The Interexaminer Reliability of Static and Motion Palpation for the Assessment of Spinal Joint Dysfunction in Healthy Infants Aged Two to Ten Weeks.

By

JULEE RALPH

A dissertation submitted to the Faculty of Health, Durban Institute of Technology, in partial compliance with the requirements for a Master's degree in Technology: Chiropractic.

I, Julee Ralph, do hereby declare that this dissertation is representative of my own work in both conception and execution.

Julee Ralph — Date

Approved for Examination

Dr H. Kretzmann M.Dip.Tech: Chiro, CCFC (S.A.) — Date
Research Supervisor

Dr J. Shaik M.Tech: Chiro (S.A.) — Date
Research Co-supervisor

ACKNOWLEDGEMENTS

To my family – Thank you for your never-ending support and encouragement through all the years.

To Caroline – Your willingness to always help is greatly appreciated.

To Dr Kretzmann and Dr Shaik – Thank you for your guidance and patience.

To Mrs Ireland – Thank you for all your help in the running of this research study.

To all the Clinicians that participated in this study – Thank you for your time and effort in this research.

To Mrs Harold and the Durban Metro primary health care clinics of the Inner West Region – Thank you for allowing us to conduct this research at your clinics.

To all the private antenatal clinics – Thank you for allowing your clinic to participate in this research.

ABSTRACT

Chiropractors are treating spinal joint dysfunction in infants that present with conditions such as infantile colic. Authors conducting research into spinal joint dysfunction in infants have used static and motion palpation to identify these spinal lesions in the infants. The reliability of static and motion palpation used in infants for the assessment of spinal joint dysfunction has not yet been established. The lack of a reliable assessment tool for spinal joint dysfunction in infants reduces the inferential validity of the research studies assessing the efficacy of chiropractic treatment in infants. It is therefore necessary to establish the interexaminer reliability of static and motion palpation in infants.

The purpose of this study was to determine the interexaminer reliability of static and motion palpation for the assessment of spinal joint dysfunction in healthy infants aged two to ten weeks. This interexaminer reliability study was based on a single-facet (one cause of error), repeated-measures design. It was conducted at public and private primary health care and antenatal clinics in the greater Durban area. The sample in this study comprised of one hundred healthy infants aged two to ten weeks and was selected by means of convenience sampling at the participating clinics. Infants were accepted onto the study if they were healthy with no history of past or present illness and only if their parents gave their informed consent.

Each participating infant received a spinal assessment by three independent chiropractic examiners. The examiners did not reveal their assessment findings to each other, as an attempt to eliminate observer bias from the study. The assessment included static and passive motion palpation of the spine and sacroiliac joints, to detect the presence of spinal joint dysfunction. Each examiner recorded the outcomes of their assessment separately. The data were statistically analyzed using percentage agreement between examiners,

Pearson’s Chi-square and Cohen’s Kappa coefficient to calculate the interexaminer reliability.

It was concluded that where only static palpation was used and where static and motion palpation were used together for the assessment of spinal joint dysfunction in healthy infants aged two to ten weeks, fair interexaminer reliability was found. Where only motion palpation was used for such assessment, slight interexaminer reliability was found. The result of “slightly to fairly reliable” does not reflect acceptable interexaminer reliability. These results did however demonstrate that static and motion palpation were not useless assessment tools for the identification of spinal joint dysfunction in healthy infants aged two to ten weeks. Further research into this area is therefore warranted.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF APPENDICES

CHAPTER ONE: INTRODUCTION

1.1 THE PROBLEM AND ITS SETTING

A review of the literature concerning spinal manipulative therapy in paediatrics has revealed that the conditions in infants that chiropractors were consulted for, were managed by identifying and treating spinal joint dysfunction in these infants (Olafsdottir et al. 2001, Holtrop 2000, Wiberg et al. 1999, Fallon 1997, Fallon and Fysh 1997, Klougart et al. 1989). Spinal joint dysfunction can be described as the lack of movement of a joint in the spine, caused by muscular spasm or an intra-articular blocking (Gatterman 1990: 408).

Authors conducting research into spinal joint dysfunction in infants have used static and motion palpation to identify spinal joint dysfunction in the infants (Olafsdottir et al. 2001, Wiberg et al. 1999, Fallon 1997, Klougart et al. 1989). However, the reliability of palpation as an assessment tool in infants has not yet been satisfactorily established. Palpation of the spine is considered to be an essential diagnostic tool by manipulative therapists (Boline et al. 1993) and according to Hawk et al. (1999), it is one of the most common methods for identifying spinal joint dysfunction.

According to Bergmann et al. (1993: 738), clinical trials involving joint dysfunction have questionable validity unless the assessment procedures used in those studies have been tested and found to be reliable. The validity of the results of the studies assessing the efficacy of chiropractic treatment in infants (Olafsdottir et al. 2001, Wiberg et al. 1999, Fallon 1997, Klougart et al. 1989) is thus questionable and it is therefore necessary to establish the interexaminer reliability of these assessment tools in infants.

The efficacy of any diagnostic procedure can in one way be established by determining its reliability (Haas 1991 - c). The reliability of a procedure is defined by Haas (1995), as the extent to which a repeated test will produce the same result when evaluating an unchanged characteristic. Haas (1991 - a) also describes reliability as a measure of consistency, concordance or repeatability of experimental outcomes. Reliability can be tested in terms of *intra*examiner and *inter*examiner reliability. *Intra*examiner reliability is described by Bergmann et al. (1993: 723) as the ability of one examiner to repeatedly identify the same findings on an individual using the same procedure during a blinded trial and *inter*examiner reliability as the ability of a series of examiners to identify the same findings on an individual using the same procedure in a blinded trial.

The reliability of the procedures used to identify indications in patients for chiropractic treatment is important as it is the outcomes of these procedures that ultimately indicate if and where manual treatment will be delivered (Hawk et al. 1999, Haas 1991 - b). Research into the reliability of static and motion palpation for the assessment of spinal joint dysfunction has been conducted in adults and has produced mixed results (French et al. 2000, Haas et al. 1995 - b, Hubka and Phelan 1994, Boline et al. 1993 and Keating et al. 1990).

One of the most important goals for any assessment tool is for that tool to demonstrate good reliability (Lewit and Liebensen 1993). According to Russell (1983), establishing the reliability of the procedures used for the assessment of manipulable lesions is a challenge but is essential in the road to establishing the validity of these procedures and ultimately the validity of the manipulable entity. Once reliability is established, research into the efficacy of manipulative therapy can be validated. It is therefore the purpose of this study to determine the interexaminer reliability of static and motion palpation for the assessment of spinal joint dysfunction in healthy infants aged two to ten weeks.

1.2 THE AIM OF THE STUDY

The aim of this study was to determine the interexaminer reliability of static and motion palpation for the assessment of spinal joint dysfunction in healthy infants aged two to ten weeks.

1.3 THE OBJECTIVES OF THE STUDY

1.3.1 Objective One

The first objective was to determine the interexaminer reliability of static palpation for the assessment of spinal joint dysfunction in healthy infants aged two to ten weeks.

1.3.2 Objective Two

The second objective was to determine the interexaminer reliability of motion palpation for the assessment of spinal joint dysfunction in healthy infants aged two to ten weeks.

1.3.3 Objective Three

The third objective was to determine the interexaminer reliability of both static and motion palpation used together for the assessment of spinal joint dysfunction in healthy infants aged two to ten weeks.

1.4 THE HYPOTHESES

1.4.1 Hypothesis One

It was hypothesized that there would be interexaminer reliability of static palpation for the assessment of spinal joint dysfunction in healthy infants aged two to ten weeks.

1.4.2 Hypothesis Two

It was hypothesized that there would be interexaminer reliability of motion palpation for the assessment of spinal joint dysfunction in healthy infants aged two to ten weeks.

1.4.3 Hypothesis Three

It was hypothesized that there would be interexaminer reliability of both static and motion palpation used together for the assessment of spinal joint dysfunction in healthy infants aged two to ten weeks.

1.5 THE BENEFITS OF THE STUDY

The results of this study will attempt to provide an answer to the question of interexaminer reliability of static and motion palpation used for the assessment of spinal joint dysfunction in infants.

Establishing a procedure that has been tested and found to be reliable for the assessment of spinal joint dysfunction in infants will serve as a basis on which further research may be validated (i.e. research evaluating the efficacy of chiropractic treatment).

A diagnostic procedure that is found to be statistically reliable can be used to monitor patient progress in research and in the clinical setting.

This research endeavours to contribute to the base of knowledge regarding the assessment for spinal joint dysfunction in infants.

CHAPTER TWO: LITERATURE REVIEW

2.1 RELIABILITY

A standard procedure for the diagnosis of a manipulable lesion has not been established for use in research regarding manipulative therapy (Haas et al. 1995 - a). The clinical usefulness of the assessment tools used by chiropractors, such as static and motion palpation, is therefore, an important area of research (Haas et al. 1995 - a).

The usefulness of any diagnostic procedure can be established by determining its reliability, validity and sensitivity (Haas et al. 1995 – a). The reliability of a procedure is defined by Haas (1995), as the extent to which a repeated test will produce the same result when evaluating an unchanged characteristic. Haas (1991 - c) also describes reliability as a measure of consistency, concordance or repeatability of an experimental outcome. It should be noted that a reliable diagnostic tool may be reliable without being valid (i.e. it may be yielding the same incorrect outcome consistently) (Russel 1983). Haas et al. (1995 - a) describes the construct validity of a diagnostic tool as the ability of that tool to perform up to the standards predicted from a theoretical point of view. Carmichael (1987) describes the *sensitivity* of a diagnostic tool as the ability of that tool to detect the presence of *abnormality* and the *specificity* of a tool as the ability of that tool to detect the presence of *normality*.

According to Russell (1983), establishing the reliability of the diagnostic procedures used for the assessment of manipulable lesions does pose a challenge, but is essential in the road to establishing the validity of these procedures and ultimately the validity of the manipulable entity. A reliable method of assessment will be easier to test for validity than a method that has not been found to be reliable. The reliability of the tools used to identify indications in patients for chiropractic treatment is important, as it is the outcomes in the use of these tools that ultimately indicate if and where manual treatment will be delivered (Hawk et al. 1999, Haas 1991 - b). Bergmann et al. (1993: 738) states that clinical trials involving joint dysfunction have questionable validity unless the assessment tools for the joint dysfunction used in those studies have been tested and found to be reliable.

Reliability can be tested in terms of *intra*examiner and *inter*examiner reliability. *Intra*examiner reliability is described by Bergmann et al. (1993: 723) as the ability of one examiner to repeatedly identify the same findings on an individual using the same procedure during a blinded trial and *inter*examiner reliability as the ability of a series of examiners to identify the same findings on an individual using the same procedure in a blinded trial. According to Haas (1995), agreement between examiners (i.e. interexaminer reliability) is a more compelling measure of performance of a procedure than examiner self-consistency (i.e. intraexaminer reliability). Intraexaminer reliability is tested by having the examiners perform additional examinations on each patient. The problems when testing both intra- and interexaminer reliability at the same time, are numerous and include: (a) an increase in the time and expense of conducting such a study; (b) it may be difficult to blind the examiners to their previous findings; (c) the additional procedures performed on the subjects may aggravate or even resolve any existing spinal joint dysfunction, making it difficult to prove reliability, and; (d) an increased patient inconvenience.

2.2 SPINAL JOINT DYSFUNCTION

Joint dysfunction is a term commonly used to describe a manipulable lesion (Anrig and Plaugher 1998: 467). Spinal joint dysfunction can be defined as the lack of movement of a joint in the spine, caused by muscular spasm or an intra-articular blocking (Gatterman 1990: 408).

It could be noted from reviewing the literature regarding spinal manipulative therapy in paediatrics, that the conditions in infants that chiropractors dealt with, were managed by identifying and treating spinal joint dysfunction in these infants (Olafsdottir et al. 2001, Holtrop 2000, Mercer 1999, Wiberg et al. 1999, Fallon 1997, Fallon and Fysh 1997, Klougart et al. 1989). An extensive review of the literature has revealed that no research has been conducted into the reliability of the assessment tools used in infants to identify spinal joint dysfunction.

According to Bergmann et al. (1993:63), the physical signs indicative of joint dysfunction are regional tenderness and abnormalities in alignment, tissue texture and joint mobility. Physical examination of the spine in infants can include visual observation, static palpation, motion palpation, range of motion and spinal percussion (Anrig and Plaugher 1998: 140). Anrig and Plaugher (1998: 331-387) also discuss radiographic analysis and instrumentation as assessment tools for spinal joint dysfunction in infants. No research has been conducted into the reliability of any of these assessment tools used in infants and so no conclusions can be drawn with regard to their validity.

Palpation of the spine is considered an essential diagnostic tool by manipulative therapists and is one of the most common methods of identifying manipulable lesions or joint dysfunction in the spine (Hawk et al. 1999 and Boline et al. 1993).

2.3 STATIC AND MOTION PALPATION

Static palpation is defined by Bergmann et al. (1993: 762) as, “palpatory diagnosis of somatic structures in a neutral static position”. It is often used to establish the anatomical landmarks of the spine and is used as a baseline for the spinal assessment (Anrig and Plaugher 1998: 140). Static palpation is used to assess for numerous signs as discussed in the review of the related literature (French et al. 2000, Hubka and Phelan 1994, Boline et al. 1993 and Keating et al. 1990).

Motion palpation is defined by Bergmann et al. (1993: 762) as, “palpatory diagnosis of passive and active segmental joint range of motion”. Motion palpation can be described in terms of active and passive motion palpation (Mootz et al. 1989). In passive motion palpation, the examiner assesses the quality of movement between adjacent articular structures without the assistance of the patient (i.e. the examiner moves the patient while the patient relaxes). In active motion palpation, the patient moves him- or herself while the examiner assesses the quality of movement between adjacent articular structures (Mootz et al. 1989).

Research into the reliability of static and motion palpation for the assessment of spinal joint dysfunction has been conducted in adults (French et al. 2000, Fjellner et al. 1999, Hawk et al. 1999, Haas et al. 1995, Hubka and Phelan 1994, Boline et al. 1993, Keating et al. 1990, Mootz et al. 1989 and Nansel et al. 1989). No research has been conducted into the reliability of static or motion palpation used in infants for the assessment of spinal joint dysfunction.

In terms of manipulative therapy, static and motion palpation have been used to assess for various signs of spinal joint dysfunction. These signs are numerous and are mentioned in the discussion below regarding reliability research conducted on adults. The validity of all the various signs has not yet been satisfactorily established in adults or in infants. This is perhaps the next step in the research to determine the usefulness of static and motion palpation for the assessment of spinal joint dysfunction.

French et al. (2000) asked examiners to use static and motion palpation among seven other assessment tools (which included visual postural analysis, pain description given by the patient, plain static erect x-ray film of the lumbar spine, leg length discrepancy, neurological tests and orthopaedic tests) to determine where manipulation would be directed in the lower back of the subjects in this study (sample size of nineteen). Static palpation was used by the examiners to determine any tenderness, muscle hypertonicity, misalignment or any other palpatory sign that could aid the examiners in their diagnosis of a manipulable lesion in the lower back. The examiners were able to use motion palpation but were not instructed on any specific technique regarding the procedure. The lack of clear instruction with regard to the exact procedures followed in this study reduces the validity of the results and makes it difficult to draw any conclusions from this study. The examiners were blinded to the decisions made by the other examiners. The interexaminer reliability of all the assessment tools combined produced results of fair agreement between examiners beyond chance (K = 0.27; $p<0.0001$). French et al. (2000) did not analyze the interexaminer reliability of each assessment tool on its own and so limited conclusions regarding the interexaminer reliability of static palpation and motion palpation can be drawn. The inadequate sample size of nineteen reduces the significance of these results.

A study conducted by Keating et al. (1990) examined the interexaminer reliability of static and motion palpation among other tools for the assessment of joint dysfunction in the lumbar spine. The other tools included visual observation (for gross asymmetry, local hyperemia, oedema and acute skin lesions) and dermathermography (for segmental temperature deviations greater or equal to 1.5° F). The examiners used static palpation to assess for misalignment of vertebrae in relation to the one below, osseous pain (pain over bony structures of each joint), soft tissue pain (pain upon probing of paraspinal tissues) and muscle tension. Active and passive motion palpation was used to assess for movement restriction and hard end feel. The examiners were blinded to the findings of the other examiners. The sample size in this study was only forty six (twenty five asymptomatic and twenty one low back pain patients), which reduces the significance of the results in this study (Haas 1991 - b). Static palpation for osseous pain revealed results of marginal to good and significant interexaminer reliability (K = 0.19 to K = 0.66, mean K = 0.48; $p<0.05$); for soft tissue pain, results of marginal to good and significant interexaminer reliability were noted (K = 0.10 to K = 0.59; mean K = 0.30; $p<0.05$) at most levels. Palpation for misalignment, muscle tension and temperature changes each produced results of zero to marginal interexaminer reliability beyond chance (mean K = 0.00, mean K = 0.07 and mean K = 0.25 respectively), and statistical significance of these values was achieved infrequently. Visual observation produced results of marginal reliability (mean K = 0.29). Active and passive motion palpation was found to have marginal interexaminer reliability beyond chance (mean K = 0.09 and K = 0.07 respectively) with infrequent statistically significant findings. The results of the individual assessment tools were combined to determine overall results of interexaminer reliability using Pearson's *r* and intraclass correlation coefficients. The scores obtained from the combination of the more reliable tools (palpation for osseous and soft tissue pain, visual observation and dermathermography) produced results of greater reliability (mean *r* = 0.48 and mean ICC = 0.46, p<0.05) than the scores obtained from the combination of all the assessment tools used together (mean *r* = 0.36 and mean ICC = 0.37, p<0.05). This finding supports the use of a combination of the most reliable assessment tools to establish an examination procedure that is reliable and can then be tested for validity. (Keating et al. 1990.)

Boline et al. (1993) further studied the interexaminer reliability of palpation for paraspinal soft tissue pain and osseous pain (pain over spinous processes) along with other assessment tools (visual observation, dermothermography and surface electromyography) for joint dysfunction in the lumbar spine. The sample size in this study was twenty-eight. This inadequate sample size reduces the significance of the results obtained in this study. The inadequate sample size was surprising given that Haas was one of the co-authors of this study. The examiners were blinded to the findings of the other examiners. Palpation for osseous pain and paraspinal soft tissue pain was found to have good to excellent interexaminer reliability (K = 0.48 to K = 0.90, p<0.01 for osseous pain and K = 0.40 to K = 0.79, p<0.05 for paraspinal soft tissue pain). The results of the palpation for osseous and soft tissue pain and the findings from visual observation (K = 0.34 to K = 0.84) were combined to determine their interexaminer reliability used together. The results ranged from no interexaminer reliability beyond chance (K = -0.30) to good interexaminer reliability beyond chance (K = 0.50) for the combined assessment. These scores are all lower than those obtained in the individual analysis of each assessment tool. This finding questions the relationship between the signs and symptoms under investigation and highlights the need for the individual and combined analysis of each assessment tool. (Boline et al. 1993.)

Hubka and Phelan (1994) tested the interexaminer reliability of static palpation used for the detection of maximum paraspinal tenderness in the cervical spine of thirty patients with neck pain. The small sample size reduces the significance of the results obtained in this study. Only patients with unilateral mechanical neck pain were included in the study. The examiners were required to palpate the painful side of the cervical spine and determine the most tender spot on that side by communicating with the patient. If more than one spot was found to be tender, the patient was required to select the spot that was most tender. The examiners were blinded from each other by conducting the evaluation independently. Results of good interexaminer reliability (K = 0.68, $p<0.001$) were reported by the authors. This study did not evaluate the reliability of palpation for tenderness using a “present” or “absent” response at each intervertebral level of the cervical spine and may therefore show elevated scores of interexaminer reliability. (Hubka and Phelan 1994.)

Haas et al. (1995 - b) studied the reliability of motion palpation of the thoracic spine. The examiners were required to use motion palpation for the assessment of end-play restriction in the thoracic spines of sixty first year chiropractic students. The examiners conducted a sitting end-play assessment of all the motion segments from T3-T4 to T12-L1 in left and right rotation. Examiners recorded the presence or absence of restricted end-play at each level. The finding of restricted end-play was required to be of sufficient magnitude to indicate treatment by spinal manipulation. The examiners were blinded to each other’s findings. Overall, interexaminer reliability was found to be poor (K = 0.14) for both segmental level and direction of end play restriction. Concordance between examiners on segmental restriction in either direction was only slightly better (K = 0.19). The authors did not state whether the kappa scores obtained in this study demonstrated statistical significance.

Haas et al. (1995 - a) conducted a randomized controlled trial to evaluate the short-term responsiveness of thoracic motion restrictions to spinal manipulation for the purpose of evaluating the construct validity of motion palpation. Sixty first year chiropractic students were included in this study. Motion palpation was conducted by two examiners using a blind, randomized, repeated-measures design. Subjects who were found to have end-play restrictions were randomly assigned to two groups, a treatment group and a control group. The treatment group received spinal manipulation (from a third chiropractor) to a target motion segment that was determined by the pre-selected examiner. The control group received no intervention. The two examiners were blinded from the intervention received. The follow-up examiner was then required to re-examine the target motion segment and indicate whether normal end-play had been restored. The response (change from a positive result to a negative result) of motion restriction to spinal manipulation was 60% compared with the 37% response in the control group that received no manipulation ($p<0.04$). The results of this study provide some evidence that chiropractors may be able to manually palpate segmental end-play restriction in human subjects using motion palpation.

Reliability studies regarding spinal joint dysfunction have revealed that static palpation for tenderness (including paraspinal soft tissue and osseous tenderness) has yielded results of good interexaminer reliability (Hubka and Phelan 1994; Boline et al. 1993; Keating et al. 1990). Motion palpation procedures have however, demonstrated poor interexaminer reliability (Boline et al. 1993; Keating et al. 1990 and Haas et al. 1995 - b). The usefulness of motion palpation is however demonstrated by the construct validity revealed in the previously discussed randomized controlled trial conducted by Haas et al. (1995 -

a). The relationship of the results regarding interexaminer reliability of static and motion palpation in adults and in infants has not been established. It cannot be extrapolated from these results that they may be similar when interexaminer reliability is tested in infants. These studies, conducted on adults, are a useful starting point when designing a similar research study using a sample of infants.

2.4 INFANTS AND CHIROPRACTIC

Lee et al. (2000) conducted a descriptive survey in Boston where paediatric care given by chiropractors was investigated. Responding chiropractors were required to answer questions regarding demographics, practice characteristics (e.g. solo or group practice, number of patients seen per week, length and frequency of visits) and the scope and content of their practice. The authors commented that the number of children visiting chiropractors was substantial and was increasing. According to this survey, common conditions observed in the chiropractors' paediatric patients included infantile colic, otitis media, asthma, allergies and enuresis.

A number of studies have tested the effectiveness of spinal manipulation in the treatment of infantile colic (Olafsdottir et al. 2001, Wiberg et al. 1999 and Klougart et al. 1989). The following research articles are discussed purely for the purpose of illustrating their use of static and motion palpation for the assessment of spinal joint dysfunction in infants. The results of these studies are not discussed.

In a randomized controlled trial conducted by Olafsdottir et al. (2001), the efficacy of spinal manipulation in the management of infantile colic was tested. All infants in the treatment group underwent an assessment for spinal joint dysfunction. The assessment was described as follows, "The participating chiropractor palpated the infant's spinal articulations with respect to areas of dysfunction. Dysfunctional articulations were manipulated or mobilised using light finger-tip pressure." (Olafsdottir et al. 2001). No further information is provided by the authors in terms of the exact techniques used to assess for spinal joint dysfunction. The briefly mentioned palpation was not adequately explained to enable the reproduction of this study and palpation had not been tested for interexaminer reliability. This reduces the inferential validity of this study.

Wiberg et al. (1999) conducted a randomized controlled clinical trial to determine if there was a short-term effect of spinal manipulation in the treatment of infantile colic. All infants in the treatment group receiving spinal manipulation were assessed by means of a routine case history and a physical examination that included motion palpation of the articulations of the spine and pelvis. Motion palpation was used to identify articulations that were restricted in movement. The authors did not specify the exact method of motion palpation used in the study and motion palpation had not been tested for interexaminer reliability. Therefore, this study has questionable validity.

Klougart et al. (1989) conducted a prospective uncontrolled study to determine the effectiveness of spinal manipulative therapy in the treatment of infantile colic. All participating infants were assessed by a chiropractor who conducted a physical examination of the infants. The authors stated that the findings from motion palpation of the articulations of the vertebral column and pelvis were of particular importance. The authors explained that the individual articular segments were moved through their normal range of movement looking for restrictions in their normal mobility. Once again, this explanation is inadequate and the reliability of this method of assessment had not been researched. This reduced the inferential validity of this study.

In a pilot study conducted by Fallon (1997) regarding spinal manipulation and otitis media in children (aged between twenty seven days to five years), static and motion palpation was used to identify spinal joint dysfunction in the participating children. Static palpation was used to assess for tissue texture changes and the sensation of fullness over the joint spaces while motion palpation was used to assess for restrictions in vertebral range of motion (Fallon 1997). No research had been conducted into the reliability of these assessment tools used in infants and so the results of this pilot study are questionable.

It should be noted that the authors (Olafsdottir et al. 2001, Wiberg et al. 1999, Fallon 1997 and Klougart et al. 1989) used static and or motion palpation on infants despite the fact that the reliability of these diagnostic methods had not yet been established. No alternative method of assessment had been tested and the authors may therefore have had no other option. The inferential validity of all these studies is reduced, as the reliability of the assessment tools used in the infants had not been satisfactorily established. Static and motion palpation are essential diagnostic tools employed by manipulative therapists and it is therefore vital to establish the reliability and validity of these tools.

This highlights the gap in the literature and it is therefore, the purpose of this investigation to evaluate the interexaminer reliability of static and motion palpation used to assess for spinal joint dysfunction in infants.

CHAPTER THREE: MATERIALS AND METHODS

2.1 OBJECTIVES OF THE STUDY

3.1.1 Objective One

The first objective was to determine the interexaminer reliability of static palpation for the assessment of spinal joint dysfunction in healthy infants aged two to ten weeks.

3.1.2 Objective Two

The second objective was to determine the interexaminer reliability of motion palpation for the assessment of spinal joint dysfunction in healthy infants aged two to ten weeks.

3.1.3 Objective Three

The third objective was to determine the interexaminer reliability of both static and motion palpation used together for the assessment of spinal joint dysfunction in healthy infants aged two to ten weeks.

3.2 STUDY DESIGN AND PROTOCOL

This interexaminer reliability study was based on a single-facet (one cause of error), repeated-measures design. A single-facet interexaminer reliability design consists of one assessment of all subjects by each of two or more examiners who are blinded to each other's observations (Haas 1995).

