THE RELIABILITY AND VALIDITY OF MANUAL EXAMINATION PROCEDURES FOR THE DETECTION OF MUSCULO SKELETAL DYSFUNCTION IN PEOPLE WITH BRONCHIAL ASTHMA

A RESEARCH PROJECT SUBMITTED IN PARTIAL REQUIREMENT FOR THE DEGREE OF MASTER OF OSTEOPATHY, UNITEC INSTITUTE OF TECHNOLOGY, 2010

HEATHER GROSSMITH
DECLARATION

Name of candidate: Heather Grossmith

This Research Project entitled “The reliability and validity of manual examination procedures for the detection of musculoskeletal dysfunction in people with bronchial asthma” is submitted in partial fulfillment for the requirements for the Unitec degree of Master of Osteopathy.

CANDIDATE’S DECLARATION

I confirm that:

• This Research Project represents my own work;

• The contribution of supervisors and others to this work was consistent with the Unitec Regulations and Policies.

• Research for this work has been conducted in accordance with the Unitec Research Ethics Committee Policy and Procedures, and has fulfilled any requirements set for this project by the Unitec Research Ethics Committee.

Research Ethics Committee Approval Number: 2009.963

Candidate Signature: _______________________ Date: ______________________

Heather Grossmith

Student number: 1225873
ACKNOWLEDGMENTS

First of all I would like to thank my partner Mark who has always had faith in me and who has supported me unconditionally throughout this course and has never complained when piles of notes and textbooks have taken over the living room!

Secondly, I would like to thank:

• Rob Moran – a big thanks for your continuous support and help throughout this project
• Craig Hilton – for all your help and support throughout this course
• Andy Stewart – for your statistics advice
• Rebecca McKay-Watts – for helping me with my data collection
• My examiners – for generously donating your time and expertise for the duration of the data collection
• All my friends who have been there through the highs and lows
• My participants – thank you for donating your time to help me out
## TABLE OF CONTENTS

### SECTION 1 – LITERATURE REVIEW

**Chapter 1: An introduction to asthma**

1.1. Search strategy

1.2. Asthma

1.2.1. Definition of asthma, prevalence and economic cost

1.2.2. Dysfunctional breathing and lung obstruction

1.2.3. Pharmacotherapy for asthma

1.2.4. Classification of asthma severity

1.3. The respiratory process

1.3.1. The act of ventilation and respiration

1.3.2. Muscles involved in ventilation

1.4. Fundamental osteopathic principles in relation to respiration

1.4.1. Structure/function relationship

1.4.2. Somatic Dysfunction definition

1.5. Review of osteopathic literature on the type of Somatic Dysfunction associated with asthma

1.6. Viscero-Somatic Reflexes

1.6.1. Research on the Viscero-Somatic reflex

1.6.2. Research on relationship between organic diseases and somatic dysfunction

**Chapter 2: An osteopathic approach to asthma**

2.1. Asthma and the osteopathic approach

2.2. Studies on the treatment of asthma by manual intervention

2.2.1. Sham treatments

2.2.2. Examination protocol

2.3. Limitations and weaknesses in asthmatic manual intervention studies

**Chapter 3: A background into understanding reliability and accuracy**

3.1. Diagnostic accuracy

3.2. Reliability

3.2.1. The kappa statistic

3.2.2. Prevalence and Bias Indices

3.3. An examination of the literature on reliability of manual palpation procedures

3.3.1. The reliability of manual assessment procedures

3.4. Factors that affect reliability
SECTION 2 – EXPERIMENTAL INVESTIGATION

Chapter 4: Introduction ............................................................................................................... 51
4.1. Asthma, somatic dysfunctions and current research on manual therapy for asthma ... 51
4.2. Previous studies that have investigated physical examination procedures to detect somatic dysfunction in people with asthma ................................................................. 53
4.3. Aims and objectives ............................................................................................................. 54

Chapter 5: Methods ....................................................................................................................... 56
5.1. Study design ......................................................................................................................... 56
5.2. Recruitment of participants ............................................................................................... 56
5.3. Participant inclusion / exclusion criteria .......................................................................... 57
5.4. Recruitment of examiners ................................................................................................. 58
5.5. Research ethics approval / Participant and examiner consent ......................................... 58
5.6. Experimental design .......................................................................................................... 59
5.6.1. Consensus training ........................................................................................................ 59
5.6.2. Data collection ............................................................................................................... 59
5.6.3. Physical examination procedures .................................................................................. 62
5.6.4. Recording findings ....................................................................................................... 67
5.7. Data analysis ....................................................................................................................... 67
5.7.1. Data reduction ................................................................................................................ 67
5.7.2. Assessment of the performance of a diagnostic procedures ....................................... 68
5.7.3. Data analysis of inter-examiner reliability .................................................................. 71
5.7.4. Normality of the asthmatic and asymptomatic group ................................................... 75
5.8. Subgroup analysis ................................................................................................................ 75
5.8.1. Subgroup 1 and 2 .......................................................................................................... 75
5.8.2. Subgroup 3 .................................................................................................................. 75
5.8.3. Subgroup 4 .................................................................................................................. 76

Chapter 6: Results ......................................................................................................................... 77
6.1. Participant characteristics ................................................................................................. 77
6.2. Comparison at baseline of the asthmatic and asymptomatic groups ............................... 78
6.3. Characteristics of asthmatic group ..................................................................................... 78
6.4. Validity of physical examination procedures .................................................................... 80
6.5. Inter-examiner reliability .................................................................................................. 86
6.6. Subgroup analysis for examination procedures with acceptable validity and reliability ................................................................................................................................. 93
6.6.1. Subgroup 1 and 2 analysis ......................................................................................... 95
6.6.2. Subgroup 3 analysis ................................................................................................... 97
6.6.3. Subgroup 4 analysis ................................................................................................... 98
Chapter 7: Discussion

7.1. Statement of principal findings

7.2. Reliability

7.2.1. Comparing inter-examiner reliability of experienced osteopaths versus student osteopaths

7.2.2. Comparing inter-examiner reliability of dysfunction components versus discomfort components of all examination procedures

7.3. Validity

7.3.1. Accuracy

7.3.2. Sensitivity and Specificity

7.3.3. Likelihood Ratios

7.3.4. Positive and Negative Predictive Values

7.4. Strengths, limitations and weaknesses

7.4.1. Internal validity

7.4.2. Sample size

7.4.3. Training period

7.4.4. Standardizing areas of palpation

7.4.5. Reducing examiner fatigue

7.4.6. Pre-warm up

7.4.7. Randomization of examiners

Chapter 8: Recommendations

Chapter 9: Conclusion

SECTION 3 - APPENDICES

Appendix A: Ethics Approval

Appendix B: Participant Information Sheet

Appendix C: Participant Consent Form

Appendix D: Examiner Information Sheet

Appendix E: Examiner Consent Form

Appendix F: Participant Medical History Form

Appendix G: Physical Examination Record Forms

Appendix H: Advertising Poster
LIST OF FIGURES

SECTION 1 – LITERATURE REVIEW

Figure 1: Respiratory muscles .............................................................. 10
Figure 2: Schematic representation of the viscero-somatic reflex .......... 17

SECTION 2 – EXPERIMENTAL INVESTIGATION

Figure 3: Active side bending of the trunk ....................................... 62
Figure 4: Chin tuck .............................................................................. 62
Figure 5: Thoracic spine rotation .................................................... 63
Figure 6: Thoracic spine extension .................................................. 63
Figure 7: Combined rotation and lateral flexion (CRLF) .................. 64
Figure 8: Chest expansion ................................................................. 64
Figure 9: Trapezius stretch ............................................................... 65
Figure 10: Suboccipital tissue texture .............................................. 65
Figure 11: Sternocleidomastoid (SCM) tissue texture ...................... 66
Figure 12: Observation of faulty breathing ...................................... 66
Figure 13: Accuracy of dysfunction components of examination procedures 1-10 (osteopathic students vs. experienced osteopaths) .... 85
Figure 14: Accuracy of discomfort components of examination procedures 1-9 (osteopathic students vs. experienced osteopaths) .... 85
Figure 15: Inter-examiner reliability (Cohen’s kappa) of dysfunction components of examination procedures 1-10 (osteopathic students vs. experienced osteopaths) ............................................ 91
Figure 16: Inter-examiner reliability (Cohen’s kappa) of discomfort components of examination procedures 1-9 (osteopathic students vs. experienced osteopaths) ............................................ 91
Figure 17: Inter-examiner reliability (adjusted kappa) of dysfunction components of examination procedures 1-10 (osteopathic students vs. experienced osteopaths) ............................................ 92
Figure 18: Inter-examiner reliability (adjusted kappa) of discomfort components of examination procedures 1-9 (osteopathic students vs. experienced osteopaths) ............................................ 92
Figure 19: Student osteopaths inter-examiner reliability against accuracy of dysfunction and discomfort components for all examination procedures .......... 93
Figure 20: Experienced osteopaths inter-examiner reliability against accuracy of dysfunction and discomfort components for all examination procedures .... 93
Figure 21: Student osteopath’s inter-examiner reliability (adjusted kappa) against accuracy of dysfunction and discomfort components for all examination procedures ............................................ 94
Figure 22: Experienced osteopaths inter-examiner reliability (adjusted kappa) against accuracy of dysfunction and discomfort components for all examination procedures ............................................ 94
LIST OF TABLES

SECTION 1 – LITERATURE REVIEW

Table 1: Classification of asthma severity in a patient with treated asthma ................... 7
Table 2: Evaluation of asthmatic intervention studies ..................................................... 31
Table 3: Example of diagnostic accuracy 2 x 2 contingency table .................................. 35
Table 4: Example of reliability 2 x 2 contingency table ............................................... 36

