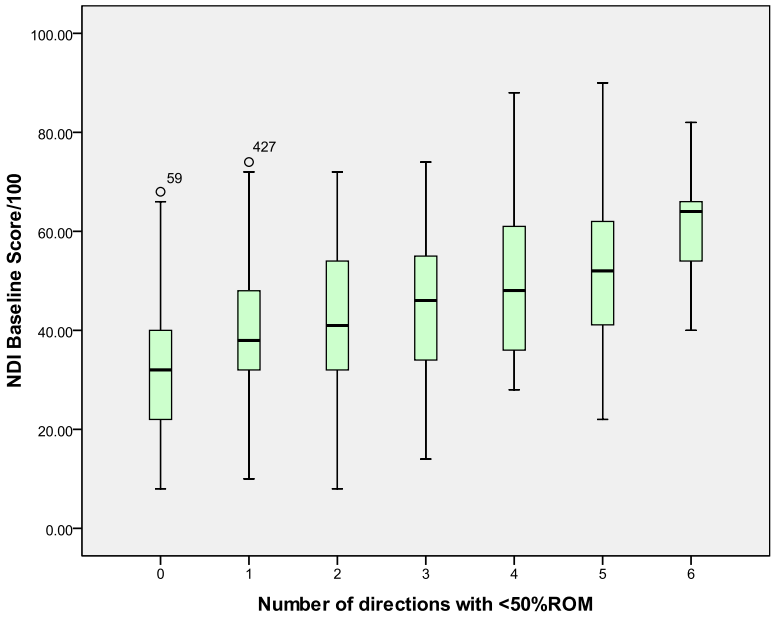
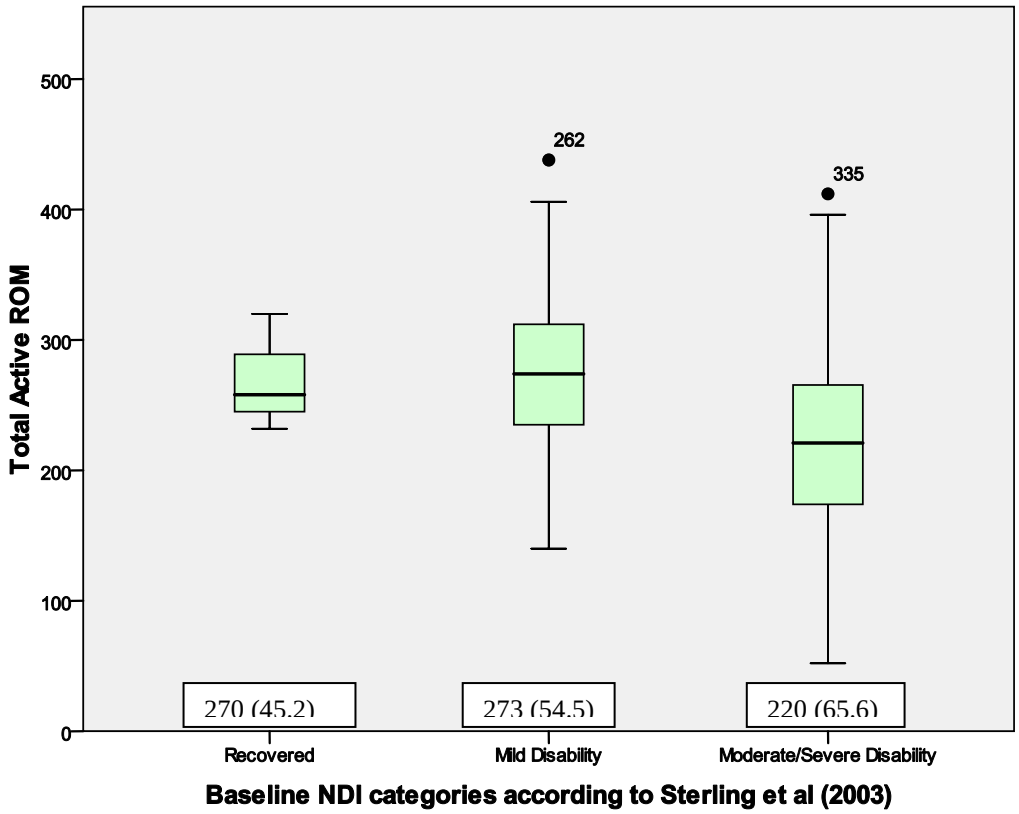


Figure 32 – Box plot of Baseline NDI against number of limited active cROM directions



Those categorised as having mild disability had greater total active (and passive) cROM (Total active cROM Mean (SD) 273 (54.5)) than those categorised as having moderate to severe disability (Total active cROM Mean (SD) 220 (65.6)) as illustrated in Figure 33.

Figure 33 – Box plot of total active cervical ROM by NDI categories according to Sterling [86]



On average, active and passive cROM were significantly higher for participants who had returned to work by the time they were assessed at the RC compared to those who had yet to return to work (M=244 vs. 212, t(479) =4.36, p<.05 and M=252vs214, t(478)=4.50,p<.05). There was no difference in the median value (3/5) of patient-rated cROM rotation between returners and non-returners. When patient-rated cROM was dichotomised into normal and

non-normal the Chi-squared test was statistically significant ($p<.05$) for both rotation and flexion-extension ratings.

7.7 MULTIVARIATE CROSS-SECTIONAL ANALYSES FOR PREDICTING BASELINE CERVICAL ROM

Physical and psychological factors that were univariately associated ($p<.05$) with the various cROM measures (clinician-measured and patient-rated) were entered into a multiple regression models. These factors are asterisked in Table 51 and 52. Table 52 to Table 55 display the results for model summaries and multiple regression results for cross sectional clinician-measured cROM analyses.

The final models explain 30% and 33% of variance for total active and passive cROM respectively. The same seven factors were present for both of these forms of cROM (pain intensity, age, Fear-avoidance beliefs questionnaire score, WAD grading, self-efficacy question score, treatment allocation and general health questionnaire score), although the importance of the factors varied between them. For both forms of cROM pain intensity explained the most variance (15% of the overall variance; so 50% of the variance explained by the model)

For patient-rated cROM, the final model included pain intensity, self-efficacy score, presence of neurological symptoms and Fear Avoidance Beliefs Questionnaire Score, explaining 17% of the overall variance of this ROM measure. Table 56 and Table 57 display the results for this analysis. Pain intensity accounted for the majority of the explanatory power of the model (11/17 %).

The final models for all of the cROM measures are a significant fit of the data (ANOVA $p<0.001$). It is likely that errors are independent as the Durbin-Watson statistics are very close to 2. Collinearity is highly unlikely to be a problem as all VIF values are under 2. Casewise diagnostics were all within the limits of normal (approximately 5% of cases were

outside the limit of ±2 Standardised Residuals, Cook's distances were all less than .05, average leverage values were all below three times the mean, Mahalanobis' distances were all below 25 and DFBeta values were all greater than one). The assumption that variance was homogeneous was not violated according to visual inspection of the scatterplots ZRESID vs. ZPRED (there was a random distribution). The histograms of residuals displayed normal distributions in the form of bell-shaped curves and the corresponding P-P plots showed straight positive diagonal lines.

Table 52 – Model summaries for cross-sectional active cervical ROM model (forward stepwise method)

Step	Variables	R Square	Adjusted R Square	Durbin-Watson
1	(Constant), Pain intensity,	0.153	0.152	
2	(Constant), Pain intensity, Age	0.208	0.205	
3	(Constant), Pain intensity, Age, FABQ-PA	0.241	0.237	
4	(Constant), Pain intensity, Age, FABQ-PA, WAD Grade	0.266	0.261	
5	(Constant), Pain intensity, Age, FABQ-PA, WAD Grade, SE	0.284	0.277	
6	(Constant), Pain intensity, Age, FABQ-PA, WAD Grade, SE, Treatment allocation	0.297	0.288	
7	(Constant), Pain intensity, Age, FABQ-PA, WAD Grade, SE, Treatment allocation, GHQ-12	0.302	0.293	2.046

Table 53 - Multiple regression results for the final cross-sectional active Range of Motion model (forward stepwise method)

Variable	***B***	***SE* B**	**β**
(Constant)	395.83	15.86	
Pain intensity	-0.92	0.16	-0.24
Age	-1.46	0.2	-0.28
FABQ-PA	-1.77	0.49	-0.15
WAD grade	-31.56	7.1	-0.17
SE	4.91	1.61	0.12
Treatment allocation	-15.75	5.03	-0.12
GHQ-12	-1.46	0.72	-0.08
			R^2=.302, p<.001

Table 54 - Model summaries for cross-sectional passive Range of Motion model (forward stepwise method)

Step	Variables	R Square	Adjusted R Square	Durbin-Watson
1	(Constant), Pain intensity,	0.179	0.177	
2	(Constant), Pain intensity, WAD grade	0.221	0.218	
3	(Constant), Pain intensity, WAD Grade, Age	0.248	0.244	
4	(Constant), Pain intensity, WAD Grade, Age, FABQ-PA	0.284	0.278	
5	(Constant), Pain intensity, WAD Grade, Age, FABQ-PA, Self-efficacy	0.307	0.3	
6	(Constant), Pain intensity, WAD Grade, Age, FABQ-PA, Self-efficacy, Treatment allocation	0.32	0.312	
7	(Constant), Pain intensity, WAD Grade, Age, FABQ-PA, Self-efficacy, Treatment allocation, GHQ-12	0.327	0.318	1.9

Table 55 - Multiple regression results for the final cross-sectional passive Range of Motion model (forward stepwise method)

Variable	*B*	*SE* B	β
(Constant)	437.44	19.54	
Pain intensity	-1.21	0.20	-0.25
WAD grade	-48.2	8.76	-0.2
Age	-1.39	0.25	-0.21
FABQ-PA	-2.39	0.61	-0.16
SE	7.14	1.98	0.14
Treatment allocation	-19.95	6.18	-0.12
GHQ-12	-2.06	0.88	-0.1
			R^2=.327, p<.001

Table 56 - Model summaries for cross-sectional patient-rated Range of Motion model (forward stepwise method)

Step	Variables	R Square	Adjusted R Square	Durbin-Watson
1	(Constant), Pain intensity,	0.107	0.106	
2	(Constant), Pain intensity, SE	0.144	0.141	
3	(Constant), Pain intensity, SE, Neuro. Sx.	0.157	0.152	
4	(Constant), Pain intensity, SE, Neuro. Sx., FABQ-PA	0.168	0.162	1.99

Table 57 - Multiple regression results for the final cross-sectional patient-rated Range of Motion model (forward stepwise method)

Variable	*B*	*SE* B	β
(Constant)	4.01	0.17	
pain intensity	-0.01	0.00	-0.23
Self-efficacy	0.09	0.02	0.18
Neuro. Sx.	-0.2	0.07	-0.12
FABQ-PA	-0.02	0.01	-0.12
			R^2=.168, p<.001

7.8 MULTIVARIATE CROSS-SECTIONAL ANALYSES FOR PREDICTING BASELINE DISABILITY

Results are presented for the two different models constructed with different groups of independent variables. The 'research clinic' model included all factors that had a significant univariate relationship ($p<.05$) with the dependent variable and the 'typical clinical assessment' model included only a limited number of these significantly associated factors that would be available to clinicians in the current typical clinical setting.

7.8.1 FACTORS SELECTED FOR THE "RESEARCH CLINIC ASSESSMENT" MODEL

It has already been demonstrated that all ROM variables significantly correlate with one another and that the cROM sum scores (total active and passive cROM) correlate most highly with the individual half-cycle measurements (Table 47).

The results of the univariate correlations (Spearman's due to non-normal distribution of NDI score) between cervical and ShAbdROM variables and baseline NDI scores are presented in Table 58. For the clinician-measured cROM values total active and total passive cROM correlated most strongly with the baseline NDI score. The average ShAbdROM score also correlated significantly. For the patient-rated cROM values, the rotation version correlated most strongly with the NDI scores. The clinician-measured cROM values correlate more strongly than the patient-rated cROM values. It was therefore decided to carry total active cervical ROM, total passive cervical ROM, patient-rated cervical rotation ROM and AvShAbdROM into the cross-sectional analyses.

Table 58 – Correlations between ROM measurement variables and baseline NDI scores

	Correlation coefficient with NDI baseline score (/100)
Patient rated cervical rotation ROM	-.369**
Patient rated cervical flex-ext ROM	-.355**
AF	-.385**
AE	-.425**
ARR	-.380**
ALR	-.368**
ARLF	-.368**
ALLF	-.319**
Total l AcROM	-.454**
PF	-.462**
PE	-.437**
PRR	-.432**
PLR	-.406**
PRLF	-.368**
PLLF	-.349**
Total PcROM	-.473**
Av ShAbdROM	-.434**
No. of limited directions	-4.38**

As described in the statistical analysis methods section, previous literature was used to limit which variables (other than ROM ones) would be evaluated for univariate and then multivariate analyses. Firstly the factors identified in the previous literature were analysed for univariate associations with the baseline NDI score (Spearman's correlations for

continuous measures and Mann-Whitney U tests for categorical variables due to non-normal distributions). The results of these analyses are presented in Table 59 and Table 60. The following variables were significantly correlated or associated with the baseline NDI score (and therefore were entered into the multivariate model):

Total active cROM, total passive cROM, AvShROM, patient-rated cROM rotation, MVK pain score, No. of Sx., coping, IES score, GHQ-12 score, FABQ score, PCS score, presence of headache, presence of Neuro. sx., treatment allocation, previous history of CWP and WAD grade.

Table 59 – Correlations between baseline factors and baseline NDI score

	Correlation coefficient with Baseline NDI score (/100)
Age	.070
Pain intensity (MVK, 0-100)	.692**
No. of Sx(0-15)	.381**
Coping (0-6)	-.253**
IES (0-75)	.405**
GHQ-12 (0-12)	.548**
FABQ-PA (0-24)	.330**
PCS (0-52)	.569**

**Correlation is significant at 0.01 level (two-tailed)

Table 60 – Results of Mann Whitney U tests for baseline categorical variables vs. baseline NDI score

Variable	n	Mann Whitney U	Z	Sig	Effect size r
Sex	589	39291.5	-0.42	0.675	-0.017
Headache	588	19597	-7.876	0.000	-0.325
Neuro. Sx.	589	29541	-5.296	0.000	-0.218
Treatment Allocation	589	36401.5	-3.374	0.001	-0.139
Pre-injury CWP	581	3026.5	-1.9	0.057	-0.079
RC WAD Gd (I vs. II/III)	589	12770	-5.73	0.000	-0.236
Previous Neck Pain	548	10997.5	-0.9	0.928	-0.038

7.8.2 RESULTS FOR "RESEARCH CLINIC ASSESSMENT" MODELS

Results for the cross-sectional (baseline) multiple regression analyses (using the forward stepwise method) are presented in Table 61 and Table 62. Nine factors were independently predictive of baseline NDI score, explaining 69% of the variance in total. Initial pain intensity (MVK score) explained the largest amount of variance (51%), followed by General Health Questionnaire-12 score (an additional 9%). Average shoulder abduction ROM was the next most important predictive factor explaining an additional 4% of the variance independent of any of the other variables. Presence of headache explains an additional 2% of the variance and the rest of the variables explain less than an additional 1% of the variance each (which included total active cervical ROM). The final model is a significant fit of the data (ANOVA $p<0.001$). It is likely that errors are independent as the Durbin-Watson statistic is very close to 2 (2.05). Collinearity is highly unlikely to be a problem for this model as all VIF values are under 2. Casewise diagnostics were all within the limits of normal (approximately 5% of cases were outside the limit of ±2 Standardised Residuals, Cook's distance were all less than .05, average leverage values were all below three times the mean, Mahalanobis' distance were all below 25 and DFBeta values were all greater than one). The assumption that variance was homogeneous was not violated according to visual inspection of the scatter plot ZRESID vs. ZPRED (there was a random distribution). The histogram of residuals displayed a normal distribution in the form of a bell-shaped curve and the corresponding P-P plot showed a straight positive diagonal line. When the analysis was re-run using a backwards stepwise method (results not shown) the model included the same variables and explained exactly the same amount of total variance (69%). When clinician-measured ROM variables were not entered into the model (i.e. subtracting AvShAbdROM and total active cROM), the FABQ-PA score, patient-rated cROM rotation and pre-injury

CWP replaced them, with this version of the final model explaining 67% of the variance (results not shown).

Table 61 – Model summaries for cross-sectional disability model (forward stepwise method)

Step	Variables	R Square	Adjusted R Square	Durbin-Watson
1	(Constant), Pain intensity	.507	.506	
2	(Constant), Pain intensity, GHQ-12	.601	.600	
3	(Constant), Pain intensity, GHQ-12, Av Sh Abd ROM	.637	.635	
4	(Constant), Pain intensity, GHQ-12, Av Sh Abd ROM, Headache	.652	.649	
5	(Constant), Pain intensity, GHQ-12, Av Sh Abd ROM, Headache, PCS	.664	.661	
6	(Constant), Pain intensity, GHQ-12, Av Sh Abd ROM, Headache, PCS, Total AcROM	.674	.670	
7	(Constant), Pain intensity, GHQ-12, Av Sh Abd ROM, Headache, PCS, Total AcROM, , WAD Grade	.679	.674	
8	(Constant), Pain intensity, GHQ-12, Av Sh Abd ROM, Headache, PCS, Total AcROM, , WAD Grade, Treatment Allocation	.683	.678	
9	(Constant), Pain intensity, GHQ-12, Av Sh Abd ROM, Headache, PCS, Total Active cROM, , WAD Grade, Treatment Allocation, No. of sx.	.687	.681	2.047

Table 62 – Multiple regression results for the final cross-sectional disability model (forward stepwise method)

Variable	*B*	*SE* B	β
(Constant)	16.39	3.65	
Pain intensity (MVK)	0.38	0.03	0.41
GHQ-12	1.03	0.13	0.24
Av Sh Abd ROM	-0.06	0.02	-0.11
Headache	3.84	1.02	0.10
PCS	0.16	0.04	0.13
Total A cROM	-0.02	0.01	-0.10
WAD Grade	3.43	1.21	0.07
Treatment Allocation	2.30	0.84	0.07
No. Of Sx.	0.39	0.17	0.07
			R^2= .687, p<.001

7.8.3 FACTORS SELECTED TO BE ENTERED INTO "TYPICAL CLINICAL ASSESSMENT"

The following variables were entered into the "typical clinical assessment" model as they can be collected in a routine therapy assessment session and were significantly associated or correlated with baseline NDI score:

Total active and passive cROM, patient-rated cROM rotation, AvShAbdROM, pain intensity (VAS), No. of Sx., presence of headache, presence of neuro. symptoms, treatment allocation, previous history of CWP, WAD grade.

Previous neck pain, sex and age were not significantly associated/ correlated with baseline NDI scores and so were not entered into the model.

7.8.4 RESULTS FOR "TYPICAL CLINICAL ASSESSMENT" MODELS

Results for the cross-sectional (baseline) multiple regression analyses (using forward stepwise method) are presented in Table 63 and Table 64.

Eight variables make up the final model which explains 59% of the variance of baseline disability. Pain intensity (VAS score) at the time of assessment explained the majority of the variance (44%). The next most important variable is total passive cROM, independently explaining a further 7%. Of the remaining six variables (headache, Av. ShAbdROM, No. of Sx., patient-rated cROM rotation, treatment allocation and WAD grade) the two cROM-related variables provide an additional 2% and 1% respectively.

The final model is a significant fit of the data ($p<0.001$). It is likely that errors are independent as the Durbin-Watson statistic is very close to 2. Collinearity is highly unlikely to be a problem for this model as all VIF values are under 2. Casewise diagnostic tests and

tests of assumptions were all within recommended limits, indicating a well-fitted model. When clinician-measured ROM factors were not entered into the model, patient-rated cervical rotation ROM replaces them, with this version of the final model explaining 54% of outcome variance.

Table 63 – Model summaries for "typical clinical assessment" cross-sectional disability model (forward stepwise method)

Step	Variables	R Square	Adjusted R Square	Durbin-Watson
1	(Constant), Pain intensity	.443	.442	
2	(Constant), Pain intensity, total P cROM	.512	.510	
3	(Constant), Pain intensity, total P cROM, headache	.549	.546	
4	(Constant), Pain intensity, total P cROM, headache, Av Sh Abd ROM	.567	.564	
5	(Constant), Pain intensity, total P cROM, headache, Av Sh Abd ROM, No. Of Sx.	.577	.573	
6	(Constant), Pain intensity, total P cROM, headache, Av Sh Abd ROM, No. Of Sx, PRcROM rot	.581	.577	
7	(Constant), Pain intensity, total P cROM, headache, Av Sh Abd ROM, No. Of Sx., PRcROM rot, Treatment allocation	.585	.580	
8	(Constant), Pain intensity, total P cROM, headache, Av Sh Abd ROM, No. Of Sx, PRcROM rot, Treatment allocation, WAD grade	.589	.583	1.915

Table 64 – Multiple regression results for the final "typical clinical assessment" cross-sectional disability model (forward stepwise method)

Variable	*B*	*SE* B	β
(Constant)	34.05	4.06	
Pain Intensity (VAS)	3.79	0.24	0.48
Total P cROM	-0.03	0.01	-0.15
Headache	5.12	1.08	0.14
AvShAbdROM	-0.08	0.02	-0.14
No. of Sx.	0.59	0.18	0.10
Patient-rated rotation cROM	-1.44	0.60	-0.07
Treatment allocation	2.20	0.90	0.07
WAD grade	2.83	1.32	0.06
			R^2= .589, p<0.001

7.9 FOLLOW-UP AND OUTCOME CHARACTERISTICS

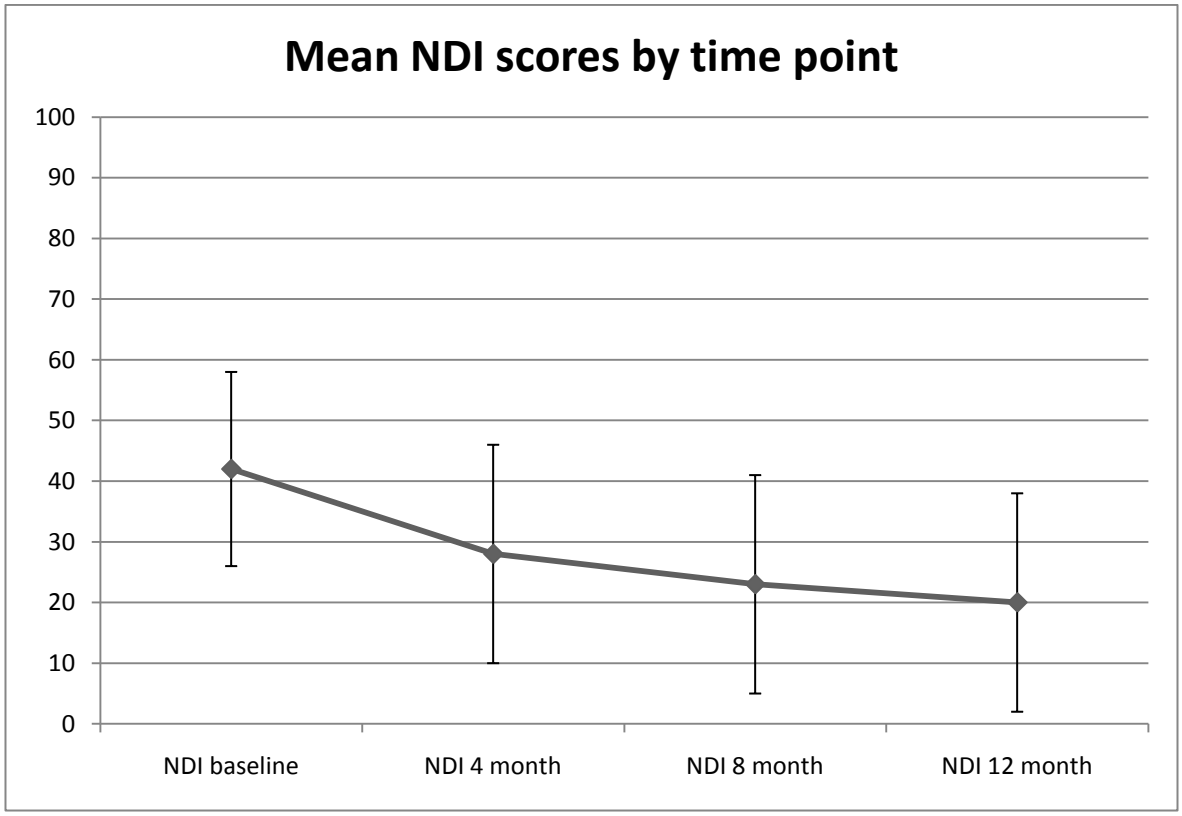
Figure 19 displays the number of participants who returned questionnaires at the various time points. At 12 months the rate of attrition was 20% (478/599). There were 17 notifications of withdrawal from the trial (reasons included moved away, no longer interested, and unhappy with trial). A proportion of the completed follow-ups were conducted over the telephone as a result of a number of participants not returning their original questionnaires.

7.9.1 OUTCOME ACCORDING TO THE NECK DISABILITY INDEX (NDI)

NDI scores reduced at each time point, with the greatest difference between baseline and 4 months. Mean (SD) NDI scores for follow-up time points were 28(17.6), 23(17.6) and 20(17.8) for 4, 8 and 12 months respectively. This improvement in disability is displayed

graphically in Figure 34. The error bars represent the standard deviation. Recovery for the majority was not complete. Two thirds of the cohort still had some long term disability.

Figure 34 - Mean (SD) NDI Score at each of the follow-up time points



According to categories derived from NDI scores published by Vernon [296] and Sterling et al [86], there was an increase in the number of participants classed as recovered at each of the three follow-up time points. Frequencies for each of the categories are presented in Table 65.

Table 65 – Frequency table for disability categories of NDI score using definitions of Sterling et al [86] by time point

Time point	Disability category					
	Recovered		**Mild**		**Moderate/Severe**	
	n(%)	Mean NDI score (SD)	n(%)	Mean NDI score (SD)	n(%)	Mean NDI score (SD)
Baseline	3(1)	0(0)	130(22)	21.6(5.3)	454(77)	47.8(13.1)
4 months	66(14)	3.7(3.0)	215(44)	19.3(5.7)	208(42)	44.4(12.7)
8 months	120(25)	3.3(3.2)	199(42)	18.3(5.6)	158(33)	43.5(12.3)
12 months	151(33)	2.5(3.0)	172(38)	17.7(5.0)	136(29)	43.5(11.6)

7.9.2 OUTCOME ACCORDING TO THE PARTICIPANT RATED RECOVERY (PRR)

Table 66 displays frequencies for each of the PRR categories at the three follow-up time points. 77% of participants reported they were better or much better four months since their baseline treatment. A small proportion reported a worsening in their condition (5% at 4 months and 9% at 8 and 12 months).

Table 66 – Participant Rated Recovery Question responses at each time point

	Participant Rated Recovery Category n(%)				
Time point	Much better	Better	Same	Worse	Much Worse
4 months	126(26)	248(51)	87(18)	18(4)	6(1)
8 months	121(25)	192(39)	132(27)	42(9)	2(0)
12 months	113(24)	164(35)	149(32)	44(9)	1(0)

7.9.3 RELATIONSHIP BETWEEN PRR AND CHANGE IN NDI SCORES

To assess the relationship between the two different outcome measures, a comparison of PRR category and change in NDI score is displayed in Table 67. On average an improvement in NDI score correlated with a positive participant rating.

Table 67 – Cross tabulation of mean change in Neck Disability Index score for each participant rated recovery category for the three follow-up time points (negative score indicates improvement in NDI)

	Mean (SD) change in NDI score for each Participant Rated Recovery Category				
Time point	Much better	Better	Same	Worse	Much Worse
4 months	-23(15.7)	-13(12.9)	-4(12.4)	4(11.8)	4(5.7)
8 months	-10(11.8)	-6(11.9)	-1(9.7)	5(13.5)	17(32.5)
12 months	-5(10.7)	-5(9.4)	-1(10.2)	5.2(11.0)	-

7.9.4 MISSING DATA - DIFFERENCES IN BASELINE VARIABLES FOR RESPONDERS AND NON-RESPONDERS

Data were investigated for differences in baseline factors for those that responded and those that were lost to follow-up at the three follow-up time-points – 4, 8 and 12 months.

Table 68 displays the characteristics of responders and those lost to follow-up for selected variables at each time point. There was no significant difference in sex, injury severity (WAD grade) and treatment allocation between responders and those lost to follow-up. There was, however, a consistent significant difference in age and baseline pain and disability scores across time points. Those that responded tended to be older and have less baseline pain and disability with mean differences approximately 7 years and 4 NDI points respectively. There was a significant difference between responders and non-responders total active cervical ROM at 4 months ($t=2.06$, $p<.05$) – on average, non-responders had less total cervical ROM. There were no statistically significant differences for values at 8 and 12 months although the trend continued for lower total cervical ROM for non-responders. Alternatively, average shoulder abduction ROM was not significantly different between the groups for 4 and 8 months but those lost to follow-up at 12 months had significantly less ROM than those that were retained.

Table 68 – Characteristics for responders and non-responders selected variables at each time point

	Time point								
	4mth			**8 mth**			**12 mth**		
Factor	Retained	Lost to F-up	sig diff*	Retained	Lost to F-up	sig diff*	Retained	Lost to F-up	sig diff*
Age, mean (SD)	40.9(13.1)	33.8(11.1)	MWU*	40.8(13.1)	34.7(11.5)	MWU*	41.3(13.1)	34.4(11.4)	MWU*
Sex, M:F	181:322	39:57:00	ChSq	177:319	43:60	ChSq	167:311	53:68	ChSq
WAD Grade, I : II/III	74:429	10:86	ChSq	74:422	10:93	ChSq	71:407	13:108	ChSq
Treatment allocation, physio : advice	255:248	44:52	ChSq	246:250	53:50	ChSq	239:239	60:61	ChSq
Baseline NDI score, mean (SD)	41.1(16.1)	44.7(16.3)	MWU*	40.9(16.2)	45.4(15.5)	MWU*	40.7(16.1)	45.6(16.1)	MWU*
MVK Pain score, mean (SD)	56.5(17.7)	60.9(15.7)	MWU*	56.5(18.0)	60.5(14.2)	MWU*	56.5(17.9)	60.1(15.1)	MWU*
Total AcROM, mean (SD)	235.0(66.3)	220(69.9)	T-test*	233.4(66.7)	228.4(69.2)	T-test	233.6(66.8)	228.3(68.1)	T-test
Av Sh Abd ROM, mean (SD)	145.6(28.3)	139.6(33.8)	MWU	145.5(29.1)	140.9(30.2)	MWU	145.7(29.0)	139.3(30.0)	MWU*

MWU = Mann Whitney U test, ChSq = Chi Squared test, *= $p<.05$

7.10 MULTIVARIATE LONGITUDINAL ANALYSIS

7.10.1 FACTORS SELECTED FOR THE LONGITUDINAL "RESEARCH CLINIC ASSESSMENT" MODELS

The results of univariate correlations (Spearman's due to non-normal distribution) between ROM variables and 4,8 and 12 month follow-up NDI scores are presented in Table 69. The results demonstrate that all ROM variables significantly correlate with all the follow-up NDI scores. It was decided that patient-rated cervical rotation ROM, total active cervical, total passive cervical and average shoulder abduction ROM scores would be used as summary measures to go into the longitudinal multivariate models because they had the strongest correlations with the follow-up NDI scores and provide distinctly separate clinical information.