The research procedure took place at the participating public and private primary health care clinics in the Greater Durban Area in Kwazulu-Natal (Appendix A). The sample comprised of one hundred (Thomas 2002) healthy infants aged two to ten weeks (Wiberg et al. 1999). According to Haas (1991 - b), the sample size should have been greater than 64 to produce statistically significant results. The sample was selected by means of convenience sampling at the participating clinics.

Healthy infants brought to the clinics by parents for the purpose of routine post-natal assessment or immunisation by the nursing sister at the participating clinics, were considered for participation in this study. Recruitment of infants for this research was by means of posters (Appendix B) and presentations at the various participating clinics to inform parents of this research project. The presentations consisted of one of the examiners explaining to the parents why the research was necessary (i.e. the benefits of the research) and what the research procedure entailed. Parents interested in their infant participating in this research were directly approached by one of the examiners who then further discussed the research procedure with the parent. The first one hundred infants who fulfilled the inclusion and exclusion criteria and whose parents, upon approving of the research study gave their informed consent, were included in the study.

Three examiners assessed each participating infant. Two of the examiners (examiner A and B) were chiropractic interns at the Durban Institute of Technology's (DIT's) Chiropractic Day Clinic while the third examiner (examiner C) was a supervising chiropractic clinician from the DIT's Chiropractic Day Clinic. The examiners did not communicate their assessment findings to each other, as an attempt to eliminate observer bias from the study.

3.2.1 Standard of Acceptance

Infants were accepted onto the study according to the inclusion and exclusion criteria. The nursing sister at the participating clinic ensured the fulfillment of these criteria by means of the infant's history, an examination by the nursing sister (as part of the post-natal assessment) and the infant's records held at that

Primary Health Care Clinic. Only infants whose parents gave their informed consent were included in the study.

3.2.2 Inclusion and Exclusion Criteria

These criteria were based on the randomised controlled trial conducted by Wiberg et al. (1999) into the effectiveness of chiropractic treatment of infantile colic.

3.2.2.1 INCLUSION CRITERIA

- The infant had to be between the ages of 2 to 10 weeks. (Wiberg et al. 1999).
- The infant had to have normal growth as mentioned by the Road to Health Chart (Appendix C).
- The infant had to be alert and healthy with no history of illness and no existing illness (Wiberg et al. 1999).
- Infants who presented with mild conditions (e.g. mild constipation, diarrhoea, rhinitis, infantile colic etc.) were permitted to participate in the study (Merenstein et al. 1994: 163, 584, 575).

3.2.2.2 EXCLUSION CRITERIA

- Infants were not included in the study if they were not accompanied to the clinic by at least one parent.
- Infants were not included in the study if they were brought to the clinic by parents under the age of eighteen.
- Infants with a history of past or present disease were excluded from the study (Wiberg et al. 1999).

3.2.3 Methodology

The infants who fulfilled the inclusion and exclusion criteria were invited to participate in this study. Parents accepting this invitation where given a letter of information (Appendix D) regarding the research study and where required to complete an informed consent form (Appendix E).

The participating infants received a spinal assessment by three independent chiropractic examiners. Two of the examiners were chiropractic interns with three years of supervised clinical experience while the third examiner was a supervising chiropractic clinician. It should be noted that the assessments by the three examiners were conducted in the same room due to space constraints at the participating clinics. The order of the examiners varied according to convenience. The three examiners did not communicate or reveal the outcome of their findings to one another. The assessments of each infant were recorded separately by each of the three examiners.

For the purpose of this study, spinal joint dysfunction was defined as the lack of movement of a joint, caused by muscular spasm or an intra-articular blocking (Gatterman 1990: 408). The assessment included static and passive motion palpation of the spine and sacro-iliac joints, to detect the presence of joint dysfunction (Shafer and Faye 1990: 47; Bergmann et al. 1993: 68, 81 and Anrig and Plaugher 1998: 140). The spinal joint dysfunction was recorded by noting the level of the spine in which it was found. For the purpose of this research, the anatomical side (i.e. left or right) of the spinal joint dysfunction was not considered. The anatomical side was excluded on the basis of trying to reduce the number of variables being tested in this study. Had this not been the case, the number of statistical tests would have doubled. The anatomical side of the spinal joint dysfunction is not always included in reliability studies (Haas 1991 – b).

The procedure of examination is described below. The outcome of the assessment was recorded in appendix F that was drawn from the assessment described below for the purpose of data capture. It should be noted that there were three copies of appendix F, one for each examiner. The author, who was also one of the examiners (examiner A), then collected the completed appendices.

Static Palpation:

The examiners used the distal aspect of the index finger to palpate the entire length of the infants' spine, sacro-iliac (SI) joints and paraspinal musculature. Static palpation was used to establish anatomical landmarks as a baseline for the spinal assessment (Anrig and Plaugher 1998: 140) and to locate any signs of spinal joint dysfunction.

One of the signs of spinal joint dysfunction is tenderness (Anrig and Plaugher 1998: 373, 362, 387). Tenderness can be noted by the infant's reaction to the palpation (Anrig and Plaugher 1998: 387), where the infant may whimper or cry. According to Klougart et al. (1989), this reaction is also seen during the traditional physical examination of the infant.

The infants were placed in the prone position on the examination table for palpation of the spine and sacro-iliac joints (Anrig and Plaugher 1998: 140). The following signs of spinal joint dysfunction were noted:

- Asymmetry, or any unusual alignment of structures (Anrig and Plaugher 1998: 140)
- Muscular spasm or asymmetric paraspinal muscle bulging (Anrig and Plaugher 1998: 140, 332, 344, 362)
- Tenderness over spinous processes, paraspinal musculature, sacro-iliac joints or around the posterior superior iliac spine (Anrig and Plaugher 1998: 373, 362, 387).

Passive Motion Palpation:

The examiners used passive motion palpation to evaluate segmental range of motion and accessory joint motion in the infant's spine and SI joints (Anrig and Plaugher 1998: 148). Any degree of restricted or aberrant movement was recorded as a sign of spinal joint dysfunction (Anrig and Plaugher 1998: 148). The motion palpation procedure used in this study is described below and is in accordance with the technique described by Anrig and Plaugher (1998: 140,148).

Cervical spine: (Anrig and Plaugher 1998: 140)

- The infant was placed supine on the examination table with the parent helping to stabilize the infant by gently supporting the chest.
- The examiner's hands embraced the infant's head whilst the first digits introduced and assessed motion and joint play with bilateral, birotational and flexion/ extension movements.

Thoracic and Lumbar spine: (Anrig and Plaugher 1998: 140, 148).

- The infant was placed prone on the examination table.
- The examiner's indifferent hand stabilized the infant and introduced motion in flexion and extension
- The index finger of the contact hand contacted the interspinous spaces to assess posterior to anterior glide of the spinous processes.
- Anrig and Plaugher (1998: 363, 373) do not recommend motion palpation in rotation and lateral flexion of the thoracic and lumbar spine for young infants.

Sacro-iliac (SI) joints: (Anrig and Plaugher 1998: 148, 387)

- The infant was placed prone on the examination table.
- The index finger of the contact hand contacted the superior medial aspect of the posterior superior iliac spine.
    - The examiner's indifferent hand raised and lowered the leg on the ipsilateral side to assess flexion and extension of the sacro-iliac joint.

3.2.4 Ethical Considerations

This research study involved static and motion palpation of one hundred healthy two to ten week old infants. The research procedure involved no more than minimal risk. The participating infants did not undergo any manipulation or treatment of any sort during the research procedure. The spinal assessments of the infants were conducted under the supervision of a chiropractic clinician.

The participation of the infants in this research study was by invitation and did not involve any coercion. The research procedure was thoroughly explained to interested parents. It was emphasized that participation was voluntary and that refusing to participate would not result in any adverse consequences. Parents were also informed that they were free to withdraw their infant's participation at any stage without fear of negative consequences. It was also mentioned that participation in this study was free. Participation in this study did not involve any financial benefits. Informed consent was gained from the infants' parents before they were allowed to participate in the study.

Parents were also assured that all personal information would remain confidential. Data collected in this study was statistically analyzed and displayed in table or chart form. No personal information was published or revealed. The data collected in this study was stored in the secure environment of the DIT's Chiropractic Day Clinic and will be shredded after a period of five years.

This study was approved by the DIT's Research Ethics Committee.

3.3 MEASUREMENT AND OBSERVATIONS

3.3.1 The Data

This study incorporated both primary and secondary data as explained below.

3.3.1.1 PRIMARY DATA

The primary data was obtained from the outcome of the assessment by the three independent examiners. The assessments were conducted at the participating private and public primary health care clinics (Appendix A). The assessment outcomes were recorded in appendix F where there were separate appendices for each examiner.

Primary clinical data was collected through information obtained via:

- Static Palpation
  - Examiner A's outcome on static palpation of the infants
  - Examiner B's outcome on static palpation of the infants
  - Examiner C's outcome on static palpation of the infants
- Motion Palpation
  - Examiner A's outcome on motion palpation of the infants
  - Examiner B's outcome on motion palpation of the infants
  - Examiner C's outcome on motion palpation of the infants

A point system was used for the purpose of statistical analysis as recommended by Thomas (2002) and suggested by Haas (1991 - b). Each sign (tenderness, spasm, asymmetry and restricted motion) received a score of one (presence of abnormality) or zero (absence of abnormality) (Boline et al. 1993; Keating et al. 1990).

The primary data collected by the examiners was analyzed to determine the interexaminer reliability of static palpation and the interexaminer reliability of motion palpation for the assessment of spinal joint dysfunction in healthy infants aged two to ten weeks. A diagnosis of spinal joint dysfunction was made on static palpation if the infant scored a minimum of two out of three possible points. Spinal joint dysfunction was diagnosed on motion palpation if the infant scored a minimum of one out of one possible points. (Boline et al. 1993, Haas 1991 - b and Keating et al. 1990.)

The data from the static palpation and the motion palpation were then pooled to determine the interexaminer reliability of the combination of static and motion palpation for the assessment of spinal joint dysfunction in infants. For the pooled data, a diagnosis of spinal joint dysfunction was made if the infant scored a minimum of three out of four possible points. (Boline et al. 1993 and Haas 1991 - b.)

3.3.1.2 SECONDARY DATA

The secondary data was obtained from indexed journals, textbooks and the Internet.

3.4 STATISTICAL ANALYSIS

The data was statistically analyzed using the software package SPSS version 9.0 (as supplied by SPSS Inc., Marketing department, 444 North Michigan Avenue, Chicago, Illinois, 60611) and is presented in the form of tables. The level of significance was set at α=0.05 and was used to make decisions regarding the null hypothesis (Chi-Squared) and to test if the values of interexaminer reliability (Kappa) demonstrated statistical significance. Statistical values were determined for each segmental level in terms of their functional units and for each examiner pair. Statistical values were also obtained for collapsed segmental levels for each examiner pair to account for the possibility that examiners may have identified the same dysfunctional segment but named it a different spinal level.

The following tests were used to examine the interexaminer reliability between the examiners for static palpation, motion palpation and static and motion palpation together:

- **Percent Agreement:**

  This is the ratio of the number of agreements between examiners to the total number of comparisons made. Percent agreement does not take the agreement between examiners that is expected to occur due to chance alone into account (Haas 1991 - a). Further testing was therefore necessary.

- **Pearson's Chi-square:**

  This is a measure of association for categorical data. Chi-square was used to test the null hypothesis that states that there is no association between the examination outcomes of the examiners. Chi-square does not evaluate the strength of concordance between examiners nor is it able to distinguish a significant relationship predominated by agreement from one predominated by disagreement. (Haas 1991 - a.)

- **Cohen's Kappa Coefficient:**

  Cohen's Kappa (K) is the statistic of choice for evaluating concordance between two examiners. It is a measure of chance corrected concordance (i.e. it corrects the observed agreement between examiners for agreement that is expected by chance) (Haas 1991 - a).

  K = 1 for perfect agreement between examiners, K = 0 for purely chance agreement and K is negative when chance concordance exceeds the observed agreement. The kappa values were tested for significance at the α=0.05 level. The guidelines used for interpreting the kappa values in this study were in accordance with those employed by French et al. (2000). These are as follows:

  - < 0.00: poor agreement
  - 0.00 – 0.20: slight agreement
  - 0.21 – 0.40: fair agreement
  - 0.41 – 0.60: moderate agreement
  - 0.61 – 0.80: substantial agreement
  - 0.81 – 1.00: almost perfect agreement

  Haas (1991 - a) described a limitation of the kappa statistic where he explained that kappa could become unstable under conditions of limited variation. Limited variation occurs when there is a large proportion of agreement and most of that agreement is limited to only one of the possible rating choices (Haas 1991 - a). Limited variation can make the difference between a study demonstrating poor reliability and excellent reliability. It is therefore necessary to consider the other statistics used in the study, such as percentage agreement (French et al. 2000).

CHAPTER FOUR: THE RESULTS

4.1 INTRODUCTION

The results of the statistical analysis are presented in this chapter. The following statistical tests were used to examine the interexaminer reliability between the examiners for static palpation, motion palpation and static and motion palpation together: Percentage Agreement, Chi-Squared and Kappa Coefficient. The level of significance was set at α=0.05 and was used to make decisions regarding the null hypothesis (Chi-Squared) and to test if the values of interexaminer reliability (Kappa) demonstrated statistical significance.

The above statistics were calculated for each segmental level in terms of their motion segments (i.e. C0/C1, C1/C2, e.t.c.), for static palpation, motion palpation and static and motion palpation together and for each examiner pair. Results were also obtained for collapsed segmental levels to account for the possibility that examiners may have identified the same dysfunctional segment but named it a different spinal level. This extensive statistical analysis is presented in the form of tables. Overall (mean) results were computed to determine the overall interexaminer reliability of the examiner pairs with respect to the segmental levels and the collapsed segments for the purpose of simplicity (Haas 1991 - a).

4.2 PREVALENCE OF SPINAL JOINT DYSFUNCTION AT EACH SPINAL LEVEL

The prevalence of spinal joint dysfunction at each spinal level and at collapsed levels in each region of the spine is presented in Tables 1 to 4. These tables illustrate the number of spinal levels out of the sample of 100 infants that were found to have spinal joint dysfunction by each examiner with static palpation, motion palpation and with the combination of static and motion palpation.

Table 1. Prevalence of Spinal Joint Dysfunction in the Cervical Spine

	Examiner A			Examiner B			Examiner C			Mean		
	Static Palpation	Motion Palpation	Static & Motion Palpation	Static Palpation	Motion Palpation	Static & Motion Palpation	Static Palpation	Motion Palpation	Static & Motion Palpation	Static Palpation	Motion Palpation	Static & Motion Palpation
C0/C1	3	1	2	2	3	2	7	6	4	4.00	3.33	2.67
C1/C2	11	15	11	5	7	5	7	8	5	7.67	10.00	7.00
C2/C3	12	12	9	19	28	17	20	19	16	17.00	19.67	14.00
C3/C4	26	30	26	18	26	18	19	20	14	21.00	25.33	19.33
C4/C5	18	12	15	3	6	3	7	7	5	9.33	8.33	7.67
C5/C6	9	18	8	3	4	2	7	8	6	6.33	10.00	5.33
C6/C7	2	6	2	0	6	0	1	3	1	1.00	5.00	1.00
C7/T1	0	0	0	0	0	0	0	0	0	0.00	0.00	0.00
Total	81	94	73	50	80	47	68	71	51	66.33	81.67	57.00

Table 2. Prevalence of Spinal Joint Dysfunction in the Thoracic Spine

	Examiner A			Examiner B			Examiner C			Mean		
	Static Palpation	Motion Palpation	Static & Motion Palpation	Static Palpation	Motion Palpation	Static & Motion Palpation	Static Palpation	Motion Palpation	Static & Motion Palpation	Static Palpation	Motion Palpation	Static & Motion Palpation
T1/T2	0	0	0	0	0	0	0	1	0	0.00	0.33	0.00
T2/T3	0	0	0	0	0	0	2	2	2	0.67	0.67	0.67
T3/T4	3	6	1	1	10	1	7	6	6	3.67	7.33	2.67
T4/T5	8	24	8	4	13	4	10	7	7	7.33	14.67	6.33
T5/T6	4	10	3	5	11	5	11	12	10	6.67	11.00	6.00
T6/T7	9	14	9	8	13	8	15	16	13	10.67	14.33	10.00
T7/T8	5	7	4	2	6	2	5	10	5	4.00	7.67	3.67
T8/T9	6	4	3	0	5	0	2	5	1	2.67	4.67	1.33
T9/T10	2	5	1	1	2	1	2	3	2	1.67	3.33	1.33
T10/T11	4	5	2	2	6	2	5	4	4	3.67	5.00	2.67
T11/T12	1	2	0	5	8	5	5	4	5	3.67	4.67	3.33
T12/L1	3	2	1	6	12	6	5	3	3	4.67	5.67	3.33
Total	45	79	32	34	86	34	69	73	58	49.33	79.33	41.33

Table 1 and 2 reveals a low prevalence of spinal joint dysfunction at the upper and lower levels of the cervical and thoracic spine. This reduces the usefulness of the Kappa values calculated at these levels due to limited variation (Haas 1991 - a). Limited variation occurs when there is a large proportion of agreement, but most agreement between the examiners is limited to only one of the possible assessment choices (e.g. examiners agree that there is no spinal joint dysfunction at these levels and therefore all assessment choices are 'negative' with no 'positive' choices) (French et al. 2000).

Table 3. Prevalence of Spinal Joint Dysfunction in the Lumbar Spine and Sacroiliac Joints

	Examiner A			**Examiner B**			**Examiner C**			**Mean**		
	Static Palpation	Motion Palpation	Static & Motion Palpation	Static Palpation	Motion Palpation	Static & Motion Palpation	Static Palpation	Motion Palpation	Static & Motion Palpation	Static Palpation	Motion Palpation	Static & Motion Palpation
L1/L2	10	15	10	9	14	9	9	14	8	9.33	14.33	9.00
L2/L3	4	7	4	5	11	5	11	17	10	6.67	11.67	6.33
L3/L4	2	2	2	1	2	1	10	11	8	4.33	5.00	3.67
L4/L5	3	4	3	1	3	1	4	7	3	2.67	4.67	2.33
SI	6	30	6	5	31	5	14	31	14	8.33	30.67	8.33
Total	25	58	25	21	61	21	48	80	43	31.33	66.33	29.67

Table 4. Prevalence of Spinal Joint Dysfunction in the Collapsed Spinal Levels

	Examiner A			Examiner B			Examiner C			Mean		
	Static Palpation	Motion Palpation	Static & Motion Palpation	Static Palpation	Motion Palpation	Static & Motion Palpation	Static Palpation	Motion Palpation	Static & Motion Palpation	Static Palpation	Motion Palpation	Static & Motion Palpation
C0-C3	20	26	17	21	34	19	29	30	23	23.33	30.00	19.67
C3-C5	34	39	32	19	27	19	25	27	18	26.00	31.00	23.00
C5-T1	11	23	10	3	10	2	8	11	7	7.33	14.67	6.33
T1-T4	3	6	1	1	10	1	8	9	8	4.00	8.33	3.33
T4-T7	20	45	19	12	31	12	25	26	21	19.00	34.00	17.33
T7-T10	11	16	8	3	10	3	8	17	8	7.33	14.33	6.33
T10-L1	6	7	3	7	16	7	9	5	7	7.33	9.33	5.67
L1-L3	11	19	11	9	18	9	14	25	13	11.33	20.67	11.00
L3-L5	5	2	5	2	5	2	11	15	9	6.00	7.33	5.33
Total	121	183	106	77	161	74	137	165	114	111.67	169.67	98.00

4.3 OBJECTIVES

The first objective was to determine the interexaminer reliability of static palpation for the assessment of spinal joint dysfunction in healthy infants aged two to ten weeks. The second objective was to determine the interexaminer reliability of motion palpation for the assessment of spinal joint dysfunction in healthy infants aged two to ten weeks. The third objective was to determine the interexaminer reliability of static and motion palpation for the assessment of spinal joint dysfunction in healthy infants aged two to ten weeks.

4.3.1 The Interexaminer Reliability of Static Palpation

Results of the statistical analysis to determine the interexaminer reliability of static palpation are presented in Tables 5 to 16.

4.3.1.1 PERCENTAGE AGREEMENT

This is the ratio of the number of agreements between examiners to the total number of comparisons made.

A. Cervical Spine

Table 5. Percentage Agreement for Static Palpation in the Cervical Spine

PERCENTAGE AGREEMENT				
	Examiner A&B	Examiner A&C	Examiner B&C	MEAN
C0/C1	95	92	91	92.667
C1/C2	88	88	90	88.667
C2/C3	81	82	81	81.333
C3/C4	80	75	73	76.000
C4/C5	83	81	94	86.000
C5/C6	92	94	94	93.333
C6/C7	98	97	99	98.000
C7/T1	100	100	100	100.000
MEAN	89.625	88.625	90.250	89.500

Table 5 presents the percentage agreement between the examiners in the cervical spine. The percentage agreement between examiners A and B ranged from 80% at C3/4 to 100% at C7/T1 while the mean percentage agreement between these examiners across the cervical spine was 89.625%. The percentage agreement between examiners A and C ranged from 75% at C3/4 to 100% at C7/T1 while the mean percentage agreement between these examiners across the cervical spine was 88.625%. The percentage agreement between examiners B and C ranged from 73% at C3/4 to 100% at C7/T1 while the mean percentage agreement between these examiners across the cervical spine was 90.25%. The mean percentage agreement between all 3 examiners ranged from 76% at C3/4 to 100% at C7/T1 while the mean percentage agreement across the cervical spine was 89.5%.

B. Thoracic Spine

Table 6. Percentage Agreement for Static Palpation in the Thoracic Spine

PERCENTAGE AGREEMENT				
	Examiner A&B	Examiner A&C	Examiner B&C	MEAN
T1/T2	100	100	100	100.000
T2/T3	100	98	98	98.667
T3/T4	98	92	94	94.667
T4/T5	90	82	90	87.333
T5/T6	93	85	86	88.000
T6/T7	91	84	87	87.333
T7/T8	97	92	95	94.667
T8/T9	94	92	98	94.667
T9/T10	99	96	97	97.333
T10/T11	96	92	94	94.000
T11/T12	96	94	92	94.000
T12/L1	95	94	91	93.333
MEAN	95.750	91.750	93.500	93.667

Table 6 presents the percentage agreement between the examiners in the thoracic spine. The percentage agreement between examiners A and B ranged from 91% at T6/7 to 100% at T1/T2 while the mean percentage agreement between these examiners across the thoracic spine was 95.75%. The percentage agreement between examiners A and C ranged from 82% at T4/5 to 100% at T1/2, while the mean percentage agreement between these examiners across the thoracic spine was 91.75%. The percentage agreement between examiners B and C ranged from 86% at T5/6 to 100% at T1/2 while the mean percentage agreement between these examiners across the thoracic spine was 93.5%. The mean percentage agreement between all 3 examiners ranged from 87.333% at T4/5 and T6/7 to 100% at T1/2 while the mean percentage agreement across the thoracic spine was 93.667%.

C. Lumbar Spine and Sacroiliac Joints

Table 7. Percentage Agreement for Static Palpation in the Lumbar Spine and Sacroiliac Joints

PERCENTAGE AGREEMENT				
	Examiner A&B	Examiner A&C	Examiner B&C	MEAN
L1/L2	93	89	94	92.000
L2/L3	97	91	92	93.333
L3/L4	97	90	91	92.667
L4/L5	98	95	95	96.000
SI	97	86	87	90.000
MEAN	96.400	90.200	91.800	92.800

Table 7 presents the percentage agreement between the examiners in the lumbar spine and sacroiliac (SI) joints. The percentage agreement between examiners A and B ranged from 93% at L1/2 to 98% at L4/5, while the mean percentage agreement between these examiners across the lumbar spine and SI joints was 96.4%. The percentage agreement between examiners A and C ranged from 86% at the SI joints to 95% at L4/5, while the mean percentage agreement between these examiners across the lumbar spine and SI joints was 90.2%. The percentage agreement between examiners B and C ranged from 87% at the SI joints to 95% at L4/5, while the mean percentage agreement between these examiners across the lumbar spine and SI joints was 91.8%. The mean percentage agreement between all 3 examiners ranged from 90% at the SI joints to 96% at L4/5 while the mean percentage agreement across the lumbar spine and SI joints was 92.8%.

D. Collapsed Spinal Levels

Table 8. Percentage Agreement for Static Palpation in the Collapsed Spinal Levels

PERCENTAGE AGREEMENT				
	Examiner A&B	Examiner A&C	Examiner B&C	MEAN
C0-C3	81	77	74	77.333
C3-C5	77	75	74	75.333
C5-T1	90	91	93	91.333
T1-T4	98	91	93	94.000
T4-T7	80	71	79	76.667
T7-T10	92	85	93	90.000
T10-L1	93	89	86	89.333
L1-L3	92	85	91	89.333
L3-L5	95	88	89	90.667
MEAN	88.667	83.556	85.778	86.000

Table 8 presents the percentage agreement between the examiners in the collapsed spinal levels. The percentage agreement between examiners A and B ranged from 77% at C3-5 to 98% at T1-4, while the mean percentage agreement between these examiners across all the collapsed segments of the spine was 88.667%. The percentage agreement between examiners A and C ranged from 71% at T4-7 to 91% at C5-T1 and T1-4, while the mean percentage agreement between these examiners across all the collapsed segments of the spine was 83.556%. The percentage agreement between examiners B and C ranged from 74% at C0-3 and C3-5 to 93% at C5-T1, T1-4 and T7-10, while the mean percentage agreement between these examiners across all the collapsed segments of the spine was 85.778%. The mean percentage agreement between all 3 examiners ranged from 75.333% at C3-5 to 94% at T1-4 while the mean percentage agreement between these examiners across all the collapsed segments of the spine was 86%.

4.3.1.2 PEARSON'S CHI-SQUARE

The Chi-Squared test was used to determine whether there was a significant relationship between the outcomes of each examiner (1993 Fisher and van Belle: 219). If the p-value calculated for Chi-Square was less than 0.05, then the null hypothesis was rejected and so we could conclude that the examination outcomes of each examiner were significantly associated. If the p-value calculated for Chi-Square was greater than 0.05, then the null hypothesis was failed to be rejected and so we could conclude that the examination outcomes of each examiner were not significantly associated.

The Null Hypothesis states that there is no association between the examination outcomes of the examiners.

The Alternative Hypothesis states that there is an association between the examination outcomes of the examiners.

Decision Rule:

If $p \geq 0.05$ (level of significance), we fail to reject the null hypothesis.

If $p < 0.05$ (level of significance), we reject the null hypothesis and so accept the alternative hypothesis.