SECTION 2 – EXPERIMENTAL INVESTIGATION

Table 5: Example of contingency table summarizing dysfunction findings ................ 68
Table 6: Example of contingency table summarizing discomfort findings .................. 68
Table 7: Formulae and descriptions of the statistics used to assess validity of physical examination procedures ................................................................. 70
Table 8: Example of contingency table summarizing Examiner A and C’s (student osteopaths) findings for dysfunction component of examination procedure 1 ................................................................................................................................. 71
Table 9: Example of contingency table summarizing examiner B and D’s (experienced osteopaths) findings for dysfunction component of examination procedure 1 ................................................................................................................................. 72
Table 10: Kappa Categorization ................................................................................... 73
Table 11: Formulae and descriptions of the statistics used to assess inter-examiner reliability ............................................................................................................. 74
Table 12: Characteristics of participants in the asthmatic and asymptomatic group .... 77
Table 13: Asthma severity ............................................................................................. 78
Table 14: Asthma duration, medication use, asthma triggers, and frequency of asthma symptoms (cough, wheezing, shortness of breath and/or chest tightness) as reported by participants .................................................................................................................. 79
Table 15: Student osteopaths – validity of each dysfunction component of examination procedures 1-10 .................................................................................................................. 81
Table 16: Experienced osteopaths – validity of each dysfunction component of procedures 1-10 .................................................................................................................. 81
Table 17: Student osteopaths – validity of each discomfort component of examination procedure 1-9 .................................................................................................................. 82
Table 18: Experienced osteopaths – validity of each discomfort component of examination procedure 1-9 .................................................................................................................. 82
Table 19: Student osteopaths – inter-examiner reliability of each dysfunction component of examination procedures 1-10 ........................................................ 87
Table 20: Experienced osteopaths – inter-examiner reliability of each dysfunction component of procedures 1-10 ........................................................ 87
Table 21: Student osteopaths – inter-examiner reliability of each discomfort component of examination procedures 1-9 ........................................................ 88
Table 22: Experienced osteopaths – inter-examiner reliability of each discomfort component of examination procedures 1-9 ................................................................. 88
Table 23: Combined clinical validity of all subgroups for both student osteopaths and experienced osteopaths......................................................................................... 96
Table 24: Clinical validity of subgroup 3 performed by the student osteopaths........ 97
Table 25: Clinical validity of subgroup 3 performed by the experienced osteopaths ... 98
Table 26: Comparison of combined clinical validity of all components of all 10 examination procedures compared to subgroup 4 ......................................................... 99
Table 27: Discomfort component of examination procedure 2 performed by the experienced osteopaths (with extremely low prevalence)................................. 104
Table 28: Discomfort component of examination procedure 2 performed by the experienced osteopaths (with zero prevalence)...................................................... 104
Table 29: Likelihood ratios and bedside estimates ......................................................... 113
OVERVIEW

The following research project is divided into three sections.

Section 1 contains the literature review which is divided into three chapters. The literature review includes an introduction to the impact of asthma on the New Zealand population, the fundamental osteopathic principles in relation to lung ventilation, the types of musculoskeletal dysfunction that are hypothesized in osteopathic literature to occur in people with asthma, manual treatment approaches that aim to increase lung function in asthmatics and the reliability and validity of manual palpatory procedures currently used in osteopathic practice to detect musculoskeletal dysfunctions.

Section 2 contains the experimental investigation which is divided into six chapters. The experimental investigation includes an introduction, methods, results, discussion, recommendations and conclusion.

Section 3 contains the appendices. The appendices include the ethics approval letter, participant and examiner information sheet and consent form, participant medical history form, physical examination record form and the advertising poster used to attract participants.
SECTION 1: LITERATURE REVIEW
SECTION 1 - TABLE OF CONTENTS

LITERATURE REVIEW

Chapter 1: An introduction to asthma .................................................................4
  1.1. Search strategy ..............................................................................................4
  1.2. Asthma ...........................................................................................................4
      1.2.1. Definition of asthma, prevalence and economic cost ......................4
      1.2.2. Dysfunctional breathing and lung obstruction ...............................5
      1.2.3. Pharmacotherapy for asthma .........................................................6
      1.2.4. Classification of asthma severity .......................................................6
  1.3. The respiratory process ..............................................................................8
      1.3.1. The act of ventilation and respiration .............................................8
      1.3.2. Muscles involved in ventilation .......................................................8
  1.4. Fundamental osteopathic principles in relation to respiration ..................12
      1.4.1. Structure/function relationship ......................................................12
      1.4.2. Somatic Dysfunction definition .....................................................13
  1.5. Review of osteopathic literature on the type of Somatic Dysfunction associated with asthma .................................................................13
  1.6. Viscero-Somatic Reflexes .........................................................................17
      1.6.1. Research on the Viscero-Somatic reflex .......................................19
      1.6.2. Research on relationship between organic diseases and somatic dysfunction 21

Chapter 2: An osteopathic approach to asthma ...............................................28
  2.1. Asthma and the osteopathic approach ......................................................28
  2.2. studies on the treatment of asthma BY Manual intervention ..................29
      2.2.1. Sham treatments ..............................................................................30
      2.2.2. examination protocol .....................................................................31
  2.3. Limitations and weaknesses in asthmatic manual intervention studies ....32

Chapter 3: A background into understanding reliability and accuracy ............35
  3.1. Diagnostic accuracy ..................................................................................35
  3.2. Reliability ..................................................................................................36
      3.2.1. The kappa statistic ........................................................................36
      3.2.2. Prevalence and Bias Indices ..........................................................37
  3.3. An examination of the literature on reliability of manual palpation procedures 37
      3.3.1. The reliability of manual assessment procedures ..........................39
  3.4. Factors that affect reliability ....................................................................41
CHAPTER 1: AN INTRODUCTION TO ASTHMA

Part one provides a background to asthma and its impact in the New Zealand population. It also discusses the physiology of pulmonary ventilation, the fundamental osteopathic principles in relation to pulmonary ventilation and the types of musculoskeletal dysfunction that are hypothesized in osteopathic literature to occur in people with asthma.

1.1. SEARCH STRATEGY

To identify literature, online databases were searched including PubMed (MEDLINE), Cochrane library, Cumulative Index to Nursing and Allied Health Literature (CINAHL), ScienceDirect and Manual Alternative and Natural Therapy System (MANTIS) using one or a combination of key words: palpation, asthma, manual therapy, manual treatment, osteopathic, chiropractic, reliability, validity, somatic dysfunction, musculoskeletal, respiration, respiratory compromise, and Chronic Obstructive Pulmonary Disease (COPD). Hand searches of the bibliographies of selected articles and textbooks were also undertaken. Standard osteopathic reference texts were also reviewed in relation to osteopathic theory on breathing and asthma. These texts included: Chaitow, Bradley & Gilbert, 2002; DiGiovanna & Schiowitz, 1991; Greenman 2003; Kuchera & Kuchera, 1994; Sammut & Searle-Barnes, 1998; Stone, 1999 and Ward, 2003.

1.2. ASTHMA

1.2.1. DEFINITION OF ASTHMA, PREVALENCE AND ECONOMIC COST

Bronchial asthma is one of the most common chronic diseases in New Zealand (Holt, Kljakovic, & Reid, 2003) with an estimated 15-20% of New Zealand children and adults suffering from asthma - this amounts to one in every six New Zealanders. In total, an estimated 600,000 New Zealanders suffer from asthma and these rates are among some of the highest prevalence rates of asthma in the world (Holt & Beasley, 2001). The economic costs related to asthma are estimated by the Asthma and Respiratory Foundation of New Zealand to be $825 million per year. Approximately $700 million of this is attributable to indirect non-medical costs: days off work, decreased quality of life, and premature death. The remaining $125 million is directly related to medication, medical devices (for example peak flow meters), primary care services, hospital inpatient care, and emergency services (Holt & Beasley, 2001).
Asthma is defined as a chronic inflammatory disorder of the airways (National Asthma Council Australia (NAC) 2006). The inflammation is associated with increased airway responsiveness to a variety of stimuli, for example, environmental irritants, allergens, cold air, viral respiratory tract infections, and physical exercise. Many cells and cellular elements play a role in increasing airway responsiveness to a variety of stimuli, which leads to swelling of the airway wall, and contraction of the airway wall’s smooth muscle (also known as ‘bronchospasm’). Swelling and bronchospasm results in excessive airway narrowing and consequent airflow obstruction. Lung hyperinflation also develops as air distal to the obstruction becomes trapped (NAC, 2006).

Indicators of airflow obstruction from asthma include breathlessness, wheezing, chest tightness, and coughing, especially in the early hours of the morning or at night. Episodes of airflow obstruction usually resolve of their own accord if the stimuli are removed or in response to pharmacotherapy (Global Initiative for Asthma (GINA)), 2009; NAC, 2006). However if left untreated, or unsuccessfully treated, over time asthma may lead to airway remodeling, bronchial wall narrowing, and a progressive loss of lung function (Pascual & Peters, 2005).

The diagnosis of asthma in New Zealand as described by the New Zealand Guidelines Group (NZGG) (2002) relies on the patient’s history of a cough, wheeze and shortness of breath as well as evidence of reversible airflow obstruction which occurs either spontaneously over time or in response to pharmacotherapy.