Table 69 – Univariate correlations between baseline ROM variables and NDI follow-up scores

	Correlation for 4 month NDI score	Correlation for 8 month NDI score	Correlation for 12 month NDI score
PR cROM rotation	-.207**	-.151**	-.199**
PR cROM flex/ext	-.179**	-.163**	-.221**
AF	-.215**	-.154**	-.167**
AE	-.214**	-.213**	-.191**
ARR	-.203**	-.174**	-.163**
ALR	-.198**	-.200**	-.178**
ARLF	-.244**	-.214**	-.186**
ALLF	-.203**	-.176**	-.179**
Total AcROM	-.260**	-.224**	-.210**
PF	-.269**	-.200**	-.187**
PE	-.224**	-.203**	-.192**
PRR	-.246**	-.245**	-.202**
PLR	-.240**	-.218**	-.206**
PRLF	-.221**	-.185**	-.198**
PLLF	-.212**	-.185**	-.214**
Total PcROM	-.272**	-.237**	-.221**
No. of limited directions of AcROM	-.221**	-.222**	-.217**
Av Sh AbdROM	-.300**	-.275**	-.224**

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

As with the cross-sectional analyses, factors identified in the previous literature were analysed for univariate associations with the NDI scores (Spearman's correlations for

continuous measures and Mann-Whitney U tests for categorical variables). The results are presented in Table 70 and Table 71. The following variables were significantly correlated or associated with the 4 month NDI score (and therefore were entered into the multivariate model along with the ROM variables):

Age, Pain intensity, No. of Sx., coping, IES score, GHQ-12 score, FABQ-PA score, PCS score, presence of headache, presence of neuro. Sx., and WAD grade.

In addition, previous history of Chronic Widespread Pain was entered into the 8 and 12 month models and history of pre-injury neck pain was entered into the 8 month model.

All longitudinal models also included adjustment of initial disability by the addition of Baseline NDI score.

Table 70 – Univariate correlations between baseline factors and follow-up NDI scores

	Correlation for 4 month NDI score	Correlation for 8 month NDI score	Correlation for 12 month NDI score
Age	.172**	.185**	.197**
Pain intensity (MVK, 0-100)	.431**	.348**	.373**
No. of Sx. (0-15)	.299**	.284**	.342**
Coping (0-6)	-.196**	-.151**	-.157**
IES (0-75)	.356**	.341**	.281**
GHQ-12 (0-12)	.366**	.344**	.310**
FABQ-PA (0-24)	.158**	.139**	.113*
PCS (0-52)	.400**	.386**	.338**

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

Table 71 - Results of Mann Whitney U tests for baseline categorical variables vs. follow-up NDI scores

4 Month					
Variable	n	Mann Whitney U	Z	Sig	Effect size r
Sex	491	26399.5	-0.876	0.381	-0.040
Headache	490	17354	-4.79	0.000	-0.216
Neuro. Sx.	491	22139.5	-3.978	0.000	-0.180
Treatment Allocation	491	29621.5	-0.323	0.747	-0.015
Pre-injury CWP	484	1865	-1.607	0.108	-0.073
RC WAD Gd (I vs. II/III)	491	10975.5	-3.83	0.000	-0.173
Previous Neck Pain	479	7710.5	-1.277	0.202	-0.058
8 month					
Variable	n	Mann Whitney U	Z	Sig	Effect size r
Sex	479	24286	-1.413	0.158	-0.065
Headache	479	17195.5	-429	0.000	-19.602
Neuro. Sx.	479	21222	-3.73	0.000	-0.170
Treatment Allocation	479	26796.5	-1.238	0.216	-0.057
Pre-injury CWP	473	1489	-3.092	0.002	-0.142
RC WAD Gd (I vs. II/III)	479	11325	-3.076	0.002	-0.141
Previous Neck Pain	468	7187.5	-2.105	0.035	-0.097
12 month					
Variable	n	Mann Whitney U	Z	Sig	Effect size r
Sex	465	22028.5	-1.874	0.061	-0.087
Headache	464	15730	-4.554	0.000	-0.211
Neuro. Sx.	465	18513	-4.865	0.000	-0.226
Treatment Allocation	465	25034.5	-1.38	0.168	-0.064

Pre-injury CWP	457	1560.5	-2.832	0.005	-0.132
RC WAD Gd (I vs. II/III)	465	10607	-2.831	0.005	-0.131
Previous Neck Pain	452	6477.5	-1.807	0.071	-0.085

7.10.2 RESULTS FOR LONGITUDINAL "RESEARCH CLINIC ASSESSMENT" MODELS

7.10.2.1 4 months

Results for the longitudinal multiple regression analyses for 4 month NDI score outcome (using forward stepwise method) are presented in Table 72 and Table 73.

Baseline NDI score explains the majority of the variance (37% out of 40%) with the following baseline measures providing small additions to the independent explanation of variance; Impact of Events scale score, Age and number of symptoms. No ROM variables were independently predictive of NDI score in this multivariate model.

The final model was a significant fit of the data (ANOVA $p<0.001$). It is likely that errors were independent as the Durbin-Watson statistic was very close to 2. Collinearity is highly unlikely to be a problem for this model as all VIF values were under 2. Casewise diagnostic tests and tests of assumptions were all within recommended limits, indicating a well-fitted model.

Table 72 - Model summaries for 4 month longitudinal disability model (forward stepwise method)

Step	Variables	R Square	Adjusted R Square	Durbin-Watson
1	(Constant), NDI Baseline Score	0.37	0.37	
2	(Constant), NDI Baseline Score, IES	0.39	0.38	
3	(Constant), NDI Baseline Score, IES, Age	0.40	0.39	
4	(Constant), NDI Baseline Score, IES, Age, No. of Sx.	0.40	0.40	1.94

Table 73 - Multiple regression results for the final 4 month longitudinal disability model (forward stepwise method)

Variable	*B*	*SE* B	β	
(Constant)	-6.00	2.70		
Baseline NDI	0.54	0.05	0.51	
IES	0.14	0.04	0.14	
Age	0.12	0.05	0.09	
No. Of Sx.	0.53	0.26	0.08	
			R^2= .402, p<.001	

7.10.2.2 8 months

Results for the longitudinal multiple regression analyses for 8 month NDI score outcome (using forward stepwise method) are presented in Table 74 and Table 75. Baseline NDI score explained majority of the variance again (36% out of 38%) with age, pre-injury chronic widespread pain, Impact of Events scale score and number of symptoms in decreasing order of contribution to the remaining 5% of variance explanation. Again no ROM variable made it into the final model.

Similar to the 4 month model, the final 8 month model was a significant fit of the data (ANOVA $p<0.001$). Tests of model fit, casewise diagnostics and assumptions were conducted as described in section 7.3.3.1 and were all within recommended limits, indicating a well-fitted, stable and generalisable model.

Table 74 - Model summaries for 8 month longitudinal models (forward stepwise method)

Step	Variables	R Square	Adjusted R Square	Durbin-Watson
1	(Constant), NDI Baseline Score	0.36	0.33	
2	(Constant), NDI Baseline Score, Age	0.36	0.35	
3	(Constant), NDI Baseline Score, Age, CWP	0.37	0.36	
4	(Constant), NDI Baseline Score, Age, CWP, IES	0.38	0.37	
5	(Constant), NDI Baseline Score, Age, CWP, IES, No. of Sx.	0.38	0.38	1.825

Table 75 - Multiple regression results for the final 8 month longitudinal disability model (forwards stepwise method)

Variable	*B*	*SE* B	β
(Constant)	-10.53	2.89	
Baseline NDI	0.52	0.05	0.48
Age	1.55	0.06	0.11
CWP	11.63	4.36	0.11
IES	0.11	0.04	0.11
No. Of Sx.	5.75	0.27	0.09
			R^2= .382, p<.001

7.10.2.3 12 months

Results for the longitudinal multiple regression analyses for 12 month NDI score outcome (using forward stepwise method) are presented in Table 76 and Table 77. As with 4 and 8 month models, Baseline NDI score explained the majority of the variance in the final model (31% out of a total of 36%). Number of symptoms, age, pre-injury chronic widespread pain and coping score all independently explained a small amount of the remaining 5% of variance. No ROM variable was present in the final model. Similar to the 4 and 8 month models there was a significant fit of the data (ANOVA p<0.001). Tests of model fit, casewise diagnostics and assumptions were conducted as described in section 7.3.3.1 and were all within recommended limits, indicating a well-fitted, stable and generalisable model.

Table 76 - Model summaries for 12 month longitudinal disability models (forward stepwise method)

Step	Variables	R Square	Adjusted R Square	Durbin-Watson
1	(Constant), NDI Baseline Score	0.31	0.31	
2	(Constant), NDI Baseline Score, No. Of Sx.	0.34	0.33	
3	(Constant), NDI Baseline Score, No. Of Sx., Age	0.35	0.35	
4	(Constant), NDI Baseline Score, No. Of Sx., Age, CWP	0.36	0.35	
5	(Constant), NDI Baseline Score, No. Of Sx., Age, CWP, Coping	0.37	0.36	1.98

Table 77 - Multiple regression results for the final 12 month longitudinal disability model (forward stepwise method)

Variable	*B*	*SE* B	β
(Constant)	-6.84	3.55	
Baseline NDI	0.47	0.05	0.44
No. Of Sx.	1.09	0.28	0.17
Age	0.16	0.06	0.11
CWP	10.08	4.44	0.09
Coping	-1.05	0.47	-0.10
			R^2= .366, p<.001

7.10.3 FACTORS SELECTED FOR THE LONGITUDINAL "TYPICAL CLINICAL ASSESSMENT"

The following variables were entered into the "typical clinical assessment" longitudinal model as they can be collected in a routine therapy assessment session and were significantly associated or correlated with NDI follow-up score:

Total active and passive cervical ROM, patient-rated cervical ROM rotation, age, average pain score (worst, average of last week and today's), number of physical symptoms, presence of headache, presence of neuro symptoms, previous history of CWP (8 and 12 month), WAD grade, average shoulder abduction ROM, pre-injury neck pain (8 month only). Sex was not significantly associated/ correlated with baseline NDI scores.

7.10.4 RESULTS FOR LONGITUDINAL "TYPICAL CLINICAL ASSESSMENT" MODELS

7.10.4.1 4 months

Results for the longitudinal multiple regression analyses for 4 month NDI score outcome (using forward stepwise method) are presented in Table 78 and Table 79.

The model explained 29% of the variance in disability in total. Pain measured at the time of baseline assessment explained the vast majority of this variance (22%). Number of physical symptoms, average shoulder abduction ROM and age all independently explain a small amount of the remaining 5% of variance each.

The model was a significant fit of the data (ANOVA $p<0.001$). Tests of model fit, casewise diagnostics and assumptions were conducted as described in section 7.3.3.1 and were all within recommended limits, indicating a well-fitted, stable and generalisable model. When

average shoulder abduction ROM was not entered into the model, total active cervical ROM was substituted into the final model and explains slightly less variance. When all clinician-measured ROM factors were not entered into the model, patient-rated cervical rotation ROM was also substituted into the final model, which only explained one percent less variance overall (28%- results not shown).

Table 78 - Model summaries for 4 month longitudinal “TCA” disability models (forward stepwise method)

Step	Variables	R Square	Adjusted R Square	Durbin-Watson
1	(Constant), Pain intensity (VAS)	0.217	0.215	
2	(Constant), Pain intensity (VAS), No. of Sx.	0.259	0.256	
3	(Constant), Pain intensity (VAS), No. of Sx., Av Sh Abd	0.280	0.276	
4	(Constant), Pain intensity (VAS), No. of Sx., Av Sh Abd , Age	0.287	0.281	1.804

Table 79 - Multiple regression results for the final 4 month longitudinal “TCA” disability model (forward stepwise method)

Variable	*B*	*SE* B	β
(Constant)	14.70	5.60	
Pain intensity (VAS)	3.17	0.35	0.38
No. Of Sx.	1.03	0.26	0.17
AvShAbd	-0.09	0.03	-0.15
Age	0.11	0.05	0.08
			R^2= .287, p<.001

7.10.4.2 8 months

Results for the longitudinal multiple regression analyses for 8 month NDI score outcome (using forward stepwise method) are presented in Table 80 and Table 81

The final model explains 25% of variance of NDI scores and contains 5 baseline factors – pain intensity at assessment, number of symptoms, pre-injury chronic widespread pain, average shoulder ROM and Age. Pain intensity explains the majority of the variance (15%) as it did in the 4 month model.

The model was a significant fit of the data (ANOVA $p<0.001$). Tests of model fit, casewise diagnostics and assumptions were conducted as described in section 7.3.3.1 and were all within recommended limits, indicating a well-fitted, stable and generalisable model. When either average shoulder abduction ROM or all clinician-measured ROM factors are not entered into the model the final model consists of four factors (pain, number of symptoms, pre-injury chronic widespread pain and age). Thus any form of cervical ROM measurement does not explain any further amount of variance.

Table 80 - Model summaries for 8 month longitudinal “TCA” disability models (forward stepwise method)

Step	Variables	R Square	Adjusted R Square	Durbin-Watson
1	(Constant), Pain intensity (VAS)	0.15	0.148	
2	(Constant), Pain intensity (VAS), No. of Sx.	0.20	0.196	
3	(Constant), Pain intensity (VAS), No. of Sx., CWP	0.23	0.22	
4	(Constant), Pain intensity (VAS), No.of Sx., CWP, Sh Abd ROM	0.24	0.235	
5	(Constant), Pain intensity (VAS), No.of Sx., CWP, Sh Abd ROM, Age	0.25	0.243	1.739

Table 81 - Multiple regression results for the final 8 month longitudinal “TCA” disability model (forward stepwise method)

Variable	*B*	*SE* B	β
(Constant)	9.40	5.90	
Pain intensity (VAS)	2.37	0.37	0.29
No. Of Sx.	1.26	0.27	0.20
CWP	15.13	4.74	0.13
Sh Abd ROM	-0.08	0.03	-0.13
Age	-0.07	0.06	0.10
			R^2= .251, p<0.001

7.10.4.3 12 month

Results for the longitudinal multiple regression analyses for 12 month NDI score outcome (using forward stepwise method) are presented in Table 82 and

Table 83.

The 12 month model explains the same amount of variance as the 8 month model (25%) but contains a different ROM factor – patient-rated cROM instead of average shoulder abduction ROM.

The model was a significant fit of the data (ANOVA $p<0.001$). Tests of model fit, casewise diagnostics and assumptions were conducted as described in section 7.3.3.1 and were all within recommended limits, indicating a well-fitted, stable and generalisable model.

Table 82- Model summaries for 12 month longitudinal "TCA" disability models (forward stepwise method)

Step	Variables	R Square	Adjusted R Square	Durbin-Watson
1	(Constant), Pain intensity (VAS)	0.15	0.15	
2	(Constant), Pain intensity (VAS), No. Of Sx.	0.22	0.21	
3	(Constant), Pain intensity (VAS), No. Of Sx., CWP	0.23	0.23	
4	(Constant), Pain intensity (VAS), No. Of Sx., CWP, Age	0.24	0.24	
5	(Constant), Pain intensity (VAS), No. Of Sx., CWP, Age, PRcROM rot	0.25	0.24	0.19

Table 83 - Multiple regression results for the final 12 month longitudinal "TCA" disability model (forward stepwise method)

Variable	*B*	*SE* B	β
(Constant)	1.28	4.99	
Pain intensity (VAS)	2.33	3.72	0.28
No. Of Sx.	1.49	0.27	0.24
Age	12.69	4.58	0.12
CWP	0.15	0.06	0.11
PRcROM rot	-2.12	0.95	-0.10
			R^2= .252, p<.001

7.11 SECONDARY ANALYSIS – PREDICTING PATIENT RATED NON-RECOVERY

Although there was a slight trend for reduced cervical ROM for participants that rated their neck problems as the same or worse at each of the three time points, all the clinician-measured values were non-significant when subjected to Mann-Whitney or independent t-tests depending on distributions (see Table 84). However, patient-rated cervical rotation range of motion was significantly different between those that reported improvement and those that did not, and this was consistent at all follow-up time points.

Therefore patient-rated cervical rotation ROM was entered into the multivariate model along with other non-ROM factors that were univariately associated with the outcome of PRR (improved vs. Same/worse) (table not shown). The variables entered into the models for the three time points were:

4 months: PRcROM rotate, NDI baseline score, Pain intensity (MVK), FABQ-PA score, coping, PCS score, treatment allocation, headaches, and WAD grade.

8 months: PRcROM rotate, NDI baseline score, FABQ-PA score, coping, previous neck pain.

12 months: PRcROM rotate, NDI baseline score, pain intensity (MVK), FABQ-PA score, GHQ-12 score, No. of Sx., age, coping, and PCS score.

Table 84 – Comparison of ROM variables between Patient Rated Recovery categories (improved vs. same/worse)

	4 months				8 months				12 months			
ROM Factor	Improved		Same or Worse		Improved		Same or Worse		Improved		Same or Worse	
	Median	IQR	Median	IQR	Median	IQR	Median	IQR	Median	IQR	Median	IQR
PRcROM rot*	3	1	3	1	3	1	3	1	4	1	3	1
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
TAcROM**	236	66	226	69	236	68	228	66	233	65	232	69
TPcROM**	243	84	230	82	242	87	234	78	241	84	236	83
Av Sh Abd ROM**	146	29	142	29	145	29	143	29	145	29	144	30

*MWU significant p<.05 at 4, 8 and 12 months

** T-test/MWU non significant

7.11.1 RESULTS

7.11.1.1 4 months

The baseline factors that were independently predictive of patient rated non-recovery at 4 months are shown in Table 85. Pain intensity, pain catastrophising score and treatment allocation combined to predict approximately 6-8% of the outcome variance.

Table 85 – Final logistic regression model for baseline predictive factors of patient rated recovery at 4 months

Baseline Factor	B	SE	Sig.	Exp (β)	95% CI
Pain intensity (MVK)	.015	.006	.014	1.016	1.00-1.03
PCS	.02	.008	.015	1.02	1.00-1.04
Rx allocation	-.55	.19	.003	.58	.40-.83
(Constant)	-1.53	.33	.000	.22	

R^2=.06 (Cox and Snell), .079(Nagelkerke), Model Chi-square=32.50, p<.001

7.11.1.2 8 months

The final model for predictors of patient rated non-recovery at 8 months is shown in Table 86. Only one baseline factor was independently predictive of outcome which was the ability to cope, explaining just 1% of outcome variance.

Table 86 - Final logistic regression model for baseline predictive factors of patient rated recovery at 8 months

Baseline Factor	B	SE	Sig.	Exp (β)	95% CI
coping	-.12	.052	.018	.884	.798-.979
(Constant)	.27	.211	.198	1.31	

R^2=.01 (Cox and Snell), .01(Nagelkerke), Model Chi-square=5.61, p<.018

7.11.1.3 12 months

The final model for predictors of patient rated non-recovery at 12 months is shown in Table 87. As with the 8 month model, only one factor was predictive of outcome, however for this time point it was NDI baseline score, although it explained just 2% of outcome variance.

Table 87 - Final logistic regression model for baseline predictive factors of patient rated recovery at 12 months

Baseline Factor	**B**	**SE**	**Sig.**	**Exp (β)**	**95% CI**
NDI baseline	.017	.005	.002	.1.017	1.006-1.028
(Constant)	.-.491	.241	.042	.612	

R^2=.02 (Cox and Snell), .02(Nagelkerke), Model Chi-square=9.58, p<.002

7.12 SECONDARY ANALYSIS – EFFECT OF TREATMENT GROUP ON PROGNOSTIC VALUE OF ROM

In order to explore whether treatment allocation moderated the prognostic value of ROM variables, the linear regression models described in section 7.8.4 and 7.10.4 were run with the additional interaction variables: Treatment allocation X total active cervical ROM, Treatment allocation X total passive cROM, Treatment allocation X shoulder abduction ROM and Treatment allocation X patient-rated rotation cROM.

There were no changes to the final models at all three follow-up time points (4, 8 and 12 months) and so treatment allocation had no moderation effect on ROM variables longitudinal prognostic value. ROM variables present in the final typical clinical assessment models were therefore defined as non-specific predictors.

7.13 DISCUSSION

7.13.1 KEY RESULTS AND INTERPRETATION

Five separate systematic reviews conducted in the last five years have all concluded that further high quality cohort studies are required to evaluate prognostic value of physical, psychological and social factors for the outcome of Whiplash Associated Disorders (WAD) [39, 40, 295, 298, 299]. This cohort study is the largest inception cohort for sub-acute whiplash injured patients completed to date.

The first aim of this study was to describe the clinical characteristics of whiplash-injured patients in this cohort. From the results it is clear that WAD do not just involve neck pain, but are a complex set of physical and psychosocial characteristics, reinforcing previous findings [87]. Not only did the vast majority of participants report symptoms in the cervical spine area (98%) but also two thirds of the cohort indicated that they were experiencing symptoms in their shoulder complex area. Neurological impairment - evaluated through neurological signs such as reduced muscle power, altered sensation and diminished reflexes- was experienced by approximately a tenth of the participants (12%). It is interesting to note that there was an increase in the proportion who displayed neurological signs from their ED visit to the research clinic assessment. It may be that there was a delayed neurological reaction to the injury or there was increased detection due to the greater detail of the assessment that is afforded within a research clinic. From a psychosocial perspective, a considerable number of participants showed signs of fear-avoidant beliefs, post-injury distress and depression as measured by the Fear-Avoidance Beliefs (FABQ), Impact of Events Scale (IES) and General Health Questionnaire (GHQ-12) questionnaires respectively, confirming previous findings from other recent cohorts [65, 87].

Clinician-measured cervical Range of Motion (ROM) was on average 30-40% less than normative values although there was considerable variation (indicated by the large Standard Deviations) with a proportion of individuals having normal amounts of cervical ROM. The average total cervical ROM is less than previous cohorts, although this may have been because more severely injured patients and those with previous chronic pain problems were excluded from these studies [65, 140]. Individual ROM measurements also appear to be lower than other study results of CROM device measurements on symptomatic individuals not from WAD populations [186, 228].

The pattern of loss of cervical ROM was fairly consistent in all directions, although extension was more reduced than other directions, similar to the findings of Kasch et al [140] although they found smaller reductions (approximately 10%) compared to matched control subjects. Significantly reduced extension should perhaps be expected in light of the mechanism of injury for the majority of participants, whose whiplash injury was a result of a rear impact to their vehicle which would have led to hyper-extension of the cervical spine and injury to the posterior structures surrounding the spine (as summarised in Chapter One, [4]).

The pattern of loss of cervical ROM observed was consistent for both active and passive measurements, with very little difference between the two forms. Although these measures have not previously been taken simultaneously in a cohort of whiplash-injured patients, this is somewhat unexpected as clinicians would usually expect passive movement to have a greater ROM or at the very least be equal to the active movements that the patient conducts themselves [147]. A greater proportion of participants reported limited passive cervical movement compared to active. Previous authors [261] have speculated that a patient's lack of understanding regarding measurement process can cause apprehension and can

compromise passive ROM results although in this study research clinicians were trained to provide education as to the measurements that would be conducted before the participants were examined. The research clinicians may well have had an effect on the results. Due to the acute nature of the injury they may have been reluctant to force passive cervical movements too far – previous work has indicated that HCP's can have fear-avoidant beliefs regarding their patients[305]. Patient-rated cervical ROM correlated with the clinician-measured equivalents indicating validity of the scales. It was interesting to find that participants that reported no limitation in cervical ROM on the patient-reported Likert scale still, on average, had less ROM than normative values. This may mean that all patients experience some reduction in ROM but this reduction may be at a sub-clinical level and not enough to affect function.

The findings that cervical ROM reduced with age and that there was no significant difference in ROM between males and females concur with previous studies [168, 306].

Similar to cervical ROM, shoulder abduction ROM was reduced in this cohort when compared to non-matched normative data. As far as the author is aware, this is the first reporting of shoulder ROM in a WAD cohort study. Previous studies have reported increased shoulder pain [82, 131] and impaired shoulder proprioception[307] but have not explicitly reported on the affect of a whiplash injury on shoulder ROM. Shoulder abduction ROM was associated with cervical spine ROM and neck pain related disability. In particular shoulder ROM correlated with cervical rotation which is justified by work by Takasaki et al [308]explaining the relationship between the two movements.

Pain appears to have a significant effect on cervical and shoulder ROM. There was a linear association between pain intensity and ROM in that the higher the pain scores, the lower the

ROM. This concurs with previous findings for cervical ROM [68]. Those that reported pain as the limiting factor had the least ROM. This may be because individuals and assessors see pain as a more serious symptom than stiffness and are more wary to try and move the neck any further. It is noteworthy however that neck pain in the month prior to injury was not significantly associated with lower cervical ROM.

Psychosocial factors are prevalent within WAD and have become increasingly important in the understanding of musculoskeletal conditions with the ascent of the biopsychosocial model [309]. One psychological model that fits within the biopsychosocial model and is receiving increased research attention is that of the fear-avoidance model (FAM) [275]. Nieto et al [284] published results of a cross-sectional study investigating the relationship between fear-avoidance, catastrophising and disability. They reported a cohort with a similar mean disability score (NDI 38/100) and found that fear-avoidance significantly mediated the relationship between catastrophising and disability in line with the FAM. There is an obvious link to the physical measurement of ROM with this model – fear avoidant beliefs would be expected to result in reduced ROM. Findings from this cohort study showed that Fear Avoidance beliefs were associated with cervical ROM, although univariately less strongly than other psychological constructs such as catastrophising, depression and distress.

When univariately associated/correlated factors were entered into a multivariate regression model for prediction of baseline cervical ROM, both physical (pain and age) and psychological (fear of movement, self-efficacy and depression) factors explained some of the variance in the amount of ROM.

Within the aim to describe the cross-sectional characteristics of this cohort, there was also the objective to conduct a multivariate analysis to explore factors that can predict the amount

of baseline disability. The results indicate that there is a moderate independent relationship between cervical ROM and disability suggesting a causal relationship, although this cannot be defined by a cross-sectional analysis. The surprising element of this model is that shoulder abduction ROM has the strongest relationship of all the ROM measurements with neck pain related disability in the form of NDI score. This measure explained a significant amount of the variance for disability at the time of measurement. When models were re-run excluding shoulder ROM and then all clinician-measured ROM factors, final models were very similar suggesting that the different measures of ROM have almost identical predictive value in cross-sectional model. Other factors that explained disability besides ROM factors were pain intensity, depression, presence of headaches, catastrophising, injury severity (WAD grade), number of physical symptoms present and treatment allocation. This model was accurate for the sample and generalisable to the population. Some of these measures are not available to the average clinician and so a simple clinical model was also constructed. This model provided almost as much predictive power as the research clinic assessment version. The final model included pain, total passive cervical ROM, presence of headaches, average shoulder abduction ROM, number of physical symptoms, patient-rated cervical rotation ROM, treatment allocation and WAD grade.

The second main aim of this cohort study was to evaluate the prognostic value of cervical ROM measurements and consequently participants were followed-up at three time points over the course of one year after their ED visit. Substantial numbers of patients had disability at 12 months irrespective of the treatment they received, with most improvement occurring early on, in agreement with other studies [298]. The mean change of 22 NDI points between baseline and 12 months, whilst clinically significant [289], meant that there were still individuals with significant amounts of residual disability. Four previous cohort studies have

used the NDI as an outcome measure [73, 75, 76, 86, 87, 168, 293] . Although these studies were inconsistent in how they reported the NDI, there appears to be consistency in the proportions of participants and levels of their disability up to 3 years following a whiplash injury and are in line with the findings of this study. Using previously published categories of the NDI score [168, 296], approximately one third were classified as recovered, a third had mild disability and a third had moderate to severe disability one year after injury.

Recovery can be defined in many ways and has a direct impact on the prognostic factors that predict this. This was demonstrated in this study by the use of two outcome measures – the NDI Score as already discussed and also the patient- rated recovery question (PRR). PRR mirrored NDI scores in that there was a reduction in the improvement reported over time (77% reported improvement at 4 months compared to 59% at 12 months). PRR asks about a change since the last time the question was asked and therefore could be subject to recall bias and therefore may need greater caution in interpretation.

The various measures of cervical ROM were not present in the final multivariate models, despite significant univariate correlation with NDI scores at all follow-up time points. This compares with recent previous work by Atherton et al [58] and Hendriks et al [65] but contrasts with other studies by Kasch et al [140, 159] and Sterling [86, 87, 168], although the latter only found prognostic value for short term outcome with mildly disabled individuals . This cohort is the largest studied in order to rigorously evaluate the prognostic value of cervical ROM and coupled with findings of previous good quality studies provide a convincing argument that other factors are more valuable in the prognosis of outcome from WAD. This work particularly highlights that psychosocial factors such as post-injury distress (measured by the Impact of Events Scale questionnaire) and negative coping beliefs are important.

However, with the simulation of the typical clinical assessment where these questionnaires would not routinely be available, ROM does then become prognostically important. The really interesting finding is that the most useful ROM measure is not regarding the cervical spine, but shoulder abduction ROM. This is not a prognostic factor that has been considered before, although it is commonly assessed in the clinical setting as part of a generalised assessment. If shoulder ROM is not measured, active, passive or patient-rated cervical spine ROM are almost as useful when attempting to predict outcome in the short term.

There were three factors that consistently displayed prognostic value for predicting recovery as defined by NDI score at the various follow-up points. These were baseline NDI score, age and number of physical symptoms. Initial disability score is the only factor that has been consistently highlighted by systematic reviews as a valuable prognostic indicator. When initial disability score is not available for information, as simulated by the typical clinical assessment models in section 7.10.4, initial pain intensity provides a useful surrogate, although with only 50-75% of the explanatory power depending on the time point. Pre-injury chronic widespread pain and number of physical symptoms had also shown promise as a prognostic factor in one cohort study [58] and is confirmed here as a worthy of consideration when attempting to predict outcomes for patients. A psychosocial factor related to ROM that has shown inconclusive evidence of prognostic value is Fear-Avoidance beliefs (FAB). This study provides evidence that FAB's do not offer significant prognostic value when entered into a comprehensive research or clinical assessment model, agreeing with Sterling et al [86, 87]. It is noted that another recent study has been published concluding the opposite [310]. They performed multivariate analyses, however with a much smaller cohort and less comprehensive multivariate model.

7.13.2 STRENGTHS AND LIMITATIONS

This is the largest prospective cohort study of sub-acute whiplash-injured patients to date and from this comes great strengths. The arguments for the value of a prospective and not a retrospective design to study prognostic factors are well versed [18]. Poor outcome following a whiplash injury is not particularly rare therefore a cohort study is a good design to have enough power and precision to answer the questions being asked.

This study was nested within a pragmatic RCT that had broad selection criteria, and therefore there was the potential to include high and low risk groups in the cohort. Also, the collection of data for 599 participants would have been unfeasible for a single PhD student.