A. Cervical Spine

Table 9. Chi-Square Values for Static Palpation in the Cervical Spine

	Examiner A&B		Examiner A&C		Examiner B&C		MEAN	
	Chi-Square	**p-value**	**Chi-Square**	**p-value**	**Chi-Square**	**p-value**	**Chi-Square**	**p-value**
C0/C1	0.063	0.802	3.294	0.07	0.154	0.695	1.170	0.522
C1/C2	4.521	0.033	7.803	0.005	1.366	0.242	4.563	0.093
C2/C3	8.515	0.004	12.524	0	15.611	0	12.217	0.001
C3/C4	18.868	0	8.647	0.003	1.099	0.294	9.538	0.099
C4/C5	4.963	0.026	3.151	0.076	16.913	0	8.342	0.034
C5/C6	12.558	0	35.818	0	16.913	0	21.763	0.000
C6/C7	(a)	(a)	0.021	0.886	(a)	(a)	0.021	0.886
C7/T1	(a)	(a)	(a)	(a)	(a)	(a)	(a)	(a)
MEAN	8.248	0.144	10.180	0.149	8.676	0.205	9.035	0.166

(a) Values could not be calculated because one or both of the examiners did not have any positive assessments at this level (i.e. no spinal joint dysfunction at this spinal level).

Table 9 presents the Pearson's Chi-Square values and their p-values calculated for the cervical spine. The p-value calculated for examiners A and B was greater than 0.05 at C0/1; we therefore fail to reject the null hypothesis at the 0.05 level of significance and so conclude that there is no association between the examiners at this spinal level for static palpation. The p-values calculated for examiners A and B were less than 0.05 at C1/2, C2/3, C3/4, C4/5 and C5/6; we therefore reject the null hypothesis at the 0.05 level of significance and so conclude that there is an association between the examiners at these spinal levels for static palpation. Values could not be calculated at C7/T1 because one or both of the examiners did not have any positive assessments at this level (i.e. no spinal joint dysfunction at this spinal level).

The p-values calculated for examiners A and C were greater than 0.05 at C0/1, C4/5 and C6/7; we therefore fail to reject the null hypothesis at the 0.05 level of significance and so conclude that there is no association between the examiners at these spinal levels for static palpation. The p-values calculated for examiners A and C were less than 0.05 at C1/2, C2/3, C3/4 and C5/6; we therefore reject the null hypothesis at the 0.05 level of significance and so conclude that there is an association between the examiners at these spinal levels for static palpation. Values could not be calculated at C7/T1 because one or both of the examiners did not have any positive assessments at this level (i.e. no spinal joint dysfunction at this spinal level).

The p-values calculated for examiners B and C were greater than 0.05 at C0/1, C1/2 and C3/4; we therefore fail to reject the null hypothesis at the 0.05 level of significance and so conclude that there is no association between the examiners at these spinal levels for static palpation. The p-values calculated for examiners B and C were less than 0.05 at C2/3, C4/5 and C5/6; we therefore reject the null hypothesis at the 0.05 level of significance and so conclude that there is an association between the examiners at these spinal levels for static palpation. Values could not be calculated at C6/7 and C7/T1 because one or both of the examiners did not have any positive assessments at these levels (i.e. no spinal joint dysfunction at these spinal levels).

The mean p-values calculated between all 3 examiners were greater than 0.05 at C0/1, C1/2 C3/4 and C6/7; we therefore fail to reject the null hypothesis at the 0.05 level of significance and so conclude that there is no association between the examiners at these spinal levels for static palpation. The mean p-values calculated between all 3 examiners were less than 0.05 at C2/3, C4/5 and C5/6; we therefore reject the null hypothesis at the 0.05 level of significance and so conclude that there is an association between the examiners at these spinal levels for static palpation. Values could not be calculated at C7/T1 because one or all of the examiners did not have any positive assessments at this level (i.e. no spinal joint dysfunction at this spinal level).

B. Thoracic Spine

Table 10. Chi-Square Values for Static Palpation in the Thoracic Spine

	Examiner A&B		Examiner A&C		Examiner B&C		MEAN	
	Chi-Square	**p-value**	**Chi-Square**	**p-value**	**Chi-Square**	**p-value**	**Chi-Square**	**p-value**
T1/T2	(a)	(a)	(a)	(a)	(a)	(a)	(a)	(a)
T2/T3	(a)	(a)	(a)	(a)	(a)	(a)	(a)	(a)
T3/T4	32.66	0	3.294	0.07	13.42	0	16.458	0.023
T4/T5	1.636	0.201	0.966	0.326	7.407	0.006	3.336	0.178
T5/T6	3.509	0.061	0.515	0.473	0.435	0.509	1.486	0.348
T6/T7	17.848	0	6.725	0.01	15.388	0	13.320	0.003
T7/T8	38.776	0	2.493	0.114	8.7	0.003	16.656	0.039
T8/T9	(a)	(a)	0.13	0.718	(a)	(a)	0.130	0.718
T9/T10	49.495	0	0.042	0.838	0.021	0.886	16.519	0.575
T10/T11	11.246	0.001	0.174	0.677	0.085	0.771	3.835	0.483
T11/T12	19.192	0	0.053	0.818	2.493	0.114	7.246	0.311
T12/L1	20.182	0	5.227	0.022	1.892	0.176	9.100	0.066
MEAN	21.616	0.029	1.962	0.407	5.538	0.274	9.705	0.237

(a) Values could not be calculated because one or both of the examiners did not have any positive assessments at these levels (i.e. no spinal joint dysfunction at these spinal levels).

Table 10 presents the Pearson's Chi-Square values and their p-values calculated for the thoracic spine. The p-values calculated for examiners A and B were greater than 0.05 at T4/5 and T5/6; we therefore fail to reject the null hypothesis at the 0.05 level of significance and so conclude that there is no association between the examiners at these spinal levels for static palpation. The p-values calculated for examiners A and B were less than 0.05 at T3/4, T6/7, T7/8, T9/10, T10/11, T11/12 and T12/L1; we therefore reject the null hypothesis at the 0.05 level of significance and so conclude that there is an association between the examiners at these spinal levels for static palpation. Values could not be calculated at T1/2, T2/3 and T8/9 because one or both of the examiners did not have any positive assessments at these levels (i.e. no spinal joint dysfunction at these spinal levels).

The p-values calculated for examiners A and C were greater than 0.05 at T3/4, T4/5, T5/6, T7/8, T8/9, T9/10, T10/11 and T11/12; we therefore fail to reject the null hypothesis at the 0.05 level of significance and so conclude that there is no association between the examiners at these spinal levels for static palpation. The p-values calculated for examiners A and C were less than 0.05 at T6/7 and T12/L1; we therefore reject the null hypothesis at the 0.05 level of significance and so conclude that there is an association between the examiners at these spinal levels for static palpation. Values could not be calculated at T1/2 and T2/3 because one or both of the examiners did not have any positive assessments at these levels (i.e. no spinal joint dysfunction at these spinal levels).

The p-values calculated for examiners B and C were greater than 0.05 at T5/6, T9/10, T10/11, T11/12 and T12/L1; we therefore fail to reject the null hypothesis at the 0.05 level of significance and so conclude that there is no association between the examiners at these spinal levels for static palpation. The p-values calculated for examiners B and C were less than 0.05 at T3/4, T4/5, T6/7 and T7/8; we therefore reject the null hypothesis at the 0.05 level of significance and so conclude that there is an association between the examiners at these spinal levels for static palpation. Values could not be calculated at T1/2, T2/3 and T8/9 because one or both of the examiners did not have any positive assessments at these levels (i.e. no spinal joint dysfunction at these spinal levels).

The mean p-values calculated between all 3 examiners were greater than 0.05 at T4/5, T5/6, T8/9, T9/10, T10/11, T11/12 and T12/L1; we therefore fail to reject the null hypothesis at the 0.05 level of significance and so conclude that there is no association between the examiners at these spinal levels for static palpation. The mean p-values calculated between all 3 examiners were less than 0.05 at T3/4, T6/7 and T7/8; we therefore reject the null hypothesis at the 0.05 level of significance and so conclude that there is an association between the examiners at these spinal levels for static palpation. Values could not be calculated at T1/2 and T2/3 because one or all of the examiners did not have any positive assessments at these levels (i.e. no spinal joint dysfunction at these spinal levels).

C. Lumbar Spine and SI Joints

Table 11. Chi-Square Values for Static Palpation in the Lumbar Spine and Sacroiliac Joints

	Examiner A&B		Examiner A&C		Examiner B&C		MEAN	
	Chi-Square	**p-value**	**Chi-Square**	**p-value**	**Chi-Square**	**p-value**	**Chi-Square**	**p-value**
L1/L2	35.287	0	13.038	0	40.158	0	29.494	0.000
L2/L3	42.982	0	17.433	0	25.595	0	28.670	0.000
L3/L4	0.021	0.886	3.628	0.057	9.091	0.003	4.247	0.315
L4/L5	32.66	0	6.93	0.008	0.042	0.837	13.211	0.282
SI	51.101	0	6.871	0.009	9.25	0.002	22.407	0.004
MEAN	32.410	0.177	9.580	0.015	16.827	0.168	19.606	0.120

Table 11 presents the Pearson's Chi-Square values and their p-values calculated for the lumbar spine and sacroiliac (SI) joints. The p-value calculated for examiners A and B was greater than 0.05 at L3/4; we therefore fail to reject the null hypothesis at the 0.05 level of significance and so conclude that there is no association between the examiners at this spinal level for static palpation. The p-values calculated for examiners A and B were less than 0.05 at L1/2, L2/3, L4/5 and the SI joints; we therefore reject the null hypothesis at the 0.05 level of significance and so conclude that there is an association between the examiners at these levels of the lumbar spine and the SI joints for static palpation.

The p-value calculated for examiners A and C was greater than 0.05 at L3/4; we therefore fail to reject the null hypothesis at the 0.05 level of significance and so conclude that there is no association between the examiners at this spinal level for static palpation. The p-values calculated for examiners A and C were less than 0.05 at L1/2, L2/3, L4/5 and the SI joints; we therefore reject the null hypothesis at the 0.05 level of significance and so conclude that there is an association between the examiners at these spinal levels for static palpation.

The p-value calculated for examiners B and C was greater than 0.05 at L4/5; we therefore fail to reject the null hypothesis at the 0.05 level of significance and so conclude that there is no association between the examiners at this spinal level for static palpation. The p-values calculated for examiners B and C were less than 0.05 at L1/2, L2/3, L3/4 and the SI joints; we therefore reject the null hypothesis at the 0.05 level of significance and so conclude that there is an association between the examiners at these spinal levels for static palpation.

The mean p-values calculated between all 3 examiners were greater than 0.05 at L3/4 and L4/5; we therefore fail to reject the null hypothesis at the 0.05 level of significance and so conclude that there is no association between the examiners at these spinal levels for static palpation. The mean p-values calculated between all 3 examiners were less than 0.05 at L1/2, L2/3 and the SI joints; we therefore reject the null hypothesis at the 0.05 level of significance and so conclude that there is an association between the examiners at these spinal levels for static palpation.

D. Collapsed Spinal Levels

Table 12. Chi-Square Values for Static Palpation in the Collapsed Spinal Levels

	Examiner A&B		Examiner A&C		Examiner B&C		MEAN	
	Chi-Square	**p-value**	**Chi-Square**	**p-value**	**Chi-Square**	**p-value**	**Chi-Square**	**p-value**
C0-C3	17.42	0	15.736	0	10.225	0.001	14.460	0.000
C3-C5	21.118	0	17.172	0	6.259	0.012	14.850	0.004
C5-T1	9.789	0.002	23.558	0	14.463	0	15.937	0.001
T1-T4	32.66	0	2.697	0.101	11.616	0.001	15.658	0.034
T4-T7	7.67	0.006	3	0.083	12.626	0	7.765	0.030
T7-T10	25.023	0	1.741	0.187	14.463	0	13.742	0.062
T10-L1	18.129	0	4.615	0.032	0.257	0.612	7.667	0.215
L1-L3	31.305	0	10.156	0.001	33.413	0	24.958	0.000
L3-L5	8.7	0.003	4.521	0.033	3.171	0.075	5.464	0.037
MEAN	19.090	0.001	9.244	0.049	11.833	0.078	13.389	0.043

Table 12 presents the Pearson's Chi-Square values and their p-values calculated for the collapsed spinal levels. The p-values calculated for examiners A and B were less than 0.05 at all the collapsed spinal levels; we therefore reject the null hypothesis at the 0.05 level of significance and so conclude that there is an association between the examiners at all the collapsed spinal levels for static palpation.

The p-values calculated for examiners A and C were greater than 0.05 at T1-4, T4-7 and T7-10; we therefore fail to reject the null hypothesis at the 0.05 level of significance and so conclude that there is no association between the examiners at these collapsed spinal levels for static palpation. The p-values calculated for examiners A and C were less than 0.05 at C0-3, C3-5, C5-T1, T10-L1, L1-3 and L3-5; we therefore reject the null hypothesis at the 0.05 level of significance and so conclude that there is an association between the examiners at these collapsed spinal levels for static palpation.

The p-values calculated for examiners B and C were greater than 0.05 at T10-L1 and L3-5; we therefore fail to reject the null hypothesis at the 0.05 level of significance and so conclude that there is no association between the examiners at these collapsed spinal levels for static palpation. The p-values calculated for examiners B and C were less than 0.05 at C0-3, C3-5, C5-T1, T1-4, T4-7, T7-10 and L1-3; we therefore reject the null hypothesis at the 0.05 level of significance and so conclude that there is an association between the examiners at these collapsed spinal levels for static palpation.

The mean p-values calculated between all 3 examiners were greater than 0.05 at T7-10 and T10-L1; we therefore fail to reject the null hypothesis at the 0.05 level of significance and so conclude that there is no association between the examiners at these collapsed spinal levels for static palpation. The mean p-values calculated between all 3 examiners were less than 0.05 at C0-3, C3-5, C5-T1, T1-4, T4-7, L1-3 and L3-5; we therefore reject the null hypothesis at the 0.05 level of significance and so conclude that there is an association between the examiners at these collapsed spinal levels for static palpation.

4.3.1.3 COHEN'S KAPPA COEFFICIENT

Cohen's Kappa (K) was used to measure the chance corrected concordance between the examiners. The kappa values were tested for significance at the α=0.05 level. The guidelines used for interpreting the kappa values in this study were in accordance with those employed by French et al. (2000), i.e.:

- < 0.00: poor agreement
- 0.00 – 0.20: slight agreement
- 0.21 – 0.40: fair agreement
- 0.41 – 0.60: moderate agreement
- 0.61 – 0.80: substantial agreement
- 0.81 – 1.00: almost perfect agreement

A. Cervical Spine

Table 13. Kappa Values for Static Palpation in the Cervical Spine

	Examiner A&B		Examiner A&C		Examiner B&C		MEAN	
	Kappa	**p-value**	**Kappa**	**p-value**	**Kappa**	**p-value**	**Kappa**	**p-value**
C0/C1	-0.025	0.802	0.165	0.07	-0.032	0.695	0.036	0.522
C1/C2	0.195	0.033	0.271	0.005	0.115	0.242	0.194	0.093
C2/C3	0.281	0.004	0.338	0	0.395	0	0.338	0.001
C3/C4	0.423	0	0.288	0.003	0.105	0.294	0.272	0.099
C4/C5	0.147	0.026	0.155	0.076	0.374	0	0.225	0.034
C5/C6	0.302	0	0.593	0	0.374	0	0.423	0.000
C6/C7	(a)	(a)	-0.014	0.886	(a)	(a)	-0.014	0.886
C7/T1	(a)	(a)	(a)	(a)	(a)	(a)	(a)	(a)
MEAN	0.221	0.144	0.257	0.149	0.222	0.205	0.233	0.166

(a) Values could not be calculated because one or both of the examiners did not have any positive assessments at these levels (i.e. no spinal joint dysfunction at these spinal levels).

Table 13 presents the Cohen's Kappa (K) scores and their p-values calculated for the cervical spine. Kappa scores calculated for examiners A and B in the cervical spine ranged from –0.025 (at C0/1) to 0.423 (at C3/4) which indicates a range of poor to moderate agreement between the examiners beyond chance. The overall (mean) Kappa score calculated for the cervical spine was 0.221, which indicates fair agreement beyond chance between these examiners for static palpation in the cervical spine. Kappa scores were statistically significant ($p<0.05$) at all spinal levels in the cervical spine besides C0/1. Kappa values could not be calculated at C6/7 and C7/T1 because one or both of the examiners did not have any positive assessments at these levels (i.e. no spinal joint dysfunction at these spinal levels).

Kappa scores calculated for examiners A and C in the cervical spine ranged from –0.014 (at C6/7) to 0.593 (at C5/6) which indicates a range of poor to moderate agreement between the examiners beyond chance. The overall (mean) Kappa score calculated for the cervical spine was 0.257, which indicates fair agreement beyond chance between these examiners for static palpation in the cervical spine. Kappa scores were statistically significant ($p<0.05$) at all spinal levels in the cervical spine besides C0/1, C4/5 and C6/7. Values could not be calculated at C7/T1 because one or both of the examiners did not have any positive assessments at these levels (i.e. no spinal joint dysfunction at these spinal levels).

Kappa scores calculated for examiners B and C in the cervical spine ranged from –0.032 (at C0/1) to 0.395 (at C2/3), which indicates a range of poor to fair agreement between the examiners beyond chance. The overall (mean) Kappa score calculated for the cervical spine was 0.222, which indicates fair agreement beyond chance between these examiners for static palpation in the cervical spine. Kappa scores were statistically significant ($p<0.05$) at all spinal levels in the cervical spine besides C0/1, C1/2 and C3/4. Kappa values could not be calculated at C6/7 and C7/T1 because one or both of the examiners did not have any positive assessments at these levels (i.e. no spinal joint dysfunction at these spinal levels).

Mean Kappa scores calculated for all examiners in the cervical spine ranged from –0.014 (at C6/7) to 0.423 (at C5/6) which indicates a range of poor to moderate agreement between the examiners beyond chance. The overall (mean) Kappa score calculated for all examiners across the cervical spine was 0.233, which indicates fair agreement beyond chance between the examiners for static palpation. Kappa scores were statistically significant ($p<0.05$) at all spinal levels in the cervical spine besides C0/1, C1/2, C3/4 and C6/7. Mean Kappa values could not be calculated at C7/T1 because one or both of the examiners did not have any positive assessments at these levels (i.e. no spinal joint dysfunction at these spinal levels).

B. Thoracic Spine

Table 14. Kappa Values for Static Palpation in the Thoracic Spine

	Examiner A&B		Examiner A&C		Examiner B&C		MEAN	
	Kappa	**p-value**	**Kappa**	**p-value**	**Kappa**	**p-value**	**Kappa**	**p-value**
T1/T2	(a)	(a)	(a)	(a)	(a)	(a)	(a)	(a)
T2/T3	(a)	(a)	(a)	(a)	(a)	(a)	(a)	(a)
T3/T4	0.492	0	0.165	0.07	0.237	0	0.298	0.023
T4/T5	0.12	0.201	-0.098	0.326	0.242	0.006	0.088	0.178
T5/T6	0.186	0.061	-0.062	0.473	0.06	0.509	0.061	0.348
T6/T7	0.422	0	0.249	0.01	0.369	0	0.347	0.003
T7/T8	0.559	0	0.158	0.114	0.265	0.003	0.327	0.039
T8/T9	(a)	(a)	-0.031	0.718	(a)	(a)	-0.031	0.718
T9/T10	0.662	0	-0.02	0.838	-0.014	0.886	0.209	0.575
T10/T11	0.315	0.001	-0.042	0.677	-0.027	0.771	0.082	0.483
T11/T12	0.332	0	-0.017	0.818	0.158	0.114	0.158	0.311
T12/L1	0.421	0	0.221	0.022	0.135	0.176	0.259	0.066
MEAN	0.390	0.029	0.052	0.407	0.158	0.274	0.200	0.237

(a) Values could not be calculated because one or both of the examiners did not have any positive assessments at these levels (i.e. no spinal joint dysfunction at these spinal levels).

Table 14 presents the Cohen's Kappa (K) scores and their p-values calculated for the thoracic spine. Kappa scores calculated for examiners A and B ranged from 0.12 (at T4/5) to 0.662 (at T9/10), which indicates a range of slight to substantial agreement between the examiners beyond chance. The overall (mean) Kappa score calculated here was 0.39, which indicates fair agreement beyond chance between these examiners for static palpation in the thoracic spine. Kappa scores were statistically significant ($p<0.05$) at all spinal levels in the thoracic spine besides T4/5 and T5/6. Kappa values could not be calculated at T1/2, T2/3 and T8/9 because one or both of the examiners did not have any positive assessments at these levels (i.e. no spinal joint dysfunction at these spinal levels).

Kappa scores calculated for examiners A and C ranged from –0.098 (at T4/5) to 0.249 (at T6/7), which indicates a range of poor to fair agreement between the examiners beyond chance. The overall (mean) Kappa score calculated here was 0.052, which indicates slight agreement beyond chance between these examiners for static palpation in the thoracic spine. Kappa scores were statistically significant ($p<0.05$) at T1/2, T2/3, T6/7 and T12/L1 while the Kappa scores at all other levels in the thoracic spine were not statistically significant ($p>0.05$). Kappa values could not be calculated at T1/2 and T2/3 because one or both of the examiners did not have any positive assessments at these levels (i.e. no spinal joint dysfunction at these spinal levels).

Kappa scores calculated for examiners B and C ranged from –0.027 (at T10/11) to 0.369 (at T6/7), which indicates a range of poor to fair agreement between the examiners beyond chance. The overall (mean) Kappa score calculated here was 0.158, which indicates slight agreement beyond chance between these examiners for static palpation in the thoracic spine. Kappa scores were statistically significant ($p<0.05$) at all spinal levels in the thoracic spine besides T5/6, T9/10, T10/11, T11/12 and T12/L1. Values could not be calculated at T1/2, T2/3 and T8/9 because one or both of the examiners did not have any positive assessments at these levels (i.e. no spinal joint dysfunction at these spinal levels).

Mean Kappa scores calculated for all examiners in the thoracic spine ranged from –0.031 (at T8/9) to 0.347 (at T6/7) which indicates a range of poor to fair agreement between the examiners beyond chance. The overall (mean) Kappa score calculated for all examiners across the thoracic spine was 0.2, which indicates slight agreement beyond chance between the examiners for static palpation. Overall Kappa scores were statistically significant ($p<0.05$) at T1/2, T2/3, T3/4, T6/7 and T7/8 while the overall Kappa scores at all other levels in the thoracic spine were not statistically significant ($p>0.05$). Kappa values could not be calculated at T1/2 and T2/3 because one or both of the examiners did not have any positive assessments at these levels (i.e. no spinal joint dysfunction at these spinal levels).

C. Lumbar Spine and SI Joints

Table 15. Kappa Values for Static Palpation in the Lumbar Spine and Sacroiliac Joints

	Examiner A&B		Examiner A&C		Examiner B&C		MEAN	
	Kappa	**p-value**	**Kappa**	**p-value**	**Kappa**	**p-value**	**Kappa**	**p-value**
L1/L2	0.593	0	0.36	0	0.634	0	0.529	0.000
L2/L3	0.651	0	0.363	0	0.463	0	0.492	0.000
L3/L4	-0.014	0.886	0.138	0.057	0.167	0.003	0.097	0.315
L4/L5	0.492	0	0.26	0.008	-0.016	0.837	0.245	0.282
SI	0.712	0	0.236	0.009	0.261	0.002	0.403	0.004
MEAN	0.487	0.177	0.271	0.015	0.302	0.168	0.353	0.120

Table 15 presents the Cohen's Kappa (K) scores and their p-values calculated for the lumbar spine and sacroiliac (SI) joints. Kappa scores calculated for examiners A and B ranged from –0.014 (at L3/4) to 0.712 (at SI joints), which indicates a range of poor to substantial agreement between the examiners beyond chance. The overall (mean) Kappa score calculated here was 0.487, which indicates moderate agreement beyond chance between these examiners for static palpation in the lumbar spine and SI joints. Kappa scores were statistically significant ($p<0.05$) at all spinal levels in the lumbar spine and SI joints besides L3/4.

Kappa scores calculated for examiners A and C ranged from 0.138 (at L3/4) to 0.363 (at L2/3), which indicates a range of slight to fair agreement between the examiners beyond chance. The overall (mean) Kappa score calculated here was 0.271, which indicates fair agreement beyond chance between these examiners for static palpation in the lumbar spine and SI joints. Kappa scores were statistically significant ($p<0.05$) at all spinal levels in the lumbar spine and SI joints besides L3/4.

Kappa scores calculated for examiners B and C ranged from –0.016 (at L4/5) to 0.634 (at L1/2), which indicates a range of poor to substantial agreement between the examiners beyond chance. The overall (mean) Kappa score calculated here was 0.302, which indicates fair agreement beyond chance between these examiners for static palpation in the lumbar spine and SI joints. Kappa scores were statistically significant ($p<0.05$) at all spinal levels in the lumbar spine and SI joints besides L4/5.

Mean Kappa scores calculated for all examiners in the lumbar spine and SI joints ranged from 0.097 (at L3/4) to 0.529 (at L1/2), which indicates a range of slight to moderate agreement between the examiners beyond chance. The overall (mean) Kappa score calculated for all examiners across the lumbar spine and SI joints was 0.353, which indicates fair agreement beyond chance between the examiners for static palpation. Overall Kappa scores were statistically significant ($p<0.05$) at all levels in the lumbar spine and SI joints besides, L3/4 and L4/5.

D. Collapsed Spinal Levels

Table 16. Kappa Values for Static Palpation in the Collapsed Spinal Levels

	Examiner A&B		Examiner A&C		Examiner B&C		MEAN	
	Kappa	**p-value**	**Kappa**	**p-value**	**Kappa**	**p-value**	**Kappa**	**p-value**
C0-C3	0.417	0	0.385	0	0.313	0.001	0.372	0.000
C3-C5	0.426	0	0.405	0	0.246	0.012	0.359	0.004
C5-T1	0.25	0.002	0.478	0	0.335	0	0.354	0.001
T1-T4	0.492	0	0.144	0.101	0.208	0.001	0.281	0.034
T4-T7	0.265	0.006	0.171	0.083	0.323	0	0.253	0.030
T7-T10	0.4	0	0.13	0.187	0.335	0	0.288	0.062
T10-L1	0.424	0	0.21	0.032	0.05	0.612	0.228	0.215
L1-L3	0.556	0	0.316	0.001	0.561	0	0.478	0.000
L3-L5	0.265	0.003	0.195	0.033	0.124	0.075	0.195	0.037
MEAN	0.388	0.001	0.270	0.049	0.277	0.078	0.312	0.043

Table 16 presents the Cohen's Kappa (K) scores and their p-values calculated for the collapsed spinal levels. Kappa scores calculated for examiners A and B ranged from 0.25 (at C5-T1) to 0.556 (at L1-3), which indicates a range of fair to moderate agreement between the examiners beyond chance. The overall (mean) Kappa score calculated here was 0.388, which indicates fair agreement beyond chance between these examiners for static palpation in the collapsed spinal levels. Kappa scores were statistically significant ($p<0.05$) at all the collapsed spinal levels.