1.2.2. DYSFUNCTIONAL BREATHING AND LUNG OBSTRUCTION

In order for asthmatic individuals to maintain adequate lung ventilation and gas exchange, adjustments occur to their breathing pattern (named dysfunctional breathing pattern) in comparison to normal asymptomatic individuals (Courtney, 2009). Dysfunctional breathing patterns can be clinically observed (Perri & Halford, 2004) and can also be measured using a self evaluation breathing questionnaire (Courtney & Greenwood, 2009). Dysfunctional breathing patterns and airflow obstruction are thought to contribute to structural changes of the biomechanics of the thorax which may further impair ventilation (Chaitow, Bradley, & Gilbert, 2002; DiGiovanna & Schiowitz, 1991; Green, 2000; Sammut & Searle-Barnes, 1998; Ward, 2003). Findings of structural biomechanical changes to the thorax are primarily based on subjective clinical palpation and observation of dysfunctional breathing (Chaitow et al., 2002; DiGiovanna & Schiowitz, 1991; Green, 2000; Sammut & Searle-Barnes, 1998; Ward, 2003). There is minimal objective research on the biomechanics of the thorax. The geometry of the thorax is complex and some components of the thorax change their geometry during breathing,
presenting many difficulties when developing a tool to objectively research it. As a consequence there has been no gold standard model developed that can objectively measure normal biomechanical movements during respiration in asymptomatic subjects and thus there is no objective comparison to measure dysfunctional breathing against. However, there are new three dimensional (3D) mathematical models of the thorax being developed which are designed to objectively evaluate breathing mechanics. Although 3D model development is in its infancy, it may prove useful in future studies with subjects with breathing dysfunctions (Behr et al., 2010).

1.2.3. PHARMACOTHERAPY FOR ASTHMA

The current pharmacological management of asthma consists of daily use of inhaled medication, colloquially known as ‘preventer’, with or without the use of ‘relievers’, and/or symptom controllers. Preventers are used to suppress bronchial inflammation. Preventers include: Inhaled Corticosteroids (ICS), Leukotriene receptor antagonists (LTRAs), and Cromones. Relievers and symptom controller’s work by decreasing the contraction of the airway wall’s smooth muscle, thus increasing the diameter of the airway and increasing airflow. Relievers include Short-action-beta-2-agonists (SABA) which give immediate relief from asthma symptoms. Symptom controllers include long-acting beta-2-agonists (LABA) which give up to 12-hours relief from asthma symptoms. For asthma that is unresponsive to preventers, relievers, and symptom controllers, oral corticosteroids may be prescribed (NAC, 2006).

1.2.4. CLASSIFICATION OF ASTHMA SEVERITY

The New Zealand Guidelines Group (NZGG) (2002) classifies asthma into three categories that reflect disease severity: mild; moderate; and severe. Severity, as classified by NZGG (2002), is based on the individual’s level of symptoms, lung function variability, and impact on activity. However, this classification system does not take into account individual responses to pharmacotherapy - how adequately a person’s symptoms are controlled\(^1\) by pharmacotherapy.

The Global Initiative for Asthma (GINA) (2005) is an international organization that works in collaboration with the World Health Organization on the prevention and control of chronic respiratory diseases. GINA (2005) have suggested that the intensity of treatment required to maintain ‘good asthma control\(^1\)’ be included in asthma classification systems because this takes into account the individuals response to treatment and would be more reflective of the subjects

\(^1\) Good asthma control is defined as: minimal symptoms during day and night minimal need for reliever medication, no exacerbations, no limitation of physical activity, normal lung function (Forced Expiratory Volume in one second (FEV\(_1\)) and/or Peak Expiratory Flow (PEF) >80% predicted or best (NAC, 2006).
current lung function. The National Asthma Council of Australia (2006) subsequently provided an asthma severity classification system (adapted from GINA 2005’s classification system) that considers the intensity of treatment and how adequately symptoms are controlled. The National Asthma Council of Australia (2006) classification system subdivides asthma into four categories: intermittent; mild persistent; moderate persistent; severe persistent. These categories are based on the level of symptoms, airflow limitation, lung function variability, and also the intensity of the treatment (type and amount of drugs to maintain good control of asthma symptoms) – see Table 1 for further detail about the classification system. A classification system that includes the intensity of treatment and how adequately symptoms are controlled has been recommended for research studies on people with asthma (GINA, 2005).

Table 1
Classification of asthma severity in a patient with treated asthma

<table>
<thead>
<tr>
<th>Daily treatment requirement</th>
<th>No ICS</th>
<th>Low dose ICS</th>
<th>Low- to Medium dose ICS and LABA</th>
<th>High-dose ICS + LABA+ other agents</th>
</tr>
</thead>
<tbody>
<tr>
<td>Any of:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• daytime symptoms occur less than once per week</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• night-time symptoms occur less than twice per month</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• exacerbations are brief</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• FEV1 between episodes is at least 80% predicted and 90% personal best</td>
<td>Intermittent</td>
<td>Mild persistent</td>
<td>Moderate persistent</td>
<td>Severe persistent</td>
</tr>
<tr>
<td>Any of:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• daytime symptoms more than once per week but not every day</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• night-time symptoms more than twice per month but not weekly</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• FEV1 between episodes is more than 80% predicted and 90% personal best</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Any of:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• daytime symptoms daily</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• night-time symptoms at least weekly</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• exacerbations affect sleep/activity</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• SABA use daily</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• FEV1 is 60–80% predicted and 70–90% personal</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Any of:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• daytime symptoms every day and restrict physical activity</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• night-time symptoms frequent</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• exacerbations are frequent</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• FEV1 is less than 60% predicted and less than 70% personal best</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Key: FEV1 = Forced Expiratory volume in one second; ICS = Inhaled Corticosteroid; LABA = Long acting beta-2 agonist; SABA = Short acting beta-2 agonist; Other agents include antibiotics or antihistamines and/or sedatives

Increase treatment and reassess severity within 3 months

If patient’s asthma has matched this category for 3 months and is stable, consider down-titration of medications and reassess within three months.

1.3. THE RESPIRATORY PROCESS

1.3.1. THE ACT OF VENTILATION AND RESPIRATION

Whilst many authors use the terms respiration and ventilation interchangeably, each term has been distinctly defined. In this study:

- Respiration is defined as the act of oxygen and carbon dioxide exchange at the pulmonary alveolar level, where gas diffuses between the lung alveolar tissues and the pulmonary capillaries (Porth, 2002).

- Ventilation is defined as the mechanical process of moving air in and out of the lungs (Ratnovsky, Elad, & Halpern, 2008; Steimle, Mogensen, Karbing, Bernardino de la Serna, & Andreassen, 2011).

Ventilation requires the coordinated activity of more than 146 components including joints, muscles and connective tissues and neural reflexes. The components are coordinated together to create an effective pumping mechanism, lengthening and shortening the thoracic cavity, allowing lung expansion and contraction, and movement of air in and out of the lungs respectively (Breslin, 1996; Chaitow et al., 2002; Kuchera & Kuchera, 1994). Whilst the primary function of ventilation is to meet the demand for oxygen (which is essential for all cellular functions) and eliminate waste products (Kuchera & Kuchera, 1994; Ward, 2003), it is also thought to:

- assist venous and lymphatic fluid movement (Chaitow et al., 2002; DiGiovanna & Schiowitz, 1991; Stone, 1999; Ward, 2003)

- enhance digestive function and visceral tissue health, through the massaging of abdominal organs resulting from rhythmic fluctuations of that occur from lung expansion and contraction (Chaitow et al., 2002; Stone, 1999; Ward, 2003)

- maintain thoracic spine mobility through movement (Chaitow et al., 2002).

1.3.2. MUSCLES INVOLVED IN VENTILATION

An array of muscles are used in breathing. Knowledge of the normal workings of the respiratory muscles enables osteopaths and other therapists to identify anatomical structures that may be dysfunctional. Many anatomy textbooks, osteopathic textbooks, and some published articles describe the muscles involved in ventilation. These muscles are conventionally considered in two groups: primary muscles of ventilation (inhalation and exhalation) and; accessory muscles of ventilation (inhalation and exhalation) (Chaitow et al., 2002; Ward, 2003). Chaitow et al., (2002) describes the function of accessory muscles of ventilation as acting “like a reserve tank” - they are recruited when the work of breathing increases during exercise or disease. Although not
significant, slight variations exist in osteopathic literature and research data on the number of muscles involved in ventilation, and also on the categorization of the various muscles into primary or accessory muscle groups. The basis of the historical categorization of primary and secondary muscles of ventilation as defined in many textbooks is somewhat limited (Gray, 1974; Marieb, 2004). There is minimal objective evidence exists to support these categories. For example the scalene muscles, originally classified as accessory muscles of ventilation have been documented through electromyography (EMG) studies to be working during every inspiratory effort (Breslin, 1996; De Troyer & Estenne, 1984). The following paragraphs outline the muscles involved in respiration as described by the relevant literature.

**Inspiratory muscles**

The diaphragm is the main primary muscle of inhalation (Chaitow et al., 2002; DiGiovanna & Schiowitz, 1991; Hruska, 1997). The diaphragm’s contribution to inhalation has been investigated by Ratnovsky & Elad (2004) using 2D quantitative analysis of respiration muscle mechanics. Ratnovsky & Elad (2004) reported that the diaphragm performs 60-80% of the total inspiratory work.

The diaphragm is a thin muscle that separates the thoracic cavity from the abdominal cavity. Its muscle fibers adhere to the lower six ribs and first three vertebrae (see Figure 1). When contracted, the diaphragm descends, lengthening the thoracic cavity and rotating the lower six ribs outwards (Ratnovsky, Elad, & Halpern, 2008).
Other primary muscles of inspiration include the external intercostal muscles (see Figure 1). The external intercostal muscles span obliquely in an inferior-anterior direction from the rib above to the rib below. When contracted, these muscles raise the rib cage (Chaitow et al., 2002; DiGiovanna & Schiowitz, 1991; Green, 2000). De Troyer & Estenne (1984) conducted EMG analysis on the parasternal intercostals which found them to be active during quiet breathing. This supports the classification of the intercostals muscles into the primary muscles of ventilation category.

Accessory muscles include the sternocleiodomastoid muscles (SCM) and scalene muscles (see Figure 1) (Chaitow et al, 2002; DiGiovanna & Schowitz., 1991; Greenman, 2003). The SCM has two bands; each runs from the mastoid process, connecting with the manubrium and the middle third of the clavicle respectively. Bilateral contraction of these muscles raises the sternum vertically.