The measurement process was comprehensive and rigorous with the use of validated industry-standard questionnaires where possible and standardised clinical assessment procedures. Another advantage of being part of a well-funded RCT was the availability of administrative staff to implement a standardised system of follow-up, including the use of core outcome telephone calls. This resulted in a relatively small proportion of participants lost to follow-up and maintenance of blinding until all data was obtained.

There is the potential that this cohort study may suffer from selection bias due to the cohort being assembled from participants of an RCT. Hendriks et al [65] also postulated this regarding a similar trial. Approximately 50% of patients attending the participating emergency departments were approached and of these around 50% agreed to participate in the MINT study. Therefore there are a significant proportion of individuals who experienced a whiplash injury that did not participate in this cohort study. It is difficult to conclude in which direction this potential bias could act. This said the characteristics of this cohort are

similar to other inception cohorts recruited not only in an emergency department but also from primary care or mixed catchments.

Attrition bias was also a possibility for this study as not all data from participants enrolled at baseline were available at the follow-up time points, although the response rate of 80% at the 12 month follow-up is commendable for postal questionnaire follow-up of acutely injured patients recruited in an emergency care setting. The fact that those who were lost to follow-up were younger and more disabled at baseline may have affected the results of analyses, however despite a maximum mean difference in baseline NDI score of 5 points being statistically significant, this is not a clinically significant difference (MCID = 10 points [289]).

There is a chance with multivariate analyses using numerous predictor variables that models can be over-fitted and findings may be as a result of chance. A-priori awareness of this led to a restriction of the number of factors used considering the sample size, and therefore the chances of this type of bias have been reduced as much as possible. The fact that the findings were consistent when alternative analytical methods were used (backwards stepwise methods not presented) is reassuring.

There is a chance that confounding variables may have been omitted due to the limitation of factors described above, however it is argued that by using a number of systematic literature reviews conducted by international experts, this chance has been minimised as far as practically possible within the limitations of a doctoral study.

With particular attention to the measurement of cervical ROM, the assessments were carried out by a number of research clinicians. This may have lead to a large variation in measurement error. Measurement of cervical ROM reliability between testers using a

standardised protocol was conducted as part of this work and is described in Chapter Six, the results of which suggest that the CROM device provides reliable measurements between testers. The author was careful to ensure that certain principles were adhered to, for example consistent positioning and fixation of body parts where possible and standardised instruction and documentation [261]. All research clinicians underwent training and were provided with a manual describing all the assessment procedures. It is possible that reliability of the CROM could be improved, for example by the resting/neutral position being referenced to gravity. Shoulder ROM reliability was not investigated within this cohort study, although previous studies indicate that reliability between testers is good [274].

It was highlighted previously in this chapter that there has been a lack of continuity in the use of outcome measures for prognostic studies for WAD and even when the same outcome measure has been used different cut points have been used to define poor outcome making comparisons and subsequent conclusions very difficult for both researchers and clinicians. The choice of the Neck Disability Index as the primary outcome measure was made due to its wide validation and most frequent use within previous cohort studies of any quality. This said a continuous measure such as this may be difficult to interpret, especially for clinicians, with regards to what constitutes a "disabled" and "non-disabled" score. To facilitate ease of interpretation the NDI has been converted into categories, but as is usually the case with this methodology; the categories are far from perfect in terms of distinct groups.

7.13.3 GENERALISABILITY

Overall the study generalisability should be good, with a large sample recruited from a wide range of UK hospitals and good representation of a range of injury severity.

Differences were small between patients who were and were not recruited into this cohort study from Step One of MINT. The population recruited included a slightly greater proportion of females, a slightly higher mean age and a higher proportion of patients categorised with WAD grade II and III injury severity. The latter point is to be expected in that those with more severe injuries would be more likely to present for further treatment.

There were some differences in those who were and were not followed up in terms of pain and disability (non-responders were significantly more disabled) although loss to follow-up was acceptable (20%). This has the small possibility of limiting generalisability of statistical modelling.

As part of the RCT, participants in this cohort study received standardised initial treatments according to the MINT protocol, consisting of either a session of advice or package of up to six sessions with a physiotherapist. Not all patients in the UK who continue to have problems following a whiplash injury will receive such treatments and this may mean the findings are less generalisable. However, patients were free to seek any types of treatment following trial treatments, which perhaps is more common within the UK.

When comparing this cohort with others studied in the UK, there appear to be similarities in characteristics and outcome, inferring a representative sample of whiplash-injured patients. A Bristol hospital cohort which has recently reported a 30 year follow-up [311], initially comprised 61 consecutive patients presenting, who reported a variety of symptoms akin to those described in this cohort. At two year follow-up [45] over 60% of the cohort still had symptoms that were affecting activities of daily living. Mayou and colleagues [72] studied a cohort recruited in an Oxford emergency department and followed participants over one year. They found 75 -85% of participants had problems with activities regarding recreation

and work. Crouch and colleagues [293] also studied a cohort recruited from a UK emergency department and concluded that two thirds had disability four to six weeks following presentation.

The results section has described how multivariate models were examined for generalisability to a wider population and these tests provided good evidence that the models should be applicable to a wider population of whiplash-injured individuals.

7.13.4 CLINICAL AND RESEARCH IMPLICATIONS

The key finding that shoulder abduction ROM is useful in characterising and predicting outcome is new and, although unlikely due to chance, should be investigated further in a cohort of whiplash-injured patients. It would also be worthwhile to investigate other planes of shoulder movement.

As far as the author is aware, patient-rated cervical ROM has not been tested previously in a cohort of WAD individuals and this measure requires further development work to investigate whether there can be any improvement made in its diagnostic and prognostic value. A suggestion might be to use either a visual analogue scale or a percentage rating instead of a limited Likert scale. Investigation into whether this correlates with clinician rated ROM more or less strongly than the format used here would be useful. It certainly offers promise as a more rapid assessment tool compared to a clinician measuring all planes of movement to calculate a sum score. Since the conduct of this study, the Movement Ability Measure (MAM) has been developed by Allen [312] and provides a comprehensive evaluation of self reported movement, assessing all the different facets of the construct such as stiffness/tightness, joint mobility and ROM. This has been developed from the Movement Continuum Theory [313]. If this had been utilised in this cohort study it may have provided

more accurate information (its reliability, validity and responsiveness have been assessed and found to be promising [312, 314]). The limitation of the measure is that it is very long and may not have been feasible as a lower priority measure within MINT.

Implications of the findings of this study for clinicians are numerous. The impairment that results from a reduction in cervical ROM has been shown to have direct effect on how disabled a patient reports they are. Clinicians may use active, passive or patient-rated cervical ROM to provide this information. There does not seem to be any huge difference in diagnostic or prognostic information between active and passive cervical movements for patients with WAD. Clinicians should also be aware the shoulder abduction ROM can be an equally, if not slightly more informative measure. Alongside measurements of ROM, other physical examination factors that reflect disability at the time of assessment include pain intensity, presence of headaches, the number of symptoms and the injury severity (WAD) grading. Psychosocial factors that should also be evaluated for a complete picture of disability are depression and catastrophic thoughts.

When attempting to predict a patient's outcome early on (approximately four weeks post injury), clinicians should predominantly take into account the patient's initial disability rating. Using the NDI in a clinical setting would not be too onerous as there are only ten questions and patients find these questions easy and quick to answer. Clinicians should also be aware of the consistent finding that older age is a risk factor for poor outcome, even though this factor is not clinically modifiable. Of the measurements currently routinely recorded by musculoskeletal clinicians, number of symptom areas and average shoulder abduction ROM were found to have prognostic value in this cohort. Measurements of these alongside asking the patient if they had pre-injury long term widespread pain (according to the Manchester definition) will enable clinicians to make estimates on how the patient will

recover and whether more intensive treatment is required. It should be noted that ROM measurements should not be used to decide whether a patient will respond to intensive physiotherapy or not according to the results of this study.

7.14 SUMMARY

This chapter has presented a large cohort study of whiplash- injured patients recruited as part of a randomised controlled trial.

Findings reinforce the belief that WAD manifests itself in both physical and psychosocial symptoms, and that a considerable proportion of patients will be affected one year after their injury. The findings suggest that measurements of ROM are useful in explaining disability at the time of measurement, and in the form of both cervical and shoulder ROM. Indeed WAD involved more than just a short period of neck pain in terms of symptom location and duration. Participant rated recovery is greatest in the early stages – up to 4 months.

When attempting to predict recovery, clinicians should consider what outcome they are evaluating as this will influence which factors will be useful. For predicting neck-pain related disability both pre-injury (age, chronic widespread pain), physical (number of symptoms, initial pain-related disability) and psychological (distress and coping) should be evaluated to provide the best information on prognosis according to this study.

The next chapter will provide an overall discussion of all the studies in this thesis and how they inter-relate.

8 CHAPTER EIGHT – SUMMARY DISCUSSION

8.1 INTRODUCTION

8.1.1 AIMS AND OBJECTIVES OF THE THESIS

The aim of this thesis was to contribute to the evidence base for assessing and treating patients with Whiplash Associated Disorders (WAD). As a physiotherapist, the author aspired to produce knowledge valuable and relevant to fellow clinicians, and as a result, patients diagnosed with WAD. Another aspiration was that this work could be coherent and easily integrated into both research and clinical settings. Ultimately, the aim was to investigate the role of cervical ROM in recovery from WAD. In order to do this, more specific objectives were:

- To systematically review literature regarding prognostic factors in order to assess the current evidence base regarding cervical ROM as a prognostic factor and to inform multivariate analyses of other appropriate variables (see Chapter Three)
- To systematically review literature in order to select the best method for measuring cervical spine ROM (see Chapter Five)
- To evaluate both within and between observer reliability for the selected device in a WAD population (see Chapter Six)
- To conduct a prospective cohort study in order to provide robust data for univariate and multivariate analyses to evaluate the diagnostic and prognostic value of cervical ROM (see Chapter Seven)

This chapter will seek to summarise the previous chapters, describing the intersections between them and highlighting the key findings and questions generated.

8.1.2 OVERVIEW OF CURRENT RESEARCH INTO WHIPLASH ASSOCIATED DISORDERS

Since commencing work for this thesis in 2005, research into all aspects of Whiplash Associated Disorders has continued apace. Most notably during this time The Bone and Joint Decade 2000-2010 Task Force on Neck Pain and Its Associated Disorders was assembled and conducted a number of research projects including a best evidence synthesis programme, epidemiological studies and intervention studies [315] .The aim of this task force was to update the work of the Quebec Task Force (QTF) [6] and make recommendations that would reduce the consequences of neck pain and its associated disorders. The Neck Pain Task Force's key findings were that the number of patient seeking health care at emergency departments for WAD globally has increased over the past three decades, head restraints to limit the whiplash mechanism of injury have a preventative effect, and that there is no one superior conservative treatment for WAD but early return to usual activities still remains the best policy. They recommended a revision to the QTF classification system for neck pain severity as follows:

Grade I: No signs or symptoms suggestive of major structural pathology and no or minor interference with activities of daily living; will likely respond to minimal intervention such as reassurance and pain control; does not require intensive investigations or ongoing treatment.

Grade II: No signs or symptoms of major structural pathology, but major interference with activities of daily living; requires pain relief and early activation/intervention aimed at preventing long-term disability

Grade III: No signs or symptoms of major structural pathology, but presence of neurologic signs such as decreased deep tendon reflexes, weakness and/or sensory deficits; might require investigation and occasionally more invasive treatments

Grade IV: Signs or symptoms of major structural pathology, such as fracture, myelopathy, neoplasm, or systematic disease; requires prompt investigation and treatment

The notable omission in the classification system related to this thesis is there is no specific mention of reduced cervical ROM as an assessment finding that can differentiate between the severity grades, unlike the original QTF grading system which has been presented in previous chapters. The findings of this thesis would seem to support this omission.

The Neck Pain Task Force concluded "Future research should be directed to assessing the impact of modifiable risk factors through innovative treatment approaches." [315] This highlights the importance placed on high quality research investigating prognostic factors in WAD, such as that presented in Chapter Seven.

In 2008, Walton [288] published a review of definitions of recovery for WAD which highlighted how exact rates of recovery are still difficult to ascertain due to the inconsistency of definitions used. It is clear however that Whiplash is still a very common injury and patients are presenting in increasing numbers to primary and secondary care in the UK [8].

It is unclear of the size of effect of an increasing frequency and cost of compensation here in the UK on reporting of WAD and recovery[316]. In MINT, over 90% of participants

reported at 12 months having pursued seeking compensation following their injury. Some research evidence where changes of insurance system have lead to reduction in the reporting and chronicity of WAD indicates that compensation may have an effect [317, 318]. This has led some researchers to advocate an evaluation of a public health approach for the management of WAD such as modification of the community environment (e.g. health care information and litigation systems) [319].

In terms of specific guidance for clinicians managing WAD patients in the UK, no new guidelines have been issued during the time of conducting this doctoral work. The only existing guidance [16, 154] was provided for physiotherapists and was in-line with the physiotherapy package provided for participants of Step Two of MINT.

8.1.3 THE UNIQUE CONTRIBUTION OF THIS THESIS TO WHIPLASH ASSOCIATED DISORDERS RESEARCH AND MANAGEMENT

There are a number of novel contributions to the evidence base that this thesis provides. Firstly, at the time of conduct and publication, the systematic review of physical prognostic factors described in Chapter Three provided an up-date on the state of evidence; indeed a systematic review had not been published for three years prior to this, during which time a considerable number of cohort studies had been published. Subsequently, four systematic reviews have been reported [295, 298, 299, 320], although they do not provide a specific review of solely physical prognostic factors.

Secondly, the systematic review of reliability and validity studies of measurement tools for cervical ROM (Chapter Five) was the first to be published for seven years and remains the only published review to include studies of both active and passive cervical ROM. In Chapter Six, intra- and inter-observer reliability studies of the CROM device were conducted

for the first time in a sub-acute WAD population, and for both active and passive forms of cervical ROM – another first. These studies provide new information to researchers and clinicians' regarding what is becoming an increasingly well-used tool in a commonly researched and treated population. Prior to these studies, the reliability of the CROM was promising but only had limited generalisability.

Finally, the preceding chapter (Chapter Seven) provides a number of unique contributions to the current evidence base; the description of cervical ROM in both active and passive forms in the same WAD cohort and their relation to other important measures of physical and psychological states. Patient-rated cervical ROM was studied for the first time in a WAD population as far as the author is aware. The size of this cohort and avoidance of conduct and reporting problems that exist in most other previous studies is advantageous. Prior to this thesis, the prognostic value of cervical ROM was uncertain. All of the studies described in this thesis aimed to be high quality in nature, something that was found lacking from previous studies.

It is also worthy to note that this thesis forms part of the documentation of the largest UK cohort of whiplash-injured individuals conducted to-date (recruited as part of the Managing Injuries of the Neck Trial) and sits side by side with the other outputs as part of the main clinical trial.

8.2 OVERVIEW OF THESIS FINDINGS

The following section provides a summary of findings for each chapter in more detail, highlighting key messages and implications, limitations and any recent related evidence.

8.2.1 CHAPTER TWO – MANAGING INJURIES OF THE NECK TRIAL (MINT)

8.2.1.1 Key messages and implications

Although the objectives of this thesis were not to discuss results of the Managing Injuries of the Neck Trial, there are some findings that are worthy of highlighting due to the obvious inter-connections between the trial and this doctoral work.

The key learning points from the findings of MINT suggest we still are unable to provide the optimum treatment for all patients who are failing to recover (evaluated in Step Two) despite implementing an evidence-based risk factor modification approach. This approach, framed within the biopsychosocial model, and evaluated as the experimental physiotherapy package included targeting impairment in cervical spine, thoracic spine and shoulder ROM. The physiotherapy package resulted in short term improvements in neck disability in comparison to the advice session with a physiotherapist, but these effects were not maintained in the long term. Coupled with the fact that 67% of participants still had some disability at one year post injury means we are still searching for improvements in the interventions for WAD patients.

Particularly relevant to this thesis is the finding from the documentation of the assessments and interventions for all Step Two participants that ROM was a consistent focus for the therapists, despite this not being an enforced part of the control arm protocol.

8.2.1.2 Limitations

Increasing cervical spine, thoracic spine and shoulder ROM is often a treatment target of physiotherapy for WAD patients. Measurement of cervical and shoulder ROM following Step Two interventions could have been a beneficial addition to the study and therefore the evidence base. The focus of MINT was primarily disability and general health-related quality

of life. Surrogates and impairment measures such as ROM were not a priority and therefore were not repeatedly measured. Repeated measures of ROM would have added considerable cost to the research budget in order to bring patients back for a clinical assessment and additional burden to the participant that some may not have accepted.

MINT evaluated a stepped care approach which involved early re-assessment approximately one month post-injury, at which point cervical ROM was measured. Historically in the NHS, cervical ROM has been measured in the very acute setting (Emergency Department) and then again when patients are received in therapy departments at a much later time point than this study. ROM measurements taken at these different time points may provide alternate diagnostic and prognostic value.

8.2.1.3 Recent evidence

The Neck Pain Task Force conducted a systematic review of interventions for WAD which was published in 2008[171]. They concluded that there was some evidence for educational videos, mobilisation and exercises being more effective than usual care or other physical modalities. This review's conclusions may have been different with the results of MINT incorporated.

8.2.2 CHAPTER THREE - PROGNOSTIC FACTORS FOR LWS

8.2.2.1 Key messages and implications

The key points for the systematic literature review presented in Chapter Three are as follows. A number of physical factors may be important in the prognosis for poor outcome following a whiplash injury. Initial neck pain intensity, initial disability and cold hyperalgesia were found to have moderate evidence to support their prognostic value. The evidence to support

cervical ROM as a prognostic factor was inconclusive, stimulating further investigation that has been achieved by the other studies that make up this thesis.

Methodological quality and reporting of the reviewed articles was variable with no high quality studies and no worthwhile meta-analysis was possible.

8.2.2.2 Limitations

If individual patient data had been sought, then it may have been possible to provided an overall estimate of prognostic value for some factors, however this amount of work was not feasible within the constraints of this doctoral work. From the findings of the systematic review in Chapter Three, it was noted that any further study evaluating prognostic factors should be rigorous in the approach to methodological aspects such as comprehensive conduct and reporting of multivariate measurement and analysis with the use of validated outcome measures.

8.2.2.3 Recent evidence

Since the systematic review of Chapter Three was conducted (searches included work published up to August 2006), five additional systematic reviews have been published, one of which the author co-authored regarding psychosocial prognostic factors for LWS [39]. The four other systematic reviews [295, 298, 299, 320] were generic in their aim to evaluate all prognostic factors for persistent problems following a whiplash injury. Carroll et al [299] conducted a systematic review as part of their work for the Neck Pain Task Force and concluded that approximately 50% of individuals with WAD will have long term problems. They concluded that initial pain and disability, coping style, depression and fear of movement were prognostic for poorer recovery. Kamper et al [298] conducted a systematic review and meta-analysis, however the authors were unable to pool results, despite only

considering the results of univariate analyses. They concluded that data regarding prognostic factors was too difficult to interpret to make definitive statements. Lakke et al [320] conducted a review of all prognostic factors for musculoskeletal pain, not just those for WAD. They concluded that there was strong evidence that older age, being female, having angular deformity of the neck, and having an acute psychological response were not prognostic factors. Finally, Walton et al [295] conducted a synthesis and meta-analysis of prognostic studies of WAD and concluded that initial neck pain intensity, WAD grade III, presence of headache, and no post-secondary education were strong predictors of poor recovery.

In addition to these reviews the author has performed an up-date of the searches for the systematic review from September 2006 until January 2011. The results of which are presented in Appendix 13. 15 articles [107, 159, 321-333] describing 13 cohorts were accepted once duplicates were removed and abstracts and full texts had been screened using the eligibility criteria previously described in Chapter Three. It was not possible to have the articles reviewed by a second reviewer due to limitations of time, so the synthesis described below was solely conducted by the author. The methodological and reporting quality was variable, and there is evidence of an increase in use of validated measures since the conduct of the last systematic review. Use of a variety of different outcome measures remained, with neck pain being used most frequently (7 cohorts) but inconsistently in its definition (VAS, dichotomised VAS, presence/absence). It is noted that the most frequently used, validated disability measure was the Neck Disability Index (two cohorts; [327, 330]). New physical factors to be investigated were palpation tenderness [112] and smooth pursuit eye movement and neck torsion testing [323]. The former were found to be independently predictive of reduced working ability 12 months after injury. Results of all other investigations of physical

factors do not provide any changes in the conclusions made at the time of the initial review. Initial neck pain intensity and disability still provide the greatest prognostic value for poor outcome following a whiplash injury. With respect to investigation of cervical ROM, two of the new cohorts evaluated this as a prognostic factor with contrary results. The overall conclusions of the review regarding cervical ROM do not change when these two new studies are incorporated into the synthesis; cervical ROM is inconclusive as a prognostic factor for poor outcome following a whiplash injury. Referring back to the issue of meta-analysis highlighted previously, with the inclusion of the new studies obtained from the up-dated search, it is still not possible that a meaningful pooling of data can be carried out. Although this systematic review only focussed on physical prognostic factors for poor outcome following a whiplash injury, it is acknowledged that psychological factors are important. The relative importance of physical and psychological factors is still not known.

8.2.3 CHAPTER FOUR – ASSESSMENT OF THE CERVICAL SPINE IN WHIPLASH ASSOCIATED DISORDERS

8.2.3.1 Key messages and implications

The key learning points from the theoretical discussions regarding cervical spine assessment of individuals experiencing WAD (Chapter Four) are as follows. The cervical spine is of primary concern following a whiplash-mechanism of injury and the resultant Whiplash-Associated Disorders (e.g. 99% complained of cervical spine problems in MINT), although other areas of the body are commonly involved. Evaluation of cervical ROM is used by clinicians as an important part of clinical assessment. Furthermore, it was argued that active and passive forms of cervical ROM provide different information and when both are measured simultaneously in the same assessment, this can facilitate diagnosis of the problem. Clinician measured cervical ROM has widely been measured in previous research

of WAD populations in prognostic and intervention studies but active and passive ROM have not been measured simultaneously.

8.2.3.2 Limitations

Chapter Four provided an argument that cervical ROM is evaluated in active and passive forms and is deemed important in the assessment of WAD. However, no direct empirical evidence was available to validly conclude that UK clinicians working with this patient group hold these beliefs. In order to do this a survey would have to be performed, which was not feasible within the limits this doctoral work considering the other studies that had to be prioritised. Such a survey could provide answers to questions such as how important range of motion is believed to be relative to other aspects of the clinical assessment process and how active and passive forms of ROM are used in the diagnostic and prognostic process in this patient group. The proposed implications of the findings of this thesis may then be refined and implementation and dissemination made more effective.

8.2.3.3 Recent evidence

Van Trijffel et al [334] performed a questionnaire study of 367 Dutch manual therapists, enquiring about the use of passive inter-vertebral movements in assessment and concluded that therapists felt that assessment of end-feel and provocation of pain were important for diagnosis and subsequent treatment decisions. The patient's rating of their own cervical ROM is less commonly evaluated and has so far received very limited research attention despite its potential benefit in the diagnostic and prognostic process.

8.2.4 CHAPTER FIVE – SYSTEMATIC REVIEW OF RELIABILITY AND VALIDITY STUDIES OF METHODS FOR THE MEASUREMENT OF ACTIVE AND PASSIVE CERVICAL ROM

8.2.4.1 Key messages and implications

In Chapter Five, a systematic literature review of reliability and validity studies of methods for measuring active and passive cervical ROM was presented. The key findings of this review were that there are a wide range of methods available, with a wide variation in classification of their clinimetric properties. It was apparent that most instruments had not rigorously been investigated. It was difficult to conclude whether reliability and validity was consistently different between active and passive forms of cervical ROM due to the small number of studies evaluating passive cervical ROM. The CROM device was deemed the most reliable and valid device for use in measuring both active and passive cervical ROM in a symptomatic population, with both the Spin-T goniometer and single inclinometer providing acceptable alternatives. It was concluded that visual estimation was the least reliable and concurrently valid method of measuring cervical ROM. As a result of these findings, the CROM device was selected for use in the cohort study described in Chapter Seven, although it was noted that the CROM had yet to undergo reliability testing in a WAD population, highlighting the need for the reliability studies described in Chapter Six.

Methodological and reporting quality was assessed as part of the systematic review process. A large range in quality assessment scores led to the conclusion that authors should attempt to adhere to standardised conduct and reporting methods set out in the STARD checklist. By providing important information such as a flow of participants through the study, description of any blinding procedures and details of any missing data and how this was handled helps to inform readers of articles of potential sources of bias. With this awareness, the author

ensured that the reliability studies described in Chapter Six adhered to the STARD guidelines as best as practicably possible (see discussion of this later).

8.2.4.2 Limitations

Another issue related to the consistency of methodology and reporting is that of meta-analysis. As with the systematic review in Chapter Three, this systematic review was unable to lead to legitimate pooling of study data to provide a single estimate of reliability/validity for the devices. This was because of the large heterogeneity of study outcomes. Unlike the systematic review of prognostic factors, is it is unlikely that data synthesis would have been possible with the availability of individual patient data. For example, for intra-observer reliability of the CROM device, the time between tests ranged from consecutively to weeks for each of the seven studies that would have been eligible for combining. This source of clinical heterogeneity is unlikely to produce a meaningful overall estimate of the reliability within a single observer.

Development of a quality assessment tool for prognostic studies was challenging with no pre-existing consensus to draw on [335]. Since the conduct of this systematic review there does not appear to be any further work on developing a quality assessment tool although other authors have suggested that the QUADAS tool be modified to suit the nature of the included studies as was done in this review [336].

Although the CROM device was chosen as the method for evaluating active and passive cervical ROM in the cohort study based on the results of the systematic review and the reliability studies of Chapter Six, it is clear that there is still considerable further research required to find a clinimetrically excellent tool. Some of this improvement should be

obtained from optimised testing protocols. This topic is discussed further in the summary section of Chapter Six below.

8.2.4.3 Recent evidence

Since conducting the systematic literature review of Chapter Five (searches were up to January 2009) the author has updated the review for the purposes of this discussion chapter. Searches were re-run from February 2009 until January 2011 using the strategy previously described in Chapter Five, The results of the searches are presented in the flow chart in Appendix 14. Eight articles [142, 337-343] describing 16 reliability and 10 validity studies were found (5 articles described a reliability and validity study in the same article). The methods evaluated in these newly acquired articles were the CROM device, the Flock of Birds device (3D electromagnetic tracking), single inclinometer, digital inclinometer, visual estimation, tape measure and universal goniometer. Results were generally in line with previous findings for these tools and the conclusions of the systematic review are unchanged when taking these into account. A notable finding was reported by Whitcroft et al [343], who reported visual estimation to have good reliability; however the methodology and reporting for this study are of questionable quality as it is uncertain on how many participants this result was based – it could have been on as little as one participant. The CROM device was evaluated in a further five reliability studies [337-339, 343] and one validity study [337] and consistently found to demonstrate good clinimetric properties. It was also the most commonly used reference device for the other recent criterion validity studies (7/10 studies). This reinforces the selection of this device for the cohort study within this thesis.

Only one of the recent articles [339] performed studies using symptomatic participants. It is vital to conduct evaluation in different symptomatic populations if conclusions are to be

generalised into the clinical setting because reliability and validity are only applicable to the population and measurement protocol used at that time.

In the discussion of Chapter Five it was stated that at the time of completion of the review there were no other systematic literature reviews that evaluated reliability and validity studies for both active and passive cervical ROM measurement tools. Having performed the searches again, no other systematic review has been located, indicating that this systematic review currently remains a novel contribution to the evidence base.

8.2.5 CHAPTER SIX – INTRA- AND INTER-OBSERVER RELIABILITY OF THE CROM DEVICE IN A WAD POPULATION

8.2.5.1 Key messages and implications

In Chapter Six, intra- and inter-observer reliability studies of the CROM device were presented. Key findings from these studies were that the CROM device is substantially reliable for measuring both active and passive cervical spine ROM in a WAD population. The findings from the variety of statistical techniques used were consistent in these findings. This provided justification for the selection of the CROM device for the prospective cohort study of Chapter Seven and therefore utilised the standardised measurement protocol that had been developed.

Confidence was taken from the fact that characteristics of the reliability study participants were very similar to those of the entire cohort described in Chapter Seven. The participants involved in the reliability studies tolerated repetition to the extent that for all assessments bar one, two measurements of each direction of ROM were able to be taken. Future studies may therefore be justified in investigating more observations/observers at a single time.

Reliability and validity of the CROM device has only been evaluated in sitting, however passive cervical ROM is often evaluated in supine in a clinical setting. Further research could evaluate the reliability of measurements with a modified CROM device used in a supine position.

8.2.5.2 Limitations

The main limitation of the reliability studies described in Chapter Six is that the inter-observer reliability study recruited only 50% of the target sample size. This may have introduced Type I or II errors (most likely Type II); however confidence intervals around the reliability estimates were not wide ranging. The Intra- and Inter-observer reliability of the CROM device was investigated for one and two of the MINT research clinicians respectively. It would have been more rigorous to include all the research clinicians for MINT and for the studies included here. Assembling all the research clinicians at once or even performing a rigorous intra-tester study for each of the 22 clinicians was logistically unfeasible for the trial and doctoral work of this size. Also, the result of using more observers is an increase in the sample size required [267]. As already mentioned, the study protocol did involve each research clinician undergoing a quality control visit to check they were performing the research clinics (and the included assessment) according to documented protocol.

It was unknown what influence certain aspects of the measurement protocol had on reliability of the device e.g. testing position, instructions, warm-up, repetitions etc. This was unable to be evaluated due to limitations of a project of this size.

8.2.5.3 Recent evidence

As summarised in the previous chapter's section, the reliability of the CROM device has been evaluated in five studies published in the last three years. All the studies found the CROM device to have good reliability, although none of them conducted evaluation in a WAD population.

The reliability studies in Chapter Six are reported according to the STARD guidelines [344], which although were not designed specifically for reliability study reporting were the best available at the time and were adapted appropriately. In January 2011, new guidelines for the reporting of reliability studies were published by Kottner et al [345]. These guidelines are welcomed to answer the calls made for more specific guidelines for reporting in Chapters Five and Six in order to improve the standards and therefore facilitate future conduct and synthesis of reliability studies.

8.2.6 CHAPTER SEVEN - PROGNOSTIC COHORT STUDY OF INDIVIDUALS WITH SUB-ACUTE WAD

8.2.6.1 Key messages and implications

In Chapter Seven, the conduct and analysis of a large prospective cohort was described and concluded that cervical ROM is not an independent prognostic factor for poor outcome following a whiplash injury.

Factors that predict the amount of cervical ROM at the time of measurement are both physical (pain and age) and psychological (fear of movement, self-efficacy and depression) in nature. There was no clinically significant difference between active and passive cervical ROM in this WAD population. There was an a-priori expectation that passive cervical ROM

would be greater. Cervical ROM does provide some significant independent explanation of cross-sectional disability.

Longitudinal findings show that cervical ROM is not an independent predictor of neck pain-related disability in a full multivariable model. Psychosocial and other physical factors are more important. Findings reinforce previous research that initial pain intensity is the most important prognostic factor for poor outcome following a whiplash injury. Research now needs to be conducted to interpret what contributes to this initial pain intensity rating. This has already been started. Work by Holm et al [346] show that sociodemographic, pre-injury and collision-related factors are associated but causation has yet to be studied. The findings of this study indicate that shoulder abduction ROM has more prognostic value than active, passive and patient-rated cervical spine ROM when predicting neck pain-related disability in the short and medium term in the absence of research questionnaire scores.