Kappa scores calculated for examiners A and C ranged from 0.13 (at T7-10) to 0.478 (at C5-T1), which indicates a range of slight to moderate agreement between the examiners beyond chance. The overall (mean) Kappa score calculated here was 0.270, which indicates fair agreement beyond chance between these examiners for static palpation in the collapsed spinal levels. Kappa scores were statistically significant ($p<0.05$) at all collapsed spinal levels besides T1-4, T4-7 and T7-10.

Kappa scores calculated for examiners B and C ranged from 0.05 (at T10-L1) to 0.561 (at L1-3), which indicates a range of slight to moderate agreement between the examiners beyond chance. The overall (mean) Kappa score calculated here was 0.277, which indicates fair agreement beyond chance between these examiners for static palpation in the collapsed spinal levels. Kappa scores were statistically significant ($p<0.05$) at all collapsed spinal levels besides T10-L1 and L3-5.

Mean Kappa scores calculated for all examiners in the collapsed spinal levels ranged from 0.195 (at L3-5) to 0.478 (at L1-3), which indicates a range of slight to moderate agreement between the examiners beyond chance. The overall (mean) Kappa score calculated for all examiners across the collapsed spinal levels was 0.312, which indicates fair agreement beyond chance between the examiners. Overall Kappa scores were statistically significant ($p<0.05$) at all collapsed spinal levels besides T7-10 and T10-L1.

4.3.2 The Interexaminer Reliability of Motion Palpation

Results of the statistical analysis to determine the interexaminer reliability of motion palpation are presented in Tables 17 to 28.

4.3.2.1 PERCENTAGE AGREEMENT

This is the ratio of the number of agreements between examiners to the total number of comparisons made.

A. Cervical Spine

Table 17. Percentage Agreement for Motion Palpation in the Cervical Spine

	Examiner A&B	Examiner A&C	Examiner B&C	MEAN
C0/C1	96	95	93	94.667
C1/C2	86	85	87	86.000
C2/C3	76	77	77	76.667
C3/C4	76	72	72	73.333
C4/C5	86	85	91	87.333
C5/C6	82	82	94	86.000
C6/C7	92	91	93	92.000
C7/T1	100	100	100	100.000
MEAN	86.750	85.875	88.375	87.000

Table 17 presents the percentage agreement between the examiners for motion palpation in the cervical spine. The percentage agreement between examiners A and B ranged from 76% at C2/3 and C3/4 to 100% at C7/T1, while the mean percentage agreement between these examiners across the cervical spine was 86.75%. The percentage agreement between examiners A and C ranged from 72% at C3/4 to 100% at C7/T1 while the mean percentage agreement between these examiners across the cervical spine was 85.875%. The percentage agreement between examiners B and C ranged from 72% at C3/4 to 100% at C7/T1 while the mean percentage agreement between these examiners across the cervical spine was 88.375%. The mean percentage agreement between all 3 examiners ranged from 73.333% at C3/4 to 100% at C7/T1 while the mean percentage agreement for all examiners across the cervical spine was 87%.

B. Thoracic Spine

Table 18. Percentage Agreement for Motion Palpation in the Thoracic Spine

	Examiner A&B	Examiner A&C	Examiner B&C	MEAN
T1/T2	100	99	99	99.333
T2/T3	100	98	98	98.667
T3/T4	86	88	88	87.333
T4/T5	77	75	86	79.333
T5/T6	87	80	81	82.667
T6/T7	87	82	85	84.667
T7/T8	91	85	90	88.667
T8/T9	91	93	90	91.333
T9/T10	95	92	95	94.000
T10/T11	95	92	91	92.667
T11/T12	92	94	90	92.000
T12/L1	90	95	87	90.667
MEAN	90.917	89.417	90.000	90.111

Table 18 presents the percentage agreement between the examiners for motion palpation in the thoracic spine. The percentage agreement between examiners A and B ranged from 77% at T4/5 to 100% at T1/T2 and T2/3, while the mean percentage agreement between these examiners across the thoracic spine was 90.917%. The percentage agreement between examiners A and C ranged from 75% at T4/5 to 99% at T1/2, while the mean percentage agreement between these examiners across the thoracic spine was 89.417%. The percentage agreement between examiners B and C ranged from 81% at T5/6 to 99% at T1/2, while the mean percentage agreement between these examiners across the thoracic spine was 90%. The mean percentage agreement between all 3 examiners ranged from 79.333% at T4/5 to 99.333% at T1/2 while the mean percentage agreement for all examiners across the thoracic spine for motion palpation was 90.111%.

C. Lumbar Spine and Sacroiliac Joints

Table 19. Percentage Agreement for Motion Palpation in the Lumbar Spine and Sacroiliac Joints

	Examiner A&B	Examiner A&C	Examiner B&C	MEAN
L1/L2	90	87	87	88.000
L2/L3	92	86	88	88.667
L3/L4	96	87	89	90.667
L4/L5	95	91	90	92.000
SI	79	72	77	76.000
MEAN	90.400	84.600	86.200	87.067

Table 19 presents the percentage agreement for motion palpation between the examiners in the lumbar spine and sacroiliac (SI) joints. The percentage agreement between examiners A and B ranged from 79% at the SI joints to 96% at L3/4, while the mean percentage agreement between these examiners across the lumbar spine and SI joints was 90.4%. The percentage agreement between examiners A and C ranged from 72% at the SI joints to 91% at L4/5, while the mean percentage agreement between these examiners across the lumbar spine and SI joints was 84.6%. The percentage agreement between examiners B and C ranged from 77% at the SI joints to 90% at L4/5, while the mean percentage agreement between these examiners across the lumbar spine and SI joints was 86.2%. The mean percentage agreement between all 3 examiners ranged from 76% at the SI joints to 90.667% at L3/4 while the mean percentage agreement for all examiners across the lumbar spine and SI joints for motion palpation was 87.667%.

D. Collapsed Spinal Levels

Table 20. Percentage Agreement for Motion Palpation in the Collapsed Spinal Levels

PERCENTAGE AGREEMENT				
	Examiner A&B	Examiner A&C	Examiner B&C	MEAN
C0-C3	76	70	76	74.000
C3-C5	74	72	74	73.333
C5-T1	81	78	89	82.667
T1-T4	86	85	85	85.333
T4-T7	66	61	73	66.667
T7-T10	82	73	83	79.333
T10-L1	87	88	81	85.333
L1-L3	87	80	81	82.667
L3-L5	97	85	84	88.667
MEAN	81.778	76.889	80.667	79.778

Table 20 presents the percentage agreement between the examiners for motion palpation in the collapsed spinal levels. The percentage agreement between examiners A and B ranged from 66% at T4-7 to 97% at L3-5, while the mean percentage agreement between these examiners across all the collapsed segments of the spine was 81.778%. The percentage agreement between examiners A and C ranged from 61% at T4-7 to 88% at T10-L1, while the mean percentage agreement between these examiners across all the collapsed segments of the spine was 76.889%. The percentage agreement between examiners B and C ranged from 73% at T4-7 to 89% at C5-T1, while the mean percentage agreement between these examiners across all the collapsed segments of the spine was 80.667%. The mean percentage agreement between all 3 examiners ranged from 66.667% at T4-7 to 88.667% at L3-5, while the mean percentage agreement between these examiners for motion palpation across all the collapsed segments of the spine was 79.778%.

4.3.2.2 PEARSON'S CHI-SQUARE

The Chi-Squared test was used to determine whether there was a significant relationship between the outcomes of each examiner (1993 Fisher and van Belle: 219). If the p-value calculated for Chi-Square was less than 0.05, then the null hypothesis was rejected and so we could conclude that the examination outcomes of each examiner were significantly associated. If the p-value calculated for Chi-Square was greater than 0.05, then the null hypothesis was failed to be rejected and so we could conclude that the examination outcomes of each examiner were not significantly associated.

The Null Hypothesis states that there is no association between the examination outcomes of the examiners.

The Alternative Hypothesis states that there is an association between the examination outcomes of the examiners.

Decision Rule:

If $p \geq 0.05$ (level of significance), we fail to reject the null hypothesis.

If $p < 0.05$ (level of significance), we reject the null hypothesis and so accept the alternative hypothesis.

A. Cervical Spine

Table 21. Chi-Square Values for Motion Palpation in the Cervical Spine

	Examiner A&B		Examiner A&C		Examiner B&C		MEAN	
	Chi-Square	**p-value**	**Chi-Square**	**p-value**	**Chi-Square**	**p-value**	**Chi-Square**	**p-value**
C0/C1	0.031	0.86	15.825	0	4.097	0.043	6.651	0.301
C1/C2	10.485	0.001	8.355	0.004	0.404	0.525	6.415	0.177
C2/C3	10.113	0.001	1.82	0.177	14.382	0	8.772	0.059
C3/C4	16.642	0	7.44	0.006	4.691	0.03	9.591	0.012
C4/C5	2.751	0.097	1.957	0.162	6.799	0.009	3.836	0.089
C5/C6	2.891	0.089	6.033	0.014	25.413	0	11.446	0.034
C6/C7	8.455	0.004	0.197	0.657	4.097	0.043	4.250	0.235
C7/T1	(a)	(a)	(a)	(a)	(a)	(a)	(a)	(a)
MEAN	7.338	0.150	5.947	0.146	8.555	0.093	7.280	0.130

(a) Values could not be calculated because one or both of the examiners did not have any positive assessments at this level (i.e. no spinal joint dysfunction at this spinal level).

Table 21 presents the Pearson’s Chi-Square values and their p-values calculated for the cervical spine. The p-values calculated for examiners A and B were greater than 0.05 at C0/1, C4/5 and C5/6; we therefore fail to reject the null hypothesis at the 0.05 level of significance and so conclude that there is no association between the examiners at this spinal level for motion palpation. The p-values calculated for examiners A and B were less than 0.05 at C1/2, C2/3, C3/4 and C6/7; we therefore reject the null hypothesis at the 0.05 level of significance and so conclude that there is an association between the examiners at these spinal levels for motion palpation. Values could not be calculated at C7/T1 because one or both of the examiners did not have any positive assessments at this level (i.e. no spinal joint dysfunction at this spinal level).

The p-values calculated for examiners A and C were greater than 0.05 at C2/3, C4/5 and C6/7; we therefore fail to reject the null hypothesis at the 0.05 level of significance and so conclude that there is no association between the examiners at these spinal levels for motion palpation. The p-values calculated for examiners A and C were less than 0.05 at C0/1, C1/2, C3/4 and C5/6; we therefore reject the null hypothesis at the 0.05 level of significance and so conclude that there is an association between the examiners at these spinal levels for motion palpation. Values could not be calculated at C7/T1 because one or both of the examiners did not have any positive assessments at this level (i.e. no spinal joint dysfunction at this spinal level).

The p-value calculated for examiners B and C was greater than 0.05 at C1/2; we therefore fail to reject the null hypothesis at the 0.05 level of significance and so conclude that there is no association between the examiners at this spinal level for motion palpation. The p-values calculated for examiners B and C were less than 0.05 at C0/1, C2/3, C3/4, C4/5, C5/6 and C6/7; we therefore reject the null hypothesis at the 0.05 level of significance and so conclude that there is an association between the examiners at these spinal levels for motion palpation. Values could not be calculated at C7/T1 because one or both of the examiners did not have any positive assessments at these levels (i.e. no spinal joint dysfunction at these spinal levels).

The mean p-values calculated between all 3 examiners were greater than 0.05 at C0/1, C1/2, C2/3, C4/5 and C6/7; we therefore fail to reject the null hypothesis at the 0.05 level of significance and so conclude that there is no association between the examiners at these spinal levels for motion palpation. The mean p-values calculated between all 3 examiners were less than 0.05 at C3/4 and C5/6; we therefore reject the null hypothesis at the 0.05 level of significance and so conclude that there is an association between the examiners at these spinal levels for motion palpation. Values could not be calculated at C7/T1 because one or all of the examiners did not have any positive assessments at this level (i.e. no spinal joint dysfunction at this spinal level).

B. Thoracic Spine

Table 22. Chi-Square Values for Motion Palpation in the Thoracic Spine

	Examiner A&B		Examiner A&C		Examiner B&C		MEAN	
	Chi-Square	**p-value**	**Chi-Square**	**p-value**	**Chi-Square**	**p-value**	**Chi-Square**	**p-value**
T1/T2	(a)	(a)	(a)	(a)	(a)	(a)	(a)	(a)
T2/T3	(a)	(a)	(a)	(a)	(a)	(a)	(a)	(a)
T3/T4	0.315	0.575	0.407	0.523	3.861	0.049	1.528	0.382
T4/T5	7.298	0.007	1.468	0.226	5.933	0.015	4.900	0.083
T5/T6	9.545	0.002	0.042	0.837	0.447	0.504	3.345	0.448
T6/T7	19.705	0	8.737	0.003	15.925	0	14.789	0.001
T7/T8	6.799	0.009	0.154	0.695	11.348	0.001	6.100	0.235
T8/T9	0.219	0.64	3.509	0.061	0.277	0.599	1.335	0.433
T9/T10	8.7	0.003	0.163	0.687	0.063	0.802	2.975	0.497
T10/T11	27.212	0	0.163	0.687	0.197	0.657	9.191	0.448
T11/T12	4.891	0.027	0.085	0.771	1.636	0.201	2.204	0.333
T12/L1	14.966	0	0.063	0.802	1.333	0.248	5.454	0.350
MEAN	9.965	0.126	1.479	0.529	4.102	0.308	5.182	0.321

(a) Values could not be calculated because one or both of the examiners did not have any positive assessments at these levels (i.e. no spinal joint dysfunction at these spinal levels).

Table 22 presents the Pearson's Chi-Square values and their p-values calculated for the thoracic spine. The p-values calculated for examiners A and B were greater than 0.05 at T3/4 and T8/9; we therefore fail to reject the null hypothesis at the 0.05 level of significance and so conclude that there is no association between the examiners at these spinal levels for motion palpation. The p-values calculated for examiners A and B were less than 0.05 at T4/5, T5/6, T6/7, T7/8, T9/10, T10/11, T11/12 and T12/L1; we therefore reject the null hypothesis at the 0.05 level of significance and so conclude that there is an association between the examiners at these spinal levels for motion palpation. Values could not be calculated at T1/2 and T2/3 because one or both of the examiners did not have any positive assessments at these levels (i.e. no spinal joint dysfunction at these spinal levels).

The p-values calculated for examiners A and C were greater than 0.05 at T3/4, T4/5, T5/6, T7/8, T8/9, T9/10, T10/11, T11/12 and T12/L1; we therefore fail to reject the null hypothesis at the 0.05 level of significance and so conclude that there is no association between the examiners at these spinal levels for motion palpation. The p-value calculated for examiners A and C was less than 0.05 at T6/7; we therefore reject the null hypothesis at the 0.05 level of significance and so conclude that there is an association between the examiners at this spinal level for motion palpation. Values could not be calculated at T1/2 and T2/3 because one or both of the examiners did not have any positive assessments at these levels (i.e. no spinal joint dysfunction at these spinal levels).

The p-values calculated for examiners B and C were greater than 0.05 at T5/6, T8/9, T9/10, T10/11, T11/12 and T12/L1; we therefore fail to reject the null hypothesis at the 0.05 level of significance and so conclude that there is no association between the examiners at these spinal levels for motion palpation. The p-values calculated for examiners B and C were less than 0.05 at T3/4, T4/5, T6/7 and T7/8; we therefore reject the null hypothesis at the 0.05 level of significance and so conclude that there is an association between the examiners at these spinal levels for motion palpation. Values could not be calculated at T1/2 and T2/3 because one or both of the examiners did not have any positive assessments at these levels (i.e. no spinal joint dysfunction at these spinal levels).

The mean p-values calculated between all 3 examiners were greater than 0.05 at T3/4, T4/5, T5/6, T7/8, T8/9, T9/10, T10/11, T11/12 and T12/L1; we therefore fail to reject the null hypothesis at the 0.05 level of significance and so conclude that there is no association between the examiners at these spinal levels for motion palpation. The mean p-value calculated between all 3 examiners was less than 0.05 at T6/7; we therefore reject the null hypothesis at the 0.05 level of significance and so conclude that there is an association between the examiners at this spinal level for motion palpation. Values could not be calculated at T1/2 and T2/3 because one or all of the examiners did not have any positive assessments at these levels (i.e. no spinal joint dysfunction at these spinal levels).

C. Lumbar Spine and SI Joints

Table 23. Chi-Square Values for Motion Palpation in the Lumbar Spine and Sacroiliac Joints

	Examiner A&B		Examiner A&C		Examiner B&C		MEAN	
	Chi-Square	**p-value**	**Chi-Square**	**p-value**	**Chi-Square**	**p-value**	**Chi-Square**	**p-value**
L1/L2	34.467	0	22.676	0	19.705	0	25.616	0.000
L2/L3	28.075	0	15.803	0	27.203	0	23.694	0.000
L3/L4	0.042	0.838	0.252	0.616	3.171	0.075	1.155	0.510
L4/L5	6.93	0.008	2.074	0.15	0.233	0.629	3.079	0.262
SI	25.488	0	11.111	0.001	20.947	0	19.182	0.000
MEAN	19.000	0.169	10.383	0.153	14.252	0.141	14.545	0.154

Table 23 presents the Pearson's Chi-Square values and their p-values calculated for the lumbar spine and sacroiliac (SI) joints. The p-value calculated for examiners A and B was greater than 0.05 at L3/4; we therefore fail to reject the null hypothesis at the 0.05 level of significance and so conclude that there is no association between the examiners at this spinal level for motion palpation. The p-values calculated for examiners A and B were less than 0.05 at L1/2, L2/3, L4/5 and the SI joints; we therefore reject the null hypothesis at the 0.05 level of significance and so conclude that there is an association between the examiners at these levels of the lumbar spine and the SI joints for motion palpation.

The p-values calculated for examiners A and C were greater than 0.05 at L3/4 and L4/5; we therefore fail to reject the null hypothesis at the 0.05 level of significance and so conclude that there is no association between the examiners at these spinal levels for motion palpation. The p-values calculated for examiners A and C were less than 0.05 at L1/2, L2/3 and the SI joints; we therefore reject the null hypothesis at the 0.05 level of significance and so conclude that there is an association between the examiners at these spinal levels for motion palpation.

The p-values calculated for examiners B and C were greater than 0.05 at L3/4 and L4/5; we therefore fail to reject the null hypothesis at the 0.05 level of significance and so conclude that there is no association between the examiners at these spinal levels for motion palpation. The p-values calculated for examiners B and C were less than 0.05 at L1/2, L2/3 and the SI joints; we therefore reject the null hypothesis at the 0.05 level of significance and so conclude that there is an association between the examiners at these spinal levels for motion palpation.

The mean p-values calculated between all 3 examiners were greater than 0.05 at L3/4 and L4/5; we therefore fail to reject the null hypothesis at the 0.05 level of significance and so conclude that there is no association between the examiners at these spinal levels for motion palpation. The mean p-values calculated between all 3 examiners were less than 0.05 at L1/2, L2/3 and the SI joints; we therefore reject the null hypothesis at the 0.05 level of significance and so conclude that there is an association between the examiners at these spinal levels for motion palpation.

D. Collapsed Spinal Levels

Table 24. Chi-Square Values for Motion Palpation in the Collapsed Spinal Levels

	Examiner A&B		Examiner A&C		Examiner B&C		MEAN	
	Chi-Square	**p-value**	**Chi-Square**	**p-value**	**Chi-Square**	**p-value**	**Chi-Square**	**p-value**
C0-C3	19.434	0	6.692	0.01	20.38	0	15.502	0.003
C3-C5	19.126	0	15.3	0	11.59	0.001	15.339	0.000
C5-T1	13.859	0	6.945	0.008	17.263	0	12.689	0.003
T1-T4	0.315	0.575	0.631	0.427	1.642	0.2	0.863	0.401
T4-T7	9.388	0.002	3.883	0.049	11.703	0.001	8.325	0.017
T7-T10	4.762	0.029	0.041	0.839	8.575	0.003	4.459	0.290
T10-L1	17.206	0	0.396	0.529	0.063	0.802	5.888	0.444
L1-L3	32.408	0	18.215	0	20.325	0	23.649	0.000
L3-L5	38.776	0	1.961	0.161	2.58	0.108	14.439	0.090
MEAN	17.253	0.067	6.007	0.225	10.458	0.124	11.239	0.139

Table 24 presents the Pearson's Chi-Square values and their p-values calculated for the collapsed spinal levels. The p-value calculated for examiners A and B was greater than 0.05 at T1-4; we therefore fail to reject the null hypothesis at the 0.05 level of significance and so conclude that there is no association between the examiners at this collapsed spinal level for motion palpation. The p-values calculated for examiners A and B were less than 0.05 at C0-3, C3-5, C5-T1, T4-7, T7-10, T10-L1, L1-3 and L3-5; we therefore reject the null hypothesis at the 0.05 level of significance and so conclude that there is an association between the examiners at these collapsed spinal levels for motion palpation.

The p-values calculated for examiners A and C were greater than 0.05 at T1-4, T7-10 and T10-L1 and L3-5; we therefore fail to reject the null hypothesis at the 0.05 level of significance and so conclude that there is no association between the examiners at these collapsed spinal levels for motion palpation. The p-values calculated for examiners A and C were less than 0.05 at C0-3, C3-5, C5-T1, T4-7 and L1-3; we therefore reject the null hypothesis at the 0.05 level of significance and so conclude that there is an association between the examiners at these collapsed spinal levels for motion palpation.

The p-values calculated for examiners B and C were greater than 0.05 at T1-4, T10-L1 and L3-5; we therefore fail to reject the null hypothesis at the 0.05 level of significance and so conclude that there is no association between the examiners at these collapsed spinal levels for motion palpation. The p-values calculated for examiners B and C were less than 0.05 at C0-3, C3-5, C5-T1, T4-7, T7-10 and L1-3; we therefore reject the null hypothesis at the 0.05 level of significance and so conclude that there is an association between the examiners at these collapsed spinal levels for motion palpation.

The mean p-values calculated between all 3 examiners were greater than 0.05 at T1-4, T7-10, T10-L1 and L3-5; we therefore fail to reject the null hypothesis at the 0.05 level of significance and so conclude that there is no association between the examiners at these collapsed spinal levels for motion palpation. The mean p-values calculated between all 3 examiners were less than 0.05 at C0-3, C3-5, C5-T1, T4-7 and L1-3; we therefore reject the null hypothesis at the 0.05 level of significance and so conclude that there is an association between the examiners at these collapsed spinal levels for motion palpation.

4.3.2.3 COHEN'S KAPPA COEFFICIENT

Cohen's Kappa (K) was used to measure the chance corrected concordance between the examiners. The kappa values were tested for significance at the α=0.05 level. The guidelines used for interpreting the kappa values in this study were in accordance with those employed by French et al. (2000), i.e.:

- < 0.00: poor agreement
- 0.00 – 0.20: slight agreement

- 0.21 – 0.40: fair agreement
- 0.41 – 0.60: moderate agreement
- 0.61 – 0.80: substantial agreement
- 0.81 – 1.00: almost perfect agreement

A. Cervical Spine

Table 25. Kappa Values for Motion Palpation in the Cervical Spine

	Examiner A&B		Examiner A&C		Examiner B&C		MEAN	
	Kappa	**p-value**	**Kappa**	**p-value**	**Kappa**	**p-value**	**Kappa**	**p-value**
C0/C1	-0.015	0.86	0.273	0	0.19	0.043	0.149	0.301
C1/C2	0.296	0.001	0.272	0.004	0.063	0.525	0.210	0.177
C2/C3	0.279	0.001	0.13	0.177	0.367	0	0.259	0.059
C3/C4	0.406	0	0.263	0.006	0.213	0.03	0.294	0.012
C4/C5	0.155	0.097	0.134	0.162	0.26	0.009	0.183	0.089
C5/C6	0.125	0.089	0.221	0.014	0.472	0	0.273	0.034
C6/C7	0.291	0.004	-0.042	0.657	0.19	0.043	0.146	0.235
C7/T1	(a)	(a)	(a)	(a)	(a)	(a)	(a)	(a)
MEAN	0.220	0.150	0.179	0.146	0.251	0.093	0.216	0.130

(a) Values could not be calculated because one or all of the examiners did not have any positive assessments at this level (i.e. no spinal joint dysfunction at this spinal level).

Table 25 presents the Cohen’s Kappa (K) scores and their p-values calculated for the cervical spine. Kappa scores calculated for examiners A and B in the cervical spine ranged from –0.015 (at C0/1) to 0.406 (at C3/4) which indicates a range of poor to moderate agreement between the examiners beyond chance. The overall (mean) Kappa score calculated for the cervical spine was 0.220, which indicates fair agreement beyond chance between these examiners for motion palpation in the cervical spine. Kappa scores were statistically significant ($p<0.05$) at all spinal levels in the cervical spine besides C0/1, C4/5 and C5/6. Values could not be calculated at C7/T1 because one or all of the examiners did not have any positive assessments at this level (i.e. no spinal joint dysfunction at this spinal level).

Kappa scores calculated for examiners A and C in the cervical spine ranged from –0.042 (at C6/7) to 0.273 (at C0/1) which indicates a range of poor to fair agreement between the examiners beyond chance. The overall (mean) Kappa score calculated for the cervical spine was 0.179, which indicates slight agreement beyond chance between these examiners for motion palpation in the cervical spine. Kappa scores were statistically significant ($p<0.05$) at all spinal levels in the cervical spine besides C2/3, C4/5 and C6/7. Values could not be calculated at C7/T1 because one or all of the examiners did not have any positive assessments at this level (i.e. no spinal joint dysfunction at this spinal level).

Kappa scores calculated for examiners B and C in the cervical spine ranged from 0.063 (at C1/2) to 0.472 (at C5/6), which indicates a range of slight to moderate agreement between the examiners beyond chance. The overall (mean) Kappa score calculated for the cervical spine was 0.251, which indicates fair agreement beyond chance between these examiners for motion palpation in the cervical spine. Kappa scores were statistically significant ($p<0.05$) at all spinal levels in the cervical spine besides C1/2. Values could not be calculated at C7/T1 because one or all of the examiners did not have any positive assessments at this level (i.e. no spinal joint dysfunction at this spinal level).

Mean Kappa scores calculated for all examiners in the cervical spine ranged from 0.149 (at C0/1) to 0.294 (at C3/4), which indicates a range of slight to fair agreement between the examiners beyond chance. The overall (mean) Kappa score calculated for all examiners across the cervical spine was 0.216, which indicates fair agreement beyond chance for motion palpation between the examiners. Kappa scores were statistically significant ($p<0.05$) at all spinal levels in the cervical spine besides C0/1, C1/2, C2/3, C4/5 and C6/7. Values could not be calculated at C7/T1 because one or all of the examiners did not have any positive assessments at this level (i.e. no spinal joint dysfunction at this spinal level).