The scalene muscles consist of three separate bands, named for their positions: anterior, posterior and middle scalene muscles. The scalene muscles extend from the cervical spine transverse processes to the first rib (anterior and middle scalene) and second rib (posterior scalene). Contraction of these muscles raises the upper ribs, assisting in the expansion of the rib.
cage (Ratnovsky et al. 2008). Although many authors classify the scalene muscles as accessory muscles (Chaitow et al, 2002; DiGiovanna & Schowitz., 1991; Greenman, 2003) their electrical activity has been documented through electromyography studies to be active during every inspiratory effort and may therefore, on the basis of experimental data, be more accurately considered as primary muscles of inspiration and not accessory to inspiration (Breslin, 1996; De Troyer & Estenne, 1984) as is commonly described (Chaitow et al., 2002; DiGiovanna & Schiowitz, 1991; Ward, 2003). As such the scalene group will be referred to as primary muscles of inspiration.

Other accessory inspiratory muscles include the pectoralis major, pectoralis minor, serratus anterior and upper trapezius (Chaitow et al., 2002; DiGiovanna & Schiowitz, 1991; Greenman, 2003; Hruska, 1997; Ward, 2003). All of these muscles influence movement of the rib cage when they contract. Due to their attachments to the thorax, they assist inspiration by raising the ribs and sternum upwards, thereby increasing the vertical diameter of the thoracic cavity and elevating the shoulders, which in turn allows a greater volume of lung expansion (Chaitow et al., 2002; Ward, 2003).

**Expiratory muscles**

In contrast to inspiration, normal quiet expiration is primarily a passive process which involves the elastic recoil of the lungs and relaxation of the diaphragm (Chaitow et al., 2002; DiGiovanna & Schiowitz, 1991; Greenman, 2003; Ward, 2003). During exercise and respiratory compromise (for example asthma), active expiration occurs. In active expiration accessory expiratory muscles, namely the internal intercostal and abdominal muscles are recruited. Both these muscles aid in expelling air out of the lungs by decreasing the dimensions of the thoracic cavity. The internal intercostal muscles run obliquely in an inferior posterior direction from the rib above to the one below (Figure 1). When contracted, they lower the rib cage decreasing the transverse diameter of the thoracic cavity (Ratnovsky et al. 2008; DiGiovanna & Schiowitz. 1991; Greeman 2003).

The abdominal muscles work by pulling the lower ribs inferiorly, deflating the ribcage as well as pulling the abdominal wall inwards, increasing intra-abdominal pressure and causing the diaphragm to move superiorly (Chaitow et al., 2002; DiGiovanna & Schiowitz, 1991; Greenman, 2003; Ward, 2003). Other accessory muscles include the quadratus lumborum and latissimus dorsi (Chaitow et al., 2002; DiGiovanna & Schiowitz, 1991; Ward, 2003).
1.4. FUNDAMENTAL OSTEOPATHIC PRINCIPLES IN RELATION TO RESPIRATION

1.4.1. STRUCTURE/FUNCTION RELATIONSHIP

One of the fundamental osteopathic principles is that structure and function are reciprocally inter-related (DiGiovanna & Schiowitz, 1991; Sammut & Searle-Barnes, 1998; Stone, 1999; Ward, 2003). Alterations in anatomical structure result in inefficient or decreased ability to function, and changes in function result in structural modifications related to the functional changes (DiGiovanna & Schiowitz, 1991; Sammut & Searle-Barnes, 1998).

Although increased respiratory work (including the recruitment of accessory muscles of ventilation) is a normal requirement during exercise, many other processes also increase ventilatory work. These processes include:

- pathological factors, for example inflammation and bronchial obstruction that occurs in people with asthma
- biochemical factors, for example change in blood pH balance
- psychosocial factors, for example stress and anxiety (Perri & Halford, 2004).

When bronchial obstruction is present the respiratory work to overcome obstruction increases. When ventilatory work is increased or prolonged, the coordinated activity of the neural reflexes, muscular contractions, and joints involved in respiration may be lost (Courtney, 2009). The accessory muscles are thought to take on more of a primary role during ventilation resulting in a faulty “upper chest” breathing pattern (Courtney, 2009; Perri & Halford, 2004). This faulty breathing pattern may be retained even when the biochemical, pathological or psychosocial stimuli has been removed (Courtney, 2009). Perri & Halford (2004) describe the typical appearance of this faulty breathing pattern including:
  (a) predominantly chest movements
  (b) ‘lifting up motion’ of the upper chest during inspiration
  (c) absence of or lifting up motion of the lateral ribs
  (d) abdominal movement is rigid or paradoxical (source: Perri & Halford, 2004).

As a result of this faulty breathing pattern, greater stresses are imposed on the neuromusculoskeletal elements involved in ventilation. Osteopathic authors have theorized that when these stresses extend beyond the normal capabilities of these neuromusculoskeletal elements, resultant dysfunctional structural modifications occur (Chaitow et al., 2002; Denslow, 1964; Educational Council on Osteopathic Principles, 2006; Sammut & Searle-Barnes, 1998).
Palpatory and observational evidence of dysfunctional structural modifications are what osteopaths term ‘Somatic Dysfunction’ (SD).

### 1.4.2. SOMATIC DYSFUNCTION DEFINITION

The concept of Somatic Dysfunction is deeply grounded in osteopathic medicine. Somatic dysfunction (SD) is defined as: “altered function/disturbance of the body framework (including skeletal, arthrodial and myofascial structures and related vascular, lymphatic and neural elements) that can be identified by observational or palpatory diagnosis” (Ward, 2003). The term 'somatic dysfunction' may be used to describe aberrant structure or function in any region of the body. In clinical settings, somatic dysfunction is identified through physical examination consisting of palpatory and observational findings of:

- tissue texture changes detected on palpation (adjectives commonly used by practitioners include: ‘tight’, ‘hypertonic’, ‘ropy’, ‘flaccid’, ‘loose’, or ‘boggy’ tissue)
- asymmetry of structure
- restriction of motion
- tenderness on palpation

The acronym ‘T.A.R.T’ is commonly used when identifying SD; this stands for tissue texture changes, asymmetry of structure, restriction of motion and tenderness (Ward, 2003). A commonly used descriptor for this dysfunction entity described by chiropractors, physiotherapists and manual therapist textbooks as cited by Fryer (2003) is ‘segmental somatic dysfunction’, also termed ‘intervertebral dysfunction', which signifies a SD that appears to involve a single spinal level. Although SD may be a common term in textbooks, findings from a survey of British osteopaths concluded that SD is not a common term used in British practice (Fryer, Johnson, & Fossum, 2010). There have been no studies to determine how common the term ‘SD’ is used in clinical practice by Australasian osteopaths.

### 1.5. REVIEW OF OSTEOPATHIC LITERATURE ON THE TYPE OF SOMATIC DYSFUNCTION ASSOCIATED WITH ASTHMA

Whilst many osteopathic authors postulate areas of the body where SD would be expected to be found in asthmatics with faulty breathing patterns, few have specifically reported on what physical examination procedures have been used to identify these findings, or the details of the specific characteristic findings that they would expect to find using their examination procedures.
The following section outlines literature from commonly used osteopathic textbooks pertaining to the relationship between SD and asthma. Many of the claims in relation to SD and asthma made by authors of osteopathic textbooks are based on clinical opinion and have not been verified by objective data. The airflow obstruction and dysfunctional breathing that occurs in people with asthma are thought to contribute to structural changes to the biomechanics of the thorax (Chaitow et al., 2002; DiGiovanna & Schiowitz, 1991; Green, 2000; Sammut & Searle-Barnes, 1998; Ward, 2003). However, there is no hard evidence to support an association between these structural biomechanical changes (named somatic dysfunction), and asthma. As outlined earlier, 3D investigations into the biomechanics of the thorax is still in its infant stages, and thus findings of somatic dysfunction thought to be associated to asthma are currently only based on observational and clinical data.

Somatic dysfunction of sternocleidomastoid and scalene muscles

Hypertonia of the sternocleidomastoid (SCM) and scalene muscles has been reported by authors of osteopathic text books to be associated with asthma (Chaitow et al., 2002; DiGiovanna & Schiowitz, 1991; Kuchera & Kuchera, 1994; Sammut & Searle-Barnes, 1998). Muscle hypertonia is described by authors of osteopathic text books as increased tone of the muscle, and is hypothesized to occur from overuse of the muscle when it is sustained in a state of muscle contraction (Chaitow et al., 2002; DiGiovanna & Schiowitz, 1991; Kuchera & Kuchera, 1994; Sammut & Searle-Barnes, 1998). Visible and/or palpable examination findings of muscle hypertonia are described by Sammut & Searle-Barnes (1998) as: tense muscle tissue, muscle tissue that has a greater resistance to passive stretch, and muscle tissue that feels ‘boggy’ when compared to ‘normal’ muscle tissue. ‘Normal’ is defined by Sammut & Searle-Barnes (1998) as “a relative term and is only recognized by comparing to abnormal tissue states” p120. The ‘boggy’ tissue is thought to result from tissue congestion that occurs secondary to muscle hypertonia (Sammut & Searle-Barnes, 1998). There is no objective evidence related to hypertonia of SCM or scalene muscles in people with asthma.