When predicting long term neck disability, it appears that patient-rated cervical ROM is more useful. However this measure does need further development – in other areas of the body there have been apparent advances e.g. Carter et al [347] have used a diagram-based patient reported cervical ROM that appears to be very accurate.

This study highlighted that it is much harder to predict patient-rated recovery than outcome derived from a validated condition specific disability questionnaire (NDI). Definition of recovery and resultant prognostic factors will vary with outcome measures – disability rating is not necessarily the same as patient reported change. Researchers need to be cognisant of this when using findings in research and clinical settings. There is still scope for the development of a measure that is more comprehensive, perhaps combining aspects of pre-existing disability and patient-rated outcome measures.

8.2.6.2 Limitations

Despite the use of the vast majority of potential prognostic factors known at the time of conducting the cohort study and use of multivariate modelling, the amount of variance explained was below 50%. Therefore there are still unknown factors that are affecting prognosis, which if known, may be able to be addressed and therefore improve treatments.

A strength of this study is that initial treatments were standardised and could therefore be controlled for in analyses as necessary and that these treatments are described in sufficient detail (see Chapter Two).

8.2.6.3 Recent evidence

Two recent cohort studies evaluated cervical ROM. Kasch et al [159] used the CROM device and found that reduced active cervical ROM was the strongest independent predictor of handicap (inability to work or prolonged sick leave) in a multivariate analysis of 625 participants with WAD recruited through ED's and GP practices. Borenstein et al [326] found that reduced active cervical ROM was associated with poor outcome (again sick leave) but was not an independent predictor when entered into a multivariate model that also included psychological symptoms, age, sex, initial pain and treatment. The findings of this latter cohort study and the cohort study described in Chapter Seven should now tip the balance of equipoise regarding this factor to conclude that cervical ROM is not a prognostic factor for poor outcome following a whiplash injury,

8.3 CONCLUSIONS

The investigations that make up this thesis have confirmed some existing knowledge (e.g. that WAD affects multiple areas of the body and recovery is multifactorial) and also provided new knowledge which could stimulate further research (prognostic value of

shoulder abduction ROM and patient-rated cervical ROM). The current evidence base suggests the most valuable prognostic factors for poor outcome following a whiplash injury are high levels of initial pain and disability. This was confirmed by a large cohort study, which also found that age, quantity of physical symptoms and psychological factors have some influence on recovery. When physical, psychological, social and demographic factors are analysed together, it is concluded that active, passive and patient-rated forms of cervical ROM have no independent prognostic value for mid and long-term disability. When only standard clinical assessment findings are available from which to make a judgment on prognosis of long term disability, active shoulder abduction ROM and patient-rated cervical ROM (depending on the outcome time point) can be used to provide a limited indication of likelihood of recovery.

There is a significant correlation between cervical ROM and disability at the time of measurement and this study has also provided insight into which physical (pain intensity, age, WAD grade) and psychological (Fear-avoidance, self-efficacy and depression) factors are associated with ROM measurements.

8.4 SUMMARY OF LIMITATIONS OF THIS THESIS

Limitations of each of the studies within this thesis have been discussed in individual chapters; however, it is valuable to revisit the main themes prior to making statements regarding the clinical and research implications of the findings.

With the distinct advantage that is afforded by recruiting participants as part of a larger RCT, there is also the disadvantage that this results in a distinct selection of patients. Large numbers of participants were recruited to the cohort study, however in order to do this, the study had to involve multiple centres and therefore a considerable number of clinicians were

involved in the assessments and data collection process which may introduce unquantified variability for the ROM measurements. Also, because patients were participating in a trial that is providing treatment, this may have inadvertently led to the introduction of bias such as the Hawthorne effect [348]. Within the reliability studies conducted with the CROM device, there is the possibility for numerous sources of variation which were out of the control of the observers, despite their best efforts at following the standardised measurement protocol.

8.5 CLINICAL IMPLICATIONS AND RECOMMENDATIONS

As intended, findings from the studies presented in this thesis have direct clinical implications that can be incorporated into the management of WAD patients.

Firstly, if clinicians require a reliable method of measuring cervical ROM then the CROM device can be recommended for use with WAD patients. It appears that visual estimation is not reliable for measuring cervical ROM generally.

Secondly, active and passive cervical ROM measurements provide almost identical value for predicting disability at the time of measurement. Clinicians should be aware that cervical ROM measurements are influenced by both physical (primarily pain, but also age and injury severity) and psychological (fear avoidance, self-efficacy and depression) factors in patients with WAD.

Results of the analyses of the prospective cohort study provide evidence that clinicians can screen patients at approximately one month post-injury for poor prognosis using a limited number of measures. Using a combination of the patient's baseline disability rating, age, rating of distress and the number of physical symptoms helps to identify which patients may have a poor outcome following a whiplash injury.

If the clinical assessment is limited by time and/or access to some of these measures, a standard clinical assessment does provide some prognostic value. Clinicians should particularly note high intensity of pain at the time of assessment, a high number of physical symptoms around the patient's body and reduction in active shoulder abduction.

The findings from the cohort study indicate that when clinicians are assessing patients with sub-acute WAD approximately one month post injury they can be confident in reassuring their patients that even though they may have reduced cervical ROM at that point in time this does not mean that they will necessarily be functionally impaired in the future when other, particularly psychosocial, factors are taken into account.

Findings from MINT indicate that it probably doesn't matter which intervention patients are offered in terms of long-term functional outcome. A sensitivity analysis of the cohort study indicates that treatment group and cervical ROM did not interact and therefore possibly cervical ROM will improve irrespective of whether patients receive an advice session or package of physiotherapy. Indeed, studies summarised in Chapter Four support this hypothesis. Further studies in which cervical ROM is measured post intervention are required to fully answer this question.

8.6 FUTURE RESEARCH

Future research areas have been discussed in each of the chapters and questions generated by work within this thesis have been discussed in the individual chapter overview sections above. Points below formalise the main areas for future research.

- Development of methodological quality assessment criteria specifically for prognostic studies
- Development of modification of the STARD checklist for the methodological quality assessment of reliability and validity studies
- Intra- and Inter-observer reliability study of the CROM device in a WAD population with increased re-test time periods e.g. one week, to determine confidence that can be assigned with using CROM device measurements to determine changes due to treatment or natural recovery
- Reliability and criterion validity studies involving other methods for measurement of cervical spine ROM in a WAD population. It would be worthwhile to validate the CROM device against another method e.g. radiography or magnetic resonance imaging
- Further investigation of the role of shoulder ROM in diagnosis and prognosis of WAD

- Development work investigating the value of patient-rated cervical ROM and in particular the optimizing of a measurement tool for this e.g. VAS or percentage score

- Survey clinician's attitudes and beliefs regarding the use of both active and passive forms of ROM and the relative emphasis placed on these measures for diagnosis and prognosis compared to other aspects of motion assessment (e.g. "quality" of movement, speed, "ease" etc.)

8.7 SUMMARY

This chapter aimed to draw together the findings from the preceding seven chapters and provide overall conclusions for this doctoral work. Key messages, implications and limitations of findings and discussion of relevant recent literature have been provided.

9 REFERENCES

1. CROWE H: **A NEW DIAGNOSTIC SIGN IN NECK INJURIES.** *Calif Med* 1964, **100**:12-13.
2. GAY JR, ABBOTT KH: **Common whiplash injuries of the neck.** *J Am Med Assoc* 1953, **152**(18):1698-1704.
3. SEVERY DM, MATHEWSON JH, BECHTOL CO: **Controlled automobile rearend collisions, an investigation of related engineering and medical phenomena.** *Can Serv Med J* 1955, **11**(10):727-759.
4. Kaneoka K, Ono K, Inami S, Hayashi K: **Motion analysis of cervical vertebrae during whiplash loading.** *Spine (Phila Pa 1976)* 1999, **24**(8):763-769; discussion 770.
5. Grauer J, Panjabi M, Cholewicki J, Nibu K, Dvorak J: **Whiplash produces an S-shaped curvature of the neck with hyperextension at lower levels.** *Spine (Phila Pa 1976)* 1997, **22**(21):2489-2494.
6. Spitzer WO, Skovron ML, Salmi LR, Cassidy JD, Duranceau J, Suissa S, Zeiss E: **Scientific monograph of the Quebec Task Force on Whiplash-Associated Disorders: redefining "whiplash" and its management.[see comment][erratum appears in Spine 1995 Nov 1;20(21):2372]**. *Spine* 1995, **20**(8 Suppl):1S-73S.
7. Quinlan KP, Annest JL, Myers B, Ryan G, Hill H: **Neck strains and sprains among motor vehicle occupants-United States, 2000.** *Accid Anal Prev* 2004, **36**(1):21-27.
8. Galasko CSB, Murray P, Stephenson W: **Incidence of whiplash-associated disorder**. *British Columbia Medical Journal* 2002, **44**(5).
9. Pennie B, Agambar L: **Patterns of injury and recovery in whiplash**. *Injury* 1991, **22**(1):57-59.
10. Karlsborg M, Smed A, Jespersen H, Stephensen S, Cortsen M, Jennum P, Herning M, Korfitsen E, Werdelin L: **A prospective study of 39 patients with whiplash injury**. *Acta Neurologica Scandinavica* 1997, **95**(2):65-72.
11. Bostick G, Ferrari R, Carroll L, Russell A, Buchbinder R, Krawciw D, Gross D: **A population-based survey of beliefs about neck pain from whiplash injury, work-related neck pain, and work-related upper extremity pain.** *Eur J Pain* 2009, **13**(3):300-304.
12. Balla JI: **The late whiplash syndrome.** *Aust N Z J Surg* 1980, **50**(6):610-614.
13. Verhagen A, Scholten-Peeters G, van Wijngaarden S, de Bie R, Bierma-Zeinstra S: **Conservative treatments for whiplash.** *Cochrane Database Syst Rev* 2007(2):CD003338.
14. Lamb SE, Williams MA, Withers E, Perry J, Gates S, Williamson E, Underwood M, Cooke M: **A national survey of clinical practice for the management of whiplash-associated disorders in UK emergency departments.** *Emerg Med J* 2009, **26**(9):644-647.
15. Lamb SE, Gates S, Underwood MR, Cooke MW, Ashby D, Szczepura A, Williams MA, Williamson EM, Withers EJ, Mt Isa S *et al*: **Managing Injuries of the Neck Trial (MINT): design of a randomised controlled trial of treatments for whiplash associated disorders.** *BMC Musculoskelet Disord* 2007, **8**:7.

16. Scholten-Peeters GG, Bekkering GE, Verhagen AP, van Der Windt DA, Lanser K, Hendriks EJ, Oostendorp RA: **Clinical practice guideline for the physiotherapy of patients with whiplash-associated disorders.[see comment]**. *Spine* 2002, **27**(4):412-422.
17. McClune T, Burton AK, Waddell G: **Whiplash associated disorders: a review of the literature to guide patient information and advice.** *Emerg Med J* 2002, **19**(6):499-506.
18. Cote P, Cassidy JD, Carroll L, Frank JW, Bombardier C: **A systematic review of the prognosis of acute whiplash and a new conceptual framework to synthesize the literature.[see comment]**. *Spine* 2001, **26**(19):E445-458.
19. Waddell G, Buton K, McClune T: **The Whiplash Book**. London: The Stationery Office; 2002.
20. McClune T, Burton AK, Waddell G: **Evaluation of an evidence based patient educational booklet for management of whiplash associated disorders.** *Emerg Med J* 2003, **20**(6):514-517.
21. Teasdale G, Jennett B: **Assessment of coma and impaired consciousness. A practical scale.** *Lancet* 1974, **2**(7872):81-84.
22. Eldridge SM, Ashby D, Feder GS: **Informed patient consent to participation in cluster randomized trials: an empirical exploration of trials in primary care.** *Clin Trials* 2005, **2**(2):91-98.
23. Vernon H, Mior S: **The Neck Disability Index: a study of reliability and validity**. *Journal of manipulative and physiological therapeutics* 1991, **14**(7):409-415.
24. Rosenfeld M, Seferiadis A, Carlsson J, Gunnarsson R: **Active intervention in patients with whiplash-associated disorders improves long-term prognosis: a randomized controlled clinical trial**. *Spine* 2003, **28**(22):2491-2498.
25. Hoving J, Koes B, de Vet H, van der Windt D, Assendelft W, van Mameren H, Devillé W, Pool J, Scholten R, Bouter L: **Manual therapy, physical therapy, or continued care by a general practitioner for patients with neck pain. A randomized, controlled trial.** *Ann Intern Med* 2002, **136**(10):713-722.
26. Ware J, Kosinski M, Keller SD: **A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity.** *Med Care* 1996, **34**(3):220-233.
27. Kind P, Dolan P, Gudex C, Williams A: **Variations in population health status: results from a United Kingdom national questionnaire survey.** *BMJ* 1998, **316**(7133):736-741.
28. Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M, Guidance MRC: **Developing and evaluating complex interventions: the new Medical Research Council guidance.** *BMJ* 2008, **337**:a1655.
29. Verhagen AP, Scholten-Peeters GG, de Bie RA, Bierma-Zeinstra SM: **Conservative treatments for whiplash**. *Cochrane database of systematic reviews (Online)* 2004, **(1):CD003338. </pre></dd> </dl> <hr>
<table cellpadding="0" cellspacing="5" width="100%"><tr><td valign="top" width="100%"><input name="uid" type="checkbox" value="15324774">2: <a href="http://www.ncbi.nlm.(TRUNCATED)**(1):CD003338.
30. Miró J, Nieto R, Huguet A: **The Catalan version of the Pain Catastrophizing Scale: a useful instrument to assess catastrophic thinking in whiplash patients.** *J Pain* 2008, **9**(5):397-406.

31. Maitland GD: **Maitland's vertebral manipulation**. Oxford: Butterworth Heinemann; 2001.
32. WILLIAMSON E, WILLIAMS M, HANSEN Z, JOSEPH S, LAMB S: **Development and delivery of a physiotherapy intervention for the early management of whiplash injuries: The Managing Injuries of Neck Trial (MINT) Intervention**. *Physiotherapy* 2009, **95**(1):15-23.
33. Stratford PW, Riddle DL, Binkley JM, Spadoni G, Westaway MD, Padfield B: **Using the Neck Disability Index to make decisions concerning individual patients**. *Physiotherapy Canada* 1999, **Spring**:107-112.
34. Cooke M, Marsh J, Clark M, Nakash R, Jarvis R, Hutton J, Szczepura A, Wilson S, Lamb S: **Treatment of severe ankle sprain: a pragmatic randomised controlled trial comparing the clinical effectiveness and cost-effectiveness of three types of mechanical ankle support with tubular bandage. The CAST trial.** *Health Technol Assess* 2009, **13**(13):iii, ix-x, 1-121.
35. Simon R, Altman D: **Statistical aspects of prognostic factor studies in oncology.** *Br J Cancer* 1994, **69**(6):979-985.
36. Altman D, Lyman G: **Methodological challenges in the evaluation of prognostic factors in breast cancer.** *Breast Cancer Res Treat* 1998, **52**(1-3):289-303.
37. **Undertaking systematic reviews of research and effectiveness**
38. Scholten-Peeters GG, Verhagen AP, Bekkering GE, van der Windt DA, Barnsley L, Oostendorp RA, Hendriks EJ: **Prognostic factors of whiplash-associated disorders: a systematic review of prospective cohort studies.[see comment]**. *Pain* 2003, **104**(1-2):303-322.
39. WILLIAMSON E, WILLIAMS M, GATES S, LAMB S: **A systematic literature review of psychological factors and the development of late whiplash syndrome**. *Pain* 2008, **135**(1-2):20-30.
40. WILLIAMS M, WILLIAMSON E, GATES S, LAMB S, COOKE M: **A systematic literature review of physical prognostic factors for the development of Late Whiplash Syndrome**. *Spine* 2007, **32**(25):E764-E780.
41. Moher D, Cook D, Eastwood S, Olkin I, Rennie D, Stroup D: **Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. Quality of Reporting of Meta-analyses.** *Lancet* 1999, **354**(9193):1896-1900.
42. Moher D, Liberati A, Tetzlaff J, Altman D: **Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement.** *BMJ* 2009, **339**:b2535.
43. Egger M, Smith GD, Altman DG: **Systematic Reviews in Health Care - Meta-analysis in context**. London: BMJ Books; 2001.
44. Bornman J: **The World Health Organisation's terminology and classification: application to severe disability**. *Disability and rehabilitation* 2004, **26**(3):182-188.
45. Norris SH, Watt I: **The prognosis of neck injuries resulting from rear-end vehicle collisions**. *Journal of Bone & Joint Surgery - British Volume* 1983, **65**(5):608-611.
46. Côté P, Hogg-Johnson S, Cassidy JD, Carroll L, Frank JW: **The association between neck pain intensity, physical functioning, depressive symptomatology and time-to-claim-closure after whiplash.** *J Clin Epidemiol* 2001, **54**(3):275-286.
47. Pearce JM: **The late whiplash syndrome: a biopsychosocial approach.** *J Neurol Neurosurg Psychiatry* 2001, **71**(6):820-821.

48. Harder S, Veilleux M, Suissa S: **The effect of socio-demographic and crash-related factors on the prognosis of whiplash**. *Journal of clinical epidemiology* 1998, **51**(5):377-384.
49. Smith B, Darzins P, Quinn M, Heller R: **Modern methods of searching the medical literature.** *Med J Aust* 1992, **157**(9):603-611.
50. Hunt DL, Jaeschke R, McKibbon KA: **Users' guides to the medical literature: XXI. Using electronic health information resources in evidence-based practice. Evidence-Based Medicine Working Group.** *JAMA* 2000, **283**(14):1875-1879.
51. Altman DG: **Systematic review of evaluations of prognostic variables**. In: *Systematic Reviews in Health Care, Meta-analysis in context.* 2nd edn. Edited by Egger M, Davey Smith G, Altman D. London: BMJ publishing group; 2001.
52. Di Stefano G, Radanov BP: **Course of attention and memory after common whiplash: a two-years prospective study with age, education and gender pair-matched patients**. *Acta Neurologica Scandinavica* 1995, **91**(5):346-352.
53. Voyvodic F, Dolinis J, Moore VM, Ryan GA, Slavotinek JP, Whyte AM, Hoile RD, Taylor GW: **MRI of car occupants with whiplash injury**. *Neuroradiology* 1997, **39**(1):35-40.
54. Warren RAaWMA: **Whiplash injury sustained in motor vehicle accidents:factors influencing time off work**. *The Journal of Orthopaedic Medicine* 2001, **23**(2):50-54.
55. Hohl M: **Soft tissue injuries of the neck**. *Clinical Orthopaedics & Related Research* 1975(109):42-49.
56. Gargan M, Bannister G, Main C, Hollis S: **The behavioural response to whiplash injury**. *The Journal of bone and joint surgeryBritish volume* 1997, **79**(4):523-526.
57. Miles KA, Maimaris C, Finlay D, Barnes MR: **The incidence and prognostic significance of radiological abnormalities in soft tissue injuries to the cervical spine**. *Skeletal radiology* 1988, **17**(7):493-496.
58. Atherton K, Wiles NJ, Lecky FE, Hawes SJ, Silman AJ, Macfarlane GJ, Jones GT: **Predictors of persistent neck pain after whiplash injury**. *Emergency Medicine Journal* 2006, **23**(3):195-201.
59. Borchgrevink GE, Smevik O, Nordby A, Rinck PA, Stiles TC, Lereim I: **MR imaging and radiography of patients with cervical hyperextension-flexion injuries after car accidents**. *Acta radiologica* 1995, **36**(4):425-428.
60. Borchgrevink G, Smevik O, Haave I, Haraldseth O, Nordby A, Lereim I: **MRI of cerebrum and cervical columna within two days after whiplash neck sprain injury**. *Injury* 1997, **28**(5-6):331-335.
61. Brison RJ, Hartling L, Pickett W: **A prospective study of acceleration-extension injuries following rear-end motor vehicle collisions... World Congress on Whiplash-Associated Disorders in Vancouver, British Columbia, Canada in February of 1999**. *Journal of Musculoskeletal Pain* 2000, **8**(1/2):97-113.
62. Hartling L, Brison RJ, Ardern C, Pickett W: **Prognostic value of the Quebec Classification of Whiplash-Associated Disorders**. *Spine* 2001, **26**(1):36-41.
63. Gargan MF, Bannister GC: **The rate of recovery following whiplash injury**. *European Spine Journal* 1994, **3**(3):162-164.
64. Gun RT, Osti OL, O'Riordan A, Mpelasoka F, Eckerwall CG, Smyth JF: **Risk factors for prolonged disability after whiplash injury: a prospective study**. *Spine* 2005, **30**(4):386-391.

65. Hendriks EJ, Scholten-Peeters GG, van der Windt DA, Neeleman-van der Steen CW, Oostendorp RA, Verhagen AP: **Prognostic factors for poor recovery in acute whiplash patients**. *Pain* 2005, **114**(3):408-416.
66. Herrstrom P, Lannerbro-Geijer G, Hogstedt B: **Whiplash injuries from car accidents in a Swedish middle-sized town during 1993-95**. *Scandinavian journal of primary health care* 2000, **18**(3):154-158.
67. Hildingsson C, Toolanen G: **Outcome after soft-tissue injury of the cervical spine. A prospective study of 93 car-accident victims**. *Acta Orthopaedica Scandinavica* 1990, **61**(4):357-359.
68. Kasch H, Bach FW, Jensen TS: **Handicap after acute whiplash injury: a 1-year prospective study of risk factors**. *Neurology* 2001, **56**(12):1637-1643.
69. Kasch H, Qerama E, Bach FW, Jensen TS: **Reduced cold pressor pain tolerance in non-recovered whiplash patients: a 1-year prospective study**. *European journal of pain (London, England)* 2005, **9**(5):561-569.
70. Kivioja J, Jensen I, Lindgren U: **Early coping strategies do not influence the prognosis after whiplash injuries**. *Injury* 2005, **36**(8):935-940.
71. Kyhlback M, Thierfelder T, Soderlund A: **Prognostic factors in whiplash-associated disorders**. *International Journal of Rehabilitation Research* 2002, **25**(3):181-187.
72. Mayou R, Bryant B: **Outcome of 'whiplash' neck injury**. *Injury* 1996, **27**(9):617-623.
73. Miettinen T, Leino E, Airaksinen O, Lindgren KA: **The possibility to use simple validated questionnaires to predict long-term health problems after whiplash injury**. *Spine* 2004, **29**(3):E47-51.
74. Minton R, Murray P, Stephenson W, Galasko CS: **Whiplash injury--are current head restraints doing their job?** *Accident; Analysis and Prevention* 2000, **32**(2):177-185.
75. Nederhand MJ, Hermens HJ, Ijzerman MJ, Turk DC, Zilvold G: **Chronic neck pain disability due to an acute whiplash injury**. *Pain* 2003, **102**(1-2):63-71.
76. Nederhand MJ, Ijzerman MJ, Hermens HJ, Turk DC, Zilvold G: **Predictive value of fear avoidance in developing chronic neck pain disability: consequences for clinical decision making**. *Archives of Physical Medicine and Rehabilitation* 2004, **85**(3):496-501.
77. Olsson I, Bunketorp O, Carlsson SG, Styf J: **Prediction of outcome in whiplash-associated disorders using West Haven-Yale multidimensional pain inventory**. *Clinical Journal of Pain* 2002, **18**(4):238-244.
78. Pettersson K, Hildingsson C, Toolanen G, Fagerlund M, Bjornebrink J: **Disc pathology after whiplash injury. A prospective magnetic resonance imaging and clinical investigation**. *Spine* 1997, **22**(3):283-287; discussion 288.
79. Radanov BP, Di Stefano G, Schnidrig A, Ballinari P: **Role of psychosocial stress in recovery from common whiplash**. *Lancet* 1991, **338**(8769):712-715.
80. Radanov BP, Di Stefano G, Schnidrig A, Sturzenegger M, Augustiny KF: **Cognitive functioning after common whiplash: A controlled follow-up study**. *Archives of Neurology* 1993, **50**(1):87-91.
81. Radanov BP, Sturzenegger M, Di Stefano G, Schnidrig A, Aljinovic M: **Factors influencing recovery from headache after common whiplash**. *British medical journal* 1993, **307**(6905):652-655.
82. Radanov BP, Sturzenegger M, De Stefano G, Schnidrig A: **Relationship between early somatic, radiological, cognitive and psychosocial findings and outcome**

during a one-year follow-up in 117 patients suffering from common whiplash. *British journal of rheumatology* 1994, **33**(5):442-448.
83. Radanov BP, Sturzenegger M, Di Stefano G: **Long-term outcome after whiplash injury: A 2-year follow-up considering features of injury mechanism and somatic, radiologic, and psychosocial findings**. *Medicine* 1995, **74**(5):281-297.
84. Sturzenegger M, Radanov BP, Di Stefano G: **The effect of accident mechanisms and initial findings on the long-term course of whiplash injury**. *Journal of neurology* 1995, **242**(7):443-449.
85. Richter M, Ferrari R, Otte D, Kuensebeck HW, Blauth M, Krettek C: **Correlation of clinical findings, collision parameters, and psychological factors in the outcome of whiplash associated disorders**. *Journal of Neurology, Neurosurgery & Psychiatry* 2004, **75**(5):758-764.
86. Sterling M, Jull G, Vicenzino B, Kenardy J, Darnell R: **Physical and psychological factors predict outcome following whiplash injury**. *Pain* 2005, **114**(1-2):141-148.
87. Sterling M, Jull G, Kenardy J: **Physical and psychological factors maintain long-term predictive capacity post-whiplash injury**. *Pain* 2006, **122**(1-2):102-108.
88. Sterner Y, Toolanen G, Gerdle B, Hildingsson C: **The incidence of whiplash trauma and the effects of different factors on recovery**. *Journal of Spinal Disorders & Techniques* 2003, **16**(2):195-199.
89. Crombez G, Vlaeyen JW, Heuts PH, Lysens R: **Pain-related fear is more disabling than pain itself: evidence on the role of pain-related fear in chronic back pain disability**. *Pain* 1999, **80**(1-2):329-339.
90. Easterbrook PJ, Berlin JA, Gopalan R, Matthews DR: **Publication bias in clinical research.[see comment]**. *Lancet* 1991, **337**(8746):867-872.
91. Foletti GB, Regli F: **Indirect cervical trauma: long-term adverse prognostic factors**. *Schweizerische Rundschau fur Medizin Praxis = Revue suisse de medecine Praxis* 1987, **76**(47):1304-1309.
92. Kageyama N, Tanaka M, Ikeda K, Nakajima S, Someda K: **[Traumatic cervical syndromes--diagnosis, therapy and prognosis]**. *No to Shinkei - Brain & Nerve* 1969, **21**(8):893-901.
93. Huber A, Beran H, Trenkler J, Hager A, Witzmann A, Fischer J: **Whiplash trauma of the cervical spine from the neurosurgical, traumatologic and psychologic viewpoint**. *Neurochirurgia* 1993, **36**(2):51-55.
94. Pujol A, Puig L, Mansilla J, Idiaquez I: **[Relevant factors in medico-legal prognosis of whiplash injury]**. *Medicina clinica* 2003, **121**(6):209-215.
95. Moog M, Quintner J, Hall T, Zusman M: **The late whiplash syndrome: a psychophysical study**. *European journal of pain (London, England)* 2002, **6**(4):283-294.
96. Bogduk N, Mercer S: **Biomechanics of the cervical spine. I: Normal kinematics.** *Clin Biomech (Bristol, Avon)* 2000, **15**(9):633-648.
97. Neumann DA: **Regional kinematics of the spine**. St Louis, Missouri: Mosby; 2002.
98. Mercer S, Bogduk N: **The ligaments and annulus fibrosus of human adult cervical intervertebral discs.** *Spine (Phila Pa 1976)* 1999, **24**(7):619-626; discussion 627-618.
99. Wyke B: **The neurology of joints.** *Ann R Coll Surg Engl* 1967, **41**(1):25-50.
100. Neumann DA: **Regional kinematics of the spine**. In: *Kinesiology of the musculoskeletal system.* edn. St Louis, Missouri: Mosby; 2002: 278.
101. Chen J, Solinger AB, Poncet JF, Lantz CA: **Meta-analysis of normative cervical motion**. *Spine* 1999, **24**(15):1571-1578.

102. Uhrenholt L, Grunnet-Nilsson N, Hartvigsen J: **Cervical spine lesions after road traffic accidents: a systematic review.** *Spine (Phila Pa 1976)* 2002, **27**(17):1934-1941; discussion 1940.
103. Taylor JR, Twomey LT: **Acute injuries to cervical joints. An autopsy study of neck sprain.** *Spine (Phila Pa 1976)* 1993, **18**(9):1115-1122.
104. Barnsley L, Lord SM, Wallis BJ, Bogduk N: **The prevalence of chronic cervical zygapophysial joint pain after whiplash.** *Spine (Phila Pa 1976)* 1995, **20**(1):20-25; discussion 26.
105. Krakenes J, Kaale BR: **Magnetic resonance imaging assessment of craniovertebral ligaments and membranes after whiplash trauma.** *Spine (Phila Pa 1976)* 2006, **31**(24):2820-2826.
106. Johansson BH: **Whiplash injuries can be visible by functional magnetic resonance imaging.** *Pain Res Manag* 2006, **11**(3):197-199.
107. Kongsted A, Sorensen JS, Andersen H, Keseler B, Jensen TS, Bendix T: **Are early MRI findings correlated with long-lasting symptoms following whiplash injury? A prospective trial with 1-year follow-up.** *Eur Spine J* 2008, **17**(8):996-1005.
108. Matsumoto M, Okada E, Ichihara D, Chiba K, Toyama Y, Fujiwara H, Momoshima S, Nishiwaki Y, Hashimoto T, Inoue T *et al*: **Prospective ten-year follow-up study comparing patients with whiplash-associated disorders and asymptomatic subjects using magnetic resonance imaging.** *Spine (Phila Pa 1976)* 2010, **35**(18):1684-1690.
109. Cagnie B, Dolphens M, Peeters I, Achten E, Cambier D, Danneels L: **Use of muscle functional magnetic resonance imaging to compare cervical flexor activity between patients with whiplash-associated disorders and people who are healthy.** *Phys Ther* 2010, **90**(8):1157-1164.
110. Gerdle B, Lemming D, Kristiansen J, Larsson B, Peolsson M, Rosendal L: **Biochemical alterations in the trapezius muscle of patients with chronic whiplash associated disorders (WAD)--a microdialysis study.** *Eur J Pain* 2008, **12**(1):82-93.
111. Falla D, Bilenkij G, Jull G: **Patients with chronic neck pain demonstrate altered patterns of muscle activation during performance of a functional upper limb task.** *Spine (Phila Pa 1976)* 2004, **29**(13):1436-1440.
112. Kasch H, Qerama E, Kongsted A, Bach FW, Bendix T, Jensen TS: **Deep muscle pain, tender points and recovery in acute whiplash patients: a 1-year follow-up study.** *Pain* 2008, **140**(1):65-73.
113. Sterling M, Jull G, Vicenzino B, Kenardy J: **Sensory hypersensitivity occurs soon after whiplash injury and is associated with poor recovery**. *Pain* 2003, **104**(3):509-517.
114. Kenneally M, Rubenach H, Elvey R: **The upper limb tension test: the SLR of the arm**. In: *Physical therapy of the Cervical and Thoracic Spine.* edn. Edited by Grant R. New York: Churchill Livingstone; 1988: 167-194.
115. Heikkilä H, Wenngren B: **Cervicocephalic kinesthetic sensibility, active range of cervical motion, and oculomotor function in patients with whiplash injury.** *Arch Phys Med Rehabil* 1998, **79**(9):1089-1094.
116. Armstrong BS, McNair PJ, Williams M: **Head and neck position sense in whiplash patients and healthy individuals and the effect of the cranio-cervical flexion action.** *Clin Biomech (Bristol, Avon)* 2005, **20**(7):675-684.