B. Thoracic Spine

Table 26. Kappa Values for Motion Palpation in the Thoracic Spine

	Examiner A&B		Examiner A&C		Examiner B&C		MEAN	
	Kappa	**p-value**	**Kappa**	**p-value**	**Kappa**	**p-value**	**Kappa**	**p-value**
T1/T2	(a)	(a)	(a)	(a)	(a)	(a)	(a)	(a)
T2/T3	(a)	(a)	(a)	(a)	(a)	(a)	(a)	(a)
T3/T4	0.054	0.575	-0.064	0.523	0.189	0.049	0.060	0.382
T4/T5	0.252	0.007	0.096	0.226	0.23	0.015	0.193	0.083
T5/T6	0.309	0.002	-0.02	0.837	0.067	0.504	0.119	0.448
T6/T7	0.443	0	0.295	0.003	0.396	0	0.378	0.001
T7/T8	0.26	0.009	0.038	0.695	0.324	0.001	0.207	0.235
T8/T9	-0.047	0.64	0.186	0.061	-0.053	0.599	0.029	0.433
T9/T10	0.265	0.003	-0.039	0.687	-0.025	0.802	0.067	0.497
T10/T11	0.519	0	-0.039	0.687	-0.042	0.657	0.146	0.448
T11/T12	0.174	0.027	-0.027	0.771	0.12	0.201	0.089	0.333
T12/L1	0.26	0	-0.025	0.802	0.09	0.248	0.108	0.350
MEAN	0.249	0.126	0.040	0.529	0.130	0.308	0.140	0.321

(a) Values could not be calculated because one or all of the examiners did not have any positive assessments at these levels (i.e. no spinal joint dysfunction at these spinal levels).

Table 26 presents the Cohen's Kappa (K) scores and their p-values calculated for the thoracic spine. Kappa scores calculated for examiners A and B ranged from –0.047 (at T8/9) to 0.519 (at T10/11), which indicates a range of poor to moderate agreement between the examiners beyond chance. The overall (mean) Kappa score calculated here was 0.249, which indicates fair agreement beyond chance between these examiners for motion palpation in the thoracic spine. Kappa scores were statistically significant (p<0.05) at all spinal levels in the thoracic spine besides T3/4 and T8/9. Values could not be calculated at T1/2 and T2/3 because one or all of the examiners did not have any positive assessments at these levels (i.e. no spinal joint dysfunction at these spinal levels).

Kappa scores calculated for examiners A and C ranged from –0.064 (at T3/4) to 0.295 (at T6/7), which indicates a range of poor to fair agreement between the examiners beyond chance. The overall (mean) Kappa score calculated here was 0.04, which indicates slight agreement beyond chance between these examiners for motion palpation in the thoracic spine. Kappa scores were statistically significant (p<0.05) at T6/7 while the Kappa scores at all other levels in the thoracic spine were not statistically significant (p>0.05). Values could not be calculated at T1/2 and T2/3 because one or all of the examiners did not have any positive assessments at these levels (i.e. no spinal joint dysfunction at these spinal levels).

Kappa scores calculated for examiners B and C ranged from –0.053 (at T8/9) to 0.396 (at T6/7), which indicates a range of poor to fair agreement between the examiners beyond chance. The overall (mean) Kappa score calculated here was 0.130, which indicates slight agreement beyond chance between these examiners for motion palpation in the thoracic spine. Kappa scores were statistically significant (p<0.05) at all spinal levels in the thoracic spine besides T5/6, T8/9, T9/10, T10/11, T11/12 and T12/L1. Values could not be calculated at T1/2 and T2/3 because one or all of the examiners did not have any positive assessments at these levels (i.e. no spinal joint dysfunction at these spinal levels).

Mean Kappa scores calculated for all examiners in the thoracic spine ranged from 0.029 (at T8/9) to 0.378 (at T6/7), which indicates a range of slight to fair agreement between the examiners beyond chance. The overall (mean) Kappa score calculated for all examiners across the thoracic spine was 0.140, which indicates slight agreement beyond chance between the examiners for motion palpation. Overall Kappa scores were statistically significant ($p<0.05$) at T6/7 while the overall Kappa scores at all other levels in the thoracic spine were not statistically significant ($p>0.05$). Values could not be calculated at T1/2 and T2/3 because one or all of the examiners did not have any positive assessments at these levels (i.e. no spinal joint dysfunction at these spinal levels).

C. Lumbar Spine and SI Joints

Table 27. Kappa Values for Motion Palpation in the Lumbar Spine and Sacroiliac Joints

	Examiner A&B		Examiner A&C		Examiner B&C		MEAN	
	Kappa	**p-value**	**Kappa**	**p-value**	**Kappa**	**p-value**	**Kappa**	**p-value**
L1/L2	0.585	0	0.476	0	0.443	0	0.501	0.000
L2/L3	0.514	0	0.352	0	0.505	0	0.457	0.000
L3/L4	-0.02	0.838	-0.035	0.616	0.124	0.075	0.023	0.510
L4/L5	0.26	0.008	0.138	0.15	-0.044	0.629	0.118	0.262
SI	0.505	0	0.333	0.001	0.458	0	0.432	0.000
MEAN	0.369	0.169	0.253	0.153	0.297	0.141	0.306	0.154

Table 27 presents the Cohen's Kappa (K) scores and their p-values calculated for the lumbar spine and sacroiliac (SI) joints. Kappa scores calculated for examiners A and B ranged from –0.02 (at L3/4) to 0.585 (at L1/2), which indicates a range of poor to moderate agreement between the examiners beyond chance. The overall (mean) Kappa score calculated here was 0.369, which indicates fair agreement beyond chance between these examiners for motion palpation in the lumbar spine and SI joints. Kappa scores were statistically significant ($p<0.05$) at all spinal levels in the lumbar spine and SI joints besides L3/4.

Kappa scores calculated for examiners A and C ranged from -0.035 (at L3/4) to 0.476 (at L1/2), which indicates a range of poor to moderate agreement between the examiners beyond chance. The overall (mean) Kappa score calculated here was 0.253, which indicates fair agreement beyond chance between these examiners for motion palpation in the lumbar spine and SI joints. Kappa scores were statistically significant ($p<0.05$) at all spinal levels in the lumbar spine and SI joints besides L3/4 and L4/5.

Kappa scores calculated for examiners B and C ranged from –0.044 (at L4/5) to 0.505 (at L2/3), which indicates a range of poor to moderate agreement between the examiners beyond chance. The overall (mean) Kappa score calculated here was 0.297, which indicates fair agreement beyond chance between these examiners for motion palpation in the lumbar spine and SI joints. Kappa scores were statistically significant ($p<0.05$) at all spinal levels in the lumbar spine and SI joints besides L3/4 and L4/5.

Mean Kappa scores calculated for all examiners in the lumbar spine and SI joints ranged from 0.023 (at L3/4) to 0.501 (at L1/2), which indicates a range of slight to moderate agreement between the examiners beyond chance. The overall (mean) Kappa score calculated for all examiners across the lumbar spine and SI joints was 0.306, which indicates fair agreement beyond chance between the examiners for motion palpation. Overall Kappa scores were statistically significant ($p<0.05$) at all levels in the lumbar spine and SI joints besides, L3/4 and L4/5.

D. Collapsed Spinal Levels

Table 28. Kappa Values for Motion Palpation in the Collapsed Spinal Levels

	Examiner A&B		Examiner A&C		Examiner B&C		MEAN	
	Kappa	**p-value**	**Kappa**	**p-value**	**Kappa**	**p-value**	**Kappa**	**p-value**
C0-C3	0.433	0	0.257	0.01	0.45	0	0.380	0.003
C3-C5	0.421	0	0.377	0	0.34	0.001	0.379	0.000
C5-T1	0.331	0	0.24	0.008	0.415	0	0.329	0.003
T1-T4	0.054	0.575	-0.078	0.427	0.128	0.2	0.035	0.401
T4-T7	0.293	0.002	0.181	0.049	0.34	0.001	0.271	0.017
T7-T10	0.211	0.029	0.02	0.839	0.28	0.003	0.170	0.290
T10-L1	0.374	0	-0.062	0.529	0.021	0.802	0.111	0.444
L1-L3	0.569	0	0.42	0	0.441	0	0.477	0.000
L3-L5	0.559	0	0.085	0.161	0.135	0.108	0.260	0.090
MEAN	0.361	0.067	0.160	0.225	0.283	0.124	0.268	0.139

Table 28 presents the Cohen’s Kappa (K) scores and their p-values calculated for the collapsed spinal levels. Kappa scores calculated for examiners A and B ranged from 0.054 (at T1-4) to 0.569 (at L1-3), which indicates a range of slight to moderate agreement between the examiners beyond chance. The overall (mean) Kappa score calculated here was 0.361, which indicates fair agreement beyond chance between these examiners for motion palpation in the collapsed spinal levels. Kappa scores were statistically significant ($p<0.05$) at all the collapsed spinal levels besides T1-4.

Kappa scores calculated for examiners A and C ranged from -0.078 (at T1-4) to 0.42 (at L1-3), which indicates a range of poor to moderate agreement between the examiners beyond chance. The overall (mean) Kappa score calculated here was 0.160, which indicates slight agreement beyond chance between these examiners for motion palpation in the collapsed spinal levels. Kappa scores were statistically significant ($p<0.05$) at all collapsed spinal levels besides T1-4, T7-10, T10-L1 and L3-5.

Kappa scores calculated for examiners B and C ranged from 0.021 (at T10-L1) to 0.45 (at C0-3), which indicates a range of slight to moderate agreement between the examiners beyond chance. The overall (mean) Kappa score calculated here was 0.283, which indicates fair agreement beyond chance between these examiners for motion palpation in the collapsed spinal levels. Kappa scores were statistically significant ($p<0.05$) at all collapsed spinal levels besides T1-4, T10-L1 and L3-5.

Mean Kappa scores calculated for all examiners in the collapsed spinal levels ranged from 0.035 (at T1-4) to 0.477 (at L1-3), which indicates a range of slight to moderate agreement between the examiners beyond chance. The overall (mean) Kappa score calculated for all examiners across the collapsed spinal levels was 0.268, which indicates fair agreement beyond chance between the examiners for motion palpation. Overall Kappa scores were statistically significant ($p<0.05$) at all collapsed spinal levels besides T1-4, T7-10, T10-L1 and L3-5.

4.3.3 Interexaminer Reliability of Static and Motion Palpation

Results of the statistical analysis to determine the interexaminer reliability of static palpation are presented in Tables 29 to 40.

4.3.3.1 PERCENTAGE AGREEMENT

This is the ratio of the number of agreements between examiners to the total number of comparisons made.

A. Cervical Spine

Table 29. Percentage Agreement for Static and Motion Palpation in the Cervical Spine

PERCENTAGE AGREEMENT				
	Examiner A&B	Examiner A&C	Examiner B&C	MEAN
C0/C1	96	96	94	95.333
C1/C2	88	90	92	90.000
C2/C3	86	85	83	84.667
C3/C4	80	76	76	77.333
C4/C5	86	84	96	88.667
C5/C6	94	92	96	94.000
C6/C7	98	97	99	98.000
C7/T1	100	100	100	100.000
MEAN	91.000	90.000	92.000	91.000

Table 29 presents the percentage agreement between the examiners in the cervical spine. The percentage agreement between examiners A and B ranged from 80% at C3/4 to 100% at C7/T1 while the mean percentage agreement between these examiners across the cervical spine was 91%. The percentage agreement between examiners A and C ranged from 76% at C3/4 to 100% at C7/T1 while the mean percentage agreement between these examiners across the cervical spine was 90%. The percentage agreement between examiners B and C ranged from 76% at C3/4 to 100% at C7/T1, while the mean percentage agreement between these examiners across the cervical spine was 92%. The mean percentage agreement between all 3 examiners ranged from 77.333% at C3/4 to 100% at C7/T1 while the mean percentage agreement across the cervical spine was 91%.

B. Thoracic Spine

Table 30. Percentage Agreement for Static and Motion Palpation in the Thoracic Spine

PERCENTAGE AGREEMENT				
	Examiner A&B	Examiner A&C	Examiner B&C	MEAN
T1/T2	100	100	100	100.000
T2/T3	100	98	98	98.667
T3/T4	100	95	95	96.667
T4/T5	90	85	91	88.667
T5/T6	94	87	87	89.333
T6/T7	91	85	88	88.000
T7/T8	98	92	94	94.667
T8/T9	97	95	98	96.667
T9/T10	100	97	97	98
T10/T11	98	94	94	95.333
T11/T12	95	95	92	94.000
T12/L1	95	96	93	94.667
MEAN	96.500	93.250	93.917	94.556

Table 30 presents the percentage agreement between the examiners in the thoracic spine. The percentage agreement between examiners A and B ranged from 90% at T4/5 to 100% at T1/T2, T2/3, T3/4 and T9/10, while the mean percentage agreement between these examiners across the thoracic spine was 96.5%. The percentage agreement between examiners A and C ranged from 85% at T4/5 and T6/7 to 100% at T1/2, while the mean percentage agreement between these examiners across the thoracic spine was 93.25%. The percentage agreement between examiners B and C ranged from 87% at T5/6 to 100% at T1/2 while the mean percentage agreement between these examiners across the thoracic spine was 93.917%. The mean percentage agreement between all 3 examiners ranged from 88% at T6/7 to 100% at T1/2 while the mean percentage agreement across the thoracic spine was 94.556%.

C. Lumbar Spine and Sacroiliac Joints

Table 31. Percentage Agreement for Static and Motion Palpation in the Lumbar Spine and Sacroiliac Joints

PERCENTAGE AGREEMENT				
	Examiner A&B	Examiner A&C	Examiner B&C	MEAN
L1/L2	93	90	93	92.000
L2/L3	97	92	93	94.000
L3/L4	97	92	93	94.000
L4/L5	98	96	96	96.667
SI	97	86	87	90.000
MEAN	96.400	91.200	92.400	93.333

Table 31 presents the percentage agreement between the examiners in the lumbar spine and sacroiliac (SI) joints. The percentage agreement between examiners A and B ranged from 93% at L1/2 to 98% at L4/5, while the mean percentage agreement between these examiners across the lumbar spine and SI joints was 96.4%. The percentage agreement between examiners A and C ranged from 86% at the SI joints to 96% at L4/5, while the mean percentage agreement between these examiners across the lumbar spine and SI joints was 91.2%. The percentage agreement between examiners B and C ranged from 87% at the SI joints to 96% at L4/5, while the mean percentage agreement between these examiners across the lumbar spine and SI joints was 92.4%. The mean percentage agreement between all 3 examiners ranged from 90% at the SI joints to 96.667% at L4/5 while the mean percentage agreement across the lumbar spine and SI joints was 93.333%.

D. Collapsed Spinal Levels

Table 32. Percentage Agreement for Static and Motion Palpation in the Collapsed Spinal Levels

PERCENTAGE AGREEMENT				
	Examiner A&B	Examiner A&C	Examiner B&C	MEAN
C0-C3	86	78	78	80.667
C3-C5	79	76	77	77.333
C5-T1	92	89	95	92.000
T1-T4	100	93	93	95.333
T4-T7	81	74	79	78.000
T7-T10	95	89	94	92.667
T10-L1	94	90	88	90.667
L1-L3	92	86	90	89.333
L3-L5	95	90	91	92.000
MEAN	90.444	85.000	87.222	87.556

Table 32 presents the percentage agreement between the examiners in the collapsed spinal levels. The percentage agreement between examiners A and B ranged from 79% at C3-5 to 100% at T1-4, while the mean percentage agreement between these examiners across all the collapsed segments of the spine was 90.444%. The percentage agreement between examiners A and C ranged from 74% at T4-7 to 93% at T1-4, while the mean percentage agreement between these examiners across all the collapsed segments of the spine was 85%. The percentage agreement between examiners B and C ranged from 77% at C3-5 and C3-5 to 95% at C5-T1, while the mean percentage agreement between these examiners across all the collapsed segments of the spine was 87.222%. The mean percentage agreement between all 3 examiners ranged from 77.333% at C3-5 to 95.333% at T1-4 while the mean percentage agreement between the examiners across all the collapsed segments of the spine was 87.556%.

4.3.3.2 PEARSON'S CHI-SQUARE

The Chi-Squared test was used to determine whether there was a significant relationship between the outcomes of each examiner (1993 Fisher and van Belle: 219). If the p-value calculated for Chi-Square was less than 0.05, then the null hypothesis was rejected and so we could conclude that the examination outcomes of each examiner were significantly associated. If the p-value calculated for Chi-Square was greater than 0.05, then the null hypothesis was failed to be rejected and so we could conclude that the examination outcomes of each examiner were not significantly associated.

The Null Hypothesis states that there is no association between the examination outcomes of the examiners.

The Alternative Hypothesis states that there is an association between the examination outcomes of the examiners.

Decision Rule:

If $p \geq 0.05$ (level of significance), we fail to reject the null hypothesis.

If $p < 0.05$ (level of significance), we reject the null hypothesis and so accept the alternative hypothesis.

A. Cervical Spine

Table 33. Chi-Square Values for static and Motion Palpation in the Cervical Spine

	Examiner A&B		Examiner A&C		Examiner B&C		MEAN	
	Chi-Square	**p-value**	**Chi-Square**	**p-value**	**Chi-Square**	**p-value**	**Chi-Square**	**p-value**
C0/C1	0.042	0.838	11.264	0.001	0.085	0.771	3.797	0.537
C1/C2	4.521	0.033	12.908	0	2.493	0.114	6.641	0.049
C2/C3	17.29	0	11.514	0.001	14.701	0	14.502	0.000
C3/C4	18.868	0	8.206	0.004	1.233	0.267	9.436	0.090
C4/C5	6.475	0.011	2.58	0.108	24.76	0	11.272	0.040
C5/C6	23.469	0	15.298	0	31.973	0	23.580	0.000
C6/C7	(a)	(a)	0.021	0.886	(a)	(a)	0.021	0.886
C7/T1	(a)	(a)	(a)	(a)	(a)	(a)	(a)	(a)
MEAN	11.778	0.147	8.827	0.143	12.541	0.192	9.893	0.229

(a) Values could not be calculated because one or both of the examiners did not have any positive assessments at this level (i.e. no spinal joint dysfunction at this spinal level).

Table 33 presents the Pearson’s Chi-Square values and their p-values calculated for the cervical spine. The p-value calculated for examiners A and B was greater than 0.05 at C0/1; we therefore fail to reject the null hypothesis at the 0.05 level of significance and so conclude that there is no association between the examiners at this spinal level for static and motion palpation. The p-values calculated for examiners A and B were less than 0.05 at C1/2, C2/3, C3/4, C4/5 and C5/6; we therefore reject the null hypothesis at the 0.05 level of significance and so conclude that there is an association between the examiners at these spinal levels for static and motion palpation. Values could not be calculated at C6/7 and C7/T1 because one or both of the examiners did not have any positive assessments at this level (i.e. no spinal joint dysfunction at this spinal level).

The p-values calculated for examiners A and C were greater than 0.05 at C4/5 and C6/7; we therefore fail to reject the null hypothesis at the 0.05 level of significance and so conclude that there is no association between the examiners at these spinal levels for static and motion palpation. The p-values calculated for examiners A and C were less than 0.05 at C0/1, C1/2, C2/3, C3/4 and C5/6; we therefore reject the null hypothesis at the 0.05 level of significance and so conclude that there is an association between the examiners at these spinal levels for static and motion palpation. Values could not be calculated at C7/T1 because one or both of the examiners did not have any positive assessments at this level (i.e. no spinal joint dysfunction at this spinal level).

The p-values calculated for examiners B and C were greater than 0.05 at C0/1, C1/2 and C3/4; we therefore fail to reject the null hypothesis at the 0.05 level of significance and so conclude that there is no association between the examiners at these spinal levels for static and motion palpation. The p-values calculated for examiners B and C were less than 0.05 at C2/3, C4/5 and C5/6; we therefore reject the null hypothesis at the 0.05 level of significance and so conclude that there is an association between the examiners at these spinal levels for static and motion palpation. Values could not be calculated at C6/7 and C7/T1 because one or both of the examiners did not have any positive assessments at these levels (i.e. no spinal joint dysfunction at these spinal levels).

The mean p-values calculated between all 3 examiners were greater than 0.05 at C0/1, C3/4 and C6/7; we therefore fail to reject the null hypothesis at the 0.05 level of significance and so conclude that there is no association between the examiners at these spinal levels for static and motion palpation. The mean p-values calculated between all 3 examiners were less than 0.05 at C1/2, C2/3, C4/5 and C5/6; we therefore reject the null hypothesis at the 0.05 level of significance and so conclude that there is an association between the examiners at these spinal levels for static and motion palpation. Values could not be calculated at C7/T1 because one or all of the examiners did not have any positive assessments at this level (i.e. no spinal joint dysfunction at this spinal level).

B. Thoracic Spine

Table 34. Chi-Square Values for Static and Motion Palpation in the Thoracic Spine

	Examiner A&B		Examiner A&C		Examiner B&C		MEAN	
	Chi-Square	**p-value**	**Chi-Square**	**p-value**	**Chi-Square**	**p-value**	**Chi-Square**	**p-value**
T1/T2	(a)	(a)	(a)	(a)	(a)	(a)	(a)	(a)
T2/T3	(a)	(a)	(a)	(a)	(a)	(a)	(a)	(a)
T3/T4	100	0	15.825	0	15.825	0	43.883	0.000
T4/T5	1.636	0.201	0.655	0.419	2.074	0.15	1.455	0.257
T5/T6	5.227	0.022	0.344	0.558	0.585	0.444	2.052	0.341
T6/T7	17.848	0	4.262	0.039	11.891	0.001	11.334	0.013
T7/T8	48.98	0	2.667	0.102	7.005	0.008	19.551	0.037
T8/T9	(a)	(a)	0.063	0.802	(a)	(a)	0.063	0.802
T9/T10	100	0	0.021	0.886	0.021	0.886	33.347	0.591
T10/T11	23.99	0	0.085	0.771	0.085	0.771	8.053	0.514
T11/T12	(a)	(a)	(a)	(a)	2.493	0.114	2.493	0.114
T12/L1	15.825	0	0.031	0.86	4.097	0.043	6.651	0.301
MEAN	39.188	0.028	2.661	0.493	4.897	0.269	12.888	0.297

(a) Values could not be calculated because one or both of the examiners did not have any positive assessments at these levels (i.e. no spinal joint dysfunction at these spinal levels).

Table 34 presents the Pearson's Chi-Square values and their p-values calculated for the thoracic spine. The p-values calculated for examiners A and B were greater than 0.05 at T4/5; we therefore fail to reject the null hypothesis at the 0.05 level of significance and so conclude that there is no association between the examiners at these spinal levels for static and motion palpation. The p-values calculated for examiners A and B were less than 0.05 at T3/4, T5/6, T6/7, T7/8, T9/10, T10/11 and T12/L1; we therefore reject the null hypothesis at the 0.05 level of significance and so conclude that there is an association between the examiners at these spinal levels for static and motion palpation. Values could not be calculated at T1/2, T2/3, T8/9 and T11/12 because one or both of the examiners did not have any positive assessments at these levels (i.e. no spinal joint dysfunction at these spinal levels).

The p-values calculated for examiners A and C were greater than 0.05 at T4/5, T5/6, T7/8, T8/9, T9/10, T10/11 and T12/L1; we therefore fail to reject the null hypothesis at the 0.05 level of significance and so conclude that there is no association between the examiners at these spinal levels for static and motion palpation. The p-values calculated for examiners A and C were less than 0.05 at T3/4 and T6/7; we therefore reject the null hypothesis at the 0.05 level of significance and so conclude that there is an association between the examiners at these spinal levels for static and motion palpation. Values could not be calculated at T1/2, T2/3 and T11/12 because one or both of the examiners did not have any positive assessments at these levels (i.e. no spinal joint dysfunction at these spinal levels).

The p-values calculated for examiners B and C were greater than 0.05 at T4/5, T5/6, T9/10, T10/11 and T11/12; we therefore fail to reject the null hypothesis at the 0.05 level of significance and so conclude that there is no association between the examiners at these spinal levels for static and motion palpation. The p-values calculated for examiners B and C were less than 0.05 at T3/4, T6/7, T7/8 and T12/L1; we therefore reject the null hypothesis at the 0.05 level of significance and so conclude that there is an association between the examiners at these spinal levels for static and motion palpation. Values could not be calculated at T1/2, T2/3 and T8/9 because one or both of the examiners did not have any positive assessments at these levels (i.e. no spinal joint dysfunction at these spinal levels).

The mean p-values calculated between all 3 examiners were greater than 0.05 at T4/5, T5/6, T8/9, T9/10, T10/11, T11/12 and T12/L1; we therefore fail to reject the null hypothesis at the 0.05 level of significance and so conclude that there is no association between the examiners at these spinal levels for static and motion palpation. The mean p-values calculated between all 3 examiners were less than 0.05 at T3/4, T6/7 and T7/8; we therefore reject the null hypothesis at the 0.05 level of significance and so conclude that there is an association between the examiners at these spinal levels for static and motion palpation. Values could not be calculated at T1/2 and T2/3 because one or all of the examiners did not have any positive assessments at these levels (i.e. no spinal joint dysfunction at these spinal levels).

C. Lumbar Spine and SI Joints

Table 35. Chi-Square Values for Static and Motion Palpation in the Lumbar Spine and Sacroiliac Joints

	Examiner A&B		Examiner A&C		Examiner B&C		MEAN	
	Chi-Square	**p-value**	**Chi-Square**	**p-value**	**Chi-Square**	**p-value**	**Chi-Square**	**p-value**
L1/L2	35.287	0	15.459	0	30.39	0	27.045	0.000
L2/L3	42.982	0	19.56	0	28.655	0	30.399	0.000
L3/L4	0.021	0.886	4.891	0.027	11.616	0.001	5.509	0.305
L4/L5	32.66	0	9.779	0.002	0.013	0.86	14.151	0.287
SI	51.101	0	6.871	0.009	9.25	0.002	22.407	0.004
MEAN	32.410	0.177	11.312	0.008	15.985	0.173	19.902	0.119

Table 35 presents the Pearson's Chi-Square values and their p-values calculated for the lumbar spine and sacroiliac (SI) joints. The p-value calculated for examiners A and B was greater than 0.05 at L3/4; we therefore fail to reject the null hypothesis at the 0.05 level of significance and so conclude that there is no association between the examiners at this spinal level for static and motion palpation. The p-values calculated for examiners A and B were less than 0.05 at L1/2, L2/3, L4/5 and the SI joints; we therefore reject the null hypothesis at the 0.05 level of significance and so conclude that there is an association between the examiners at these levels of the lumbar spine and the SI joints for static and motion palpation.

The p-values calculated for examiners A and C were less than 0.05 at all levels of the lumbar spine and the SI joints; we therefore reject the null hypothesis at the 0.05 level of significance and so conclude that there is an association between the examiners at these spinal levels for static and motion palpation.

The p-value calculated for examiners B and C was greater than 0.05 at L4/5; we therefore fail to reject the null hypothesis at the 0.05 level of significance and so conclude that there is no association between the examiners at this spinal level for static and motion palpation. The p-values calculated for examiners B and C were less than 0.05 at L1/2, L2/3, L3/4 and the SI joints; we therefore reject the null hypothesis at the 0.05 level of significance and so conclude that there is an association between the examiners at these spinal levels for static and motion palpation.