Prolonged overuse of the scalene muscles and accessory muscles of ventilation may also result in muscle hypertrophy (palpable as enlarged muscle), shortened muscle tissue, and/or muscle fibrosis (Chaitow et al., 2002; DiGiovanna & Schiowitz, 1991; Sammut & Searle-Barnes, 1998). Fibrotic muscle tissue has a larger percentage of fibrous connective tissue, which reduces the ability of the muscle to contract and relax (Chaitow et al., 2002), but saves the muscle from continual contraction (Sammut & Searle-Barnes, 1998). When palpated, fibrotic muscle is
described as ‘rope like’ and frequently tender (Sammut & Searle-Barnes, 1998). Fibrotic respiratory muscles are thought to negatively influence ventilation as their efficiency to overcome airway resistance/obstruction is reduced (Chaitow et al., 2002). Although fibrosis of the lung tissue causes decreased lung function (Panos, Mortenson, Niccoli, & King, 1990) there are no studies which have examined whether respiratory muscle fibrosis of the scalene muscles and/or accessory muscles of ventilation, alter lung function. There are also no studies examining the relationship between the scalene muscles and/or accessory muscle fibrosis and asthma.

Other somatic dysfunction of the accessory muscles claimed to be associated with asthma include hypertonicity of the:

- internal intercostals muscles (Sammut & Searle-Barnes, 1998)
- upper trapezius (Chaitow et al., 2002; Janda, 1983)
- levator scapulae (Chaitow et al., 2002; Janda, 1983)
- pectoralis major and minor (Janda, 1983).

**Upper Cross Syndrome**

Upper Cross Syndrome (UCS) is a postural pattern characterized by rounded shoulders, an increased curvature of the upper thoracic spine, and a forward head posture. It occurs when certain muscles (agonists) shorten and tighten, and their opposing muscles (antagonists) become lengthened and weak. Janda (1986) claims that one of the factors that can cause this syndrome is a faulty breathing pattern which places increased stress on the accessory muscles of breathing, particularly the upper trapezius, pectoralis, sternocleidomastoid and levator scapular, which become hypertonic and shorten. Whilst these muscles shorten, inhibition (via neural reflexes) of their opposition muscles (antagonists) occurs. As a result, weakening of the lower trapezius, middle trapezius, serratus anterior and rhomboids occurs. Because the length/tension relationships of the agonist and antagonist muscles change, so too does the position of the neck, head and shoulders. The shoulders become rounded; weakness of the deep neck flexors develops, as well as increased tone in the suboccipital muscles. An increase in the upper cervical lordosis also occurs, which places stress on the lower cervical and upper thoracic vertebral (down to T4²). These changes may negatively influence breathing function due to the slumped upper body position (Janda 1986). The UCS pattern is based on clinical opinion and has not been verified by objective data. Although Moore (2004) cites from Janda that UCS may be a possible causative factor for supraspinatus degeneration, there appears to be an absence of data on UCS.

---

² T4 = fourth thoracic vertebrae
being linked with other musculoskeletal pain/complaints. There is also an absence of research on diagnostic procedures used to detect UCS.

**Cervical Somatic Dysfunction**

Cervical SD has been claimed by Janda (1983), Sammut & Searle-Barnes (1998) and Chaitow (2002) to be associated with respiratory compromise; however no studies have examined the relationship. Chaitow (2002) reports on findings of cervical spine rigidity associated with upper chest breathing. Additionally, an association was noted between observational findings of increased cervical lordosis and asthma (Sammut & Searle-Barnes, 1998) and upper chest breathing (Janda, 1983). An increased upper cervical lordosis is also thought to result in shortening and increased tone of the suboccipital muscles (Janda, 1983). Again, none of these statements has been supported by objective evidence.

**Rib Somatic Dysfunction**

Texts by DiGiovanna & Schiowitz (1991) & Ward, (2003) note associations between asthma and dysfunctions/restrictions in the upper rib cage, particularly in the first rib. Rib dysfunctions are thought to result from fatigue of the respiratory muscles, in particular the scalene group which insert into the first and second rib (Ward, 2003). Other associations have been noted between acute asthma and restricted forth rib on the right side (Ward, 2003). However there is an absence of research on diagnostic procedures used to detect rib dysfunctions in people with asthma. In addition, there is no research which supports the association between these types of SD and asthma.

**Thoracic Somatic Dysfunction**

Somatic Dysfunction of the thoracic spine has been hypothesized by osteopathic authors to routinely correspond with asthma and respiratory compromise. An association has been noted between asthma and thoracic immobility (DiGiovanna & Schiowitz, 1991), increased thoracic kyphosis (DiGiovanna & Schiowitz, 1991; Sammut & Searle-Barnes, 1998), and paraspinal muscle rigidity, tension and tenderness (Ward, 2003). Other associations have been noted between acute asthma and segmental somatic dysfunction of T2-T4 (Ward, 2003). Somatic Dysfunction of the thoracic spine, paraspinal muscles and overlying tissue from asthma can result from two distinctly different origins/processes: directly from mechanical dysfunctions of the musculoskeletal system, and from viscero-somatic reflexes from the lungs to the spinal segments. Viscero-somatic reflexes are discussed in the next section. There is a small amount of
research which investigates the association between thoracic SD and people with asthma or respiratory compromise. The association between thoracic SD and asthma is discussed further in section 1.6.2.

1.6. VISCERO-SOMATIC REFLEXES

The concept of viscero-somatic (VS) reflexes is widely described in many osteopathic text books (DiGiovanna & Schiowitz, 1991; Sammut & Searle-Barnes, 1998; Stone, 1999; Ward, 2003). The physiologic process involved in VS reflexes is depicted in Figure 2. The theory is that afferent nerve impulses from the disturbed organ (e.g. lungs) get relayed to interconnecting neurons in the spinal cord. These interconnecting neurons then relay impulses to the sympathetic and peripheral motor efferent’s of the same spinal level as the organ (e.g.T1-4) resulting in sensory and motor changes of the somatic tissue at the level (e.g. T1-4) (Beal, 1985). The characteristic findings of SD from VS reflexes are almost indistinguishable from mechanical SD. The same diagnostic ‘T.A.R.T’ criteria pertains to diagnosing VS reflexes, however specific characteristic findings indicative of VS reflexes include the following.

- Skin changes: increased atrophic skin (DiGiovanna & Schiowitz, 1991).
- Tissue texture changes: firm, dry or spongy tissue (DiGiovanna & Schiowitz, 1991).
- A more restricted and ‘fixed’ range of motion than usual mechanical dysfunction (DiGiovanna & Schiowitz, 1991).
- SD findings in two or more adjacent spinal segments (Beal, 1985).

Figure 2: Schematic representation of the viscero-somatic reflex

Although osteopathic literature supports the existence of VS reflexes (DiGiovanna & Schiowitz, 1991; Sammut & Searle-Barnes, 1998; Stone, 1999; Ward, 2003), there is no objective evidence to support the concept that afferent impulses from a disturbed organ cause SD (at the same spinal level as the organ).

The VS reflex as a valid concept has not been established by scientific literature and is not widely accepted by the medical profession. To date there is little evidence to support the idea that segmental spinal SD is associated to organic disease states. Even if the VS reflex concept was plausible, detection of purely visceral related SD via palpatory diagnostic procedures is not. The characteristic differences in findings between SD originating from VS reflex and SD related to other factors are so subtle it would seem impossible for any clinician to discern one from the other using manual palpation procedures. This is especially the case when current research suggests that there is a lack of evidence for the use of manual palpation procedures which detect three of the four components of SD: asymmetry of structure, restriction of motion and/or tissue texture changes (Haneline, Cooperstein, Young, & Birkeland, 2008; Haneline & Young, 2009; Hollerwöger, 2006; Huijbregts., 2002; Hestboek & Leboeuf-Yde, 2000).

Even if palpable findings of SD were linked to asthma, the clinical usefulness of this knowledge would be limited. SD from VS reflex occur down-stream from the originating visceral dysfunction. For an organ to stimulate afferent receptors and cause a reflex segmental spinal SD, one would presume that the level of visceral dysfunction has reached a critical level where the organ is functionally challenged, resulting in objective signs and symptoms of the visceral disease. The majority of predictors of visceral dysfunction pertaining to asthma (SOB, wheezing, coughing) would be recorded during the case history of an osteopathic consultation. If signs and symptoms pertaining to undiagnosed asthma were apparent, the patient would be referred to a general practitioner for further investigation. However, even if the use of a segmental palpatory examination was accurate, and reliable at detecting SD related to asthma, findings of SD may only provide supplementary information to what is already gained from the consultation. It would not change the clinician’s decision to refer and thus provides little use in this context.

There is also a lack of objective evidence to support the tenet that the treatment of segmental SD by manipulation will interrupt the VS reflex arc, leading to a decrease in the segmental SD, and also a decrease in the visceral efferent stimulus, and adverse events associated with the viscera as stated by Beal (1985). Although this tenet is widely accepted as part of osteopathic philosophy, there is no formal evidence to suggest that patients with visceral diseases
undergoing segmental spinal manipulation have a decrease in symptoms and signs related to their visceral disease state.

1.6.1. RESEARCH ON THE VISCERO-SOMATIC REFLEX

During the 1970s and 1980s, researchers in the osteopathic profession sought to investigate the VS model. Research focused on establishing associations between organic diseases and segmentally related subjective ‘palpable’ changes to paraspinal muscle tone (Beal, 1983; Beal & Morlock 1984; Tarr et al., 1987; Nicholas et al., 1985), and/or objective measurements of myoelectric activity in segmentally related paraspinal muscle tissue (Schoen & Finn, 1978., Gwirtz et al., 2007).