117. Treleaven J, Jull G, Sterling M: **Dizziness and unsteadiness following whiplash injury: characteristic features and relationship with cervical joint position error.** *J Rehabil Med* 2003, **35**(1):36-43.
118. Twomey LT, Taylor JR: **The Maitland Concept: Assessment, Examination, and Treatment by Passive Movement**: Churchill Livingstone; 1987.
119. Koes BW, van Tulder MW, Ostelo R, Kim Burton A, Waddell G: **Clinical guidelines for the management of low back pain in primary care: an international comparison.** *Spine (Phila Pa 1976)* 2001, **26**(22):2504-2513; discussion 2513-2504.
120. Cyriax J: **Cyriax's illustrated manual of orthopaedic medicine**, 2nd edn. Oxford: Butterworth-Heinemann; 1996.
121. Evans RC: **Illustrated orthopedic physical assessment**, 3rd edn. St. Louis: Mosby/Elsevier; 2009.
122. Petty N, Moore A. In: *Neuromusculoskeletal Examination and Assessment: A Handbook for Therapists.* 1st edn. Edinburgh: Churchill Livingstone; 1998: 47.
123. Petty N, Moore A: **Neuromusculoskeletal Examination and Assessment: A Handbook for Therapists**. In., 1st edn.: Churchill Livingstone; 1998: 47.
124. Seffinger M, Najm W, Mishra S, Adams A, Dickerson V, Murphy L, Reinsch S: **Reliability of spinal palpation for diagnosis of back and neck pain: a systematic review of the literature.** *Spine* 2004, **29**(19):E413-425.
125. Norkin CCaWDJ: **Measurement of Joint Motion - A guide to Goniometry**. Philadelphia, USA: F.A. Davis Company; 2003.
126. Maitland GD: **Maitland's Vertebral Manipulation**. In., Sixth edn. Oxford: Butterworth-Heinemann; 2001: 5.
127. AMA: **Guides to the evaluation of permanent impairment**. Chicago: AMA; 1993.
128. Toomingas A, Németh G, Alfredsson L: **Self-administered examination versus conventional medical examination of the musculoskeletal system in the neck, shoulders, and upper limbs. The Stockholm MUSIC I Study Group.** *J Clin Epidemiol* 1995, **48**(12):1473-1483.
129. Maher CG, Simmonds M, Adams R: **Therapists' conceptualization and characterization of the clinical concept of spinal stiffness.** *Phys Ther* 1998, **78**(3):289-300.
130. Drottning M, Staff P, Sjaastad O: **Cervicogenic headache (CEH) after whiplash injury.** *Cephalalgia* 2002, **22**(3):165-171.
131. Borchgrevink G, Kaasa A, McDonagh D, Stiles T, Haraldseth O, Lereim I: **Acute treatment of whiplash neck sprain injuries. A randomized trial of treatment during the first 14 days after a car accident.** *Spine (Phila Pa 1976)* 1998, **23**(1):25-31.
132. Bennett S, Schenk R, Simmons E: **Active range of motion utilized in the cervical spine to perform daily functional tasks.** *J Spinal Disord Tech* 2002, **15**(4):307-311.
133. Bible JE, Biswas D, Miller CP, Whang PG, Grauer JN: **Normal functional range of motion of the cervical spine during 15 activities of daily living.** *J Spinal Disord Tech* 2010, **23**(1):15-21.
134. Klaber-Moffett J, Hughes I, Griffiths P: **Measurement of cervical spine movements using a simple inclinometer**. *Physiotherapy* 1989, **75**(6):309-312.
135. Petersen C, Johnson R, Schuit D: **Reliability of cervical range of motion using the OSI CA 6000 spine motion analyser on asymptomatic and symptomatic subjects.** *Man Ther* 2000, **5**(2):82-88.

136. Antonaci F, Bulgheroni M, Ghirmai S, Lanfranchi S, Dalla Toffola E, Sandrini G, Nappi G: **3D kinematic analysis and clinical evaluation of neck movements in patients with whiplash injury**. *Cephalalgia : an international journal of headache* 2002, **22**(7):533-542.
137. Dall'Alba PT, Sterling MM, Treleaven JM, Edwards SL, Jull GA: **Cervical range of motion discriminates between asymptomatic persons and those with whiplash**. *Spine* 2001, **26**(19):2090-2094.
138. Bergman G, Knoester B, Assink N, Dijkstra P, Winters J: **Variation in the cervical range of motion over time measured by the "flock of birds" electromagnetic tracking system.** *Spine* 2005, **30**(6):650-654.
139. Barnsley L, Lord S, Bogduk N: **Whiplash injury.** *Pain* 1994, **58**(3):283-307.
140. Kasch H, Stengaard-Pedersen K, Arendt-Nielsen L, Staehelin Jensen T: **Headache, neck pain, and neck mobility after acute whiplash injury: a prospective study**. *Spine* 2001, **26**(11):1246-1251.
141. Turk DC, Robinson JP, Sherman JJ, Burwinkle T, Swanson K: **Assessing fear in patients with cervical pain: development and validation of the Pictorial Fear of Activity Scale-Cervical (PFActS-C).** *Pain* 2008, **139**(1):55-62.
142. Reynolds J, Marsh D, Koller H, Zenenr J, Bannister G: **Cervical range of movement in relation to neck dimension.** *Eur Spine J* 2009, **18**(6):863-868.
143. Castro WH, Sautmann A, Schilgen M, Sautmann M: **Noninvasive three-dimensional analysis of cervical spine motion in normal subjects in relation to age and sex. An experimental examination.** *Spine (Phila Pa 1976)* 2000, **25**(4):443-449.
144. Castro WH, Schilgen M, Meyer S, Weber M, Peuker C, Wörtler K: **Do "whiplash injuries" occur in low-speed rear impacts?** *Eur Spine J* 1997, **6**(6):366-375.
145. Strimpakos N: **The assessment of the cervical spine. Part 1: Range of motion and proprioception.** *J Bodyw Mov Ther* 2011, **15**(1):114-124.
146. Fernández-de-las-Peñas C, Alonso-Blanco C, Cuadrado ML, Pareja JA: **Forward head posture and neck mobility in chronic tension-type headache: a blinded, controlled study.** *Cephalalgia* 2006, **26**(3):314-319.
147. Christensen H, Nilsson N: **The reliability of measuring active and passive cervical range of motion: an observer-blinded and randomized repeated-measures design.** *J Manipulative Physiol Ther* 1998, **21**(5):341-347.
148. Strimpakos N, Sakellari V, Gioftsos G, Papathanasiou M, Brountzos E, Kelekis D, Kapreli E, Oldham J: **Cervical spine ROM measurements: optimizing the testing protocol by using a 3D ultrasound-based motion analysis system**. *Cephalalgia : an international journal of headache* 2005, **25**(12):1133-1145.
149. Cagnie B, Cools A, De Loose V, Cambier D, Danneels L: **Reliability and normative database of the Zebris cervical range-of-motion system in healthy controls with preliminary validation in a group of patients with neck pain.** *J Manipulative Physiol Ther* 2007, **30**(6):450-455.
150. Lantz CA, Chen J, Buch D: **Clinical validity and stability of active and passive cervical range of motion with regard to total and unilateral uniplanar motion**. *Spine* 1999, **24**(11):1082-1089.
151. Nilsson N: **Measuring passive cervical motion: a study of reliability**. *Journal of Manipulative & Physiological Therapeutics* 1995, **18**(5):293-297.
152. Nilsson N, Christensen HW, Hartvigsen J: **The interexaminer reliability of measuring passive cervical range of motion, revisited**. *Journal of Manipulative & Physiological Therapeutics* 1996, **19**(5):302-305.

153. Stiell IG, Wells GA, Vandemheen KL, Clement CM, Lesiuk H, De Maio VJ, Laupacis A, Schull M, McKnight RD, Verbeek R *et al*: **The Canadian C-spine rule for radiography in alert and stable trauma patients.** *JAMA* 2001, **286**(15):1841-1848.
154. Moore A, Jackson A, Jordan J, Hammersley S, Hill J, Mercer C, Smith C, Thompson J, Woby S, Hudson A: **Clinical guidelines for the physiotherapy management of Whiplash Associated Disorder**. London: Chartered Society of Physiotherapy; 2005.
155. Motor, Accidents, Authority: **Guidelines for the Management of Whiplash-Associated Disorders**. In., 2nd edn: Motor Accidents Authority of NSW; 2007.
156. Stewart MJ, Maher CG, Refshauge KM, Herbert RD, Bogduk N, Nicholas M: **Randomized controlled trial of exercise for chronic whiplash-associated disorders.** *Pain* 2007, **128**(1-2):59-68.
157. Dvir Z, Gal-Eshel N, Shamir B, Prushansky T, Pevzner E, Peretz C: **Cervical motion in patients with chronic disorders of the cervical spine: a reproducibility study**. *Spine* 2006, **31**(13):E394-399.
158. Dvir Z, Prushansky T: **Reproducibility and instrument validity of a new ultrasonography-based system for measuring cervical spine kinematics**. *Clinical biomechanics (Bristol, Avon)* 2000, **15**(9):658-664.
159. Kasch H, Qerama E, Kongsted A, Bendix T, Jensen TS, Bach FW: **Clinical assessment of prognostic factors for long-term pain and handicap after whiplash injury: a 1-year prospective study.** *Eur J Neurol* 2008, **15**(11):1222-1230.
160. Antonaci F, Ghirmai S, Bono G, Nappi G: **Current methods for cervical spine movement evaluation: a review**. *Clinical & Experimental Rheumatology* 2000, **18**(2 Suppl 19):S45-52.
161. Bono G, Antonaci F, Ghirmai S, D'Angelo F, Berger M, Nappi G: **Whiplash injuries: clinical picture and diagnostic work-up.** *Clin Exp Rheumatol* 2000, **18**(2 Suppl 19):S23-28.
162. Highland T, Dreisinger T, Vie L, Russell G: **Changes in isometric strength and range of motion of the isolated cervical spine after eight weeks of clinical rehabilitation.** *Spine* 1992, **17**(6 Suppl):S77-82.
163. Kaale B, Krakenes J, Albrektsen G, Wester K: **Active range of motion as an indicator for ligament and membrane lesions in the upper cervical spine after a whiplash trauma.** *J Neurotrauma* 2007, **24**(4):713-721.
164. Klein G, Mannion A, Panjabi M, Dvorak J: **Trapped in the neutral zone: another symptom of whiplash-associated disorder?** *Eur Spine J* 2001, **10**(2):141-148.
165. Osterbauer PJ, Long K, Ribaudo TA, Petermann EA, Fuhr AW, Bigos SJ, Yamaguchi GT: **Three-dimensional head kinematics and cervical range of motion in the diagnosis of patients with neck trauma.** *J Manipulative Physiol Ther* 1996, **19**(4):231-237.
166. Ovadia D, Steinberg EL, Nissan MN, Dekel S: **Whiplash injury--a retrospective study on patients seeking compensation.** *Injury* 2002, **33**(7):569-573.
167. Ryan GA, Taylor GW, Moore VM, Dolinis J: **Neck strain in car occupants: injury status after 6 months and crash-related factors.** *Injury* 1994, **25**(8):533-537.
168. Sterling M, Jull G, Vicenzino B, Kenardy J, Darnell R: **Development of motor system dysfunction following whiplash injury**. *Pain* 2003, **103**(1-2):65-73.

169. Kjellman GV, Skargren EI, Öberg BE: **A critical analysis of randomised clinical trials on neck pain and treatment efficacy. A review of the literature.** . *Scandinavian Journal of Rehabilitation Medicine* 1999, **31**(3):139-152.
170. Nordin M, Carragee E, Hogg-Johnson S, Weiner S, Hurwitz E, Peloso P, Guzman J, van der Velde G, Carroll L, Holm L *et al*: **Assessment of neck pain and its associated disorders: results of the Bone and Joint Decade 2000-2010 Task Force on Neck Pain and Its Associated Disorders.** *Spine* 2008, **33**(4 Suppl):S101-122.
171. Hurwitz EL, Carragee EJ, van der Velde G, Carroll LJ, Nordin M, Guzman J, Peloso PM, Holm LW, Côté P, Hogg-Johnson S *et al*: **Treatment of neck pain: noninvasive interventions: results of the Bone and Joint Decade 2000-2010 Task Force on Neck Pain and Its Associated Disorders.** *Spine (Phila Pa 1976)* 2008, **33**(4 Suppl):S123-152.
172. Aigner N, Fialka C, Radda C, Vecsei V: **Adjuvant laser acupuncture in the treatment of whiplash injuries: a prospective, randomized placebo-controlled trial.** *Wien Klin Wochenschr* 2006, **118**(3-4):95-99.
173. Bonk A, Ferrari R, Giebel G, Edelmann M, Huser R: **Prospective, Randomized, Controlled Study of Activity versus Collar, and the Natural History for Whiplash Injury, in Germany**. *Journal of Musculoskeletal Pain* 2000, **8**(1-2):123-132.
174. Bunketorp L, Lindh M, Carlsson J, Stener-Victorin E: **The effectiveness of a supervised physical training model tailored to the individual needs of patients with whiplash-associated disorders--a randomized controlled trial.** *Clin Rehabil* 2006, **20**(3):201-217.
175. Crawford JR, Khan RJ, Varley GW: **Early management and outcome following soft tissue injuries of the neck-a randomised controlled trial.** *Injury* 2004, **35**(9):891-895.
176. Fialka V, Preisinger E, Böhler A: **Zur physikalischen diagnostik und physikalischer therapie der distorsio columnae vertebralis cervicalis.** 1989, **18**:390-397.
177. Foley-Nolan D, Moore K, Codd M, Barry C, O'Connor P, Coughlan RJ: **Low energy high frequency pulsed electromagnetic therapy for acute whiplash injuries. A double blind randomized controlled study.** *Scand J Rehabil Med* 1992, **24**(1):51-59.
178. Hendriks O, Horgan A: **Ultra-reiz current as an adjunct to standard physiotherapy treatment of the acute whiplash patient**. *Physiotherapy Ireland* 1996, **7**(1):3-7.
179. McKinney LA, Dornan JO, Ryan M: **The role of physiotherapy in the management of acute neck sprains following road-traffic accidents**. *Archives of Emergency Medicine* 1989, **6**(1):27-33.
180. Mealy K, Brennan H, Fenelon GC: **Early mobilization of acute whiplash injuries.** *Br Med J (Clin Res Ed)* 1986, **292**(6521):656-657.
181. Provinciali L, Baroni M, Illuminati L, Ceravolo MG: **Multimodal treatment to prevent the late whiplash syndrome.** *Scand J Rehabil Med* 1996, **28**(2):105-111.
182. Rosenfeld M, Gunnarsson R, Borenstein P: **Early intervention in whiplash-associated disorders: a comparison of two treatment protocols.** *Spine (Phila Pa 1976)* 2000, **25**(14):1782-1787.

183. Soderlund A, Olerud C, Lindberg P: **Acute whiplash-associated disorders (WAD): the effects of early mobilization and prognostic factors in long-term symptomatology**. *Clinical rehabilitation* 2000, **14**(5):457-467.
184. Söderlund A, Lindberg P: **An integrated physiotherapy/cognitive-behavioural approach to the analysis and treatment of chronic whiplash associated disorders, WAD.** *Disabil Rehabil* 2001, **23**(10):436-447.
185. Thuile C, Walzl M: **Evaluation of electromagnetic fields in the treatment of pain in patients with lumbar radiculopathy or the whiplash syndrome.** *NeuroRehabilitation* 2002, **17**(1):63-67.
186. Youdas JW, Carey JR, Garrett TR: **Reliability of measurements of cervical spine range of motion--comparison of three methods**. *Physical Therapy* 1991, **71**(2):98-104.
187. Jordan K, Dziedzic K, Jones PW, Ong BN, Dawes PT: **The reliability of the three-dimensional FASTRAK measurement system in measuring cervical spine and shoulder range of motion in healthy subjects**. *Rheumatology (Oxford, England)* 2000, **39**(4):382-388.
188. WILLIAMS M, MCCARTHY C, CHORTI A, COOKE M, GATES S: **A SYSTEMATIC REVIEW OF RELIABILITY AND VALIDITY STUDIES OF METHODS FOR MEASURING ACTIVE AND PASSIVE CERVICAL RANGE OF MOTION**. *Journal of Manipulative and Physiological Therapeutics* 2010, **33**(2):138-155.
189. Streiner D, Norman G: **Health measurement scales: A practical guide to their development and use**, 3rd edn. New York: Oxford University Press; 2003.
190. McDowell I: **Measuring Health: A guide to rating scales and questionnaires**, 3rd edn. New York: Oxford University Press; 2006.
191. Sim J, Wright C: **The kappa statistic in reliability studies: use, interpretation, and sample size requirements.** *Phys Ther* 2005, **85**(3):257-268.
192. Bravo G, Potvin L: **Estimating the reliability of continuous measures with Cronbach's alpha or the intraclass correlation coefficient: toward the integration of two traditions.** *J Clin Epidemiol* 1991, **44**(4-5):381-390.
193. Baumgarter TA: **Norm-referenced measurement: reliability**. In: *Measurement concepts in Physical Education and Exercise Science.* edn. Edited by Safrit MJ, Wood TM. Illinois: Champaign; 1989: 45-72.
194. Bland J, Altman D: **Statistical methods for assessing agreement between two methods of clinical measurement.** *Lancet* 1986, **1**(8476):307-310.
195. Bruton A, Conway JH, Holgate ST: **Reliability: What is it, and how is it measured?** *Physiotherapy* 2000, **86**(2):94-99.
196. Shrout PE: **Measurement reliability and agreement in psychiatry**. *Statistical methods in medical research* 1998, **7**(3):301-317.
197. Jordan K: **Assessment of published reliability studies for cervical spine range-of-motion measurement tools**. *Journal of Manipulative & Physiological Therapeutics* 2000, **23**(3):180-195.
198. de Koning C, van den Heuvel S, Staal J, Smits-Engelsman B, Hendriks E: **Clinimetric evaluation of active range of motion measures in patients with non-specific neck pain: a systematic review.** *Eur Spine J* 2008, **17**(7):905-921.
199. van der Wurff P, Hagmeijer RH, Meyne W: **Clinical tests of the sacroiliac joint. A systematic methodological review. Part 1: Reliability.** *Man Ther* 2000, **5**(1):30-36.

200. Stochkendahl MJ, Christensen HW, Hartvigsen J, Vach W, Haas M, Hestbaek L, Adams A, Bronfort G: **Manual examination of the spine: a systematic critical literature review of reproducibility**. *J Manipulative Physiol Ther* 2006, **29**(6):475-485, 485 e471-410.
201. van Trijffel E, Anderegg Q, Bossuyt P, Lucas C: **Inter-examiner reliability of passive assessment of intervertebral motion in the cervical and lumbar spine: a systematic review.** *Man Ther* 2005, **10**(4):256-269.
202. van der Wurff P, Meyne W, Hagmeijer R: **Clinical tests of the sacroiliac joint.** *Man Ther* 2000, **5**(2):89-96.
203. Hoppenbrouwers M, Eckhardt M, Verkerk K, Verhagen A: **Reproducibility of the measurement of active and passive cervical range of motion.** *J Manipulative Physiol Ther* 2006, **29**(5):363-367.
204. Alund M, Larsson S: **Three-dimensional analysis of neck motion. A clinical method.** *Spine* 1990, **15**(2):87-91.
205. Tousignant M, Duclos E, Lafleche S, Mayer A, Tousignant-Laflamme Y, Brosseau L, O'Sullivan JP: **Validity study for the cervical range of motion device used for lateral flexion in patients with neck pain**. *Spine* 2002, **27**(8):812-817.
206. Devillé W, Buntinx F, Bouter L, Montori V, de Vet H, van der Windt D, Bezemer P: **Conducting systematic reviews of diagnostic studies: didactic guidelines.** *BMC Med Res Methodol* 2002, **2**:9.
207. Deeks J: **Systematic reviews of evaluations of diagnostic and screening tests**. In: *Systematic Reviews in Health Care.* edn. Edited by Egger M, Davey Smith G, Altman D. London: BMJ Publishing Group; 2001: 263.
208. Swinkels R, Bouter L, Oostendorp R, van den Ende C: **Impairment measures in rheumatic disorders for rehabilitation medicine and allied health care: a systematic review.** *Rheumatol Int* 2005, **25**(7):501-512.
209. **Kappa Calculator: www.dmi.columbia.edu/homepages/chuangj/kappa/calculator.htm** [www.dmi.columbia.edu/homepages/chuangj/kappa/calculator.htm]
210. Landis J, Koch G: **The measurement of observer agreement for categorical data.** *Biometrics* 1977, **33**(1):159-174.
211. Hoving JL, Pool JJ, van Mameren H, Deville WJ, Assendelft WJ, de Vet HC, de Winter AF, Koes BW, Bouter LM: **Reproducibility of cervical range of motion in patients with neck pain**. *BMC musculoskeletal disorders* 2005, **6**:59.
212. Love S, Gringmuth RH, Kazemi M, Cornacchia P, Schmolke M: **Interexaminer and intraexaminer reliability of cervical passive range of motion using the CROM and Cybex 320 EDI**. *J Can Chiropr Assoc* 1998, **42**(4):222-228.
213. Tousignant M, Boucher N, Bourbonnais J, Gravelle T, Quesnel M, Brosseau L: **Intratester and intertester reliability of the Cybex electronic digital inclinometer (EDI-320) for measurement of active neck flexion and extension in healthy subjects**. *Manual therapy* 2001, **6**(4):235-241.
214. Zwart J: **Neck mobility in different headache disorders.** *Headache* 1997, **37**(1):6-11.
215. Amiri M, Jull G, Bullock-Saxton J: **Measuring range of active cervical rotation in a position of full head flexion using the 3D Fastrak measurement system: an intra-tester reliability study**. *Manual therapy* 2003, **8**(3):176-179.
216. Jordan K, Haywood K, Dziedzic K, Garratt A, Jones P, Ong B, Dawes P: **Assessment of the 3-dimensional Fastrak measurement system in measuring range of motion in ankylosing spondylitis.** *J Rheumatol* 2004, **31**(11):2207-2215.

217. Sterling M, Jull G, Carlsson Y, Crommert L: **Are cervical physical outcome measures influenced by the presence of symptomatology?** *Physiother Res Int* 2002, **7**(3):113-121.
218. Assink N, Bergman GJ, Knoester B, Winters JC, Dijkstra PU, Postema K: **Interobserver reliability of neck-mobility measurement by means of the flock-of-birds electromagnetic tracking system**. *Journal of manipulative and physiological therapeutics* 2005, **28**(6):408-413.
219. Morphett A, Crawford C, Lee D: **The use of electromagnetic tracking technology for measurement of passive cervical range of motion: a pilot study.** *J Manipulative Physiol Ther*, **26**(3):152-159.
220. Cleland J, Childs J, Fritz J, Whitman J: **Interrater reliability of the history and physical examination in patients with mechanical neck pain.** *Arch Phys Med Rehabil* 2006, **87**(10):1388-1395.
221. Maksymowych W, Mallon C, Richardson R, Conner-Spady B, Jauregui E, Chung C, Zappala L, Pile K, Russell A: **Development and validation of a simple tape-based measurement tool for recording cervical rotation in patients with ankylosing spondylitis: comparison with a goniometer-based approach.** *J Rheumatol* 2006, **33**(11):2242-2249.
222. Pellecchia G, Bohannon R: **Active lateral neck flexion range of motion measurements obtained with a modified goniometer: reliability and estimates of normal.** *J Manipulative Physiol Ther* 1998, **21**(7):443-447.
223. Tucci S, Hicks J, Gross E, Campbell W, Danoff J: **Cervical motion assessment: a new, simple and accurate method.** *Arch Phys Med Rehabil* 1986, **67**(4):225-230.
224. Agarwal S, Allison G, Singer K: **Reliability of the spin-T cervical goniometer in measuring cervical range of motion in an asymptomatic Indian population.** *J Manipulative Physiol Ther* 2005, **28**(7):487-492.
225. Haynes M, Edmondston S: **Accuracy and reliability of a new, protractor-based neck goniometer.** *J Manipulative Physiol Ther* 2002, **25**(9):579-586.
226. Peolsson A, Hednlund R, Ertzgaard S, Oberg B: **Intra- and inter-tester reliability and range of motion of the neck**. *Physiotherapy Canada* 2000, **Summer**:233-242.
227. Olson S, O'Connor D, Birmingham G, Broman P, Herrera L: **Tender point sensitivity, range of motion, and perceived disability in subjects with neck pain.** *J Orthop Sports Phys Ther* 2000, **30**(1):13-20.
228. Rheault W, Albright B, Byers C, Franta M, Johnson A, Skowronek M, Dougherty J: **Intertester reliability of the cervical range of motion device**. *Journal of Orthopaedic and Sports Physical Therapy* 1992, **15**(3):147-150.
229. Youdas JW, Garrett TR, Suman VJ, Bogard CL, Hallman HO, Carey JR: **Normal range of motion of the cervical spine: an initial goniometric study**. *Physical Therapy* 1992, **72**(11):770-780.
230. Lee H, Nicholson LL, Adams RD: **Cervical range of motion associations with subclinical neck pain**. *Spine* 2004, **29**(1):33-40.
231. Hole DE, Cook JM, Bolton JE: **Reliability and concurrent validity of two instruments for measuring cervical range of motion: effects of age and gender**. *Manual therapy* 1995, **1**(1):36-42.
232. Malmstrom EM, Karlberg M, Melander A, Magnusson M: **Zebris versus Myrin: a comparative study between a three-dimensional ultrasound movement analysis and an inclinometer/compass method: intradevice reliability, concurrent validity, intertester comparison, intratester reliability, and intraindividual variability**. *Spine* 2003, **28**(21):E433-440.

233. Piva S, Erhard R, Childs J, Browder D: **Inter-tester reliability of passive intervertebral and active movements of the cervical spine.** *Man Ther* 2006, **11**(4):321-330.
234. Bush KW, Collins N, Portman L, Tillett N: **Validity and Intertester Reliability of Cervical Range of Motion Using Inclinometer Measurements**. *The Journal of Manual & Manipulative Therapy* 2000, **8**(2):52-61.
235. Bulgheroni M, Antonaci F, Ghirmai S, Sandrini G, Nappi G, Pedotti A: **A 3D kinematic method for evaluating voluntary movements of the cervical spine in humans.** *Funct Neurol* 1998, **13**(3):239-245.
236. Petersen C, Schuit D, Johnson R, Knecht H, Levine P: **Agreement of measures obtained radiographically and by the OSI CA-6000 Spine Motion Analyzer for cervical spinal motion.** *Man Ther* 2007.
237. Mannion A, Klein G, Dvorak J, Lanz C: **Range of global motion of the cervical spine: intraindividual reliability and the influence of measurement device.** *Eur Spine J* 2000, **9**(5):379-385.
238. Chiu T, Sing K: **Evaluation of cervical range of motion and isometric neck muscle strength: reliability and validity.** *Clin Rehabil* 2002, **16**(8):851-858.
239. Haywood K, Garratt A, Jordan K, Dziedzic K, Dawes P: **Spinal mobility in ankylosing spondylitis: reliability, validity and responsiveness.** *Rheumatology (Oxford)* 2004, **43**(6):750-757.
240. Wolfenberger V, Bui Q, Batenchuk G: **A comparison of methods of evaluating cervical range of motion.** *J Manipulative Physiol Ther* 2002, **25**(3):154-160.
241. Bertilson B, Grunnesjö M, Strender L: **Reliability of clinical tests in the assessment of patients with neck/shoulder problems-impact of history.** *Spine* 2003, **28**(19):2222-2231.
242. Viikari-Juntura E: **Interexaminer reliability of observations in physical examinations of the neck.** *Phys Ther* 1987, **67**(10):1526-1532.
243. Fjellner A, Bexander C, Faleij R, Strender L: **Interexaminer reliability in physical examination of the cervical spine.** *J Manipulative Physiol Ther* 1999, **22**(8):511-516.
244. Pool J, Hoving J, de Vet H, van Mameren H, Bouter L: **The interexaminer reproducibility of physical examination of the cervical spine.** *J Manipulative Physiol Ther* 2004, **27**(2):84-90.
245. Van Suijlekom H, De Vet H, Van Den Berg S, Weber W: **Interobserver reliability in physical examination of the cervical spine in patients with headache.** *Headache* 2000, **40**(7):581-586.
246. Pile K, Laurent M, Salmond C, Best M, Pyle E, Moloney R: **Clinical assessment of ankylosing spondylitis: a study of observer variation in spinal measurements.** *Br J Rheumatol* 1991, **30**(1):29-34.
247. Mayer T, Brady S, Bovasso E, Pope P, Gatchel R: **Noninvasive measurement of cervical tri-planar motion in normal subjects.** *Spine* 1993, **18**(15):2191-2195.
248. Syed F, Oza A, Vanderby R, Heiderscheit B, Anderson P: **A method to measure cervical spine motion over extended periods of time.** *Spine* 2007, **32**(19):2092-2098.
249. Hermann KM, Reese CS: **Relationships among selected measures of impairment, functional limitation, and disability in patients with cervical spine disorders**. *Physical Therapy* 2001, **81**(3):903-914.
250. Agarwal S, Allison G, Singer K: **Validation of the spin-T goniometer, a cervical range of motion device.** *J Manipulative Physiol Ther* 2005, **28**(8):604-609.