The mean p-values calculated between all 3 examiners were greater than 0.05 at L3/4 and L4/5; we therefore fail to reject the null hypothesis at the 0.05 level of significance and so conclude that there is no association between the examiners at these spinal levels for static and motion palpation. The mean p-values calculated between all 3 examiners were less than 0.05 at L1/2, L2/3 and the SI joints; we therefore reject the null hypothesis at the 0.05 level of significance and so conclude that there is an association between the examiners at these spinal levels for static and motion palpation.

D. Collapsed Spinal Levels

Table 36. Chi-Square Values for Static and Motion Palpation in the Collapsed Spinal Levels

	Examiner A&B		Examiner A&C		Examiner B&C		MEAN	
	Chi-Square	**p-value**	**Chi-Square**	**p-value**	**Chi-Square**	**p-value**	**Chi-Square**	**p-value**
C0-C3	27.802	0	10.368	0.001	11.629	0.001	16.600	0.001
C3-C5	23.759	0	16.32	0	5.642	0.018	15.240	0.006
C5-T1	18.367	0	9.029	0.003	27.114	0	18.170	0.001
T1-T4	100	0	11.616	0.001	11.616	0.001	41.077	0.001
T4-T7	8.515	0.004	3.549	0.06	6.913	0.009	6.326	0.024
T7-T10	35.567	0	4.328	0.037	16.913	0	18.936	0.012
T10-L1	16.913	0	0.233	0.629	0.614	0.433	5.920	0.354
L1-L3	31.305	0	11.51	0.001	25.185	0	22.667	0.000
L3-L5	8.7	0.003	6.176	0.013	4.189	0.041	6.355	0.019
MEAN	30.103	0.001	8.125	0.083	12.202	0.056	16.810	0.046

Table 36 presents the Pearson's Chi-Square values and their p-values calculated for the collapsed spinal levels. The p-values calculated for examiners A and B were less than 0.05 at all the collapsed spinal levels; we therefore reject the null hypothesis at the 0.05 level of significance and so conclude that there is an association between the examiners at all the collapsed spinal levels for static and motion palpation.

The p-values calculated for examiners A and C were greater than 0.05 at T4-7 and T10-L1; we therefore fail to reject the null hypothesis at the 0.05 level of significance and so conclude that there is no association between the examiners at these collapsed spinal levels for static and motion palpation. The p-values calculated for examiners A and C were less than 0.05 at C0-3, C3-5, C5-T1, T1-4, T7-10, L1-3 and L3-5; we therefore reject the null hypothesis at the 0.05 level of significance and so conclude that there is an association between the examiners at these collapsed spinal levels for static and motion palpation.

The p-values calculated for examiners B and C were greater than 0.05 at T10-L1; we therefore fail to reject the null hypothesis at the 0.05 level of significance and so conclude that there is no association between the examiners at these collapsed spinal levels for static and motion palpation. The p-values calculated for examiners B and C were less than 0.05 at C0-3, C3-5, C5-T1, T1-4, T4-7, T7-10, L1-3 and L3-5; we therefore reject the null hypothesis at the 0.05 level of significance and so conclude that there is an association between the examiners at these collapsed spinal levels for static and motion palpation.

The mean p-values calculated between all 3 examiners were greater than 0.05 at T10-L1; we therefore fail to reject the null hypothesis at the 0.05 level of significance and so conclude that there is no association between the examiners at these collapsed spinal levels for static and motion palpation. The mean p-values calculated between all 3 examiners were less than 0.05 at C0-3, C3-5, C5-T1, T1-4, T4-7, T7-10, L1-3 and L3-5; we therefore reject the null hypothesis at the 0.05 level of significance and so conclude that there is an association between the examiners at these collapsed spinal levels for static and motion palpation.

4.3.3.3 COHEN'S KAPPA COEFFICIENT

Cohen's Kappa (K) was used to measure the chance corrected concordance between the examiners. The kappa values were tested for significance at the α=0.05 level. The guidelines used for interpreting the kappa values in this study were in accordance with those employed by French et al. (2000), i.e.:

- < 0.00: poor agreement
- 0.00 – 0.20: slight agreement
- 0.21 – 0.40: fair agreement
- 0.41 – 0.60: moderate agreement

- 0.61 – 0.80: substantial agreement
- 0.81 – 1.00: almost perfect agreement

A. Cervical Spine

Table 37. Kappa Values for Static and Motion Palpation in the Cervical Spine

	Examiner A&B		Examiner A&C		Examiner B&C		MEAN	
	Kappa	**p-value**	**Kappa**	**p-value**	**Kappa**	**p-value**	**Kappa**	**p-value**
C0/C1	-0.02	0.838	0.315	0.001	-0.027	0.771	0.089	0.537
C1/C2	0.195	0.033	0.329	0	0.158	0.114	0.227	0.049
C2/C3	0.39	0	0.322	0.001	0.383	0	0.365	0.000
C3/C4	0.423	0	0.267	0.004	0.11	0.267	0.267	0.090
C4/C5	0.181	0.011	0.135	0.108	0.481	0	0.266	0.040
C5/C6	0.38	0	0.387	0	0.485	0	0.417	0.000
C6/C7	(a)	(a)	-0.014	0.886	(a)	(a)	-0.014	0.886
C7/T1	(a)	(a)	(a)	(a)	(a)	(a)	(a)	(a)
MEAN	0.258	0.147	0.249	0.143	0.265	0.192	0.231	0.229

(a) Values could not be calculated because one or all of the examiners did not have any positive assessments at these levels (i.e. no spinal joint dysfunction at these spinal levels).

Table 37 presents the Cohen's Kappa (K) scores and their p-values calculated for the cervical spine. Kappa scores calculated for examiners A and B in the cervical spine ranged from –0.02 (at C0/1) to 0.423 (at C3/4) which indicates a range of poor to moderate agreement between the examiners beyond chance. The overall (mean) Kappa score calculated for the cervical spine was 0.258, which indicates fair agreement beyond chance between these examiners for static and motion palpation in the cervical spine. Kappa scores were statistically significant ($p<0.05$) at all spinal levels in the cervical spine besides C0/1. Values could not be calculated at C6/7 and C7/T1 because one or all of the examiners did not have any positive assessments at these levels (i.e. no spinal joint dysfunction at these spinal levels).

Kappa scores calculated for examiners A and C in the cervical spine ranged from –0.014 (at C6/7) to 0.387 (at C5/6) which indicates a range of poor to fair agreement between the examiners beyond chance. The overall (mean) Kappa score calculated for the cervical spine was 0.249, which indicates fair agreement beyond chance between these examiners for static and motion palpation in the cervical spine. Kappa scores were statistically significant ($p<0.05$) at all spinal levels in the cervical spine besides C4/5 and C6/7. Values could not be calculated at C7/T1 because one or all of the examiners did not have any positive assessments at this level (i.e. no spinal joint dysfunction at this spinal level).

Kappa scores calculated for examiners B and C in the cervical spine ranged from -0.027 (at C0/1) to 0.485 (at C5/6), which indicates a range of poor to moderate agreement between the examiners beyond chance. The overall (mean) Kappa score calculated for the cervical spine was 0.265, which indicates fair agreement beyond chance between these examiners for static and motion palpation in the cervical spine. Kappa scores were statistically significant ($p<0.05$) at all spinal levels in the cervical spine besides C0/1, C1/2 and C3/4. Values could not be calculated at C6/7 and C7/T1 because one or all of the examiners did not have any positive assessments at this level (i.e. no spinal joint dysfunction at this spinal level).

Mean Kappa scores calculated for all examiners in the cervical spine ranged from –0.014 (at C6/7) to 0.417 (at C5/6) which indicates a range of poor to moderate agreement between the examiners beyond chance. The overall (mean) Kappa score calculated for all examiners across the cervical spine was 0.231, which indicates fair agreement beyond chance for static and motion palpation between the examiners. Kappa scores were statistically significant ($p<0.05$) at all spinal levels in the cervical spine besides C0/1, C3/4 and C6/7. Values could not be calculated at C7/T1 because one or all of the examiners did not have any positive assessments at these levels (i.e. no spinal joint dysfunction at these spinal levels).

B. Thoracic Spine

Table 38. Kappa Values for Static and Motion Palpation in the Thoracic Spine

	Examiner A&B		Examiner A&C		Examiner B&C		MEAN	
	Kappa	**p-value**	**Kappa**	**p-value**	**Kappa**	**p-value**	**Kappa**	**p-value**
T1/T2	(a)	(a)	(a)	(a)	(a)	(a)	(a)	(a)
T2/T3	(a)	(a)	(a)	(a)	(a)	(a)	(a)	(a)
T3/T4	1	0	0.273	0	0.273	0	0.515	0.000
T4/T5	0.12	0.201	-0.081	0.419	0.138	0.15	0.059	0.257
T5/T6	0.221	0.022	-0.048	0.558	0.071	0.444	0.081	0.341
T6/T7	0.422	0	0.204	0.039	0.336	0.001	0.321	0.013
T7/T8	0.658	0	0.16	0.102	0.227	0.008	0.348	0.037
T8/T9	(a)	(a)	-0.025	0.802	(a)	(a)	-0.025	0.802
T9/T10	1	0	-0.014	0.886	-0.014	0.886	0.324	0.591
T10/T11	0.49	0	-0.027	0.771	-0.027	0.771	0.145	0.514
T11/T12	(a)	(a)	(a)	(a)	0.158	0.114	0.158	0.114
T12/L1	0.273	0	-0.015	0.86	0.19	0.043	0.149	0.301
MEAN	0.523	0.028	0.047	0.493	0.150	0.269	0.208	0.297

(a) Values could not be calculated because one or all of the examiners did not have any positive assessments at these levels (i.e. no spinal joint dysfunction at these spinal levels).

Table 38 presents the Cohen's Kappa (K) scores and their p-values calculated for the thoracic spine. Kappa scores calculated for examiners A and B ranged from 0.12 (at T4/5) to 1 (at T3/4 and T9/10), which indicates a range of slight to almost perfect agreement between the examiners beyond chance. The overall (mean) Kappa score calculated here was 0.523, which indicates moderate agreement beyond chance between these examiners for static and motion palpation in the thoracic spine. Kappa scores were statistically significant ($p<0.05$) at all spinal levels in the thoracic spine besides T4/5. Values could not be calculated at T1/2, T2/3, T8/9 and T11/12 because one or all of the examiners did not have any positive assessments at these levels (i.e. no spinal joint dysfunction at these spinal levels).

Kappa scores calculated for examiners A and C ranged from –0.081 (at T4/5) to 0.273 (at T3/4), which indicates a range of poor to fair agreement between the examiners beyond chance. The overall (mean) Kappa score calculated here was 0.047, which indicates slight agreement beyond chance between these examiners for static and motion palpation in the thoracic spine. Kappa scores were statistically significant ($p<0.05$) at T3/4 and T6/7 while the Kappa scores at all other levels in the thoracic spine were not statistically significant ($p>0.05$). Values could not be calculated at T1/2, T2/3 and T11/12 because one or all of the examiners did not have any positive assessments at these levels (i.e. no spinal joint dysfunction at these spinal levels).

Kappa scores calculated for examiners B and C ranged from –0.027 (at T10/11) to 0.336 (at T6/7), which indicates a range of poor to fair agreement between the examiners beyond chance. The overall (mean) Kappa score calculated here was 0.150, which indicates slight agreement beyond chance between these examiners for static and motion palpation in the thoracic spine. Kappa scores were statistically significant ($p<0.05$) at all spinal levels in the thoracic spine besides T4/5, T5/6, T9/10, T10/11 and T11/12. Values could not be calculated at T1/2, T2/3 and T8/9 because one or all of the examiners did not have any positive assessments at these levels (i.e. no spinal joint dysfunction at these spinal levels).

Mean Kappa scores calculated for all examiners in the thoracic spine ranged from –0.025 (at T8/9) to 0.515 (at T3/4) which indicates a range of poor to moderate agreement between the examiners beyond chance. The overall (mean) Kappa score calculated for all examiners across the thoracic spine was 0.208, which indicates slight agreement beyond chance between the examiners for static and motion palpation. Overall Kappa scores were statistically significant ($p<0.05$) at T3/4, T6/7 and T7/8 while the overall Kappa scores at all other levels in the thoracic spine were not statistically significant ($p>0.05$). Values could not be calculated at T1/2 and T2/3 because one or all of the examiners did not have any positive assessments at these levels (i.e. no spinal joint dysfunction at these spinal levels).

C. Lumbar Spine and SI Joints

Table 39. Kappa Values for Static and Motion Palpation in the Lumbar Spine and Sacroiliac Joints

	Examiner A&B		Examiner A&C		Examiner B&C		MEAN	
	Kappa	**p-value**	**Kappa**	**p-value**	**Kappa**	**p-value**	**Kappa**	**p-value**
L1/L2	0.593	0	0.39	0	0.55	0	0.511	0.000
L2/L3	0.651	0	0.394	0	0.5	0	0.515	0.000
L3/L4	-0.014	0.886	0.174	0.027	0.208	0.001	0.123	0.305
L4/L5	0.492	0	0.313	0.002	-0.015	0.86	0.263	0.287
SI	0.712	0	0.236	0.009	0.261	0.002	0.403	0.004
MEAN	0.487	0.177	0.301	0.008	0.301	0.173	0.363	0.119

Table 39 presents the Cohen's Kappa (K) scores and their p-values calculated for the lumbar spine and sacroiliac (SI) joints. Kappa scores calculated for examiners A and B ranged from –0.014 (at L3/4) to 0.712(at the SI joints), which indicates a range of poor to substantial agreement between the examiners beyond chance. The overall (mean) Kappa score calculated here was 0.487, which indicates moderate agreement beyond chance between these examiners for static and motion palpation in the lumbar spine and SI joints. Kappa scores were statistically significant ($p<0.05$) at all spinal levels in the lumbar spine and SI joints besides L3/4.

Kappa scores calculated for examiners A and C ranged from 0.174 (at L3/4) to 0.394 (at L2/3), which indicates a range of slight to fair agreement between the examiners beyond chance. The overall (mean) Kappa score calculated here was 0.301, which indicates fair agreement beyond chance between these examiners for static and motion palpation in the lumbar spine and SI joints. Kappa scores were statistically significant ($p<0.05$) at all spinal levels in the lumbar spine and SI joints.

Kappa scores calculated for examiners B and C ranged from –0.015 (at L4/5) to 0.55 (at L1/2), which indicates a range of poor to moderate agreement between the examiners beyond chance. The overall (mean) Kappa score calculated here was 0.301, which indicates fair agreement beyond chance between these examiners for static and motion palpation in the lumbar spine and SI joints. Kappa scores were statistically significant ($p<0.05$) at all spinal levels in the lumbar spine and SI joints besides L4/5.

Mean Kappa scores calculated for all examiners in the lumbar spine and SI joints ranged from 0.123 (at L3/4) to 0.515 (at L1/2), which indicates a range of slight to moderate agreement between the examiners beyond chance. The overall (mean) Kappa score calculated for all examiners across the lumbar spine and SI joints was 0.363, which indicates fair agreement beyond chance between the examiners for static and motion palpation. Overall Kappa scores were statistically significant ($p<0.05$) at all levels in the lumbar spine and SI joints besides L3/4 and L4/5.

D. Collapsed Spinal Levels

Table 40. Kappa Values for Static and Motion Palpation in the Collapsed Spinal Segments

	Examiner A&B		Examiner A&C		Examiner B&C		MEAN	
	Kappa	**p-value**	**Kappa**	**p-value**	**Kappa**	**p-value**	**Kappa**	**p-value**
C0-C3	0.526	0	0.316	0.001	0.339	0.001	0.394	0.001
C3-C5	0.459	0	0.376	0	0.237	0.018	0.357	0.006
C5-T1	0.31	0	0.295	0.003	0.427	0	0.344	0.001
T1-T4	1	0	0.208	0.001	0.208	0.001	0.472	0.001
T4-T7	0.281	0.004	0.188	0.06	0.249	0.009	0.239	0.024
T7-T10	0.525	0	0.207	0.037	0.374	0	0.369	0.012
T10-L1	0.374	0	-0.044	0.629	0.078	0.433	0.136	0.354
L1-L3	0.556	0	0.338	0.001	0.491	0	0.462	0.000
L3-L5	0.265	0.003	0.237	0.013	0.154	0.041	0.219	0.019
MEAN	0.477	0.001	0.236	0.083	0.284	0.056	0.332	0.046

Table 40 presents the Cohen's Kappa (K) scores and their p-values calculated for the collapsed spinal levels. Kappa scores calculated for examiners A and B ranged from 0.265 (at L3-5) to 0.556 (at L1-3), which indicates a range of fair to moderate agreement between the examiners beyond chance. The overall (mean) Kappa score calculated here was 0.477, which indicates moderate agreement beyond chance between these examiners for static and motion palpation in the collapsed spinal levels. Kappa scores were statistically significant ($p<0.05$) at all the collapsed spinal levels.

Kappa scores calculated for examiners A and C ranged from -0.044 (at T10-L1) to 0.376 (at C3-5), which indicates a range of poor to fair agreement between the examiners beyond chance. The overall (mean) Kappa score calculated here was 0.236, which indicates fair agreement beyond chance between these examiners for static and motion palpation in the collapsed spinal levels. Kappa scores were statistically significant ($p<0.05$) at all collapsed spinal levels besides T4-7 and T10-L1.

Kappa scores calculated for examiners B and C ranged from 0.078 (at T10-L1) to 0.491 (at L1-3), which indicates a range of slight to moderate agreement between the examiners beyond chance. The overall (mean) Kappa score calculated here was 0.284, which indicates fair agreement beyond chance between these examiners for static and motion palpation in the collapsed spinal levels. Kappa scores were statistically significant ($p<0.05$) at all collapsed spinal levels besides T10-L1.

Mean Kappa scores calculated for all examiners in the collapsed spinal levels ranged from 0.136 (at T10-L1) to 0.472 (at T1-4), which indicates a range of slight to moderate agreement between the examiners beyond chance. The overall (mean) Kappa score calculated for all examiners across the collapsed spinal levels was 0.332, which indicates fair agreement beyond chance between the examiners for static and motion palpation. Overall Kappa scores were statistically significant ($p<0.05$) at all collapsed spinal levels besides T10-L1.

CHAPTER FIVE: DISCUSSION OF THE RESULTS

5.1 THE PREVALENCE OF SPINAL JOINT DYSFUNCTION

The results of this study reveal that the prevalence of spinal joint dysfunction in the cervical and thoracic spine in infants is lowest at the upper and lower segmental levels and highest in the mid segmental levels. This trend of a low prevalence of spinal joint dysfunction at transitional areas is consistent with the findings of French et al. (2000). French et al. (2000) found a low prevalence of manipulable lesions in the lower thoracic and upper lumbar spine. In the lumbar spine, the prevalence of spinal joint dysfunction is greatest at the upper two segmental levels. The prevalence of spinal joint dysfunction at the sacroiliac (SI) joints is greater than that at any other segmental level in the lower back.

A low prevalence of spinal joint dysfunction at certain levels reduces the usefulness of the Kappa values calculated at these levels due to limited variation (Haas 1991 - a). Kappa values for segmental levels with a low prevalence of spinal joint dysfunction should therefore be interpreted with caution.

The negative effect of limited variation on the results of this interexaminer reliability study adds to the reasons for analysing collapsed segmental levels. When segmental levels are collapsed, the prevalence of spinal joint dysfunction at each collapsed segment is increased. This makes the results of the Kappa statistic more realistic.

5.2 INTEREXAMINER RELIABILITY OF STATIC PALPATION

According to Haas (1991 - a), the statistic of choice when analysing nominal data for interexaminer reliability is the Cohen's Kappa statistic. Haas (1991 - a) further states that percentage agreement should also be included.

5.2.1 CERVICAL SPINE

With the Pearson's Chi-Square analysis of the data, we concluded that there was no association between the examiners at C0/1, C1/2 C3/4 and C6/7 and that there was an association between the examiners at C2/3, C4/5 and C5/6 for static palpation. Chi-square does not evaluate the strength of concordance between examiners nor is it able to distinguish a significant association predominated by agreement from one predominated by disagreement (Haas 1991 - a). We therefore rely more on the Kappa statistic for an indication of interexaminer reliability.

Mean Kappa scores calculated for all examiners in the cervical spine indicated a range of poor to moderate agreement between the examiners beyond chance. The overall (mean) Kappa score calculated for all examiners across the cervical spine indicated fair agreement beyond chance (K = 0.233) between the examiners for static palpation. The mean percentage agreement between the mean scores calculated for all 3 examiners ranged from 76% to 100% while the mean percentage agreement across the cervical spine was 89.5%.

5.2.2 THORACIC SPINE

With the Pearson's Chi-Square analysis of the data, we concluded that there was no association between the examiners at T4/5, T5/6, T8/9, T9/10, T10/11, T11/12 and T12/L1, and that there was an association between the examiners at T3/4, T6/7 and T7/8 for static palpation. Chi-square does not evaluate the strength of concordance between examiners nor is it able to distinguish a significant association predominated by agreement from one predominated by disagreement (Haas 1991 - a). We therefore rely more on the Kappa statistic for an indication of interexaminer reliability.

Mean Kappa scores calculated for all examiners in the thoracic spine indicated a range of poor to fair agreement between the examiners beyond chance. The overall (mean) Kappa score calculated for all examiners across the thoracic spine indicated slight agreement beyond chance (K = 0.2) between the examiners for static palpation. The mean percentage agreement between the mean scores calculated for all 3 examiners ranged from 87.333% to 100% while the mean percentage agreement across the thoracic spine was 93.667%.

5.2.3 LUMBAR SPINE AND SACROILIAC JOINTS

With the Pearson's Chi-Square analysis of the data, we concluded that there was no association between the examiners at L3/4 and L4/5, and that there was an association between the examiners at L1/2, L2/3 and the SI joints for static palpation. Chi-square does not evaluate the strength of concordance between examiners nor is it able to distinguish a significant association predominated by agreement from one predominated by disagreement (Haas 1991 - a). We therefore rely more on the Kappa statistic for an indication of interexaminer reliability.

Mean Kappa scores calculated for all examiners in the lumbar spine and SI joints indicated a range of slight to moderate agreement between the examiners beyond chance. The overall (mean) Kappa score calculated for all examiners across the lumbar spine and SI joints indicated fair agreement beyond chance (K = 0.353) between the examiners for static palpation. The mean percentage agreement between the mean scores calculated for all 3 examiners ranged from 90% to 96% while the mean percentage agreement across the lumbar spine and SI joints was 92.8%.

5.2.4 COLLAPSED LEVELS

With the Pearson's Chi-Square analysis of the data, we concluded that there was no association between the examiners at T7-10 and T10-L1, and that there was an association between the examiners at C0-3, C3-5, C5-T1, T1-4, T4-7, L1-3 and L3-5 for static palpation. Chi-square does not evaluate the strength of concordance between examiners nor is it able to distinguish a significant association predominated by agreement from one predominated by disagreement (Haas 1991 - a). We therefore rely more on the Kappa statistic for an indication of interexaminer reliability.

Mean Kappa scores calculated for all examiners in the collapsed spinal levels indicated a range of slight to moderate agreement between the examiners beyond chance. The overall (mean) Kappa score calculated for all examiners across the collapsed spinal levels indicated fair agreement beyond chance (K = 0.312) between the examiners for static palpation. The mean percentage agreement between the mean scores calculated for all 3 examiners ranged from 75.333% to 94% while the mean percentage agreement between these examiners across all the collapsed segments of the spine was 86%.

5.3 INTEREXAMINER RELIABILITY OF MOTION PALPATION

5.3.1 CERVICAL SPINE

With the Pearson's Chi-Square analysis of the data, we concluded that there was no association between the examiners at C0/1, C1/2, C2/3, C4/5 and C6/7 and that there was an association between the examiners at C3/4 and C5/6 for motion palpation. Chi-square does not evaluate the strength of concordance between examiners nor is it able to distinguish a significant association predominated by agreement from one predominated by disagreement (Haas 1991 - a). We therefore rely more on the Kappa statistic for an indication of interexaminer reliability.

Mean Kappa scores calculated for all examiners in the cervical spine indicated a range of slight to fair agreement between the examiners beyond chance. The overall (mean) Kappa score calculated for all examiners across the cervical spine indicated fair agreement beyond chance (K = 0.216) between the examiners for motion palpation. The mean percentage agreement between the mean scores calculated for all 3 examiners ranged from 73.333% to 100% while the mean percentage agreement for all examiners across the cervical spine was 87%.

5.3.2 THORACIC SPINE

With the Pearson's Chi-Square analysis of the data, we concluded that there was no association between the examiners at T3/4, T4/5, T5/6, T7/8, T8/9, T9/10, T10/11, T11/12 and T12/L1 and that there was an association between the examiners at T6/7 for motion palpation. Chi-square does not evaluate the strength of concordance between examiners nor is it able to distinguish a significant association predominated by agreement from one predominated by disagreement (Haas 1991 - a). We therefore rely more on the Kappa statistic for an indication of interexaminer reliability.

Mean Kappa scores calculated for all examiners in the thoracic spine indicated a range of slight to fair agreement between the examiners beyond chance. The overall (mean) Kappa score calculated for all examiners across the thoracic spine indicated slight agreement beyond chance (K = 0.140) between the examiners for motion palpation. The mean percentage agreement between the mean scores calculated for all 3 examiners ranged from 79.333% to 99.333% while the mean percentage agreement for all examiners across the thoracic spine for motion palpation was 90.111%.

5.3.3 LUMBAR SPINE AND SACROILIAC JOINTS

With the Pearson's Chi-Square analysis of the data, we concluded that there was no association between the examiners at L3/4 and L4/5 and that there was an association between the examiners at L1/2, L2/3 and the SI joints for motion palpation. Chi-square does not evaluate the strength of concordance between examiners nor is it able to distinguish a significant association predominated by agreement from one predominated by disagreement (Haas 1991 - a). We therefore rely more on the Kappa statistic for an indication of interexaminer reliability.

Mean Kappa scores calculated for all examiners in the lumbar spine and the SI joints indicated a range of slight to moderate agreement between the examiners beyond chance. The overall (mean) Kappa score calculated for all examiners across the lumbar spine and SI joints indicated fair agreement beyond chance (K = 0.306) between the examiners for motion palpation. The mean percentage agreement between the mean scores calculated for all 3 examiners ranged from 76% to 90.667% while the mean percentage agreement for all examiners across the lumbar spine and SI joints for motion palpation was 87.667%.

5.3.4 COLLAPSED LEVELS

With the Pearson's Chi-Square analysis of the data, we concluded that there was no association between the examiners at T1-4, T7-10, T10-L1 and L3-5 and that there was an association between the examiners at C0-3, C3-5, C5-T1, T4-7 and L1-3 for motion palpation. Chi-square does not evaluate the strength of concordance between examiners nor is it able to distinguish a significant association predominated by agreement from one predominated by disagreement (Haas 1991 - a). We therefore rely more on the Kappa statistic for an indication of interexaminer reliability.