To date, research has investigated myoelectric changes in segmentally related paraspinal muscle activity in animals before, during and post induced heart ischemia or Myocardial Infarction (MI) (Schoen & Finn, 1978; Gwirtz et al., 2007). Ischemia is characterized by decreased blood supply to cardiac muscle and in humans this causes symptoms of chest pain and shortness of breath (commonly known as angina). Symptoms are temporary and resolve once sufficient blood flow is returned. Myocardial Infarction however, causes permanent damage to heart tissue. It is characterized by an interruption to the blood supply to the heart which if left untreated for a period of time causes oxygen deprivation to the heart muscle which leads to death to heart tissue (infarction). Inducing heart ischemia or MI in animals is done by narrowing the coronary artery (the main arterial supply to the heart) by 60 percent, or by occluding blood flow of the coronary artery respectively. Two studies have demonstrated increased myoelectric activity of segmentally related paraspinal muscle tissue during induced heart ischemia or MI (Schoen & Finn, 1978; Gwirtz et al., 2007). These studies however have been small or preliminary in nature, and they have only assessed VS reflexes related to the heart. To date, no studies have assessed lung disorders and segmentally related paraspinal myoelectric activity. The relationship between increased paraspinal myoelectric activity and findings of SD is also weak. There is also insufficient numbers of quality studies which objectively demonstrate that myoelectric changes can be detected via palpation.

In a preliminary study, Schoen & Finn (1978) sought to initiate the VS reflex by inducing MI in one cat and measured segmentally related myoelectric readings in the cutaneous maximus muscle before, during and post intervention, as well as reported any visual changes (e.g. fasciculations) observed in this muscle. The cutaneous maximus muscle originates from the linea alba (median line of the abdomen) and wraps around the cats trunk and inserts into the dermis.
(deep layer of the skin). The cutaneous maximus muscle shares the same segmental supply as the heart (T3-T5). Schoen & Finn (1978) reported observing intense fasciculations’ of the cutaneous maximus muscle, localized over three segments T3-T5 on the left side. They also reported corresponding increases in myoelectrical activity during the intervention, with higher frequency myoelectric activity observed at T3-T5 on the left side. However, the methodological quality of this study was generally poor - the results section was poorly described, it lacked both statistical analysis and examiner blinding, and therefore the findings are questionable.

In a similar, but more recent study, Gwirtz et al (2007) investigated the relationship between induced heart ischemia in 9 dogs, and increased palpatory muscular tone over T2-T5 on the left side. A cross-over design was utilized. In this cross-over design, each dog was randomly allocated to receive either the intervention or no intervention (control). After a specified time, the dogs were ‘crossed over’ and allocated to receive no intervention (control) or the intervention respectively. Using this cross-over design each dog served as their own control. Blinded Manual Palpatory Assessments (MPA) of muscle tissue over the transverse spinal processes at right and left segments of T2–T5, and T11–T12 (control segments), were conducted. myoelectric activity was also measured over the segments that were palpated. Gwirtz et al (2007) claim that MPA revealed findings of increased muscle tension/firmness during ischemia in the T2–T5 segments on the left, but not on the right or in the control segments. Correlated with these findings was an increase in myoelectric activity in T4-T5.

Although Gwirtz et al (2007) study is the first study that has reported a correlation between segmental myoelectric activity and palpable findings of increased muscle tension/firmness, there was limited control for examiner bias, and therefore caution must be used when interpreting the results. It is unclear how relevant visceral SD is to patients. It also seems implausible that treatment of visceral SD using manipulation of body tissues would interrupt the VS reflex arc, leading to better visceral functioning.

In a side study, Gwirtz et al (2007) sought to determine if the sympathetic nerves were involved in the VS reflex, as proposed by osteopathic literature. They investigated this by dissecting the sympathetic nerve in the hearts of three dogs prior to performing the intervention (ischemia). In these dogs, researchers found no palpatory changes of increased tone of segmentally related muscle tissue or changes in myoelectric activity. The authors suggested sympathetic neural stimulus may function as the mechanism for the visceral-somatic reflex (Gwirtz, 2007). This is one of the first attempts to investigate the pathophysiological mechanisms involved in the VS reflex.
reflex. Although examiners were blinded to the protocol being used in both studies, the examiners were still aware of the objectives and hypothesis of the study. There was also limited control for examiner bias (for example covering up heart rate monitors, chest tubes and equipment that could potentially cause bias results). Gwirtz et al (2007) also neglected to report on the number of examiners used to conduct the MPA, their training, clinical experience, and whether the examiners involved were blinded to each other’s findings.

1.6.2. RESEARCH ON RELATIONSHIP BETWEEN ORGANIC DISEASES AND SOMATIC DYSFUNCTION

Several authors have researched the association of palpatory findings in populations with certain disease states, particularly gastrointestinal, cardiovascular and respiratory disorders. However, results from early studies (pre-1990) considering associations between segmental SD of the spine and disease entities are uniformly of low rigor. Many of these early studies (Beal, 1983; Beal & Morlock, 1984; Nicolas et al., 1985 Tarr et al.,1987) lack in methodological quality, blinding, the use of control groups, and the use of multiple examiners to determine the reliability of the diagnostic procedures used. For physical findings to have any value they must be consistently reliable. It is also essential to use a control group in order to develop a baseline of SD in ‘normal’ population for comparative purposes with disease states (Fritz & Wainner, 2001).

Beal (1983) conducted three studies to investigate the relationship between segmental SD and various disease states (Cardiovascular (CV) and Gastrointestinal (GI)). In Beal’s first study, 108 cardiac patients were examined for evidence of palpatory cervical and/or thoracic SD. The patients were recruited from the cardiovascular service of the Chicago College of Osteopathic Medicine. Subjects were recruited from those attending cardiac referral, cardiac rehabilitation clinics, and those admitted to the teaching hospitals. Palpatory examination revealed an association between cardiac disease (Coronary Artery Disease (CAD), ischemia, and/or CV disease with hypertension) and SD of thoracic spinal segments of T1-T5 on the left, with T2 and T3 most frequently involved.
In a second study by Beal (1983), 19 patients with a range of GI diseases were examined to determine if they had segmentally related SD associated with their disease state. Findings of segmental SD were more frequently found in T5-T12, the automatic innovation of the GI organs.

In a third study, Beal (1983) investigated the predictive value of segmentally related palpable findings for CV or GI disease. Palpable findings included paraspinal muscle tension and reduction of costovertebral motion. Twenty-five patients with either CV or GI disease participated in the study. Examiners were blinded to patient’s diagnosis to determine if they could identify by palpation each patient’s diagnosis (GI or CV disease). In this study, agreement of either GI or CV disease was made by examiners in 76% of the cases. However, examiners only had to decide on one of two states (GI or CV disease) for each subject, therefore credit for agreement should only be given in excess of chance. There are several other additional shortcomings apparent in these three studies by Beal (1983).

Firstly, Beal (1983) neglected to report on the number of examiners used to conduct the palpatory examination, their training, and clinical experience. No examiner blinding occurred in the first two studies, and the study populations were examined from the cardiac or gastrointestinal service (therefore examiners may have had an expectation of what pattern of SD would be found). Even though examiners were blinded in the third study, the subjects were recruited from a hospital population, and hence are more likely than a non hospitalized population to exhibit objective signs related to their disease states for example a patient with heart disease may have signs of labored breathing and swollen ankles. These cues may inform the examiner of what disease process is occurring in the patient and thus what segmentally related spinal level they would expect to find SD in, causing examiner bias. Additionally, no asymptomatic control group was used, despite the fact that it is essential to use a control group in order to develop a baseline of SD in ‘normal’ population for comparative purposes with disease states (Fritz & Wainner, 2001).

In a subsequent study, Beal & Morlock (1984) sought to identify the incidence and location of SD in 40 patients with pulmonary disease (including Chronic Obstructive Pulmonary Disease (COPD), asthma, emphysema, bronchitis, pulmonary embolus, carcinoma of the lung, and pulmonary tuberculosis). All patients were selected from the pulmonary department of Michigan Hospital. Beal & Morlock (1984) concluded that all patients showed evidence of SD in their thoracic spine, with a higher prevalence in the T2-T7 region (spinal levels T2-T7 share the same
autonomic nerve supply as the lungs). Several weaknesses are apparent in this study. No inclusion/exclusion criteria were reported in the article. Half of the participants had reported a history of back problems during their lifetime; therefore the SD findings could be attributed to structural changes resulting from a history of back pain. Beal & Morlock (1984) also neglected to report on the number of examiners used to conduct the palpatory examination, their training and clinical experience. Also absent was the use of a control group, blinding of examiners, and verification through inter-examiner reliability tests. The evidence drawn from this study therefore cannot entirely be attributed to palpatory findings alone.

In light of the previous studies where observation clues (e.g. coughing or wheezing of a subject with bronchitis) may have lead to examiner bias, Tarr et al (1987) conducted a study to compare palpation versus observation accuracy. Tarr et al (1987) sought to determine if five osteopaths could diagnose three groups of patients (31 with asthma in remission, 22 with GI disorders, and 47 control subjects with no history of asthma or GI disorders) by palpation and/or observation alone. Examiners were blinded to patient history and diagnosis. One group of three examiners observed the patients, whilst the other group consisting of two examiners palpated and observed the subjects. Both the palpating and observational examiners could not categorize the subjects accurately, and no increase in diagnostic accuracy was seen in the palpating group versus the observational (non-palpating) group (Tarr et al., 1987).

In contrast to Tarr et al (1987) study, where no segmental SD were correlated with asthma or GI disorders, a study by Nicholas et al. (1985) reported associations between palpable segmental SD and MI. Nicholas et al., (1985) found a higher incidence (67%) of SD in the thoracic spine (T1-T8) in 25 patients with MI compared to a control group of 22 subjects without MI, who had only a 29% incidence of SD (Nicholas et al., 1985). Also, statistically significant differences were found in T1-T4 on the left, and T4 on the right, of the MI group when compared to the same sites in the control group (where SD was distributed throughout T1-T8). Nicholas et al (1985) concluded that MI is accompanied by paravertebral changes detected by palpation. Examiners were blinded to the patients’ diagnosis and all examinations were conducted at the hospital. Steps were undertaken to match visual cues of the MI group with the control group; these clues included monitors, gowns, identity bracelets, intravenous lines, and chart covers. Although these steps were undertaken, the five physicians used to conduct the examination all had previous experience with cardiac patients, therefore they may be more aware of other observational signs of cardiac disease that may have biased their results. The initial observation of patients propped up in bed may also have given away clues to the examiner that the patient
may have been suffering from heart related orthopnea. Research assistants and patients visual expressions may have also given clues away to what the patient was presenting with. As such, associations between SD and MI made in this study cannot be definitively attributed to palpatory findings alone.