251. Littlewood C, May S: **Measurement of range of movement in the lumbar spine - what methods are valid? A systematic review**. *Physiotherapy* 2007, **93**:201-211.
252. May S, Littlewood C, Bishop A: **Reliability of procedures used in the physical examination of non-specific low back pain: a systematic review.** *Aust J Physiother* 2006, **52**(2):91-102.
253. Dvir Z, Werner V, Peretz C: **The effect of measurement protocol on active cervical motion in healthy subjects.** *Physiother Res Int* 2002, **7**(3):136-145.
254. Dvir Z, Penso-Zabludowski E: **The effects of protocol and test situation on maximal vs. submaximal cervical motion: medicolegal implications.** *Int J Legal Med* 2003, **117**(6):350-355.
255. Tousignant M, de Bellefeuille L, O'Donoughue S, Grahovac S: **Criterion validity of the cervical range of motion (CROM) goniometer for cervical flexion and extension**. *Spine* 2000, **25**(3):324-330.
256. Tousignant M, Smeesters C, Breton AM, Breton E, Corriveau H: **Criterion validity study of the cervical range of motion (CROM) device for rotational range of motion on healthy adults**. *The Journal of orthopaedic and sports physical therapy* 2006, **36**(4):242-248.
257. Dunn G: **Design and analysis of reliability studies.** *Stat Methods Med Res* 1992, **1**(2):123-157.
258. Potter M, Gordon S, Hamer P: **The physiotherapy experience in private practice: the patients' perspective.** *Aust J Physiother* 2003, **49**(3):195-202.
259. Dvir Z, Prushansky T, Peretz C: **Maximal versus feigned active cervical motion in healthy patients: the coefficient of variation as an indicator for sincerity of effort.** *Spine (Phila Pa 1976)* 2001, **26**(15):1680-1688.
260. Johansson CA, Kent BE, Shepard KF: **Relationship between verbal command volume and magnitude of muscle contraction.** *Phys Ther* 1983, **63**(8):1260-1265.
261. Lea R, Gerhardt J: **Range-of-motion measurements.** *J Bone Joint Surg Am* 1995, **77**(5):784-798.
262. Boone DC, Azen SP, Lin CM, Spence C, Baron C, Lee L: **Reliability of goniometric measurements.** *Phys Ther* 1978, **58**(11):1355-1360.
263. Solinger AB, Chen J, Lantz CA: **Standardized initial head position in cervical range-of-motion assessment: reliability and error analysis.** *J Manipulative Physiol Ther* 2000, **23**(1):20-26.
264. MacFarlane G, Croft P, Schollum J, Silman A: **Widespread pain: is an improved classification possible?** *J Rheumatol* 1996, **23**(9):1628-1632.
265. Altman DG: **Preparing to analyse data**. In: *Practical Statistics for Medical Research.* edn. London: Chapman and Hall; 1991: 122-131.
266. Rankin G, Stokes M: **Reliability of assessment tools in rehabilitation: an illustration of appropriate statistical analyses.** *Clin Rehabil* 1998, **12**(3):187-199.
267. Walter SD, Eliasziw M, Donner A: **Sample size and optimal designs for reliability studies**. *Statistics in medicine* 1998, **17**(1):101-110.
268. von Elm E, Altman D, Egger M, Pocock S, Gøtzsche P, Vandenbroucke J, Initiative S: **The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies.** *J Clin Epidemiol* 2008, **61**(4):344-349.
269. **Scientific approach to the assessment and management of activity-related spinal disorders. A monograph for clinicians. Report of the Quebec Task Force on Spinal Disorders.** *Spine (Phila Pa 1976)* 1987, **12**(7 Suppl):S1-59.

270. Jensen M, Turner J, Romano J: **What is the maximum number of levels needed in pain intensity measurement?** *Pain* 1994, **58**(3):387-392.
271. Underwood M, Barnett A, Vickers M: **Evaluation of two time-specific back pain outcome measures.** *Spine (Phila Pa 1976)* 1999, **24**(11):1104-1112.
272. BenDebba M, Heller J, Ducker T, Eisinger J: **Cervical spine outcomes questionnaire: its development and psychometric properties.** *Spine (Phila Pa 1976)* 2002, **27**(19):2116-2123; discussion 2124.
273. Skolasky R, Riley Lr, Albert T: **Psychometric properties of the Cervical Spine Outcomes Questionnaire and its relationship to standard assessment tools used in spine research.** *Spine J*, **7**(2):174-179.
274. Mullaney MJ, McHugh MP, Johnson CP, Tyler TF: **Reliability of shoulder range of motion comparing a goniometer to a digital level.** *Physiother Theory Pract* 2010, **26**(5):327-333.
275. Vlaeyen JW, Kole-Snijders AM, Boeren RG, van Eek H: **Fear of movement/(re)injury in chronic low back pain and its relation to behavioral performance.** *Pain* 1995, **62**(3):363-372.
276. Waddell G, Newton M, Henderson I, Somerville D, Main C: **A Fear-Avoidance Beliefs Questionnaire (FABQ) and the role of fear-avoidance beliefs in chronic low back pain and disability.** *Pain* 1993, **52**(2):157-168.
277. George S, Fritz J, Erhard R: **A comparison of fear-avoidance beliefs in patients with lumbar spine pain and cervical spine pain.** *Spine (Phila Pa 1976)* 2001, **26**(19):2139-2145.
278. Cleland J, Fritz J, Childs J: **Psychometric properties of the Fear-Avoidance Beliefs Questionnaire and Tampa Scale of Kinesiophobia in patients with neck pain.** *Am J Phys Med Rehabil* 2008, **87**(2):109-117.
279. Soderlund A, Lindberg P: **Whiplash-associated disorders--predicting disability from a process-oriented perspective of coping**. *Clinical rehabilitation* 2003, **17**(1):101-107.
280. Williamson E: **personal communication regarding development of a brief self-efficacy measure**. In.; 2010.
281. Turner JA, Aaron LA: **Pain-related catastrophizing: what is it?** *Clin J Pain* 2001, **17**(1):65-71.
282. Sullivan M, Stanish W, Sullivan M, Tripp D: **Differential predictors of pain and disability in patients with whiplash injuries.** *Pain Res Manag* 2002, **7**(2):68-74.
283. Miró J, Nieto R, Huguet A: **Predictive factors of chronic pain and disability in whiplash: A Delphi poll.** *Eur J Pain* 2008, **12**(1):30-47.
284. Nieto R, Miró J, Huguet A: **The fear-avoidance model in whiplash injuries.** *Eur J Pain* 2009, **13**(5):518-523.
285. Horowitz M, Wilner N, Alvarez W: **Impact of Event Scale: a measure of subjective stress.** *Psychosom Med* 1979, **41**(3):209-218.
286. Goldberg D, Blackwell B: **Psychiatric illness in general practice. A detailed study using a new method of case identification.** *Br Med J* 1970, **1**(5707):439-443.
287. Pevalin D: **Multiple applications of the GHQ-12 in a general population sample: an investigation of long-term retest effects.** *Soc Psychiatry Psychiatr Epidemiol* 2000, **35**(11):508-512.
288. Walton D: **A review of the definitions of 'recovery' used in prognostic studies on whiplash using an ICF framework.** *Disabil Rehabil* 2009, **31**(12):943-957.
289. Vernon H: **The Neck Disability Index: state-of-the-art, 1991-2008.** *J Manipulative Physiol Ther* 2008, **31**(7):491-502.

290. Pietrobon R, Coeytaux R, Carey T, Richardson W, DeVellis R: **Standard scales for measurement of functional outcome for cervical pain or dysfunction: a systematic review.** *Spine* 2002, **27**(5):515-522.
291. Cleland J, Fritz J, Whitman J, Palmer J: **The reliability and construct validity of the Neck Disability Index and patient specific functional scale in patients with cervical radiculopathy.** *Spine* 2006, **31**(5):598-602.
292. MacDermid J, Walton D, Avery S, Blanchard A, Etruw E, McAlpine C, Goldsmith C: **Measurement properties of the neck disability index: a systematic review.** *J Orthop Sports Phys Ther* 2009, **39**(5):400-417.
293. Crouch R, Whitewick R, Clancy M, Wright P, Thomas P: **Whiplash associated disorder: incidence and natural history over the first month for patients presenting to a UK emergency department**. *Emergency Medicine Journal* 2006, **23**(2):114-118.
294. Riddle D, Stratford P: **Use of generic versus region-specific functional status measures on patients with cervical spine disorders.** *Phys Ther* 1998, **78**(9):951-963.
295. Walton D, Pretty J, MacDermid J, Teasell R: **Risk factors for persistent problems following whiplash injury: results of a systematic review and meta-analysis.** *J Orthop Sports Phys Ther* 2009, **39**(5):334-350.
296. Vernon H: **The Neck Disability Index: patient assessment and outcome monitoring in whiplash.** *Journal of Musculoskeletal Pain*, **4**:95-104.
297. Field A: **Sample size in regression**. In: *Discovering statistics using SPSS.* Third edn. London: Sage; 2009: 222-223.
298. Kamper S, Rebbeck T, Maher C, McAuley J, Sterling M: **Course and prognostic factors of whiplash: a systematic review and meta-analysis.** *Pain* 2008, **138**(3):617-629.
299. Carroll L, Holm L, Hogg-Johnson S, Côté P, Cassidy J, Haldeman S, Nordin M, Hurwitz E, Carragee E, van der Velde G *et al*: **Course and prognostic factors for neck pain in whiplash-associated disorders (WAD): results of the Bone and Joint Decade 2000-2010 Task Force on Neck Pain and Its Associated Disorders.** *Spine* 2008, **33**(4 Suppl):S83-92.
300. Myers R: **Classical and modern regression with applications**, Second edn. Boston, MA: Duxbury; 1990.
301. Field A: **Checking assumptions**. In: *Discovering statistics using SPSS.* Third edn. London: Sage; 2009: 247-251.
302. Kraemer HC, Wilson GT, Fairburn CG, Agras WS: **Mediators and moderators of treatment effects in randomized clinical trials.** *Arch Gen Psychiatry* 2002, **59**(10):877-883.
303. Pocock SJ, Assmann SE, Enos LE, Kasten LE: **Subgroup analysis, covariate adjustment and baseline comparisons in clinical trial reporting: current practice and problems.** *Stat Med* 2002, **21**(19):2917-2930.
304. Ackelman B, Lindgren U: **Validity and reliability of a modified version of the neck disability index.** *J Rehabil Med* 2002, **34**(6):284-287.
305. Coudeyre E, Rannou F, Tubach F, Baron G, Coriat F, Brin S, Revel M, Poiraudeau S: **General practitioners' fear-avoidance beliefs influence their management of patients with low back pain.** *Pain* 2006, **124**(3):330-337.
306. Lind B, Sihlbom H, Nordwall A, Malchau H: **Normal range of motion of the cervical spine.** *Arch Phys Med Rehabil* 1989, **70**(9):692-695.

307. Sandlund J, Djupsjöbacka M, Ryhed B, Hamberg J, Björklund M: **Predictive and discriminative value of shoulder proprioception tests for patients with whiplash-associated disorders.** *J Rehabil Med* 2006, **38**(1):44-49.
308. Takasaki H, Hall T, Kaneko S, Iizawa T, Ikemoto Y: **Cervical segmental motion induced by shoulder abduction assessed by magnetic resonance imaging.** *Spine (Phila Pa 1976)* 2009, **34**(3):E122-126.
309. Waddell G: **The Back Pain Revolution**. Edinburgh: Churchill-Livingstone; 1998.
310. Landers M, Creger R, Baker C, Stutelberg K: **The use of fear-avoidance beliefs and nonorganic signs in predicting prolonged disability in patients with neck pain.** *Man Ther* 2008, **13**(3):239-248.
311. Rooker J, Bannister M, Amirfeyz R, Squires B, Gargan M, Bannister G: **Whiplash injury: 30-year follow-up of a single series.** *J Bone Joint Surg Br* 2010, **92**(6):853-855.
312. Allen D: **Validity and reliability of the movement ability measure: a self-report instrument proposed for assessing movement across diagnoses and ability levels.** *Phys Ther* 2007, **87**(7):899-916.
313. Cott C, Finch E, Gasner D, Yoshida K: **The Movement Continuum Theory of Physical Therapy**. *Physiotherapy Canada* 1995, **47**(2):87-95.
314. Allen D: **Responsiveness of the movement ability measure: a self-report instrument proposed for assessing the effectiveness of physical therapy intervention.** *Phys Ther* 2007, **87**(7):917-924.
315. Haldeman S, Carroll L, Cassidy JD, Schubert J, Nygren A, Disorders BaJD-TFoNPaIA: **The Bone and Joint Decade 2000-2010 Task Force on Neck Pain and Its Associated Disorders: executive summary.** *Spine (Phila Pa 1976)* 2008, **33**(4 Suppl):S5-7.
316. Chappuis G, Soltermann B, CEA, AREDOC, CEREDOC: **Number and cost of claims linked to minor cervical trauma in Europe: results from the comparative study by CEA, AREDOC and CEREDOC.** *Eur Spine J* 2008, **17**(10):1350-1357.
317. Cassidy JD, Carroll LJ, Cote P, Lemstra M, Berglund A, Nygren A: **Effect of eliminating compensation for pain and suffering on the outcome of insurance claims for whiplash injury.[see comment]**. *New England Journal of Medicine* 2000, **342**(16):1179-1186.
318. Cameron ID, Rebbeck T, Sindhusake D, Rubin G, Feyer AM, Walsh J, Schofield WN: **Legislative change is associated with improved health status in people with whiplash.** *Spine (Phila Pa 1976)* 2008, **33**(3):250-254.
319. Cassidy JD, Côté P: **Is it time for a population health approach to neck pain?** *J Manipulative Physiol Ther* 2008, **31**(6):442-446.
320. Lakke SE, Soer R, Takken T, Reneman MF: **Risk and prognostic factors for non-specific musculoskeletal pain: a synthesis of evidence from systematic reviews classified into ICF dimensions.** *Pain* 2009, **147**(1-3):153-164.
321. Cobo EP, Mesquida ME, Fanegas EP, Atanasio EM, Pastor MB, Pont CP, Prieto CM, Gómez GR, Cano LG: **What factors have influence on persistence of neck pain after a whiplash?** *Spine (Phila Pa 1976)* 2010, **35**(9):E338-343.
322. Buitenhuis J, de Jong PJ, Jaspers JP, Groothoff JW: **Work disability after whiplash: a prospective cohort study.** *Spine (Phila Pa 1976)* 2009, **34**(3):262-267.
323. Kongsted A, Jørgensen LV, Leboeuf-Yde C, Qerama E, Korsholm L, Bendix T: **Are altered smooth pursuit eye movements related to chronic pain and disability following whiplash injuries? A prospective trial with one-year follow-up.** *Clin Rehabil* 2008, **22**(5):469-479.

324. Ichihara D, Okada E, Chiba K, Toyama Y, Fujiwara H, Momoshima S, Nishiwaki Y, Hashimoto T, Ogawa J, Watanabe M *et al*: **Longitudinal magnetic resonance imaging study on whiplash injury patients: minimum 10-year follow-up.** *J Orthop Sci* 2009, **14**(5):602-610.
325. Pape E, Brox JI, Hagen KB, Natvig B, Schirmer H: **Prognostic factors for chronic neck pain in persons with minor or moderate injuries in traffic accidents.** *Accid Anal Prev* 2007, **39**(1):135-146.
326. Borenstein P, Rosenfeld M, Gunnarsson R: **Cognitive symptoms, cervical range of motion and pain as prognostic factors after whiplash trauma.** *Acta Neurol Scand* 2010, **122**(4):278-285.
327. Vetti N, Kråkenes J, Eide GE, Rørvik J, Gilhus NE, Espeland A: **Are MRI high-signal changes of alar and transverse ligaments in acute whiplash injury related to outcome?** *BMC Musculoskelet Disord* 2010, **11**:260.
328. Berglund A, Bodin L, Jensen I, Wiklund A, Alfredsson L: **The influence of prognostic factors on neck pain intensity, disability, anxiety and depression over a 2-year period in subjects with acute whiplash injury.** *Pain* 2006, **125**(3):244-256.
329. Rebbeck T, Sindhusake D, Cameron I, Rubin G, Feyer A, Walsh J, Gold M, Schofield W: **A prospective cohort study of health outcomes following whiplash associated disorders in an Australian population.** *Inj Prev* 2006, **12**(2):93-98.
330. Gabel CP, Burkett B, Neller A, Yelland M: **Can long-term impairment in general practitioner whiplash patients be predicted using screening and patient-reported outcomes?** *Int J Rehabil Res* 2008, **31**(1):79-80.
331. Obermann M, Nebel K, Riegel A, Thiemann D, Yoon MS, Keidel M, Stude P, Diener H, Katsarava Z: **Incidence and predictors of chronic headache attributed to whiplash injury.** *Cephalalgia* 2010, **30**(5):528-534.
332. Kivioja J, Jensen I, Lindgren U: **Neither the WAD-classification nor the Quebec Task Force follow-up regimen seems to be important for the outcome after a whiplash injury. A prospective study on 186 consecutive patients.** *Eur Spine J* 2008, **17**(7):930-935.
333. Wynne-Jones G, Jones G, Wiles N, Silman A, Macfarlane G: **Predicting new onset of widespread pain following a motor vehicle collision.** *J Rheumatol* 2006, **33**(5):968-974.
334. van Trijffel E, Oostendorp RA, Lindeboom R, Bossuyt PM, Lucas C: **Perceptions and use of passive intervertebral motion assessment of the spine: a survey among physiotherapists specializing in manual therapy.** *Man Ther* 2009, **14**(3):243-251.
335. Whiting P, Harbord R, Kleijnen J: **No role for quality scores in systematic reviews of diagnostic accuracy studies.** *BMC Med Res Methodol* 2005, **5**:19.
336. Leeflang MM, Deeks JJ, Gatsonis C, Bossuyt PM, Group CDTAW: **Systematic reviews of diagnostic test accuracy.** *Ann Intern Med* 2008, **149**(12):889-897.
337. Audette I, Dumas JP, Côté JN, De Serres SJ: **Validity and between-day reliability of the cervical range of motion (CROM) device.** *J Orthop Sports Phys Ther* 2010, **40**(5):318-323.
338. Florêncio LL, Pereira PA, Silva ER, Pegoretti KS, Gonçalves MC, Bevilaqua-Grossi D: **Agreement and reliability of two non-invasive methods for assessing cervical range of motion among young adults.** *Rev Bras Fisioter* 2010, **14**(2):175-181.

339. Fletcher JP, Bandy WD: **Intrarater reliability of CROM measurement of cervical spine active range of motion in persons with and without neck pain.** *J Orthop Sports Phys Ther* 2008, **38**(10):640-645.
340. Gelalis I, DeFrate L, Stafilas K, Pakos E, Kang J, Gilbertson L: **Three-dimensional analysis of cervical spine motion: reliability of a computer assisted magnetic tracking device compared to inclinometer.** *Eur Spine J* 2009, **18**(2):276-281.
341. Prushansky T, Deryi O, Jabarreen B: **Reproducibility and validity of digital inclinometry for measuring cervical range of motion in normal subjects.** *Physiother Res Int* 2010, **15**(1):42-48.
342. Salo PK, Häkkinen AH, Kautiainen H, Ylinen JJ: **Quantifying the effect of age on passive range of motion of the cervical spine in healthy working-age women.** *J Orthop Sports Phys Ther* 2009, **39**(6):478-483.
343. Whitcroft KL, Massouh L, Amirfeyz R, Bannister G: **Comparison of methods of measuring active cervical range of motion.** *Spine (Phila Pa 1976)* 2010, **35**(19):E976-980.
344. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM, Moher D, Rennie D, de Vet HC, Lijmer JG *et al*: **The STARD statement for reporting studies of diagnostic accuracy: explanation and elaboration.** *Ann Intern Med* 2003, **138**(1):W1-12.
345. Kottner J, Gajewski BJ, Streiner DL: **Guidelines for Reporting Reliability and Agreement Studies (GRRAS).** *Int J Nurs Stud* 2011.
346. Holm LW, Carroll LJ, Cassidy JD, Ahlbom A: **Factors influencing neck pain intensity in whiplash-associated disorders.** *Spine (Phila Pa 1976)* 2006, **31**(4):E98-104.
347. Carter CW, Levine WN, Kleweno CP, Bigliani LU, Ahmad CS: **Assessment of shoulder range of motion: introduction of a novel patient self-assessment tool.** *Arthroscopy* 2008, **24**(6):712-717.
348. McCarney R, Warner J, Iliffe S, van Haselen R, Griffin M, Fisher P: **The Hawthorne Effect: a randomised, controlled trial.** *BMC Med Res Methodol* 2007, **7**:30.

10 APPENDICES

10.1 APPENDIX 1- PUBLISHED MANUSCRIPT OF THE MINT PROTOCOL

Available at http://www.ncbi.nlm.nih.gov/pubmed/17257408

10.2 APPENDIX 2 - MINT 2 WEEK QUESTIONNAIRE

PARTICIPANTS ID:

Managing Injuries of the Neck Trial
M.I.N.T.

Recovery after Neck Injury Questionnaire

Please complete this questionnaire and return it in the reply-paid envelope provided.

If you are still experiencing problems because of your injury we would like you to contact us, as you may be eligible for further treatment as part of the study. Please contact the research team to arrange an appointment on 024 7657 4648.

If you have any difficulties or questions relating to the questionnaire, please contact:

Mark Williams
Warwick Emergency Care and Rehabilitation
FREEPOST MID18033
Room A1.03 Medical School Building
University of Warwick
Gibbet Hill Campus
Coventry
CV4 7AL
Freephone : 0800 634 0439
Fax : 024 7657 4657
E-mail : mint@warwick.ac.uk

THE UNIVERSITY OF WARWICK

NHS

ISRCTN 3302125

Please read all the instructions before completing the questionnaire

Thank you for agreeing to take part in this study. The answers you give in this questionnaire will help us find out if the treatments you get are helpful for your neck problems.

Please answer all the questions. Although it may seem that questions are asked more than once, it is still important that you answer every one.

Please follow the instructions for each section carefully.

For each section, if you are asked to put a cross in the box, please use a cross rather than a tick.

For example in the following question, if your answer to the question is yes, you should place a cross firmly in the box next to yes.

Do you drive a car?	Yes ☒	No ☐

Please use a BLACK or BLUE pen. Please do not use a pencil.

Please check that you have completed all sections.

Please return your completed questionnaire to us in the envelope provided.

We will send another questionnaire in 4 months. Please keep a record of any days off work, and hospital or medical procedures you undergo as a result of your neck injury.

Please write any notes you have for us on the back page.

Section 1 This section is to find out some information about you.

Please complete the following information.

1. Date you are completing this questionnaire:

D	D	M	M	Y	Y

2. Date of your injury:

D	D	M	M	Y	Y

3. Date of Birth:

D	D	M	M	Y	Y

4. Sex

Male	Female
☐ 1	☐ 2

5. To which of these ethnic groups do you consider you belong? (Please tick one box)

White	☐ 1	Mixed	☐ 2
Indian	☐ 3	Pakistani	☐ 4
Bangladeshi	☐ 5	Black or Black British	☐ 6
Chinese or Other Ethnic Group	☐ 7		

6. Employment Status

Full Time Employed	☐ 1	Part Time Employed	☐ 2
Self Employed	☐ 3	Unpaid Work	☐ 4
Unemployed	☐ 5	Full Time Student	☐ 6
Retired/Looking after Home/Inactive	☐ 7		

Section 2 - This section is to find out how you felt **before** you injured your neck.

1. In the month **BEFORE** your injury did you have neck pain?

Yes ☐ 1 No ☐ 2

If yes, how troublesome was this pain?

Not at all troublesome	Slightly troublesome	Moderately troublesome	Very troublesome	Extremely troublesome
☐ 1	☐ 2	☐ 3	☐ 4	☐ 5

Section 3 - These questions ask about what problems you feel you have developed **as a result of your recent injury**.

1. **Since** your injury, have you had neck stiffness?

Yes ☐ 1 No ☐ 2

2. **Since** your injury, have you had headaches?

Yes ☐ 1 No ☐ 2

3. **Since** your injury, have you had numbness, weakness, or tingling in your arms or hands?

Yes ☐ 1 No ☐ 2

4. **Since** your injury, have you had neck pain?

Yes ☐ 1 No ☐ 2

If yes, how troublesome was this pain?

Not at all troublesome	Slightly troublesome	Moderately troublesome	Very troublesome	Extremely troublesome
☐ 1	☐ 2	☐ 3	☐ 4	☐ 5

Section 4 – This section asks for your views about your health. This information will help keep track of how you feel and how well you are able to do your usual activities.
For each of the following questions please place a cross in the one box that best describes your answer.

1. In general, would you say your health is:

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

2. The following questions are about activities you might do during a typical day.
Does your health now limit you in these activities? If so, how much?

	Yes, limited a lot	Yes, limited a little	No, not limited at all
a. **Moderate activities**, such as moving a table, pushing a vacuum cleaner, bowling or playing golf	☐ 1	☐ 2	☐ 3
b. Climbing **several** flights of stairs	☐ 1	☐ 2	☐ 3

3. In the last week, have you had any of the following problems with your work or other regular daily activities as a result of your physical health?

	Yes	No
a. **Accomplished less** than you would like	☐ 1	☐ 2
b. Were limited in the **kind** of work or other activities	☐ 1	☐ 2

4. In the last week, have you had any of the following problems with your work or other regular daily activities as a result of any emotional problems (such as feeling depressed or anxious)?

	Yes	No
a. **Accomplished less** than you would like	☐ 1	☐ 2
b. Didn't do work or other activities as carefully as usual	☐ 1	☐ 2

5. In the last week, how much did pain interfere with your normal work (including both work outside the home and housework)?

Not at all	A little bit	Moderately	Quite a bit	Extremely
☐ 1	☐ 2	☐ 3	☐ 4	☐ 5

6. These questions are about how you feel and how things have been with you in the last week.
For each question, please give the one answer that comes closest to the way you have been feeling.
How much of the time in the last week:

	All of the time	Most of the time	A good bit of the time	A little of the time	None of the time
a. Have you felt calm and peaceful?	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
b. Did you have a lot of energy?	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
c. Have you felt downhearted and low?	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5

7. In the last week, how much of the time has your physical health OR emotional problems interfered with your social activities (like visiting with friends, relatives, etc.)?

All of the time	Most of the time	Some of the time	A little of the time	None of the time
☐ 1	☐ 2	☐ 3	☐ 4	☐ 5

Section 5 - The following questions are to ask about your general health state **at the moment**. By placing a cross in one box in each group below, please indicate which statement best describes your own health state **today**.

Please cross **one** box for each question

1. Mobility

- I have no problems in walking about ☐ 1
- I have some problems in walking about ☐ 2
- I am confined to bed ☐ 3

2. Self Care

- I have no problems with self-care ☐ 1
- I have some problems washing or dressing myself ☐ 2
- I am unable to wash or dress myself ☐ 3

3. Usual Activities (e.g. work, study, housework, family or leisure activities)?

- I have no problems with performing my usual activities ☐ 1
- I have some problems with performing my usual activities ☐ 2
- I am unable to perform my usual activities ☐ 3

4. Pain / Discomfort

- I have no pain or discomfort ☐ 1
- I have moderate pain or discomfort ☐ 2
- I have extreme pain or discomfort ☐ 3

5. Anxiety / Depression

- I am not anxious or depressed ☐ 1
- I am moderately anxious or depressed ☐ 2
- I am extremely anxious or depressed ☐ 3

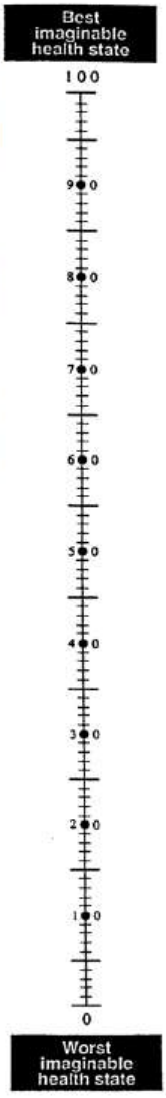
Your own health state today

To help people say how good or bad a health state is, we have drawn a scale (rather like a thermometer) on which the best state you can imagine is marked by 100 and the worst state you can imagine is marked by 0.

We would like you to indicate on this scale **how good or bad is your own health today, in your opinion.**

Please do this by drawing a line from the box below, to whichever point on the scale indicates how good or bad your current health state is **today**.

Your own health state TODAY



Section 6 - This section is to determine the costs of neck injuries to individuals and to the government. Information on these costs will help to support improved services for those who have been injured.
For each of the following questions please place a cross in the one box that best describes your answer.

1. Have you had to take any days off sick from work **since and as a result of** your injury?

Yes	No	Not Applicable
☐ 1	☐ 2	☐ 3

If **Yes**, how many days for ☐☐

2. If working, please indicate average gross annual income (before tax) £ _____

3. Are you receiving any state benefits?

Yes	No
☐ 1	☐ 2

If **Yes**, type of benefits:

Housing and Homeless Assistance	☐ 1	Attendance Allowance	☐ 2
Severe Disablement Allowance	☐ 3	Carer's Allowance	☐ 4
Disability Living Allowance	☐ 5	Income Support	☐ 6
Working Tax Credit	☐ 7	Social Fund Payment	☐ 8
Council Tax Benefit	☐ 9	Child Benefit	☐ 10
Income based Jobseekers' Allowance	☐ 11	Housing Benefits	☐ 12
Pension Credit	☐ 13	Child Tax Credit	☐ 14
Other (please specify) _____	☐ 15		

4. Please indicate your approximate annual household income (i.e. all the incomes coming into your home) £ _____

Section 7 – Finally, we would like to find out what you thought of the treatment you received when you visited A&E.

1. How satisfied were you with the treatment you received in A&E for your whiplash injury?

Very dissatisfied	☐ 1
Somewhat dissatisfied	☐ 2
Neither satisfied nor dissatisfied	☐ 3
Somewhat satisfied	☐ 4
Very satisfied	☐ 5

Please write any notes you have for us in the space below.

Thank you very much for your time.

10.3 APPENDIX 3 - MINT RESEARCH CLINIC QUESTIONNAIRE

PARTICIPANT'S ID:

Managing Injuries of the Neck Trial
MINT

Research Clinic Questionnaire

We are interested to see how well you are recovering since your injury.
Your answers will be kept strictly confidential and your name will not appear anywhere on the questionnaire.

If you have any difficulties or questions relating to the questionnaire then please ask for assistance in completing the questions.

THE UNIVERSITY OF WARWICK

NHS

ISRCTN 3302125

Please read all the instructions before completing the questionnaire.

Thank you for agreeing to take part in this study.

The answers you give in this questionnaire will help us find out if the treatments you get are helpful for your neck problems.

Please answer all the questions.

Although it may seem that questions are asked more than once, it is still important that you answer every one.

Please follow the instructions for each section carefully.

For each section, if you are asked to put a cross in the box, please use a cross rather than a tick.

For example in the following question, if your answer to the question is yes, you should place a cross firmly in the box next to yes.

Do you drive a car?	Yes	No
	☒	☐

If you are asked to circle a number, please use a circle rather than underlining a number.

For example, in the following question if you are asked 'how happy are you today?' where '1' is 'very unhappy' and '5' is 'very happy'. If you feel neither happy or unhappy you may wish to answer '3'. You do this by clearly circling the number 3.

1 2 (3) 4 5

Please use a BLACK or BLUE pen. Please do not use a pencil.

Section 1—

Please place a cross in one box only.

1. Although you will be given one of the treatments by chance, if you could choose which treatment to have what would be your preference?

- One session of advice with a physiotherapist ☐ 1
- I don't mind which treatment I receive ☐ 2
- A course of physiotherapy ☐ 3

2. How long do you think it will take for you to recover from your neck injury?

- In the next 2 weeks ☐ 1
- 2—8 weeks ☐ 2
- 2—6 months ☐ 3
- 6—12 months ☐ 4
- More than a year ☐ 5
- I am not sure I will recover ☐ 6

3. Have you returned to work?

- Yes—proceed to the next section ☐ 1
- Not in paid employment—proceed to the next section ☐ 2
- No ☐ 3

If you answered "No", when do you expect to return to work?