Mean Kappa scores calculated for all examiners in the collapsed spinal levels indicated a range of slight to moderate agreement between the examiners beyond chance. The overall (mean) Kappa score calculated for all examiners across the collapsed spinal levels indicated fair agreement beyond chance (K = 0.268) between the examiners for motion palpation. The mean percentage agreement between the mean scores calculated for all 3 examiners ranged from 66.667% to 88.667%, while the mean percentage agreement between these examiners for motion palpation across all the collapsed segments of the spine was 79.778%.

5.4 INTEREXAMINER RELIABILITY OF STATIC AND MOTION PALPATION

5.4.1 CERVICAL SPINE

With the Pearson's Chi-Square analysis of the data, we concluded that there was no association between the examiners at C0/1, C3/4 and C6/7 and that there was an association between the examiners at C1/2, C2/3, C4/5 and C5/6 for static and motion palpation. Chi-square does not evaluate the strength of concordance between examiners nor is it able to distinguish a significant association predominated by agreement from one predominated by disagreement (Haas 1991 - a). We therefore rely more on the Kappa statistic for an indication of interexaminer reliability.

Mean Kappa scores calculated for all examiners in the cervical spine indicated a range of poor to moderate agreement between the examiners beyond chance. The overall (mean) Kappa score calculated for all examiners across the cervical spine indicated fair agreement beyond chance (K = 0.231) between the examiners for static and motion palpation. The mean percentage agreement between the mean scores calculated for all 3 examiners ranged from 77.333% to 100% while the mean percentage agreement across the cervical spine was 91%.

5.4.2 THORACIC SPINE

With the Pearson's Chi-Square analysis of the data, we concluded that there was no association between the examiners at T4/5, T5/6, T8/9, T9/10, T10/11, T11/12 and T12/L1 and that there was an association between the examiners at T3/4, T6/7 and T7/8 for static and motion palpation. Chi-square does not evaluate the strength of concordance between examiners nor is it able to distinguish a significant association predominated by agreement from one predominated by disagreement (Haas 1991 - a). We therefore rely more on the Kappa statistic for an indication of interexaminer reliability.

Mean Kappa scores calculated for all examiners in the thoracic spine indicated a range of poor to moderate agreement between the examiners beyond chance. The overall (mean) Kappa score calculated for all examiners across the thoracic spine indicated slight agreement beyond chance (K = 0.208) between the examiners for static and motion palpation. The mean percentage agreement between the mean scores calculated for all 3 examiners ranged from 88% to 100% while the mean percentage agreement across the thoracic spine was 94.556%.

5.4.3 LUMBAR SPINE AND SACROILIAC JOINTS

With the Pearson's Chi-Square analysis of the data, we concluded that there was no association between the examiners at L3/4 and L4/5 and that there was an association between the examiners at L1/2, L2/3 and the SI joints for static and motion palpation. Chi-square does not evaluate the strength of concordance between examiners nor is it able to distinguish a significant association predominated by agreement from one predominated by disagreement (Haas 1991 - a). We therefore rely more on the Kappa statistic for an indication of interexaminer reliability.

Mean Kappa scores calculated for all examiners in the lumbar spine and SI joints indicated a range of slight to moderate agreement between the examiners beyond chance. The overall (mean) Kappa score calculated for all examiners across the lumbar spine and SI joints indicated fair agreement beyond chance (K = 0.363) between the examiners for static and motion palpation. The mean percentage agreement between the mean scores calculated for all 3 examiners ranged from 90% to 96.667% while the mean percentage agreement across the lumbar spine and SI joints was 93.333%.

5.4.4 COLLAPSED LEVELS

With the Pearson's Chi-Square analysis of the data, we concluded that there was no association between the examiners at T10-L1 and that there was an association between the examiners at C0-3, C3-5, C5-T1, T1-4, T4-7, T7-10, L1-3 and L3-5 for static and motion palpation. Chi-square does not evaluate the strength of concordance between examiners nor is it able to distinguish a significant association predominated by agreement from one predominated by disagreement (Haas 1991 - a). We therefore rely more on the Kappa statistic for an indication of interexaminer reliability.

Mean Kappa scores calculated for all examiners in the collapsed spinal levels indicated a range of slight to moderate agreement between the examiners beyond chance. The overall (mean) Kappa score calculated for all examiners across the collapsed spinal levels indicated fair agreement beyond chance (K = 0.332) between the examiners for static and motion palpation. The mean percentage agreement between the mean scores calculated for all 3 examiners ranged from 77.333% to 95.333% while the mean percentage agreement between the examiners across all the collapsed segments of the spine was 87.556%.

5.5 Discussion of the Results

This study has produced overall (mean) results of slight to fair interexaminer reliability across all areas of the spine and between all three examiner pairs for static palpation, motion palpation and the combination of static and motion palpation.

5.5.1 CERVICAL SPINE

The results of all three methods of assessment are very similar in the cervical spine, with motion palpation producing Kappa values that are slightly lower than static palpation and static and motion palpation together. The mean Kappa statistics in the cervical spine for static, motion, and static and motion palpation used together all indicate fair agreement between the examiners beyond chance. The result of fair agreement between examiners in the cervical spine is not adequate to indicate that the assessment procedures used in this study are reliable. Fair interexaminer reliability leaves a vast area for improvement where the possible reasons for the inadequate reliability are discussed below.

5.5.2 THORACIC SPINE

The results of all three methods of assessment are very similar in the thoracic spine, with motion palpation producing Kappa values that are slightly lower than static palpation and static and motion palpation together. The mean Kappa statistics in the thoracic spine for static, motion, and static and motion palpation used together all indicate slight agreement between the examiners beyond chance. The result of slight agreement between examiners in the thoracic spine is not adequate to indicate that the assessment procedures used in this study are reliable. Slight interexaminer reliability leaves a vast area for improvement where the possible reasons for the inadequate reliability are discussed below.

5.5.3 LUMBAR SPINE AND SACROILIAC JOINTS

The results of all three methods of assessment are very similar in the lumbar spine and SI joints, with motion palpation producing Kappa values that are slightly lower than static palpation and static and motion palpation together. The mean Kappa statistics in the lumbar spine and SI joints for static, motion, and static and motion palpation used together all indicate fair agreement between the examiners beyond chance. The result of fair agreement between examiners in the lumbar spine and SI joints is not adequate to indicate that the assessment procedures used in this study are reliable. Fair interexaminer reliability leaves a vast area for improvement where the possible reasons for the inadequate reliability are discussed below.

5.5.4 COLLAPSED SPINAL LEVELS

The results of all three methods of assessment are very similar in the collapsed spinal segments, with motion palpation producing Kappa values that are slightly lower than static palpation and static and motion palpation together. The mean Kappa statistics in the collapsed spinal segments for static, motion, and static and motion palpation used together all indicate fair agreement between the examiners beyond chance. The result of fair agreement between examiners in the collapsed spinal segments is not adequate to indicate that the assessment procedures used in this study are reliable. Fair interexaminer reliability leaves a vast area for improvement where the possible reasons for the inadequate reliability are discussed below.

It is apparent that collapsing the data into two to three spinal segments has not resulted in improved results of interexaminer reliability. It was perceived that collapsing the data would improve the results of interexaminer reliability by reducing the examiner's inability to identify the exact spinal segment found to have spinal joint dysfunction. This examiner error could be explained when the small size of a two to ten week old infant's cervical, thoracic or lumbar spine is observed. The Kappa statistic at the SI joints produced results of fair (K = 0.403) to moderate (K = 0.432) agreement between the examiners. These results are an improvement on the mean results obtained from the cervical, thoracic and lumbar spine. This could perhaps be explained by the fact that the SI joints are easier to locate without error than the individual spinal segments which are very close together in a two to ten week old infant.

It was also perceived that collapsing the data would reduce the effect of limited variation on the results. It should be noted that it was possible to compute results for all collapsed spinal levels where this was not the case in the cervical and thoracic spines (due to limited variation).

5.5.5 STATIC, MOTION, AND STATIC AND MOTION PALPATION USED TOGETHER

The overall Kappa results computed in this study have demonstrated a trend between each assessment procedure tested for interexaminer reliability. Motion palpation has consistently produced slightly lower Kappa values as compared to static and static and motion palpation together. This small difference was not substantial and did not change the interpretation of the Kappa values in any of the spinal regions.

It is interesting to note that the results obtained in all assessment procedures were very similar and can therefore be said to be related in terms of the assessment for spinal joint dysfunction. This relationship between different examination procedures is not always apparent in the reliability studies. This aspect is evident in the study of Keating et al. (1990), where the aim of the study was to evaluate the interexaminer reliability of eight different assessment methods. The results of the different assessment procedures varied, which questioned the relationship between them.

5.5.6 POSSIBLE REASONS FOR INADEQUATE RELIABILITY

The results of this study indicate that the outcomes of static, motion and static and motion palpation analyzed together are slightly to fairly reproducible between examiners. These results indicate inadequate interexaminer reliability. This result does not contraindicate the use of static and motion palpation for the assessment of spinal joint dysfunction in infants aged two to ten weeks. Further research is required to determine whether the differences between the examiners of this study can be attributed to sources of error in the research design.

Possible reasons for the slight to fair interexaminer reliability may be numerous. Limited variation is a major contributor to poor Kappa results. Limited variation is evident where the prevalence of spinal joint dysfunction is low at a spinal segment and the Kappa statistic calculated at that segment is also particularly low. Limited variation should be suspected when the corresponding percentage agreement between the examiners is substantially high. When calculating the overall results, all Kappa statistics were given equal value (i.e. there was no weighting of any values). Significantly low values (perhaps due to limited variation) may have caused a substantial reduction in the overall results.

A low Kappa value with a corresponding high percentage agreement can be attributed to the fact that the Kappa statistic represents the proportion of agreement between examiners after agreement that is expected by chance is excluded. If most of the agreement between examiners is limited to only one of the possible choices for the assessment outcome (e.g. most spinal segments found to be negative for joint dysfunction), then many of the agreements between the examiners may be expected by chance. This would increase the proportion of chance agreement and so reduce the Kappa value resulting in a poorer level of interexaminer reliability.

The examiner's experience may have played a role in the results of this study, where their prior experience in the examination of infants was limited. This may have resulted in numerous examiner errors including the inability to correctly identify a specific spinal segment as previously discussed. The examiners may have had a different idea to each other as to what constituted a positive finding (e.g. how much resistance is required to constitute a restriction in movement). The actual examination may have been a contributing factor in examiner disagreement. The examination technique employed by the examiners to assess static and motion palpation may have been inadequate. Adequate rehearsals prior to the commencement of this study were not implemented with all participating examiners. Examiners may therefore have interpreted the examination protocol differently and thereby conducted it differently. This could have increased examiner disagreement.

The nature of a two to ten week old infant must be considered when interexaminer reliability is evaluated. It should be noted that two to ten week old infants do not keep still for the examination. This continuous movement makes it difficult to conduct an accurate assessment on the infants and so may have contributed to examiner error. The small size of a two to ten week old infant's spine must also be considered when evaluating the results. The tip of an examiner's finger could cover more than three vertebral levels at once. This may also have contributed to examiner error.

Another possible reason for the inadequate interexaminer reliability may be explained by its definition as stated by Haas (1995), "The reliability of a procedure is defined as the extent to which a repeated test will produce the same result when evaluating an unchanged characteristic.". An infant cannot be described as an unchanged characteristic. The repetitive examination of an infant's spine may lead to a change in the presence or absence of spinal joint dysfunction, or in its level of severity. This aspect of variability in the human spine may have contributed to the results of this study.

Chiropractic does boast the ability to manipulate specific spinal segments in the treatment of spinal joint dysfunction Haas et al. (1995 – a). It is yet to be established if this is actually the case because chiropractic research has not yet determined that the examination procedures adopted by the profession demonstrate sufficient interexaminer reliability. The importance of the application of manipulation to specific spinal segments as opposed to a spinal segment in the same region has not been established. This aspect could be demonstrated by controlled clinical trials producing good results when spinal manipulation is being used to treat different conditions in patients, and the examination procedure used in that trial to identify the manipulable lesion has not been found to be reliable. This poses the question of how important is spinal specificity if spinal manipulation is producing good results in the treatment of conditions such as infantile colic ($p = 0.04$; Wiberg et al. 1999).

CHAPTER SIX: CONCLUSIONS AND RECOMMENDATIONS

6.1 CONCLUSIONS

6.1.1 Interexaminer Reliability of Static, Motion, and Static and Motion Palpation Used Together

From the results of this study, the following conclusions can be drawn:

- The interexaminer reliability of *static palpation* used to assess for spinal joint dysfunction in healthy infants aged two to ten weeks is as follows: in the cervical spine, lumbar spine and sacroiliac joints, and in the collapsed spinal segments, interexaminer reliability is fair; while in the thoracic spine, interexaminer reliability is slight.
- The interexaminer reliability of *motion palpation* used to assess for spinal joint dysfunction in healthy infants aged two to ten weeks is as follows: in the cervical spine, lumbar spine and sacroiliac joints, and in the collapsed spinal segments, interexaminer reliability is fair; while in the thoracic spine, interexaminer reliability is slight.
- The interexaminer reliability of *static and motion palpation* used to assess for spinal joint dysfunction in healthy infants aged two to ten weeks is as follows: in the cervical spine, lumbar spine and sacroiliac joints, and in the collapsed spinal segments, interexaminer reliability is fair; while in the thoracic spine, interexaminer reliability is slight.

It can be concluded that the outcomes of static palpation and static and motion palpation used together are fairly reproducible between the examiners used in this study, and that the outcomes of motion palpation are only slightly reproducible between the examiners used in this study. The inference of these conclusions is that static and motion palpation can be seen by practitioners as able to provide fairly reliable information regarding the presence or absence of spinal joint dysfunction in the spines of two to ten week old infants. The result of "fairly reliable" does not reflect acceptable interexaminer reliability. This result does demonstrate that static and motion palpation are not useless assessment tools used for the identification of spinal joint dysfunction in infants aged two to ten weeks. The reasons for this are as follows: Firstly, the research methodology employed in this study was not flawless and the results may therefore be questionable. Secondly, the results of "slightly to fairly reliable" do signal that static and motion palpation do have some reliability. It should also be considered that the percentage agreement between examiners in this study was good. Static and motion palpation are therefore not useless assessment tools. Further research into this area is therefore warranted.

6.2 RECOMMENDATIONS

- If this study were to be repeated there are a number of issues that must be considered.
  - It is recommended that the sample of infants be more clearly defined in terms of their health status. The inclusion criteria of "only healthy infants", was contradicted by another inclusion criteria of "mild conditions such as infantile colic". It should have been clearly explained that the term "healthy" actually should have read "essentially healthy". The terms "healthy" and "symptomatic" should be clearly defined when designing such a study and when dealing with spinal joint dysfunction. The term "healthy" was used in this study mainly for ethical purposes. It should also be considered that it is an assumption that symptomatic infants would have an increased prevalence of spinal joint dysfunction. The sample should have been categorized into groups of either healthy infants, or those with mild conditions such as infantile colic. The analysis of this data may have produced more meaningful results.

- It is recommended that a similar study be conducted with some improvements on the research design.

- It is recommended that the sample population of infants be more representative of the clinical setting (i.e. use symptomatic infants; e.g. infants with infantile colic). This may result in an increased prevalence of spinal joint dysfunction, which may also reduce the effects of limited variation. It is an assumption that symptomatic infants would have an increased prevalence of spinal joint dysfunction that also requires verification through research.
- It is recommended that examiners with more experience in the examination of infants be used in the study. This may reduce the possibility of examiner error.
- It is recommended that examiners conduct their assessment out of the view of other examiners. This would ensure more efficient blinding between examiners and so reduce the possibility of observer bias affecting the results of the study.
- It is recommended that the author of the research does not participate in the study as an examiner because participation may introduce an element of bias into the study.
- It is recommended that extensive rehearsals of the research procedure be conducted prior to the commencement of the study. It is important that all participating examiners are well rehearsed in the examination protocol and have a uniform approach to the assessment of infants. Examiners must conduct the assessment in exactly the same way. It is also important for examiners to be in agreement with each other in terms of the exact constitution of a positive finding of spinal joint dysfunction in an infant.
- It is recommended that a similar study only focus on one region of the spine. This would reduce the extensive nature of the statistical analysis and so simplify the interpretation of the results. It should be noted that reliability studies conducted on adults usually test reliability in only one region of the spine per study.
- In the analysis of the data in this study, all positive findings were given an equal score of one point. It is recommended that when the scores are

pooled to determine the interexaminer reliability of static and motion palpation used together, a weighted score of two points be allocated to the positive finding of restricted motion and only one point to the other positive findings (tenderness, asymmetry and spasm). This will provide a more realistic presentation of spinal joint dysfunction as described by the working definition of spinal joint dysfunction stated in this research study.

- Cohen's Kappa is the statistic of choice for analysis of nominal data to determine interexaminer reliability. It is therefore recommended that the use of the Pearson's chi-square statistic be excluded. This would reduce the number of statistical tests needed to analyze the data and so simplify the presentation of the results.
- It is recommended that the findings of the study be compared with the findings of the admittedly very few studies on infants and spinal joint dysfunction. Prevalence and patterns of spinal joint dysfunction may be compared.

▪ It is also recommended that once adequate reliability is established, future research be conducted into the validity of static and motion palpation in terms of the signs of spinal joint dysfunction.

REFERENCES

Anrig C. and Plaugher G. eds 1998. Pediatric Chiropractic. Baltimore, USA: Williams & Wilkins.

Bergmann, T.F., Peterson, D.H. and Lawrence, D.J. 1993. Chiropractic Technique: Principles and Procedures. New York, USA: Churchill Livingstone.

Boline, P.D., Haas, M., Meyer, J.J., Kassak, K., Nelson, C. and Keating, J.C. 1993. Interexaminer Reliability of Eight Evaluative Dimensions of Lumbar Segmental Abnormality: Part II. Journal of Manipulative and Physiological Therapeutics. 16(6): 363-374.

Carmichael, J.P. 1987. Inter- and Intra-Examiner Reliability of Palpation for Sacroiliac joint Dysfunction. Journal of Manipulative and Physiological Therapeutics. 10: 164-171.

Fallon, J.M. 1997. The Role of the Chiropractic Adjustment in the Care and Treatment of 332 Children with Otitis Media. Journal of Clinical Chiropractic Paediatrics. 2(2): 167-183.

Fallon, J.M. and Fysh, P.N. 1997. Chiropractic Care of the Newborn with Congenital Torticollis. Journal of Clinical Chiropractic Paediatrics. 2(1): 116-121.

Fisher, L.D. and van Belle, G. 1993. Biostatics, A Methodology for the Health Sciences. U.S.A.: Wiley-Interscience.

Fjellner, A., Bexander, C., Faleij, R. and Strender, L. 1999. Interexaminer Reliability in Physical Examination of the Cervical Spine. Journal of Manipulative and Physiological Therapeutics. 22(8): 511-516.

French, S.D., Green, S. and Forbes, A. 2000. Reliability of Chiropractic Methods Commonly Used to Detect Manipulable Lesions in Patients with Chronic Low-Back pain. Journal of Manipulative and Physiological Therapeutics. 23(4): 231-238.

Gatterman, M.I. 1990. Chiropractic Management of Spine Related Disorders. Baltimore: Williams and Wilkins.

Haas, M. 1991. (a) Statistical Methodology for Reliability Studies. Journal of Manipulative and Physiological Therapeutics. 14(2): 119-132.

Haas, M. 1991. (b) Interexaminer Reliability for Multiple Diagnostic Test Regimens. Journal of Manipulative and Physiological Therapeutics. 14(2): 95-103.

Haas, M. 1991. (c) The Reliability of Reliability. Journal of Manipulative and Physiological Therapeutics. 14(3): 199- 208.

Haas, M. 1995. How to Evaluate Intraexaminer Reliability Using an Interexaminer Reliability Study Design. Journal of Manipulative and Physiological Therapeutics. 18(1): 10-15.

Haas, M., Panzer, D., Peterson, D. and Raphael, R. 1995. (a) Short-Term Responsiveness of Manual Thoracic End-Play Assessment to Spinal manipulation: A Randomized Controlled Trial of Construct Validity. Journal of Manipulative and Physiological Therapeutics. 18(9): 582-589.

Haas, M., Raphael, R., Panzer, D. and Peterson, D. 1995. (b) Reliability of Manual End-Play Palpation of the Thoracic spine. Chiropractic Technique. 7(4): 120-124.

Hawk, C., Phongphua, C., Bleecker, J., Swank, L., Lopez, D. and Rubley, T. 1999. Preliminary Study of the Reliability of Assessment Procedures for Indications for Chiropractic Adjustments of the Lumbar Spine. Journal of Manipulative and Physiological Therapeutics. 22(6): 382-389.

Holtrop, D.P, 2000. Resolution of Suckling Intolerance in a 6-Month-Old Chiropractic Patient. Journal of Manipulative and Physiological Therapeutics. 23(9): 615-618.

Hubka, M.J. and Phelan, S.P. 1994. Interexaminer Reliability of Palpation for Cervical Spine Tenderness. Journal of Manipulative and Physiological Therapeutics. 17(9): 591-595.

Keating, J.C., Bergmann, T.F., Jacobs, G.E., Finer, B.A. and Larson, K. 1990. Interexaminer Reliability of Eight Evaluative Dimensions of Lumbar Segmental Abnormality. Journal of Manipulative and Physiological Therapeutics. 13(8): 463-470.

Klougart, N., Nilsson, N. and Jacobsen, J. 1989. Infantile Colic Treated by Chiropractors: A Prospective Study of 316 Cases. Journal of Manipulative and Physiological Therapeutics. 12(4): 281-288.

Lee, A.C.C., Li, D.H. and Kemper, K.J. 2000. Chiropractic Care For Children. Arch Pediatr Adolesc Med. 154: 401-407.

Lewit, K. and Liebenson, C. 1993. Palpation- Problems and Implications. Journal of Manipulative and Physiological Therapeutics. 16(9): 586-590.

Mercer, C.A. 1999. A Study to Determine the Efficacy of Chiropractic Spinal Adjustments as a Treatment Protocol in the Management of Infantile Colic. A dissertation submitted to the Faculty of Health in partial compliance with the requirements for the M.Tech: Chiropractic (SA), Technikon Natal, Durban.

Merenstein, G.B., Kaplan, D.W. and Rosenberg, A.A. 1994. Handbook of Pediatrics. 17th ed. Connecticut, USA; Appleton and Lange.

Mootz, R.D., Keating, J.C., Kontz, H.P., Milus, T.B. and Jacobs, G.E. 1989. Intra- and Interobserver Reliability of Passive Motion Palpation of the Lumber Spine. Journal of Manipulative and Physiological Therapeutics. 12(6): 440-445.

Nansel, D.D., Peneff, A.L., Jansen, R.D. and Cooperstein, R. 1989. Interexaminer Concordance in Detecting Joint-Play Asymmetries in the Cervical Spines of Otherwise Asymptomatic Subjects. Journal of Manipulative and Physiological Therapeutics. 12(6): 428-433.

Olafsdottir E, Forshei S, Fluge G and Markestad T. 2001. Randomised controlled trial of infantile colic treated with chiropractic spinal manipulation. Arch Dis Child. 84(2): 138-41.

Russell, R. 1983. Diagnostic Palpation of the Spine: A Review of Procedures and Assessment of their Reliability. Journal of Manipulative and Physiological Therapeutics. 6(4): 181-183.

Schafer, R.C. and Faye, L.J. 1990. Motion Palpation and Chiropractic Technic – Principles of Dynamic Chiropractic. 2nd edition. California, USA: The Motion Palpation Institute.

Thomas, K. N. 2002. Statistician. Personal communication to, J. Ralph, June 2002.

Wiberg, J.M., Nordsteen, J. and Nilsson, N. 1999. The Short-Term Effect of Spinal Manipulation in the Treatment of Infantile Colic: A Randomized Controlled Clinical Trial With a Blinded Observer. Journal of Manipulative and Physiological Therapeutics. 22(8): 517-522.

Appendix A: List of Participating Clinics

Clinic Name	Region	Contact	Tel No:	Cell No:
Public Primary Health Care Clinics: Chappel Street Clinic New Germany Clinic Reservoir Hills Clinic Pinetown Clinic Dassenhoek Clinic Mpola Clinic	Innerwest, Pinetown	Sister Kate Koekemoer	7182621	
Cheryl's Gentle Births	Westville	Cheryl Rowe	2669952	0832888203
Westville Hospital Cradle Club	Westville	Des Meyer	2650911	
The Baby Wise Clinic	Winston Park	Margie Ireland	7674008	0824635687
People's Chemist: Preventative and Promotive Medicine	Westville	Elizabeth de Chazal,	2665366, 2089081	0827843409
Gentle Touch Clinic	Hillcrest	Arlette Blaylock	7655460	0837655460
Parklands Hospital: The Mom and Babycare Clinic	Berea	Judy Edy	5644916	0827737533
The Mothers Clinic	La Lucia	Sheila Francis	5628933	0823768368
Antenatal Clinic	Gillits	Glenda Patterson	7672077	0731418658

Appendix B: Poster

IS YOUR INFANT HEALTHY AND BETWEEN THE AGES OF TWO TO TEN WEEKS?

CHIROPRACTIC RESEARCH IS CURRENTLY BEING CONDUCTED INTO THE EXAMINATION OF A BABY'S SPINE.

THE EXAMINATION WILL BE LOOKING FOR LACK OF MOVEMENT IN THE BABY'S SPINE WHICH IS OFTEN CAUSED BY MUSCLE SPASM.

IF YOUR BABY IS HEALTHY AND BETWEEN THE AGES OF TWO TO TEN WEEKS, HE OR SHE MAY BE INCLUDED IN THIS RESEARCH.

THE PROCEDURE WILL INVOLVE FEELING YOUR BABY'S SPINE AND MUSCLES FOR SIGNS OF SPASM AND LIMITED MOVEMENT.

NO MANIPULATION OR TREATMENT WILL BE GIVEN TO YOUR BABY DURING THE RESEARCH.

IF YOUR BABY HAS ANY ABNORMAL FINDINGS OF THE SPINE, HE OR SHE WILL RECEIVE A REFERRAL TO AN APPRORIATE HEALTH-CARE PROVIDER.