Licciardone, Fulda, Stoll, Gamber, & Cage (2007) conducted a study to investigate if there is an association between Type II diabetes and segmentally related palpatory findings of the spine. The study population consisted of 60 diabetics and 32 non-diabetics. Six examiners took part in the study, all of whom were medical students who undertook an additional year of osteopathic manipulative medicine training. Examiners were grouped into three sets of pairs, and all examiners were blinded to the patients’ history. The authors concluded that there is an association between tissue texture changes at levels T11-L2 on the right and Type II diabetes. Licciardone et al., (2007) suggest that their findings may be attributed to the VS reflex. However, inter-examiner reliability of tissue texture changes was very poor (kappa = 0.09). For examinations to be clinically useful they must be both accurate and reliable (Fritz & Wainner, 2001).

In spite of the suggestive early research on VS reflexes, the physiological concept of afferent nerve impulses from the disturbed organ getting relayed to sympathetic and peripheral motor efferents of the same spinal level as the organ resulting in palpable changes in somatic tissue at that level cannot be considered to be objectively and definitively demonstrated. A number of critical questions are yet to be addressed:

- Does the severity of organic disease increase the VS reflex and thus lead to more palpable changes? Or does the severity of organic diseases increase likelihood of that subject exhibiting more observational signs of the disease and thus increase examiner bias?
- Do all organs and all tissue types exhibit a VS reflex?
- Could segmental SD be related to biomechanical changes related to the patient’s condition rather than the VS reflex?
- In relation to asthma, how would the knowledge of a visceral SD further advance the utility of diagnosis, or perhaps prognosis?

It is unknown whether the severity of a disease is related to an increase in the VS reflex. It is possible, but not yet investigated, that the VS reflex is absent in subjects where their disorder is in remission, minimal in subjects where the severity of their condition is mild and increases exponentially with the severity of their condition. Most studies that have investigated the
incidence of segmentally related SD have recruited outpatients from hospital (Beal, 1983; Beal 
& Morlock, 1984; Nicolas et al., 1985) where the severity of their disease is likely to be greater 
than subjects recruited from the general public. However when the severity of a disorder is 
greater, there is also a greater likelihood of that subject exhibiting observational signs of the 
disease (e.g. pursed lip breathing in a patient with Chronic Obstructive Pulmonary Disease) that 
are being inadvertently observed by clinical examiners in these studies. It is possible that 
examiners may subliminally take on these observational signs which may be incorporated into 
their palpation findings. It is unknown whether segmental palpatory changes in subjects with 
severe disorders are related to the VS reflex or related to examiner bias.

Two studies have reported increased myoelectric activity of segmentally related paraspinal 
muscle tissue during induced animal heart ischemia or MI (Schoen & Finn, 1978; Gwirtz et al., 
2007). However the magnitude of an association between increased myoelectric activity and 
altered paraspinal muscle tissue on the onset of MI or heart ischemia is small, and also 
questionable, due to a lack of methodologically rigorous studies. There is also insufficient 
numbers of quality studies demonstrating an association between increased paraspinal 
myoelectric activity and segmental palpable changes. As such, the usability of palpable 
segmentally related SD is doubtful.

No studies have investigated the relationship between bronchial obstruction from asthma and 
changes in paraspinal EMG activity. The heart and bronchial tube are made up of completely 
different tissues: the heart is comprised of cardiac muscle while the bronchus is comprised of 
smooth muscle. Autonomic regulation differs for different tissues. Smooth muscles are regulated 
mainly by hormones in addition to neurotransmitters (mainly acetylcholine) and cardiac muscle 
is regulated mainly by neurotransmitters (epinephrine and norepinephrine) (Guyton & Hall, 
2006). Although it is not within the scope of this study to investigate the differences between 
these two tissues in detail, it is plausible that the VS reflex may be dissimilar in the bronchial 
tube than the heart. Although of poor quality there is a small amount of research supporting the 
concept of VS reflexes originating from the heart (Schoen & Finn, 1978; Gwirtz et al., 2007), 
however, there is a lack of evidence to support the concept of a VS reflex originating from 
smooth muscle of the bronchial tube.
It is possible that studies which have reported palpatory findings of segmentally related SD associated with visceral disease are actually palpating SD that is a consequence of biomechanical changes resulting from the disease - rather than being mediated through a putative viscerosomatic reflex. For example, MI patients may suffer from shortness of breath, therefore increasing the use of the accessory muscles of breathing. They may also suffer from decreased exercise tolerance. An inability to exercise may result in general muscular de-conditioning. It may also increase the likelihood of the patient spending time in a sedentary posture, which in turn results in postural related changes to their musculoskeletal tissues (e.g. muscle atrophy, osteopenia etc). It is therefore plausible that SD of the upper thoracic spine of patients with MI may result from mechanical related factors, rather than aberrant VS reflexes. It is possible that studies which find segmentally related SD associated to pulmonary disease are identifying SD that are largely a consequence of biomechanical changes related to abnormal respiration. Theoretically these changes could occur from decreased soft tissue compliance and articular restrictions which directly impinge on lymph vessels (e.g. intercostal lymph vessels), reducing capillary and lymphatic drainage, and leading to alterations in tissue texture (Doll, 2008).

In conclusion, in an osteopathic consultation the diagnosis of a palpable visceral segmental SD is unlikely to change the normal clinical reasoning process. An isolated visceral segmental SD does not appear to be problematic and there is no evidence to suggest that such an entity compromises patient outcomes. Although osteopathic literature supports the use of manipulation for the treatment of segmental SD (Nelson, 2007; Ward, 2003), it is unclear whether manipulation can abolish VS reflexes and lead to favorable changes in visceral function. Consequently, the pursuit to study visceral segmental SD seems unjustified. For segmental SD to provide clinically meaningful benefit to patient care, visceral segmental SD would have to be accurately and reliably diagnosable, and manual interventions would need to show cost-benefit advantages such as a decrease in disease severity, decrease in symptomatic state, improved patient satisfaction with treatment and cost savings in the therapeutic treatment of these diseases.

Despite the novel VS concept described by Beal in the 1980’s, advancement of objective EMG research on VS reflexes over the last three decades has been slow and have attracted little contemporary interest amongst biomedical researchers or clinicians. Current osteopathic textbooks still uncritically present the ideas proposed by Beal in the early 1980s. While there has been one recent study (Licciardone et al., 2007) that indicates palpable spinal SD at specific spinal levels may be related to Type II diabetes, the reliability of procedures used in this study to
detect spinal SD was poor. Without reliable procedures to detect visceral SD, no-one can claim that segmental visceral SD and organic disease state are associated because no one can with any certainty detect segmental visceral SD. Even if visceral SD and disease entities are related, they may not be of substantial importance to patient outcomes to warrant the resources involved in conducting research on this topic.
CHAPTER 2: AN OSTEOPATHIC APPROACH TO ASTHMA

Chapter two focuses on manual therapeutic approaches for increasing lung function in asthmatics. It highlights the weaknesses and limitations in studies that have investigated manual and manipulative therapy in relation to asthma, with an emphasis on evidence-based diagnosis.

2.1. ASTHMA AND THE OSTEOPATHIC APPROACH

Whilst pharmacologic therapy has been proven to be effective in decreasing asthma symptoms by targeting bronchi inflammation and constriction (Bateman, 2008), the musculoskeletal component of ventilation is clearly not addressed by pharmacologic therapy. The musculoskeletal system is an integral component of the respiratory mechanism and, given the interdependent relationships of structure and function that exist between the lungs and the musculoskeletal structures, any dysfunction related to the musculoskeletal structure involved in the respiratory pump will inevitably affect the functional capacity and efficiency of the pumping process (Chaitow et al., 2002). Authors of osteopathic textbooks have emphasized that a person is a unit in which structure and function are interdependent (DiGiovanna & Schiowitz, 1991; Stone, 1999; Ward, 2003). From this structure/function interdependency concept stems the rationale that osteopathic treatment of asthma to eliminate or reduce SD will optimize function of the structures used for respiration, which in turn may improve pulmonary function (DiGiovanna & Schiowitz, 1991; Kuchera & Kuchera, 1994; Sammut & Searle-Barnes, 1998).

Another rationale for the use of osteopathic treatment includes the assumption that manipulative treatment of spinal segmental SD that have putatively resulted from visceral disease may interrupt the VS reflex arc, reducing the visceral efferent stimulus and thus decreasing the adverse effects associated with repetitive visceral stimulation, as well as decreasing the reaction of the paraspinal tissues at the related segmental spinal level (Beal, 1985). However, the scientific basis is not definitively established and clinical studies to support this concept are lacking. In the case of asthma, there appears to be no evidence to suggest that possible compromises in VS reflex may lead to an increase in bronchoconstriction, and thus contribute to lung obstruction.
2.2. STUDIES ON THE TREATMENT OF ASTHMA BY MANUAL INTERVENTION

Many researchers have investigated whether manual therapy has a role in increasing respiratory function (Balon et al., 1998; Bockenhauer, Julliard, Lo, Huang, & Sheth, 2002; Field et al., 1998; Guiney, Chou, Vianna, & Lovenheim, 2005; Nielsen, Bronfort, Bendix, Madsen, & Weeke, 1995; Noll, Degenhardt, Johnson, & Burt, 2008; Wheatley, Gosling, & Gibbons, 2000). However, rigorous supporting evidence for the use of manual therapy remains scarce. There appears to be a lack of randomized controlled trials (RCTs) with adequate methodological quality. In a systematic review on Manual Therapy for Asthma, Hondras et al (2008) state “there is insufficient evidence to warrant widespread use of manual therapies for asthmatic patients” (Hondras, Linde, & Jones, 2008). Only three of the 59 studies evaluated by Hondras et al (2008) met the inclusion criteria which consisted of: studies that were randomized, examined one or more types of manual therapy, included clinical outcomes with observation periods of two or more weeks, and included asthmatic children or adults (Hondras et al., 2008).