- In the next 2 weeks ☐ 1
- 2—8 weeks ☐ 2
- 2—6 months ☐ 3
- 6—12 months ☐ 4
- More than a year ☐ 5
- I am not sure I will be able to return to work ☐ 6

Section 2 – This section is designed to find out how you got your whiplash injury.

For each question, please place a cross in one box only.

1. Were you injured whilst in a vehicle?

Yes ☐ 1 No ☐ 2

If **No**, please go to section 3

2. What type of vehicle were you in?

- Car ☐ 1
- MPV / 4WD ☐ 2
- Van ☐ 3
- Lorry ☐ 4
- Bus/Coach ☐ 5
- Other (please specify)_____ ☐ 6

3. If in a car, what was the:

- Make _____
- Model _____
- Year of manufacture _____

4. Where were you in the vehicle?

- Driving ☐ 1
- Front seat passenger ☐ 2
- Rear seat passenger ☐ 3
- Other (please specify)_____ ☐ 4

5. Where was your vehicle hit?

- Front ☐ 1
- Driver's side ☐ 2
- Passenger's side ☐ 3
- Rear ☐ 4

6. Were you aware you were about to have a collision?

Yes ☐ 1 No ☐ 2

7. What speed were you travelling at the time of impact?

_____mph

8. Was the vehicle driveable after the accident?

Yes ☐ 1 No ☐ 2

9. How would you rate the severity of your accident?

Very Low ☐ 1

Low ☐ 2

Moderate ☐ 3

High ☐ 4

Very High ☐ 5

Section 3——This section helps us to understand how much your neck pain has affected your ability to manage everyday activities. We realise that you may feel that more than one statement may relate to you, but please place a cross in the box for the one choice which most closely describes your problem right now.

1. Pain Intensity

A. I have no pain at the moment ☐1

B. The pain is mild at the moment ☐2

C. The pain comes and goes and is moderate ☐3

D. The pain is moderate and does not vary much ☐4

E. The pain is severe but comes and goes ☐5

F. The pain is severe and does not vary much ☐6

2. Personal Care (washing, dressing etc.)

A. I can look after myself without causing extra pain ☐1

B. I can look after myself normally but it causes extra pain ☐2

C. It is painful to look after myself and I am slow and careful ☐3

D. I need some help, but manage most of my personal care ☐4

E. I need help every day in most aspects of self-care ☐5

F. I do not get dressed, I wash with difficulty and stay in bed ☐6

3. Lifting

A. I can lift heavy weights without extra pain ☐1

B. I can lift heavy weights, but it causes extra pain ☐2

C. Pain prevents me from lifting heavy weights off the floor but I can if they are conveniently positioned, for example on a table ☐3

D. Pain prevents me from lifting heavy weights, but I can manage light to medium weights if they are conveniently positioned ☐4

E. I can lift very light weights ☐5

F. I cannot lift or carry anything at all ☐6

4. Reading

A. I can read as much as I want to with no pain in my neck ☐ 1

B. I can read as much as I want with slight pain in my neck ☐ 2

C. I can read as much as I want with moderate pain in my neck ☐ 3

D. I cannot read as much as I want because of moderate pain in my neck ☐ 4

E. I cannot read as much as I want because of severe pain in my neck ☐ 5

F. I cannot read at all ☐ 6

5. Headache

A. I have no headaches at all ☐ 1

B. I have slight headaches which come infrequently ☐ 2

C. I have moderate headaches which come infrequently ☐ 3

D. I have moderate headaches which come frequently ☐ 4

E. I have severe headaches which come frequently ☐ 5

F. I have headaches almost all the time ☐ 6

6. Concentration

A. I can concentrate fully when I want to with no difficulty ☐ 1

B. I can concentrate fully when I want to with slight difficulty ☐ 2

C. I have a fair degree of difficulty in concentrating when I want to ☐ 3

D. I have a lot of difficulty in concentrating when I want to ☐ 4

E. I have a great deal of difficulty in concentrating when I want to ☐ 5

F. I cannot concentrate at all ☐ 6

7. Work

A. I can do as much work as I want to ☐ 1

B. I can only do my usual work, but no more ☐ 2

C. I can do most of my usual work, but no more ☐ 3

D. I cannot do my usual work ☐ 4

E. I can hardly do any work at all ☐ 5

F. I cannot do any work at all ☐ 6

8. Driving

A. I can drive my car without neck pain ☐ 1

B. I can drive my car as long as I want with slight pain in my neck ☐ 2

C. I can drive my car as long as I want with moderate pain in my neck ☐ 3

D. I cannot drive my car as long as I want because of moderate pain in my neck ☐ 4

E. I can hardly drive my car at all because of severe pain in my neck ☐ 5

F. I cannot drive my car at all ☐ 6

9. Sleeping

A. I have no trouble sleeping ☐ 1

B. My sleep is slightly disturbed (less than 1 hour sleepless) ☐ 2

C. My sleep is mildly disturbed (1—2 hours sleepless) ☐ 3

D. My sleep is moderately disturbed (2—3 hours sleepless) ☐ 4

E. My sleep is greatly disturbed (3—5 hours sleepless) ☐ 5

F. My sleep is completely disturbed (5—7 hours sleepless) ☐ 6

10. Recreation

A. I am able to engage in all recreational activities with no pain in my neck at all ☐ 1

B. I am able to engage in all recreational activities with some pain in my neck ☐ 2

C. I am able to engage in most, but not all recreational activities because of pain in my neck ☐ 3

D. I am able to engage in a few of my usual recreational activities because of pain in my neck ☐ 4

E. I can hardly do any recreational activities because of pain in my neck ☐ 5

F. I cannot do any recreational activities at all ☐ 6

Section 4——This section is about the severity of your neck problem:

For the next three questions please circle the number which represents the severity of your neck pain.

For example:

0 1 2 3 4 (5) 6 7 8 9 10

1. **In the last week**, how bad has your worst neck pain been on a scale of 0-10 where 0 is 'no pain' and 10 is 'as bad as a pain could be'?

   0 1 2 3 4 5 6 7 8 9 10

2. **In the last week**, on average how bad has your neck pain been on a scale of 0-10 where 0 is 'no pain' and 10 is 'as bad as a pain could be'?

   0 1 2 3 4 5 6 7 8 9 10

3. How would you rate your neck pain **today** on a scale of 0-10 where 0 is 'no pain' and 10 is 'as bad as a pain could be'?

   0 1 2 3 4 5 6 7 8 9 10

Section 5——This section is about your neck movement:

For the next two questions please circle the number which represents your ability to move your neck.

For example:

1 2 (3) 4 5

1. How much are you able to turn your neck **today** on a scale of 1 to 5 where 1 is 'unable to turn' and 5 is 'able to turn normally'?

   1 2 3 4 5

2. How much are you able to look up or down **today** on a scale of 1 to 5 where 1 is 'unable to look up or down' and 5 is 'able to look up or down normally'?

   1 2 3 4 5

Section 6—This section is to determine what you believe about your neck pain. Here are some of the things which other patients have told us about their pain.

For each statement please circle the number from 0 to 6 to say how much physical activities such as bending, lifting, walking or driving affect or would affect your neck pain.

Please circle one number for each line

	Completely disagree			Unsure			Completely Agree
1. Physical activity makes my pain worse	0	1	2	3	4	5	6
2. Physical activity might harm my neck	0	1	2	3	4	5	6
3. I should not do physical activities which (might) make my pain worse	0	1	2	3	4	5	6
4. I cannot do physical activities which (might) make my pain worse	0	1	2	3	4	5	6

For each statement please circle the number from 0 to 6 to say how much you agree with these statements.

Please circle one number for each line

	Completely disagree			Unsure			Completely Agree
5. When you have injured your neck it is best to exercise your neck to help you to recover	0	1	2	3	4	5	6
6. I feel I am able to cope with my neck problem even when it is painful	0	1	2	3	4	5	6

Section 7—We are interested in the types of thoughts and feelings that you have when you are in pain. Listed below are thirteen statements describing different thoughts and feelings that may be associated with pain. Using the following scale, please indicate the degree to which you have these thoughts and feelings when you are experiencing pain.

When I'm in pain ...	Not at all	To a slight degree	To a moderate degree	To a great degree	All the time
1. I worry all the time about whether the pain will end	0	1	2	3	4
2. I feel I can't go on	0	1	2	3	4
3. It's terrible and I think it's never going to get any better	0	1	2	3	4
4. It's awful and I feel that it overwhelms me	0	1	2	3	4
5. I feel I can't stand it anymore	0	1	2	3	4
6. I become afraid that the pain will get worse	0	1	2	3	4
7. I keep thinking of other painful events	0	1	2	3	4
8.I anxiously want the pain to go away	0	1	2	3	4
9. I can't seem to keep it out of my mind	0	1	2	3	4
10. I keep thinking about how much it hurts	0	1	2	3	4
11. I keep thinking about how badly I want the pain to stop	0	1	2	3	4
12. There's nothing I can do to reduce the intensity of the pain	0	1	2	3	4
13. I wonder whether something serious may happen	0	1	2	3	4

Section 8—Individuals who experience pain have developed a number of ways to cope, or deal with their pain. These include saying things to themselves when they experience pain, or engaging in different activities. Below is a list of things that people have reported doing when they feel pain. For each activity, please indicate, using the scale below, how much you engage in that activity when you feel pain. A 3 indicates you sometimes do that when you are experiencing pain, and a 6 indicates you always do it when you are experiencing pain. Remember, you can use any point along the scale.

When I feel pain ...	Never do			Sometimes do that			Always do that
1. I think of things I enjoy doing	0	1	2	3	4	5	6
2. I just think of it as some other sensation, such as numbness	0	1	2	3	4	5	6
3. It is terrible and I feel it is never going to get any better	0	1	2	3	4	5	6
4. I don't pay any attention to it	0	1	2	3	4	5	6
5. I pray for the pain to stop	0	1	2	3	4	5	6
6. I tell myself I can't let the pain stand in the way of what I want to do	0	1	2	3	4	5	6
7. I do something active, like household chores or projects	0	1	2	3	4	5	6

Section 9—Below is a list of comments made by people after stressful life events. Please check each item indicating how frequently these comments were true for you in regard to your recent neck injury during the last 7 days.

If they did not occur during that time, please mark the "not at all" column.

	Not at all	Rarely	Sometimes	Often
1. I thought about it when I didn't mean to	0	1	2	3
2. I avoided letting myself get upset when I thought about it or was reminded of it	0	1	2	3
3. I tried to remove it from memory	0	1	2	3
4. I had trouble falling asleep or staying asleep, because of the pictures or thoughts about it that came into my head	0	1	2	3
5. I had waves of strong feelings about it	0	1	2	3
6. I had dreams about it	0	1	2	3
7. I stayed away from reminders of it	0	1	2	3
8. I felt as if it hadn't happened or it wasn't real	0	1	2	3
9. I tried not to talk about it	0	1	2	3
10. Pictures about it popped into my head	0	1	2	3
11. Other things kept making me think about it	0	1	2	3
12. I was aware that I still had a lot of feelings about it, but I didn't try to deal with them	0	1	2	3
13. I tried not to think about it	0	1	2	3
14. Any reminder brought back feelings about it	0	1	2	3
15. My feelings about it were kind of numb	0	1	2	3

Section 10—We are interested in how you feel about the following statements. Read each statement carefully. Indicate how you feel about each statement.

	Very strongly disagree	Strongly disagree	Mildly disagree	Neutral	Mildly agree	Strongly agree	Very strongly agree
1. There is a special person who is around when I am in need	1	2	3	4	5	6	7
2. There is a special person with whom I can share my joys and sorrows	1	2	3	4	5	6	7
3. My family really tries to help me	1	2	3	4	5	6	7
4. I get the emotional help and support I need from my family	1	2	3	4	5	6	7
5. I have a special person who is a real source of comfort to me	1	2	3	4	5	6	7
6. My friends really try to help me	1	2	3	4	5	6	7
7. I can count on my friends when things go wrong	1	2	3	4	5	6	7
8. I can talk about my problems with my family	1	2	3	4	5	6	7
9. I have friends with whom I can share my joys and sorrows	1	2	3	4	5	6	7
10. There is a special person in my life who cares about my feelings	1	2	3	4	5	6	7
11. My family is willing to help me make decisions	1	2	3	4	5	6	7
12. I can talk about my problems with my friends	1	2	3	4	5	6	7

Section 11—We should like to know if you have had any medical complaints, and how your health has been in general, *over the past few weeks*. Please answer ALL the questions simply by placing a cross in the box which you think most nearly applies to you. Remember that we want to know about present and recent complaints, not those you had in the past. It is important that you try to answer ALL the questions.

1. Have you recently been able to concentrate on whatever you're doing?

Better than usual	☐ 1
Same as usual	☐ 2
Less than usual	☐ 3
Much less than usual	☐ 4

2. Have you recently lost much sleep over worry?

Not at all	☐ 1
No more than usual	☐ 2
Rather more than usual	☐ 3
Much more than usual	☐ 4

3. Have you recently felt that you are playing a useful part in things?

More so than usual	☐ 1
Same as usual	☐ 2
Less useful than usual	☐ 3
Much less useful	☐ 4

4. Have you recently felt capable of making decisions about things?

More so than usual	☐ 1
Same as usual	☐ 2
Less so than usual	☐ 3
Much less capable	☐ 4

5. Have you recently felt constantly under strain?

Not at all	☐ 1
No more than usual	☐ 2
Rather more than usual	☐ 3
Much more than usual	☐ 4

6. Have you recently felt you couldn't overcome your difficulties?

- Not at all ☐ 1
- No more than usual ☐ 2
- Rather more than usual ☐ 3
- Much more than usual ☐ 4

7. Have you recently been able to enjoy your normal day-to-day activities?

- More so than usual ☐ 1
- Same as usual ☐ 2
- Less so than usual ☐ 3
- Much less than usual ☐ 4

8. Have you recently been able to face up to your problems?

- More so than usual ☐ 1
- Same as usual ☐ 2
- Less able than usual ☐ 3
- Much less able ☐ 4

9. Have you recently been feeling unhappy and depressed?

- Not at all ☐ 1
- No more than usual ☐ 2
- Rather more than usual ☐ 3
- Much more than usual ☐ 4

10. Have you recently been losing confidence in yourself?

- Not at all ☐ 1
- No more than usual ☐ 2
- Rather more than usual ☐ 3
- Much more than usual ☐ 4

11. Have you recently been thinking of yourself as a worthless person?

- Not at all ☐ 1
- No more than usual ☐ 2
- Rather more than usual ☐ 3
- Much more than usual ☐ 4

12. Have you recently been feeling reasonably happy, all things considered?

More so than usual	☐ 1
About the same as usual	☐ 2
Less so than usual	☐ 3
Much less than usual	☐ 4

Section 12—This question is to be completed after you have been told what treatment you will be receiving:

1. Please rate on a scale of 0—10 how confident you are that this treatment will be successful in reducing the symptoms due to your recent injury where 0 is 'no confidence at all' and 10 is 'complete confidence'?

0 1 2 3 4 5 6 7 8 9 10

Thank you very much for your time.

10.4 APPENDIX 4 - MINT RESEARCH CLINIC ASSESSMENT FORM

PARTICIPANTS ID:

Managing Injuries of the Neck Trial
MINT

Research Clinic Examination

To be completed by the Research Physiotherapist

Once complete please return to Warwick Emergency Care and Rehabilitation

THE UNIVERSITY OF WARWICK

Please read all the instructions before completing the questionnaire.

NHS

ISRCTN 3302125

Section 1

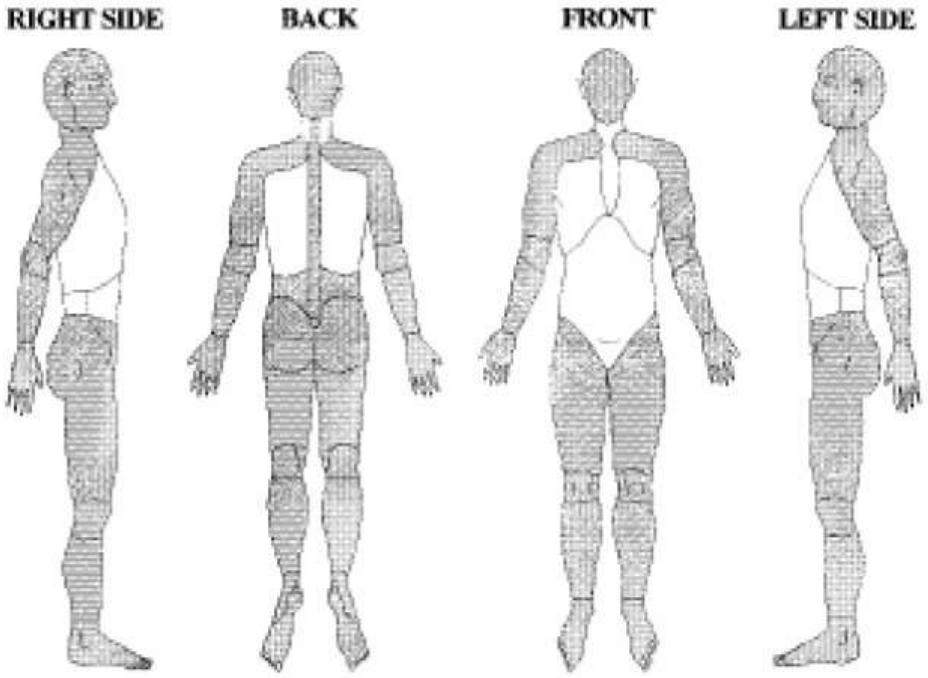
1. Please identify all areas where the patient is experiencing their symptoms since their injury

Skull	☐ 1	Neck	☐ 2		
Right Shoulder	☐ 3	Right arm above elbow	☐ 4	Right arm below elbow	☐ 5
Right hand	☐ 6				
Left Shoulder	☐ 7	Left arm above elbow	☐ 8	Left arm below elbow	☐ 9
Left hand	☐ 10				

2. Please ask the patient the following questions :	Yes	No
a. Do you have any difficulty swallowing?	☐ 1	☐ 2
b. Do you have headaches which you feel are related to your neck condition?	☐ 1	☐ 2
c. Do you have numbness, weakness, or tingling in your arms or hands?	☐ 1	☐ 2
d. Do you have difficulty grasping, picking up, or holding things in your hands?	☐ 1	☐ 2
e. Do you have numbness, clumsiness, or weakness in your legs?	☐ 1	☐ 2

3. Widespread chronic pain (Manchester definitions):

Please mark on the body chart all areas where the patient experienced pain for <u>at least 3 months</u> prior to their whiplash injury:



Manchester definition of chronic widespread pain
For subjects to satisfy the Manchester definition of chronic widespread pain, pain must be reported in at least two sections of two contralateral limbs and in the axial skeleton, and have been present for at least 3 months.

Following completion of symptoms on the body chart, did the patient experience:

	Yes	No
a. Pain in two sections of two contralateral limbs	☐ 1	☐ 2
b. Pain in the axial skeleton	☐ 1	☐ 2
c. This pain was present for at least 3 months prior to neck injury	☐ 1	☐ 2
d. Did the patient fulfil the Manchester Definition of Chronic Widespread Pain prior to their injury If Yes—Must have answered "Yes" to all 3 criteria listed above	☐ 1	☐ 2

Section 2 : Physical Examination

Cervical Range of Motion

Direction of Motion	Range of Motion/ Degrees	Limitation Pain / Stiffness / Spasm
Active Flexion		
Active Extension		
Active Right Rotation		
Active Left Rotation		
Active Right Lateral Flexion		
Active Left Lateral Flexion		
Passive Flexion		
Passive Extension		
Passive Right Rotation		
Passive Left Rotation		
Passive Right Lateral Flexion		
Passive Left Lateral Flexion		

Active Shoulder Range of Motion

Direction of Motion	Range of Motion/ Degrees	Limitation Pain / Stiffness
Right Abduction		
Left Abduction		

10.5 APPENDIX 5 - MULTICENTRE RESEARCH ETHICS COMMITTEE

APPROVAL LETTER FOR MINT AND COHORT STUDY

Trent Multi-centre Research Ethics Committee

Chairman: Dr Robert Bang
Administrator: Jill Marshall

RECEIVED 2 9 APR 2004.

Derwent Shared Services
Laurie House
Colyear Street
Derby
DE1 1LJ

Telephone: 01332 868805
Fax: 01332 868930
Email: Jill.Marshall@derwentshareddservices.nhs.uk

Your Ref:

28 April 2004

Professor Sarah Lamb
Professor of Rehabilitation
University of Warwick Rm
104 Avon Building
Westwood Campus
COVENTRY CV47AL

Dear Professor Lamb

MREC/04/4/003 - *please quote this number on all correspondence*
Managing Injuries of the Neck (MINT)
Funder's ref no: 02/35/02

Thank you for your letter of 7 April 2004, responding to the Committee's request for further information on the above research and submitting revised documentation.

The further information has been considered on behalf of the Committee by the Chairman.

Confirmation of ethical opinion

On behalf of the Committee, I am pleased to confirm a favourable ethical opinion for the above research on the basis described in the finalised application form, protocol and supporting documentation.

Multi-site studies requiring site-specific assessment

You should now arrange for site-specific assessment to be carried out for all sites at which Principal Investigators are to be appointed to conduct the research locally. *(In future under the new REC system you will be able to apply for SSA earlier on in the process, once you have received notice of validation of your application).*

Part C of the application form (complete with all signatures) together with a copy of the Principal Investigator's CV, should be sent to the relevant Local Research Ethics Committee (LREC) for each site. No further documents need to be submitted. Site-specific assessment is confined to an assessment of the suitability of the local investigators, support staff, site and facilities.

The local assessor will be either the LREC itself or another assessor approved for the site by the relevant Office for Research Ethics Committees. Local assessors have 30 days in which to notify this Committee whether or not there is any objection on site-specific grounds. We will then confirm the favourable ethical opinion for each site in writing to you.

Conditions of approval

The favourable opinion is given provided that you comply with the conditions set out in the attached document. You are advised to study the conditions carefully.

The Central Office for Research Ethics Committees is responsible for the operational management of Multi-centre Research Ethics Committees

Approved documents

The final list of documents reviewed and approved by the Committee is as follows:

. Application form dated 5.11.03
. Patient information sheet Version 2 dated 7.7.04
. Patient consent form Version 1 dated 5.11.03
. Health Questionnaire - SF36
. Hospital Anxiety and Depression Scale (HAD)
. EuroQol Quality of Life Questionnaire
. GP letter Version 1 dated 5.11.03
. Protocol Version 2 dated 6.4.04
. Confirmation of funding/peer review from NHCCT A dated 13 October 2003
. Article: Validity and Reliability of a modified version of the Neck Disability Index - J Rehab Med 2002; 34, 284-287
. Article: Cervical Spines Outcome Questionnaire: SPINE Vo127, No 19, pp 2116-2124
. Additional proposal to include randomised sub-study of incentives to promote the return of postal follow up questionnaires, embedded within the whiplash study - Prof Lamb's email of 5.1.04
. Method of initial recruitment to study
. Payments to researcher
. Provision of expenses for subjects
. Compensation arrangements for subjects
. Indemnity for investigators
. Chief Investigator's CV- Professor Sarah Lamb

Management approval - multi-site studies requiring site-specific assessment

If you are the Chief Investigator as well as the Principal Investigator for the lead site, you should obtain final management approval from your host organisation before commencing this research.
The study should not commence at any site until the local Principal Investigator has obtained final management approval from the relevant host organisation.

Notification of other bodies

We shall notify the research sponsor that the study has a favourable ethical opinion.

Statement of compliance

The Committee is constituted in accordance with the Governance Arrangements for Research Ethics Committees (July 2001) and complies fully with the Standard Operating Procedures for Research Ethics Committees in the UK.

Yours sincerely

Jill Marshal!
Trent MREC Administrator
on behalf of Dr Robert Bing, Chairman

Enclosures Standard approval conditions [SL-AC1]

MREC/04/4/003

10.6 APPENDIX 6 - PUBLISHED MANUSCRIPT OF SYSTEMATIC REVIEW OF PHYSICAL PROGNOSTIC FACTORS

Available at

http://www.ncbi.nlm.nih.gov/pubmed?term=mark%20Williams%20AND%20physical%20prognostic%20factors

10.7 APPENDIX 7 - PUBLISHED MANUSCRIPT OF SYSTEMATIC REVIEW OF RELIABILITY AND VALIDITY STUDIES OF METHODS FOR MEASURING ACTIVE AND PASSIVE CERVICAL ROM

Available at

http://www.ncbi.nlm.nih.gov/pubmed?term=SYSTEMATIC%20REVIEW%20OF%20RELIABILITY%20AND%20VALIDITY%20STUDIES%20OF%20METHODS%20FOR%20MEASURING%20ACTIVE%20AND%20PASSIVE%20CERVICAL%20ROM

10.8 APPENDIX 8 - QUALITY ASSESSMENT SCORES FOR RELIABILITY STUDIES IN SYSTEMATIC REVIEW OF RELIABILITY AND VALIDITY STUDIES OF METHODS FOR MEASURING ACTIVE AND PASSIVE CERVICAL ROM

		QA Criteria - Item Number													
Method	**First Author**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**No. of +ve items**
Digital inclinometer	Hoving	1	1	0	0	1	1	1	1	1	1	1	1	1	11
	Tousignant, b	1	1	1	0	0	n/a	1	1	1	1	1	0	1	9
	Zwart	0	0	0	0	n/a	n/a	1	0	0	n/a	0	0	0	1
Electromagnetic motion analysis	Jordan, 2000	1	1	1	1	1	1	1	1	1	1	1	1	1	13
	Assink	1	1	0	0	1	1	1	0	0	1	1	1	1	9
	Jordan, 2004	1	1	1	0	0	0	1	0	1	0	1	1	1	8
	Morphett	0	0	0	1	0	0	1	0	0	1	1	1	1	6
	Sterling	1	1	0	0	n/a	0	1	0	0	n/a	1	1	1	6
	Amiri	1	0	0	0	0	0	1	0	0	0	1	1	1	5
Goniometry	Haynes	1	1	0	0	0	1	1	1	1	1	1	1	1	10
	Cleland	1	1	0	n/a	0	0	1	1	1	1	1	1	1	9
	Pellecchia	1	0	0	0	0	0	1	1	0	1	1	1	0	6
	Agarwal, a	1	0	0	0	0	0	1	0	0	0	0	1	1	4
Gravity-plus-compass goniometer	Pool	1	0	0	0	1	0	1	1	1	1	1	1	0	8
	Hole	1	0	0	0	0	1	1	1	1	1	0	1	0	7
	Love	0	1	0	n/a	0	0	1	1	1	1	1	0	1	7
	Peolsson	1	1	0	1	1	n/a	1	0	0	1	1	0	0	7
	Youdas, 1991	1	1	0	0	0	0	1	1	1	1	0	0	0	6
	Lee	1	1	0	1	0	n/a	0	0	0	0	1	0	1	5

	Youdas, 1992	0	0	0	0	0	n/a	1	1	1	1	1	0	0	5
	Malmstrom	1	0	0	0	0	1	1	0	0	0	1	0	0	4
	Nilsson	1	0	0	0	0	n/a	1	0	0	0	1	1	0	4
	Olson	1	0	0	0	0	n/a	0	0	0	0	0	1	1	3
	Rheault	0	0	0	0	0	0	1	0	0	1	0	1	0	3
Inclinometry	Pile	1	0	0	0	1	0	0	0	1	1	0	0	0	4
	Tucci	0	0	0	0	0	0	1	0	0	1	1	1	0	4
	Bush	0	0	0	0	0	0	1	0	0	0	1	1	0	3
Optical Motion Analysis	Antonaci	1	1	0	0	0	0	1	0	0	0	1	1	1	6
	Bulgheroni	0	0	0	0	n/a	0	1	0	0	n/a	1	1	0	3
Potentiometer	Christensen	1	1	0	n/a	1	1	1	1	1	1	1	1	1	11
	Petersen, 2000	1	1	0	0	1	0	1	1	0	0	1	1	1	8
	Chiu	1	1	1	n/a	n/a	0	1	0	0	n/a	1	1	1	7
	Lantz	0	0	0	0	0	0	1	0	0	0	1	0	1	3
Tape measure	Maksymowych	1	1	0	0	1	1	1	0	0	1	0	1	1	8
	Haywood	1	1	0	0	1	0	1	0	0	0	1	1	1	7
Ultrasound motion analysis	Strimpakos	1	1	0	0	1	n/a	1	0	1	1	1	1	1	9
	Dvir	1	1	0	0	0	0	1	0	0	0	1	1	1	6
	Cagnie	1	1	0	0	0	0	1	0	0	0	1	0	1	5
	Mannion	1	0	0	1	n/a	n/a	1	0	0	n/a	1	1	0	5
Visual Estimation	Bertilson	1	1	0	1	1	1	1	1	1	1	1	1	1	12
	Hoppenbrouwers	1	1	0	1	1	1	1	1	1	1	1	1	1	12
	Piva	1	1	1	0	0	0	1	1	1	1	0	1	1	9
	Fjellner	0	1	0	0	1	0	1	1	0	1	1	1	1	8
	Van Suijlekom	1	1	0	0	1	0	0	0	0	1	1	0	1	6
	Viikari-Juntura	1	0	0	1	0	0	0	1	1	0	0	1	0	5
Miscellaneous	Petersen, 2007	1	1	0	0	n/a	n/a	1	0	0	0	1	0	1	5

10.9 APPENDIX 9 - QUALITY ASSESSMENT SCORES FOR VALIDITY STUDIES IN SYSTEMATIC REVIEW OF RELIABILITY AND VALIDITY STUDIES OF METHODS FOR MEASURING ACTIVE AND PASSIVE CERVICAL ROM

Method	First Author	QA Criteria - Item Number														No. of +ve items
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	
Digital inclinometry	Wolfenberger	0	1	0	1	0	0	1	1	0	0	0	1	1	0	6
	Syed	0	0	0	0	n/a	n/a	1	1	0	0	n/a	1	0	1	4
	Mayer	0	0	0	0	0	n/a	0	0	0	0	n/a	1	0	1	2
	Alund	0	0	0	1	0	0	1	0	0	0	0	1	1	1	5
Electromagnetic motion analysis	Jordan	1	1	1	1	0	0	1	1	0	1	0	1	0	1	9
	Morphett	0	0	0	0	n/a	n/a	1	1	0	0	n/a	1	1	1	5
Goniometry	Herrmann	0	0	0	1	0	0	1	1	0	1	0	1	1	1	7
	Agarwal, b	1	0	0	n/a	0	0	0	1	0	0	0	1	0	0	3
Gravity-plus-compass goniometry	Tousignant, c	1	1	0	1	1	0	1	1	1	1	0	1	1	1	11
	Tousignant, b	0	1	0	0	n/a	n/a	1	1	1	1	1	1	1	1	9
	Peolsson	1	1	0	0	1	1	1	1	0	0	1	0	0	1	8
	Tousignant, a	1	1	0	0	n/a	n/a	1	1	0	0	1	1	1	1	8
	Hole	1	0	0	n/a	0	1	1	1	1	1	1	0	0	0	7
Inclinometry	Tucci	0	0	0	0	0	0	1	1	0	0	1	1	1	1	6
	Bush	0	0	0	0	0	0	1	1	0	0	0	0	1	0	3
Potentiometry	Petersen	1	1	0	0	0	n/a	1	1	0	0	n/a	0	1	0	5
	Lantz	0	0	0	0	n/a	n/a	1	1	0	0	0	1	0	1	4
Ultrasound motion analysis	Mannion	1	0	0	1	n/a	n/a	1	1	0	0	1	1	1	1	8
	Strimpakos	1	1	0	0	n/a	n/a	1	1	0	0	n/a	1	1	1	7

Malmstrom	1	0	0	0	0	0	1	1	0	0	1	1	1	0	6
Wang	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

10.10 APPENDIX 10 - CERVICAL ROM MEASUREMENT PROTOCOL

Measurements will be completed in a single 15 minute session.