IF YOUR BABY DOES HAVE INFANTILE COLIC, HE OR SHE WILL BE OFFERED

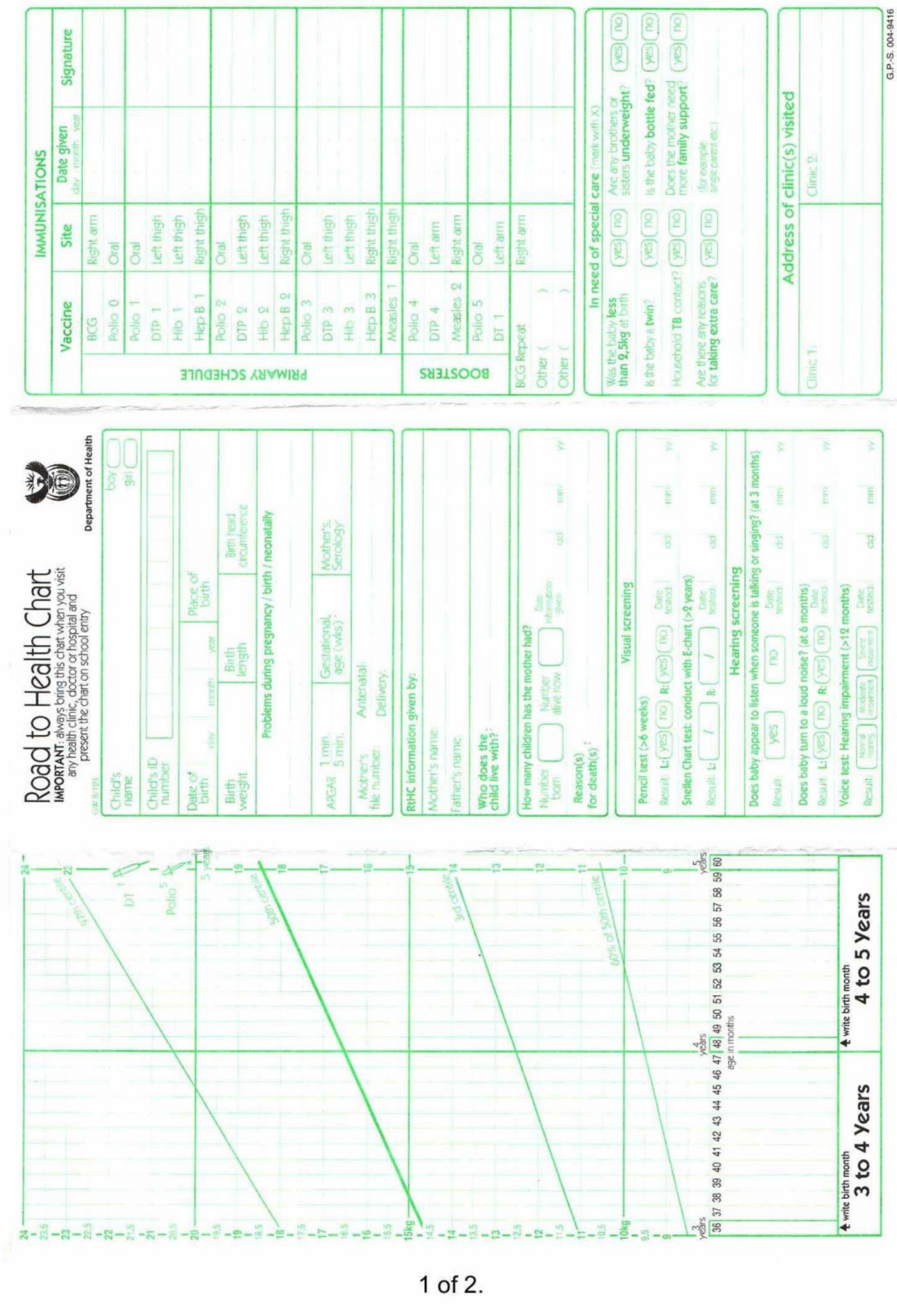
FOUR FREE TREATMENTS

OVER A TWO WEEK PERIOD AT THE DURBAN INSTITUTE OF TECHNOLOGY'S CHIROPRACTIC DAY CLINIC.

IF YOU ARE INTERESTED IN YOUR BABY PARTICIPATING IN THIS STUDY PLEASE CONTACT THE FOLLOWING PEOPLE:

JULEE RALPH 083 6544 993
CAROLINE VAN LINGEN 082 4693 143

Appendix C : Road to Health Chart



Road to Health Chart

IMPORTANT: always bring this chart when you visit any health clinic, doctor or hospital and present the chart on school entry

Department of Health

Child's name		boy / girl	
Child's ID number			
Date of birth	day month year	Place of birth	
Birth weight	Birth length	Birth head circumference	
Problems during pregnancy / birth / neonatally			
APGAR 1 min. 5 min.	Gestational age (wks)	Mother's Serology	
Mother's file number	Antenatal: Delivery:		

RtHC information given by:

Mother's name:

Father's name:

Who does the child live with?

How many children has the mother had? Number born / Number alive now / Date information given

Reason(s) for death(s):

Visual screening

Pencil test (>6 weeks) Result: L: yes no R: yes no Date tested dd mm yy

Snellen Chart test: conduct with E-chart (>2 years) Result: L: / R: / Date tested dd mm yy

Hearing screening

Does baby appear to listen when someone is talking or singing? (at 3 months) Result: yes no Date tested dd mm yy

Does baby turn to a loud noise? (at 6 months) Result: L: yes no R: yes no Date tested dd mm yy

Voice test: Hearing impairment (>12 months) Result: Normal hearing / Moderate impairment / Severe impairment Date tested dd mm yy

IMMUNISATIONS

	Vaccine	Site	Date given (day month year)	Signature
	BCG	Right arm		
	Polio 0	Oral		
PRIMARY SCHEDULE	Polio 1	Oral		
	DTP 1	Left thigh		
	Hib 1	Left thigh		
	Hep B 1	Right thigh		
	Polio 2	Oral		
	DTP 2	Left thigh		
	Hib 2	Left thigh		
	Hep B 2	Right thigh		
	Polio 3	Oral		
	DTP 3	Left thigh		
	Hib 3	Left thigh		
	Hep B 3	Right thigh		
	Measles 1	Right thigh		
BOOSTERS	Polio 4	Oral		
	DTP 4	Left arm		
	Measles 2	Right arm		
	Polio 5	Oral		
	DT 1	Left arm		
	BCG Repeat	Right arm		
	Other ( )			
	Other ( )			

In need of special care (mark with X)

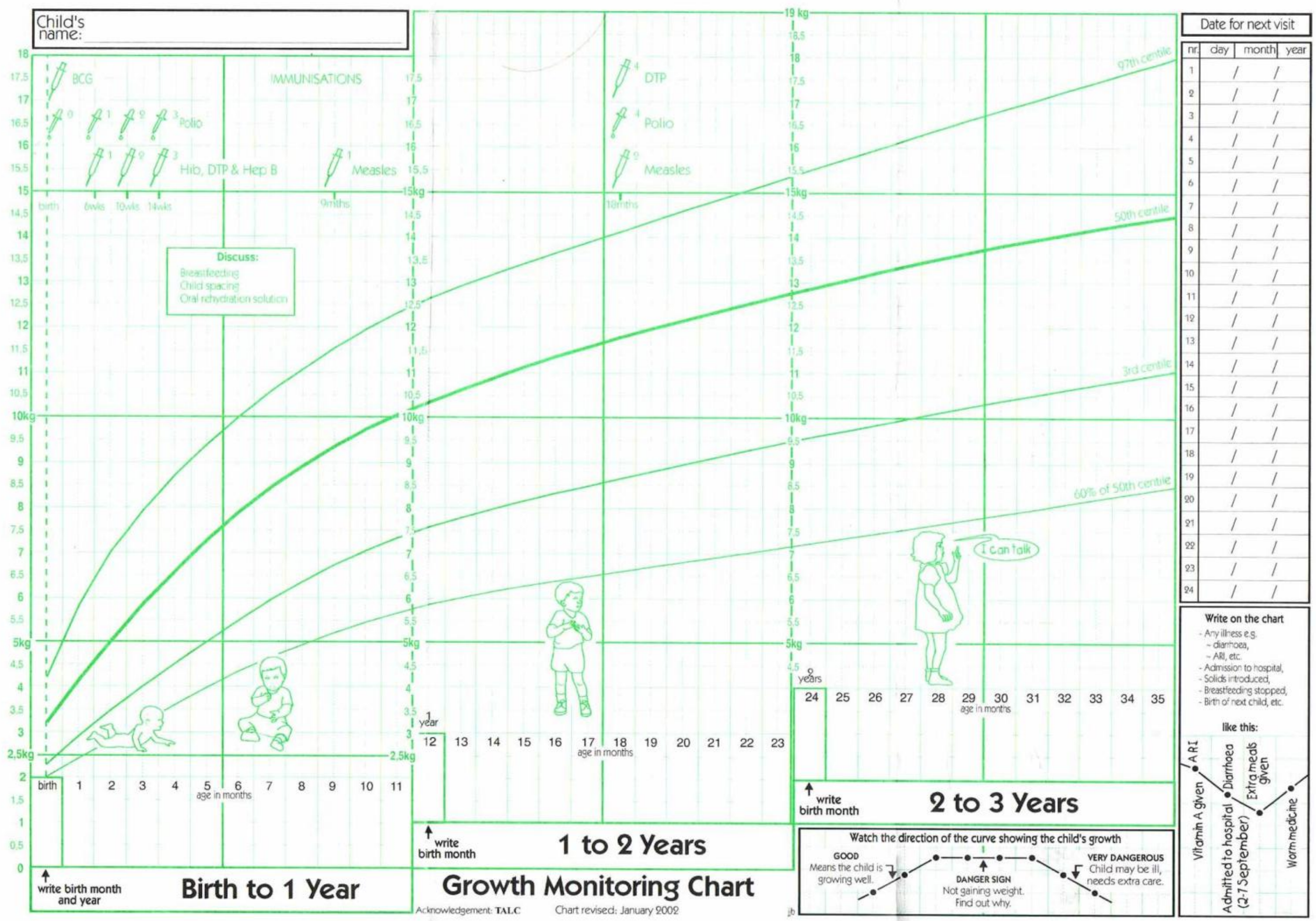
Was the baby less than 2,5kg at birth	yes no	Are any brothers or sisters underweight?	yes no
Is the baby a twin?	yes no	Is the baby bottle fed?	yes no
Household TB contact?	yes no	Does the mother need more family support? (for example: single parent etc.)	yes no
Are there any reasons for taking extra care?	yes no		

Address of clinic(s) visited

Clinic 1:

Clinic 2:

G.P.-S. 004-9416


Growth Monitoring Chart
Child's name:
IMMUNISATIONS
BCG
Polio
Hib, DTP & Hep B
Measles
DTP
Polio
Measles
birth
6wks
10wks
14wks
9mths
18mths
Discuss:
Breastfeeding
Child spacing
Oral rehydration solution
97th centile
50th centile
3rd centile
60% of 50th centile
I can talk
age in months
write birth month and year
Birth to 1 Year
write birth month
1 to 2 Years
write birth month
2 to 3 Years
Watch the direction of the curve showing the child's growth
GOOD
Means the child is growing well.
DANGER SIGN
Not gaining weight.
Find out why.
VERY DANGEROUS
Child may be ill, needs extra care.
Date for next visit
nr. day month year
Write on the chart
- Any illness e.g.
- diarrhoea,
- ARI, etc.
- Admission to hospital,
- Solids introduced,
- Breastfeeding stopped,
- Birth of next child, etc.
like this:
Vitamin A given
ARI
Admitted to hospital (2-7 September)
Diarrhoea
Extra meals given
Worm medicine
Acknowledgement: TALC
Chart revised: January 2002

Appendix D: English Letter of information

DURBAN INSTITUTE *of* TECHNOLOGY

Dear Parent

Your baby is invited to participate in a research project to determine whether three examiners are able to identify the same findings when examining a baby's spine. The process of examination will involve feeling your baby's spine, while it is kept still and while it is moved about, looking for any signs of muscle tightness, limited movement and associated pain. It is important that the tools used to find problems in a baby's spine are researched and are found to be accurate.

The researcher is currently a senior intern at the Durban Institute of Technology Chiropractic Day Clinic (formerly known as Technikon Natal). This research project forms part of a Master's Degree in Chiropractic. The number of babies participating in this study will be one hundred, where all babies must be healthy and between the ages of two and ten weeks.

The research procedure is very simple and will only take a short period of your time. There is no charge for participating in the study and your baby is not obliged to take part. The assessment that your baby will undergo is safe and risk-free. If you do decide to allow your baby to participate in this study, you will be required to sign an informed consent form. This will give us written permission to examine your baby and means that you understand the research procedure and are voluntarily allowing your baby to participate in the study.

The research procedure will begin with one examiner gently feeling your baby's spine for any signs of muscle tightness, limited movement and pain and then recording her findings on a sheet of paper. Two other examiners will then assess your baby in exactly the same way as the first examiner and record their findings on a separate sheet of paper. Your baby's participation in the study will then be complete. The researcher will then compare the findings of each examiner to see how similar they are. The personal information collected will remain confidential.

The research assessment will serve as a screening procedure where infants with any abnormal findings will receive a referral to an appropriate health-care provider. If your baby is found to have infantile colic, he or she will be offered four free treatments over a two week period at the Durban Institute of Technology's Chiropractic Day Clinic.

It is important that you understand that your baby's participation is voluntary and that you are under no obligation what so ever to allow your baby to take part in this research. If you do not want your baby to participate, or if you feel uncomfortable with any of the research procedures, you are free to withdraw from taking part in the study at any stage without fear of any negative consequences.

If you have any concerns or wish to lodge a complaint, please contact my supervisor or co-supervisor at the under-mentioned contact details. If you have any queries or would like to communicate with the relevant authorities (see below) at the Durban Institute of Technology's Department of Chiropractic, who can inspect the study's records, you may contact them at any stage.

Your baby's participation in this research study will aid in contributing knowledge about the reliability of our spinal examination in infants.

Thank you for your interest in this research project.

Yours Sincerely
Julee Ralph
Chiropractic Intern

Durban Institute of Technology Department of Chiropractic:

Principal Investigator:	Julee Ralph	0836544993
Co-Investigator:	Caroline van Lingen	0824693143
Research Supervisor:	Dr H. Kretzmann	(031) 2042205
Research Co-supervisor:	Dr J. Shaik	(031) 2042094
Chiropractic Clinic Director:	Dr C. Myburgh	(031) 2042094
Head of Department:	Mrs K. Roodt	(031) 2042094

Appendix D: Zulu Letter Of Information

Mzali

Umntanakho uyamenywa ukuba abe yingxenye yenhlolovo ukuze kubonakale ukuthi abacwaningi ababili bangathola yini umphumela ofanayo uma behlola umhlandla womntwana. Ekuhlolweni komhlandla womntwana, uzothintwathintwa upotopotozwe ngesineke, kancane umntwana angahlukumezeki. Lokhu kuzokwenziwa ngenkathi elele ethule kubuye kwenziwe futhi uma enyakaza, ngenhloso yokuthola ukuthi ukhona yini umsipha olukhuni ongenza ukuba umntwana anganyakazi kahle noma ezwe ubuhlungu uma enyakaza.

Umcwaningi walenhlolovo ungumfundi oqeqeshiwe oseqophelweni eliphezulu ophikweni lwezobuchwepheshe ekliniki ehlola iphinde izame ukuthola indlela yokwelapha izigulo ezibangwa ukuphazamiseka kokuxhumana kwamathambo omgogodla Lenhlolovo iyinxenye yokuhlolwa okufanele lomcwaningi ayenze emfundweni yakhe iphakeme (Master's Degree) ngokuhlola umuntu onganyakazi aphinde amhlole noma enyakaza ngendlela yokupotoza ngenhloso yokuthola ukuthi kukhona yini ukuphazamiseka noma ukuqina kwemisipha, egxile ekuhloleni umgogodla. Lababantwana bazoba yikhulu sebebonke. Bonke bazobe bangabantwana abaphile kahle uma ubabuka, bephakathi wamasonto amabili kuya kwayishumi okuzalwa.

Indlela ezolandelwa imfushane, izosheshe futhi. Awuzukhokhiswa mali ngokubandakanyeka komntanakho kuloluhlelo, kanti futhi awuphoqiwe ukuba yinxenye yalenhlolovo. Ukuhlolwa komntanakho ngeke kumhlukumeze, kuphephile, futhi akukho ngozi engamvelela. Uma uvuma ukuthi umntanakho abe yinxenye yalenhlolovo, uzocelwa ukuba usayine ifomu eshoyo ukuthi uyavuma. Lokhu kuzosinikeza imvume yokuhlola umntanakho, nokuthi uyaqonda ukuthi kwenziwani kulenhlolovo, futhi uyavuma ukuthi awuphoqiwe ukuletha umntanakho ukuthi abambe iqhaza kulenhlolovo.

Kuzoqala ngokuba umhloli apotopotoze, ngesineke, umhlandla womntwana ukuze ezwe ukuthi ikhona yini imisipha eqinile, noma ukuphazamiseka kokunyakaza, noma inkomba yobuhlungu uma epotoza, bese ebhala phansi ephepheni lokho akubonayo. Nowesibili umhloli uzolandela enze njengowokuqala, naye abhale phansi akutholayo kwelinye iphepha. Kobe sekuqediwe njalo. Umcwaningi uyobe esehlanganisa okutholwe yilababahloli, aqhathanise ukuthi babona izinto ezifanayo yini noma qha. Okuyotholakala ngalababantwana kuyimfihlo. Angeke kwaziswe wonke umuntu ngaphandle komzali womntwana.

Uma kutholakala ukuphazamiseka noma ukuqina kwemisipha kumntwana noma ukunganyakazi kahle noma ubuhlungu uma ethintwa noma enyakaza emhlandleni, okubonakala ukuthi angasizwa ngokutholakala kwemishanguzo yokuphazamiseka komhlandla noma kwemisipha, umntwana uyokwelashwa mahhala izikhathi ezine ekliniki i-Durban Institue of Technology Chiropractic Day Clinic.

Kusemqoka ukuba uqonde ukuthi ukuhlolwa komntwana akusiyo impoqo. Awuphoqiwe ukuletha umntwana wakho azoba yinxenye yalenhlolovo. Uma kungukuthi awuthandi noma awukhululeki kahle ngalenqubo ezolandelwa, ukhululekile ukumkhipha umntanakho, ungesabi ukuthi kukhona okubi okungenziwa kuwe.

Uma unemibuzo noma ufisa ukuthintana nabaphathi besikhungo senfundo yobunzululwazi, ophikweni locwaningo ngemisipha, okuyibona abangahlola amarekhodi ocwaningo noma ngabe kukusiphi isigaba socwaningo, ungabathola kulezizindawo ezibhalwe ngezansi. Uma unezikhalo, ufuna ukuziqondisa ophikweni lwenfundo yobunzululwazi eThekwini, ungathintana nomphathi wami noma umsizi wakhe kulezizindawo ezibhalwe ngezansi.

Ukuhlolwa komntanakho kulenhlolovo, kuzosiza ukuba sazi ukuthi ukuhlolwa komhlandla kubantwana kuthembeke kangakanani.

Siyabonga ukukhombisa kwakho uthando kulenhlolovo.

Yimina
Julee Ralph
Chiropractic Intern

i-Durban Institute of Technology - i-Department of Chiropractic:

Principal Investigator:	Julee Ralph:	083 6544 993
Co-investigator:	Caroline van Lingen:	082 4693 143
Research Supervisor:	Udokotela H. Kretzmann:	(031) 204 2205
Research Co-supervisor:	Udokotela J. Shaik:	(031) 204 2094
Chiropractic Clinic Director:	Udokotela Myburgh:	(031) 204 2094
Head of Department:	Inkosikazi K. Roodt:	(031) 204 2094

Appendix E: English Informed Consent Form

DURBAN INSTITUTE of TECHNOLOGY

INFORMED CONSENT FORM

Date: ______________________________

<u>Title of research project:</u>
The interexaminer reliability of static and motion palpation for the assessment of spinal joint dysfunction in infants aged two to ten weeks.

Name of supervisor:	**Dr Kretzmann**	**(031) 2042205**
Name of co-supervisor:	**Dr Shaik**	**(031) 2042094**
Name of research student:	**Julee Ralph**	**0836544993**

<u>Please circle the appropriate answer</u> YES NO

1. Have you read the research information sheet?	Yes	No
2. Have you had an opportunity to ask questions regarding this study?	Yes	No
3. Have you received satisfactory answers to your questions?	Yes	No
4. Have you had an opportunity to discuss this study?	Yes	No
5. Have you received enough information about this study?	Yes	No
6. Do you understand the implications of your involvement in this study?	Yes	No
7. Do you understand that you are free to withdraw from this study? at any time without having to give any a reason for withdrawing, and without affecting your future health care.	Yes	No
8. Do you agree to voluntarily participate in this study	Yes	No
9. Who have you spoken to? ______________________________		

Please ensure that the researcher completes each section with you.
If you have answered NO to any of the above, please obtain the necessary information before signing.

<u>Please Print in block letters:</u>

Infant's Name:______________________________

Parent's Name:______________________________Signature:___________

Parent's Relationship to infant:_________________

Witness Name:______________________________Signature:___________

Research Student Name:______________________Signature:___________

APPENDIX E: Zulu Informed Consent Form

DURBAN INSTITUTE *of* TECHNOLOGY

Usuku:________________________________

<u>Isihloko socwaningo</u>: **Ukuhlolwa kwangaphakathi ngokwethembeka kokuhlolwa komntwana olele onganyakazi usebenzisa ukuthinta noma ukupotoza ngesineke, nokuhlola futhi, ohlolwayo uma enyakaza, ukuze kuhlaziywe ukuthi kukhona yini ukuphazamiseka kokuxhumana kwamathambo omgogodla, lokhukuhlolwa kwenziwa ebantwaneni abaphile kahle, ababudala bungamasonto amabili okuzalwa kuya emasontweni ayishumi okuzalwa.**

Igama lika Supervisor:	**Dr Kretzman**	**(031) 2042205**
Igama lika Co-supervisor:	**Dr Shaik**	**(031) 204 2094**
Igama lomfundi ongumcwaningi:	**Julee Ralph**	**0836544993**

<u>Uyacelwa ukuba ukhethe impendulo</u>	**Yebo**	**Cha**
1. Ulifundile yini iphepha elinolwazi ngocwaningo?	Yebo	Cha
2. Ube naso yini isikhathi sokubuza imibuzo mayelana nocwaningo?	Yebo	Cha
3. Wanelisekile yini izimpendulo ozitholile emibuzweni yakho?	Yebo	Cha
4. Ube nalo yini ithuba lokuthola kabanzi ngocwaningo?	Yebo	Cha
5. Uyithole yonke imininingwane eyanele ngalolucwaningo?	Yebo	Cha
6. Uyayiqonda imiphumela yokuzimbhandakanya kwakho kulolucwaningo?	Yebo	Cha
7. Uyaqonda ukuthi ukhululekile ukuyeka lolucwaningo? noma inini ngaphandle kokunika isizathu sokuyeka ngaphandle kokubeka impilo yakho ebungozini	Yebo	Cha
8. Uyavuma ukuvolontiya kulolucwaningo?	Yebo	Cha
9. .Ukhulume nobani? ________________________		

Ngicela wenze isiqiniseko sokuthi umcwaningi usigcwalisa nawe lesisiqephu. Uma uphendule ngokuthi cha kokungaphezulu, sicela uthole ulwazi ngaphambi kokusayina.

<u>BHALA NGAMAGAMA AMAKHULU:</u>

Igama lomntwana:___________________________

Umzali:____________________________________Sayina: __________________________

Nizwana noma nitholana kanjani nengane yakho ?__________________________________

Igama lomfakazi:____________________________Sayina: __________________________

Igama lomfundi ongumcwaningi:_______________Sayina: __________________________

LEV	**STATIC PALPATION**								**MOTION PALPATION**									
	LEFT	**PNTS**	**Dx**	**RIGHT**	**PNTS**	**Dx**	**TOT**	**TDx**	**LEFT**	**PNTS**	**Dx**	**RIGHT**	**PNTS**	**Dx**	**TOT**	**TDx**	**GTOT**	**GDx**
C0 - C1	Asymmetry Spasm Tenderness			Asymmetry Spasm Tenderness					Restriction			Restriction						
C1 - C2	Asymmetry Spasm Tenderness			Asymmetry Spasm Tenderness					Restriction			Restriction						
C2 - C3	Asymmetry Spasm Tenderness			Asymmetry Spasm Tenderness					Restriction			Restriction						
C3 - C4	Asymmetry Spasm Tenderness			Asymmetry Spasm Tenderness					Restriction			Restriction						
C4 - C5	Asymmetry Spasm Tenderness			Asymmetry Spasm Tenderness					Restriction			Restriction						
C5 - C6	Asymmetry Spasm Tenderness			Asymmetry Spasm Tenderness					Restriction			Restriction						
C6 - C7	Asymmetry Spasm Tenderness			Asymmetry Spasm Tenderness					Restriction			Restriction						
C7 - T1	Asymmetry Spasm Tenderness			Asymmetry Spasm Tenderness					Restriction			Restriction						
T1 - T2	Asymmetry Spasm Tenderness			Asymmetry Spasm Tenderness					Restriction			Restriction						
T2 - T3	Asymmetry Spasm Tenderness			Asymmetry Spasm Tenderness					Restriction			Restriction						
T3 - T4	Asymmetry Spasm Tenderness			Asymmetry Spasm Tenderness					Restriction			Restriction						
T4 - T5	Asymmetry Spasm Tenderness			Asymmetry Spasm Tenderness					Restriction			Restriction						
T5 - T6	Asymmetry Spasm Tenderness			Asymmetry Spasm Tenderness					Restriction			Restriction						

LEV	STATIC PALPATION								MOTION PALPATION									
	LEFT	**PNTS**	**Dx**	**RIGHT**	**PNTS**	**Dx**	**TOT**	**TDx**	**LEFT**	**PNTS**	**Dx**	**RIGHT**	**PNTS**	**Dx**	**TOT**	**TDx**	**GTOT**	**GDx**
T6 - T7	Asymmetry Spasm Tenderness			Asymmetry Spasm Tenderness					Restriction			Restriction						
T7 - T8	Asymmetry Spasm Tenderness			Asymmetry Spasm Tenderness					Restriction			Restriction						
T8 - T9	Asymmetry Spasm Tenderness			Asymmetry Spasm Tenderness					Restriction			Restriction						
T10 - T11	Asymmetry Spasm Tenderness			Asymmetry Spasm Tenderness					Restriction			Restriction						
T11 - T12	Asymmetry Spasm Tenderness			Asymmetry Spasm Tenderness					Restriction			Restriction						
T12 - L1	Asymmetry Spasm Tenderness			Asymmetry Spasm Tenderness					Restriction			Restriction						
L1 - L2	Asymmetry Spasm Tenderness			Asymmetry Spasm Tenderness					Restriction			Restriction						
L2 - L3	Asymmetry Spasm Tenderness			Asymmetry Spasm Tenderness					Restriction			Restriction						
L3 - L4	Asymmetry Spasm Tenderness			Asymmetry Spasm Tenderness					Restriction			Restriction						
L4 - L5	Asymmetry Spasm Tenderness			Asymmetry Spasm Tenderness					Restriction			Restriction						
SI JNT	Asymmetry Spasm Tenderness			Asymmetry Spasm Tenderness					Restriction			Restriction						