Before you take these measurements explain to the participant:

"I am going to use this device to measure the movements of your neck, firstly with you doing the movements then with me moving your head for you."

Documenting ROM:

The range of movement in degrees should be recorded in the Research Clinic Examination (Form 5) after each movement.

It is also necessary to record the limiting factor for each movement.

Active ROM:

When the patient has reached their perceived limit, ask the participant:

"What is stopping you from taking your head any further? Is it pain or is it stiffness?"

Record their answer as either Pain or Stiffness.

Passive ROM:

Record what you assess to be the main limiting factor of each movement from the following: Pain, Stiffness or Spasm.

Position:

Ensure the chair/couch is positioned so the participant's left shoulder will point due north. Also, ensure you have room to get behind the patient.

Get the patient to sit down.

Remove any jewellery or clothing that may obstruct ROM.

Ensure the participant is sat with hips and knees at 90 degrees and feet flat on the floor.

Ensure the participant's arms rest on their lap.

Ensure the participant appears to be in a neutral pelvic position and their back is "straight".

Place the magnetic yoke over the participant's shoulders, ensuring it is equal anteriorly and posteriorly.

Place the CROM device carefully on the participant's head so that the nosepiece is on the bridge of the nose and the frame rests on the patient's ears. Tighten the Velcro strap across the back of the participant's head.

Place the rotational compass into the frame.

Correction of position:

It is important that the participant maintains the same thoracic and lumbar spinal position throughout the ROM assessment. You may use your hands to correct the position of their thoracic and lumbar spine. Ensure you inform the patient you will be doing this (see below for prompts).

It may also be that the participant has difficulty producing movements in the 3 distinct planes and uses substitution patterns or combined movements.

E.g. You ask them to perform lateral flexion, but they also rotate their cervical spine and elevate their shoulder.

If you assess they are not moving in the specified plane, explain and demonstrate what they are doing wrong and then ask them to repeat the movement ensuring they do not use a substitution pattern or combined movement.

Do not passively correct them as they are performing the movement.

Active ROM
Explain to the participant:

"I am going to get you to do a series of movements of your neck. I would like you to keep your back and shoulders as still as possible. I would like you to perform each movement steadily and move your head as far as you feel able. You will need to pause at the end of each movement so I can read the dial. I will ask you at this point what is stopping you from taking your head any further"

Flexion
Stand to the left of the participant.

"I want you to bend your head forward as far as you feel possible, like this [demonstrate flexion]. I want you to hold this position whilst I read the dial. Make sure you don't let your head twist or drop to the side as you do this and keep sitting up straight. I will place my hand on your chest to correct this if necessary." Read the dial on the left side of the patient's head.

Extension
"I want you to look up to the ceiling as far as you feel possible, like this [demonstrate extension]. Make sure your mouth is closed. I want you to hold this position whilst I read the dial. Make sure you don't let your head twist or drop to the side as you do this and keep

sitting up straight. I will place my hand on your back to correct this if necessary." Read the dial on the left side of the patient's head.

Rotation

Stand behind the participant.

The arrow on the magnetic yoke should be pointing north. This should be in the same direction as the compass goniometer.

Ensure the compass goniometer is level; adjust the position of the subject's head so that both gravity inclinometers read 0. After levelling the compass inclinometer, turn the rotation meter on the compass inclinometer until the pointer is 0.

"I want you to look over your right shoulder as far as you feel possible, like this [demonstrate rotation]. Imagine you are tracing a straight horizontal line with your eyes as you turn to look over your shoulder. I want you to hold this position whilst I read the dial. Try not to turn your shoulders or your back. I will place my hand on your shoulder to correct this if necessary."

Repeat this for left rotation.

Lateral Flexion
Stand in front of the participant.

"I want you to take your right ear to your right shoulder as far as you feel possible, like this [demonstrate right lateral flexion]. I want you to hold this position whilst I read the dial. Try not to let your head twist or let your shoulders hitch up. I will place my hand on your shoulder to correct this if necessary." Read the dial above the patient's forehead.

Repeat this for left lateral flexion.

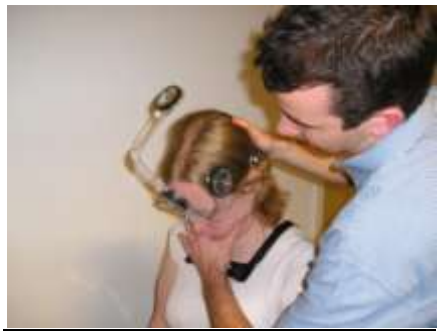
Passive ROM

"We are going to repeat the series of movements you have just done, but this time I am going to move your head for you. I want you to try and relax and do not try to help me. Please raise your arm if you want me to stop moving."

Hyperalgesia has been demonstrated in whiplash-injured subjects very soon after injury (Sterling et al, 2002; Sterling et al, 2003; Sterling et al, 2004) so some whiplash-injured patients may present with highly irritable symptoms.

We are asking you to use your clinical judgement regarding how far to passively move their cervical spine.

Flexion (figure 1)



Stand to the left of the participant.

"I am going to bend your head forward first. I don't want you to help me, try and relax. Please raise your arm if you want me to stop."

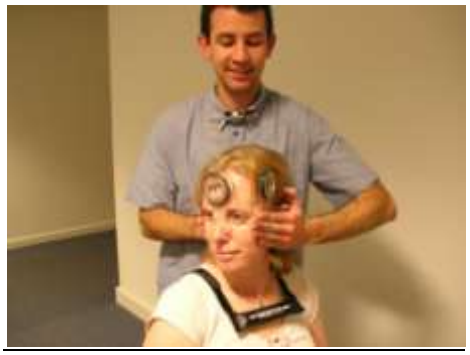
Put one hand on the back of the participant's head and, with the other hand, hold the subjects chin. Gently take the patient to the end of the available range (resistance is felt in the tissues preventing you from taking the movement further or until the patient signals you to stop the movement.)

Extension (figure 2)

"I am going to take your head backwards. I don't want you to help me, try and relax. Please raise your arm if you want me to stop."

Put one hand on the back of the participant's head and, with the other hand, hold the subject's chin. Gently take the patient to the end of the available range (resistance is felt in the tissues preventing you from taking the movement further or until the patient signals you to stop the movement.)

Rotation (figure 3)



Stand behind the participant.

Perform the same checks on the compass goniometer as for the AROM.

"I am going to turn your head to the right. I don't want you to help me, try and relax. Please raise your arm if you want me to stop."

Grasp the subject's head so both hands rest over the patient's ears. Gently take the patient to the end of the available range (resistance is felt in the tissues preventing you from taking the movement further or until the patient signals you to stop the movement.)

Repeat for left rotation.

Lateral Flexion (figure 4)

Stand in front of the participant.

"I am going to take your right ear to your right shoulder. I don't want you to help me, try and relax. Please raise your arm if you want me to stop."

Grasp the subject's head so both hands rest over the patient's the ears. Gently take the patient to the end of the available range (resistance is felt in the tissues preventing you from taking the movement further or until the patient signals you to stop the movement.)

Repeat for left lateral flexion.

To finish
Remove the rotational compass from the frame.

Remove the CROM device from the patient's head.

Ensure the CROM device is packed away securely.

10.11 APPENDIX 11 - MULTICENTRE RESEARCH ETHICS COMMITTEE

APPROVAL LETTER FOR RELIABILITY STUDIES

NHS
National Research Ethics Service
Trent Research Ethics Committee
Derwent Shared Services
Laurie House
Colyear Street
Derby
DE1 1LJ

Tel: 01332 868765
Fax: 01332 868785

10 December 2007

Professor Sarah Lamb
Professor of Rehabilitation
Room 104, Avon Building
Westwood Campus
Coventry
CV4 7AL

Dear Professor Lamb

Study title: managing injuries of the Neck (MINT)
REC reference: 04/4/003
Amendment number: Amendment 6
Amendment date: 20 November 2007

The above amendment was reviewed at the meeting of the Sub-Committee of the REC held on 06 December 2007.

Ethical opinion

The members of the Committee present gave a favourable ethical opinion of the amendment on the basis described in the notice of amendment form and supporting documentation.

Approved documents

The documents reviewed and approved at the meeting were:

Document	Version	Date
Participant Information Sheet	Reliability Study, V 2	20 November 2007
Participant Consent Form: Reliability Study	V 2	20 November 2007
Notice of Substantial Amendment (non-CTIMPs)	Amendment 6	20 November 2007
Covering Letter		22 November 2007

Membership of the Committee

The members of the Committee who were present at the meeting are listed on the attached sheet.

This Research Ethics Committee is an advisory committee to East Midlands Strategic Health Authority.
The National Research Ethics Service (NRES) represents the NRES Directorate within the National Patient Safety Agency and Research Ethics Committees in England.

10.12 APPENDIX 12 - MINT 12 MONTH FOLLOW-UP QUESTIONNAIRE

PARTICIPANTS ID:

Managing Injuries of the Neck Trial
M.I.N.T.

Recovery after Neck Injury Questionnaire

12 MONTH FOLLOW-UP

Please complete this questionnaire and return it in the reply-paid envelope provided.

If you have any difficulties or questions relating to the questionnaire, please contact:

Mark Williams
Warwick Emergency Care and Rehabilitation
FREEPOST MID18033
Room A1.03 Medical School Building
University of Warwick
Gibbet Hill Campus
Coventry
CV4 7AL
Freephone : 0800 634 0439
Fax : 024 7657 4657
E-mail : mint@warwick.ac.uk

THE UNIVERSITY OF WARWICK

NHS

ISRCTN 3302125

Please read all the instructions before completing the questionnaire

Thank you for agreeing to take part in this study. The answers you give in this questionnaire will help us find out if the treatments you get are helpful for your neck problems.

Please answer all the questions. Although it may seem that questions are asked more than once, it is still important that you answer every one.

Please follow the instructions for each section carefully.

For each section, if you are asked to put a cross in the box, please use a cross rather than a tick.

For example in the following question, if your answer to the question is yes, you should place a cross firmly in the box next to yes.

Do you drive a car?	Yes ☒	No ☐

If you are asked to circle a number, please use a circle rather than underlining a number.

For example, in the following question if you are asked 'how happy are you today?' Where '1' is 'very unhappy' and '5' is 'very happy'. If you feel neither happy or unhappy you may wish to answer '3'. You do this by clearly circling the number 3.

1 2 (3) 4 5

Please use a BLACK or BLUE pen. Please do not use a pencil.

Please check that you have completed all sections.

Please return your completed questionnaire to us in the envelope provided.

Please write any notes you have for us on the back page.

Please enter the date you completed this questionnaire:

D	D	M	M	Y	Y

Section 1—Neck Disability Index—This questionnaire is designed to enable us to understand how much your neck pain has affected your ability to manage everyday activities. Please answer each section by putting a cross in the box by the ONE CHOICE that most applies to you. We realise that you may feel that more than one statement may relate to you, but please place a cross in the box for the choice which closely describes your problem *right now*.

1. Pain Intensity

A. I have no pain at the moment ☐1

B. The pain is mild at the moment ☐2

C. The pain comes and goes and is moderate ☐3

D. The pain is moderate and does not vary much ☐4

E. The pain is severe but comes and goes ☐5

F. The pain is severe and does not vary much ☐6

2. Personal Care (Washing, Dressing etc.)

A. I can look after myself without causing extra pain ☐1

B. I can look after myself normally but it causes extra pain ☐2

C. It is painful to look after myself and I am slow and careful ☐3

D. I need some help, but manage most of my personal care ☐4

E. I need help every day in most aspects of self-care ☐5

F. I do not get dressed, I wash with difficulty and stay in bed ☐6

3. Lifting

A. I can lift heavy weights without extra pain ☐1

B. I can lift heavy weights, but it causes extra pain ☐2

C. Pain prevents me from lifting heavy weights off the floor but I can if they are conveniently positioned, for example on a table ☐3

D. Pain prevents me from lifting heavy weights, but I can manage light to medium weights if they are conveniently positioned ☐4

E. I can lift very light weights ☐5

F. I cannot lift or carry anything at all ☐6

4. Reading

A. I can read as much as I want to with no pain in my neck ☐1

B. I can read as much as I want with slight pain in my neck ☐2

C. I can read as much as I want with moderate pain in my neck ☐3

D. I cannot read as much as I want because of moderate pain in my neck ☐4

E. I cannot read as much as I want because of severe pain in my neck ☐5

F. I cannot read at all ☐6

5. Headache

A. I have no headaches at all ☐1

B. I have slight headaches which come infrequently ☐2

C. I have moderate headaches which come infrequently ☐3

D. I have moderate headaches which come frequently ☐4

E. I have severe headaches which come frequently ☐5

F. I have headaches almost all the time ☐6

6. Concentration

A. I can concentrate fully when I want to with no difficulty ☐1

B. I can concentrate fully when I want to with slight difficulty ☐2

C. I have a fair degree of difficulty in concentrating when I want to ☐3

D. I have a lot of difficulty in concentrating when I want to ☐4

E. I have a great deal of difficulty in concentrating when I want to ☐5

F. I cannot concentrate at all ☐6

7. Work

A. I can do as much work as I want to ☐1

B. I can only do my usual work, but no more ☐2

C. I can do most of my usual work, but no more ☐3

D. I cannot do my usual work ☐4

E. I can hardly do any work at all ☐5

F. I cannot do any work at all ☐6

8. Driving

A. I can drive my car without neck pain ☐ 1

B. I can drive my car as long as I want with slight pain in my neck ☐ 2

C. I can drive my car as long as I want with moderate pain in my neck ☐ 3

D. I cannot drive my car as long as I want because of moderate pain in my neck ☐ 4

E. I can hardly drive my car at all because of severe pain in my neck ☐ 5

F. I cannot drive my car at all ☐ 6

9. Sleeping

A. I have no trouble sleeping ☐ 1

B. My sleep is slightly disturbed (less than 1 hour sleepless) ☐ 2

C. My sleep is mildly disturbed (1—2 hours sleepless) ☐ 3

D. My sleep is moderately disturbed (2—3 hours sleepless) ☐ 4

E. My sleep is greatly disturbed (3—5 hours sleepless) ☐ 5

F. My sleep is completely disturbed (5—7 hours sleepless) ☐ 6

10. Recreation

A. I am able to engage in all recreational activities with no pain in my neck at all ☐ 1

B. I am able to engage in all recreational activities with some pain in my neck ☐ 2

C. I am able to engage in most, but not all recreational activities because of pain in my neck ☐ 3

D. I am able to engage in a few of my usual recreational activities because of pain in my neck ☐ 4

E. I can hardly do any recreational activities because of pain in my neck ☐ 5

F. I cannot do any recreational activities at all ☐ 6

Section 2 – This section asks for your views about your health. This information will help keep track of how you feel and how well you are able to do your usual activities.
For each of the following questions please place a cross in the one box that best describes your answer.

1. In general, would you say your health is:

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

2. The following questions are about activities you might do during a typical day.
Does your health now limit you in these activities? If so, how much?

	Yes, limited a lot	Yes, limited a little	No, not limited at all
a. **Moderate activities**, such as moving a table, pushing a vacuum cleaner, bowling or playing golf	☐ 1	☐ 2	☐ 3
b. Climbing **several** flights of stairs	☐ 1	☐ 2	☐ 3

3. In the last week, have you had any of the following problems with your work or other regular daily activities as a result of your physical health?

	Yes	No
a. **Accomplished less** than you would like	☐ 1	☐ 2
b. Were limited in the **kind** of work or other activities	☐ 1	☐ 2

4. In the last week, have you had any of the following problems with your work or other regular daily activities as a result of any emotional problems (such as feeling depressed or anxious)?

	Yes	No
a. **Accomplished less** than you would like	☐ 1	☐ 2
b. Didn't do work or other activities as carefully as usual	☐ 1	☐ 2

5. In the last week, how much did pain interfere with your normal work (including both work outside the home and housework)?

Not at all	A little bit	Moderately	Quite a bit	Extremely
☐ 1	☐ 2	☐ 3	☐ 4	☐ 5

6. These questions are about how you feel and how things have been with you in the last week.
For each question, please give the one answer that comes closest to the way you have been feeling.
How much of the time in the last week:

	All of the time	Most of the time	A good bit of the time	Some of the time	A little of the time	None of the time
a. Have you felt calm and peaceful?	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6
b. Did you have a lot of energy?	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6
c. Have you felt downhearted and low?	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐

7. In the last week, how much of the time has your physical health OR emotional problems interfered with your social activities (like visiting with friends, relatives, etc.)?

All of the time	Most of the time	Some of the time	A little of the time	None of the time
☐ 1	☐ 2	☐ 3	☐ 4	☐ 5

Section 3 - The following questions are to ask about your general health state **at the moment**. By placing a cross in one box in each group below, please indicate which statement best describes your own health state **today**.

Please cross **one** box for each question

1. Mobility

- I have no problems in walking about ☐ 1
- I have some problems in walking about ☐ 2
- I am confined to bed ☐ 3

2. Self Care

- I have no problems with self-care ☐ 1
- I have some problems washing or dressing myself ☐ 2
- I am unable to wash or dress myself ☐ 3

3. Usual Activities (e.g. work, study, housework, family or leisure activities)?

- I have no problems with performing my usual activities ☐ 1
- I have some problems with performing my usual activities ☐ 2
- I am unable to perform my usual activities ☐ 3

4. Pain / Discomfort

- I have no pain or discomfort ☐ 1
- I have moderate pain or discomfort ☐ 2
- I have extreme pain or discomfort ☐ 3

5. Anxiety / Depression

- I am not anxious or depressed ☐ 1
- I am moderately anxious or depressed ☐ 2
- I am extremely anxious or depressed ☐ 3

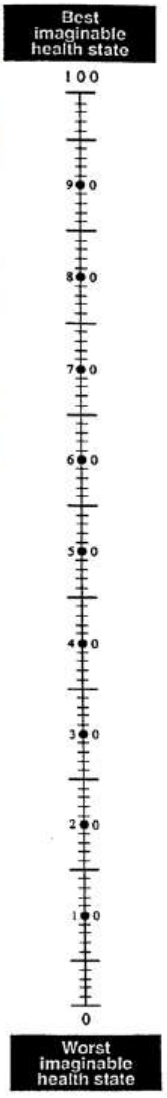
Your own health state today

To help people say how good or bad a health state is, we have drawn a scale (rather like a thermometer) on which the best state you can imagine is marked by 100 and the worst state you can imagine is marked by 0.

We would like you to indicate on this scale **how good or bad is your own health today, in your opinion.**

Please do this by drawing a line from the box below, to whichever point on the scale indicates how good or bad your current health state is **today**.

Your own health state TODAY



Section 4

This section is about health care you have received for your neck injury since your last questionnaire. There are separate parts for NHS treatment, private treatment, products/ equipment, normal activities and any benefits/entitlements. Please read each question carefully. For each question, if you have had no treatments or visits, please enter '0'.

1. Is your neck better, just the same or worse since your last questionnaire?

Much Better	Better	Same	Worse	Much Worse
☐ 1	☐ 2	☐ 3	☐ 4	☐ 5

NHS Treatment

2. **Since your last questionnaire**, how often have your attended the following **NHS Services** for your neck injury? (*Please do not include any sessions or treatments that you attended as part of the study).*

	Number of times	
Your GP or another GP	☐☐	If none enter '0'
Nurse	☐☐	If none enter '0'
Physiotherapist	☐☐	If none enter '0'
Doctor/nurse in an accident and emergency department (Casualty)	☐☐	If none enter '0'
Hospital specialist (consultant or one of his/her team)	☐☐	If none enter '0'
Psychologist/Counsellor	☐☐	If none enter '0'
Orthopaedic Clinic	☐☐	If none enter '0'
Other NHS Service (please specify) _____	☐☐	If none enter '0'

3. **Since your last questionnaire** have you been admitted to an **NHS hospital** because of your neck injury?

	Yes	No
	☐ 1	☐ 2
If **Yes**, in total, how many days were you in hospital?	☐☐	
Was this admission for surgery on your neck?	☐ 1	☐ 2

4. **Since your last questionnaire** have you had any of the following tests in a **NHS hospital** in relation to your neck injury?

	Number of times	
X-ray		If none enter '0'
CT Scan		If none enter '0'
MRI Scan		If none enter '0'
Blood test (count all blood tests done on one day, as one test)		If none enter '0'
Other (please specify) _____		If none enter '0'

5. **Since your last questionnaire** has your **doctor** prescribed any of the following medications because of your neck injury?

	Number of prescriptions	
Pain killers		If none enter '0'
Anti-inflammatory drugs (for example ibuprofen)		If none enter '0'
Gels/Creams (for example ibuleve or movelat)		If none enter '0'
Sleeping pills		If none enter '0'
Anti-depressants		If none enter '0'

6. Do you qualify for free prescriptions?	Yes ☐ 1	No ☐ 2
7. Do you have a pre-payment certificate (prescription passport) ?	Yes ☐ 1	No ☐ 2

Private Treatment

8. **Since your last questionnaire** please detail total treatment costs **you paid for yourself; or paid for by private insurance;** please do not include any treatments paid for by the NHS. Please round the amounts to the nearest pound.

	Number of times	Medical Insurance Contribution	Personal Contribution
Private physiotherapist		£	£
Private hospital specialist (consultant)		£	£
Private osteopath		£	£
Private chiropractor		£	£
Private psychologist/counsellor		£	£
Other (please specify) _____		£	£

9. **Since your last questionnaire**, have you been admitted to a **private hospital** because of your neck injury?

Yes ☐ 1 No ☐ 2

If **Yes**, in total, **how many days were you in hospital?** ____

Was this admission for surgery on your neck? Yes ☐ 1 No ☐ 2

If **Yes**, what were the **total costs borne by medical insurance?**
Please round the total costs to the nearest pound. £ _____

If **Yes**, what were the **total costs borne by you?**
Please round the total costs to the nearest pound. £ _____

10. **Since your last questionnaire** have you had any of the following tests in a **private hospital** in relation to your neck injury?

	Number of times	
X-ray	____	If none enter '0'
CT Scan	____	If none enter '0'
MRI Scan	____	If none enter '0'
Blood test (count all blood tests done on one day, as one test)	____	If none enter '0'
Other (please specify)	____	If none enter '0'

_____

If **Yes**, what were the **total costs borne by medical insurance?** £ _____
Please round the total costs to the nearest pound.

If **Yes**, what were the **total costs borne by you?** £ _____
Please round the total costs to the nearest pound.

11. **Since your last questionnaire have you bought (other than by a prescription)** any of the following treatments for your neck injury?

	Number of times	Total cost to you
Pain killers	____	£ ____
Anti-inflammatory drugs (for example ibuprofen)	____	£ ____
Gels/Creams (for example ibuleve or movelat)	____	£ ____
Sleeping pills	____	£ ____
Anti-depressants	____	£ ____

Products / Equipment

12. **Since your last questionnaire**, have you bought items such as collars, herbal remedies, orthopaedic devices or any other products or equipment **because of your neck injury**? (please list the item below and enter the cost to the nearest pound)

Item Bought

1. _____ £ |___|___|___|___|
2. _____ £ |___|___|___|___|
3. _____ £ |___|___|___|___|
4. _____ £ |___|___|___|___|
5. _____ £ |___|___|___|___|

13. Have you had to take any days off sick from work **since your last questionnaire** due to your neck injury?

Yes ☐ 1 No ☐ 2 Not Applicable ☐ 3

If **Yes**, how many days |___|___|

14. Have your hours of employment fallen **since your last questionnaire** because of your neck injury?

Yes ☐ 1 No ☐ 2 Not Applicable ☐ 3

If **Yes**, by how many hours per week as your employment fallen? |___|___|

15. Has your benefit status changed **since your last questionnaire**?

Yes ☐ 1 No ☐ 2

If **Yes**, please tick all benefits that you are currently receiving.

Housing and Homeless Assistance	☐ 1	Attendance Allowance	☐ 2
Severe Disablement Allowance	☐ 3	Carer's Allowance	☐ 4
Disability Living Allowance	☐ 5	Income Support	☐ 6
Working Tax Credit	☐ 7	Social Fund Payment	☐ 8
Council Tax Benefit	☐ 9	Child Benefit	☐ 10
Income based Jobseekers' Allowance	☐ 11	Housing Benefits	☐ 12
Pension Credit	☐ 13	Child Tax Credit	☐ 14
Other (please specify) _____	☐ 15		

Section 5 This section is to determine what you believe about your neck pain. Here are some of the things which other patients have told us about their pain.

For each statement please circle the number from 0 to 6 to say how much physical activities such as bending lifting walking or driving affect or would affect your neck pain.

For each statement please circle the number from 0 to 6 to say how much you agree with these statements.

Please circle one number for each line

	Completely disagree			Unsure			Completely Agree
1. Physical activity makes my pain worse	0	1	2	3	4	5	6
2. Physical activity might harm my neck	0	1	2	3	4	5	6
3. I should not do physical activities which (might) make my pain worse	0	1	2	3	4	5	6
4. I cannot do physical activities which (might) make my pain worse	0	1	2	3	4	5	6

Please circle one number for each line

	Completely disagree			Unsure			Completely Agree
5. When you have injured your neck it is best to exercise your neck to help you to recover	0	1	2	3	4	5	6
6. I feel I am able to cope with my neck problem even when it is painful.	0	1	2	3	4	5	6

Section 6 This section is to find out about any compensation claims you may have made.

1. Are you pursuing a personal injury claim for compensation?

Yes ☐ 1 No ☐ 2

If **Yes,** through whom are you pursuing :

Solicitor ☐ 1 Directly with Insurer ☐ 2 Other (specify)_____ ☐ 3

2.. Has the Claim been settled?

Yes ☐ 1 No ☐ 2

If **Yes,** on what date was the claim settled : ___ ___ ___

What amount of compensation did you receive? _____

3. Did the insurer ask if you had received NHS treatment and at which hospital?

Yes ☐ 1 No ☐ 2

If **Yes,** were any costs recovered by the NHS from the insurance company? :

Yes ☐ 1 No ☐ 2 Don't Know ☐ 3

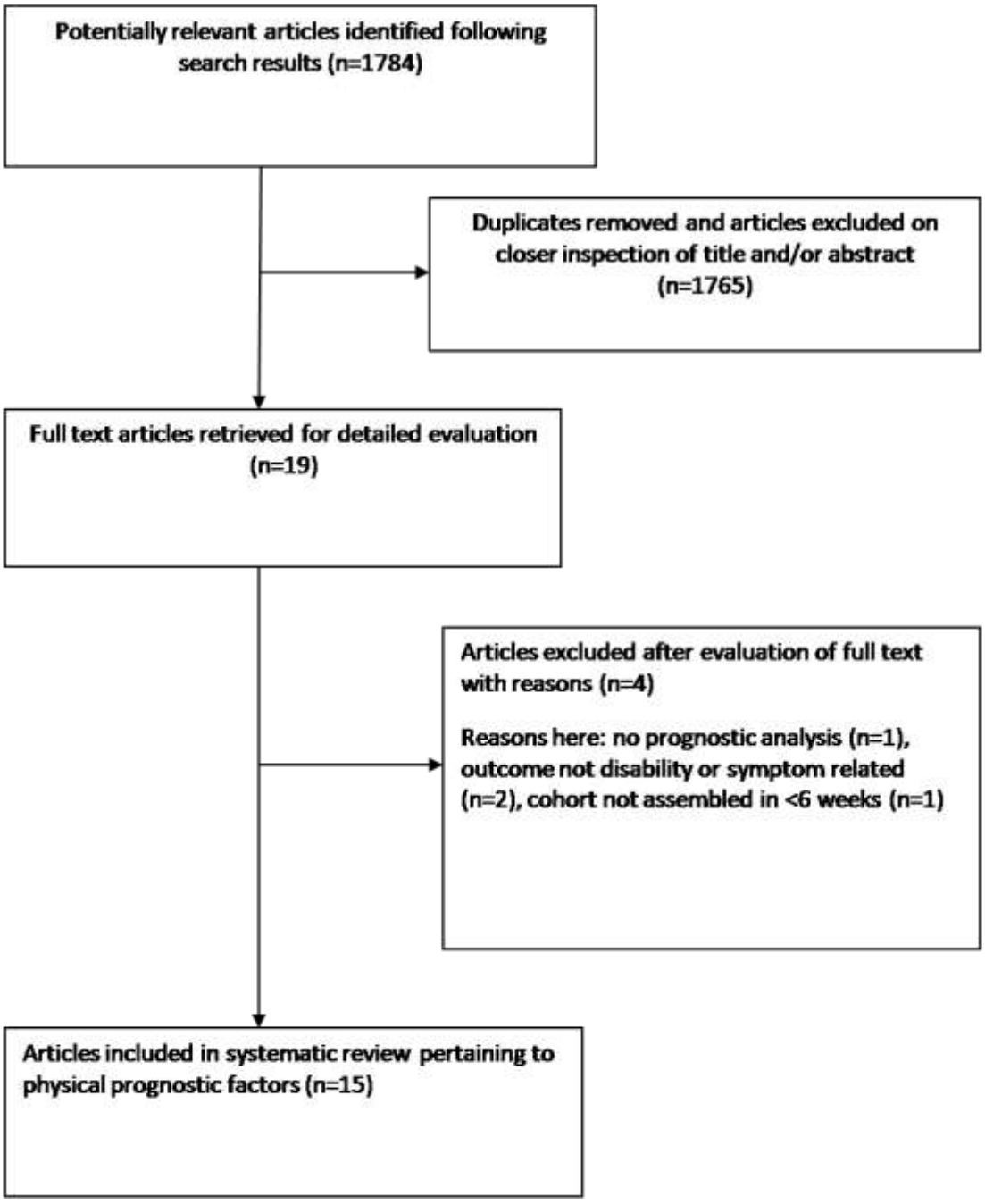
Please check that you have completed all sections.

Please return your completed questionnaire to us in the envelope provided.

Please write any notes you have for us in the space below.

Thank you very much for your time.

10.13 APPENDIX 13 - FLOW CHART OF UPDATED SEARCH RESULTS FOR SYSTEMATIC REVIEW OF PHYSICAL PROGNOSTIC FACTORS



10.14 APPENDIX 14 - FLOW CHART OF UPDATED SEARCH RESULTS FOR SYSTEMATIC REVIEW OF RELIABILITY AND VALIDITY STUDIES OF METHODS FOR MEASURING ACTIVE AND PASSIVE CERVICAL ROM