Of the three studies that did met the inclusion criteria, one study compared massage therapy versus relaxation (Field et al., 1998); the other two were chiropractic studies which investigated the effects of spinal manipulation versus a sham procedure (Balon et al., 1998; Nielsen et al., 1995). There is clearly a lack of good quality research on the manual treatment of asthma; as evidenced by the fact that only three out of 59 studies met Hondras et al (2008) inclusion criteria.

Four osteopathic manual intervention studies on the treatment of Chronic Obstructive Pulmonary Disease (COPD) or asthma have been published over the last ten years. None of these made the systematic review by Hondras et al (2008) because all assessed only the immediate effects of manual treatment (Bockenhauer et al., 2002; Guiney et al., 2005; Noll et al., 2008; Wheatley et al., 2000).

Studying the immediate effects of manual treatment on people with asthma holds little clinical use for asthmatic patients. It is likely that people with asthma aren’t concerned if their functional breathing capacity increases for five minutes after treatment, but rather whether treatment reduces the use of their symptom controller medication, and reduces their asthma exacerbations over a sustained period. Investigating the short-term effects of treatment is only useful for researchers trying to further provide proof of their concepts, and/or develop explanatory theories.
Two themes are apparent in the four osteopathic manual intervention studies on the treatment of Chronic Obstructive Pulmonary Disease (COPD) or asthma, and the three studies that met the inclusion criteria in Hondras et al (2008) review:

1. Studies using sham treatments may have inadvertently treated respiratory muscles, and/or
2. Studies are not using a broad examination procedure that assesses for all areas of SD hypothesized in osteopathic literature to be related to asthma.

These will be discussed below.

2.2.1. SHAM TREATMENTS

In the Neilsen et al (1995) and the Balon et al (1998) studies, both concluded that that when comparing manipulation of the spine to a sham treatment, no superiority existed between the two interventions. Balon et al (1998) also concluded that manipulation does not provide any adjunctive benefit to children with mild or moderate asthma (Balon et al., 1998).

However, both the manipulation and sham groups demonstrated positive changes. In Balon et al (1998) study, both groups demonstrated small increases in Peak Expiratory Flow (PEF), decreased use of beta-agonists (reliever medication), and symptoms decreased (thereby increasing quality of life). Although no changes in objective lung functions were found in the Neilson et al (1995) study, both groups showed improvements in bronchial hyper-reactivity (by 36%), and patient rated asthma severity decreased (by 34%). The positive effect observed may be due to undefined events not linked with the ‘treatment’ or it may be that the sham treatment inadvertently provided a treatment effect. Neilsen et al (1995) sham treatment consisted of “gentle manual pressure over the spine, and a non-direct type thrust” (p84). Balon et al (1998) sham procedure consisted of soft tissue massage to the paraspinal muscles and shoulders. Both sham groups used manual techniques applied to muscles that are involved in ventilation. It is therefore possible that the sham treatment had inadvertently treated SD of respiratory muscles in the upper thorax (in Balons study) and paraspinal muscles (in both studies), thereby providing a positive effect, and consequently minimizing the differences observed when comparing the two treatment groups.
2.2.2. **EXAMINATION PROTOCOL**

Many asthmatic manual intervention studies are non-specific in nature. Non-specific in nature can be defined as a treatment protocol that does not directly target SD found upon examination. Instead a blanket treatment protocol is applied regardless of whether the target area is exhibiting a dysfunctional state. It appears this non-specific treatment approach may be occurring because a broad assessment procedure to detect SD associated to asthma is lacking.

Evaluation of asthmatic intervention studies reveals at least four criteria by which their quality may be assessed. These criteria and the rationale behind them are outlined in Table 2.

Table 2: Evaluation of asthmatic intervention studies

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Rationale</th>
</tr>
</thead>
<tbody>
<tr>
<td>1) Use of an examination procedure to detect SD</td>
<td>If studies do not use an examination procedure then we cannot ascertain if subject has dysfunction related to their asthma and therefore cannot directly target dysfunctions found upon examination</td>
</tr>
<tr>
<td>2) Assess for more than one SD</td>
<td>Many studies typically only assess for spinal segmental SD and fail to include neck, shoulder and ribcage regions. An examination protocol should assess for all SD in all regions hypothesized in osteopathic literature to occur in people with asthma</td>
</tr>
<tr>
<td>3) Report on the assessment procedures used to detect SD</td>
<td>To repeat the study, clear instructions on the procedures used to detect SD are needed. The assessment procedures must also be listed to determine what areas of the body were originally investigated</td>
</tr>
<tr>
<td>4) Report on the prevalence of SD in the asthmatic and control groups</td>
<td>Without reporting on the prevalence of SD in both the control group and the asthmatic group it cannot be assumed that the SD found upon examination are directly attributable to asthma and thus will be amenable to treatment</td>
</tr>
</tbody>
</table>

Noll et al (2008) conducted a study on the immediate effects of manipulative treatment in elderly patients with Chronic Obstructive Pulmonary Disease. No assessment procedure was used prior to the intervention (breached criterion 1, 2, 3 & 4).

Wheatley et al (2000) investigated the immediate effects of rib raising on asthmatics. They did not utilize an assessment procedure (breached criterion 1, 2, 3 & 4).

Balon et al (1998) & Nielsen et al (1995) both investigated the long-term effects of spinal manipulation. Both used an examination procedure, however only spinal segmental SD were assessed (breached criterion 2).

Bockenhauer et al (2002) conducted a study on the quantifiable effects of osteopathic manipulative techniques on patients with chronic asthma. Participants were referred into the study by physicians who had diagnosed the subject with chronic asthma from their case history.
and physical examination. There was no description of the physical examination procedures or of their examination findings (breached criterion 3 & 4).

Guiney et al (2005) conducted a study on the effects of osteopathic manipulative treatment on pediatric patients with and without asthma. As stated by Guiney et al (2005) “a brief personal medical history and structural examination were obtained for all patients in both study groups. Somatic dysfunction was identified in the rib and thoracic regions.” (p.9). There was no mention of what the assessment procedure comprised of, therefore it is unknown whether only two examination procedures were used to identify rib and thoracic dysfunctions or if a wider array of procedures was used but only dysfunctions located to the rib and thoracic regions were found. Also the prevalence of these findings in the asthmatic or non-asthmatic group was not reported (breached criterion 3 & 4).

Without an examination procedure that assesses for all areas of SD hypothesized in osteopathic literature and without reporting on the prevalence of SD in both the control group and the asthmatic group it cannot be assumed that the SD found upon examination are directly attributable to asthma and thus will be amenable to treatment.

2.3. LIMITATIONS AND WEAKNESSES IN ASTHMATIC MANUAL INTERVENTION STUDIES

A major limitation with asthmatic manual intervention studies is that in the asthmatic population there may be:

- those with asthma and SD that are clinically relevant and the SD responds well to treatment
- those with asthma and SD that are not clinically relevant (i.e. related to something else), and the SD doesn’t respond to treatment
- those with asthma and no SD (Lucas & Moran, 2005).

In the Hondras et al (2008) systemic review, the author states “there is insufficient evidence to support or refute the use of manual therapy for patients with asthma” (p.1). As such, reassessment of the role of the association between SD and asthma seems justified. The prevalence of SD in the asthmatic population is unknown; not all asthmatics may have overt findings of SD, therefore not all asthmatic patients may respond to the treatment intervention.

Lessons about a suitable approach to study spectrum disorders may be drawn from studies on low back pain (LBP). Wand & O’Connell (2008) discuss an approach that uses ‘sub-groups’ that
represent different types of LBP. This sub-group approach may be useful when investigating asthma.

A ‘sub-group’ refers to group with 1) a unique etiological mechanism of the underlying disorder, or 2) a group of patients with certain clinical profiles that are likely to benefit from treatment (Wand & O’Connell, 2008). Wand & O’Connell (2008) state that many clinicians feel that the effects of an intervention are diluted when applied to a diverse group of patients with chronic LBP with varied treatment needs. The same argument applies to manual interventions for asthma. For example, in a population of asthmatic patients that have SD there may be a sub-group that respond well to treatment. At present those with asthma and SD that is not clinically relevant cannot be differentiated from those with asthma and SD that is clinically relevant. However, people with asthma and no SD may be differentiated from those with asthma and SD if an examination procedure is developed that can accurately and reliably detect SD related to asthma. If an examination procedure is established, then asthmatic manual intervention studies could determine if those with SD and asthma are greater responders to manual treatment compared to those asthmatics without SD.

Adequate diagnosis of SD related to ventilatory compromise should be at the forefront of any manual intervention study. However, in order for any intervention study to have maximal efficacy, a method for accurately and reliably diagnosing SD related to asthma is needed, diagnosis being the cornerstone to any intervention study (Fritz & Wainner, 2001). With only anecdotal evidence supporting osteopathic literature on the relationship between asthma and SD, the following questions have yet to be adequately addressed:

• is there a meaningful association between asthma and SD? If so, then:
• are there physical examination procedures that can reliably and accurately identify SD associated with asthma?

If a meaningful association between SD and asthma can be recognized and physical examination procedures can reliability and accurately identify SD associated with asthma then a treatment protocol could be implemented to treat those SD found upon examination. In an unpublished thesis, Russell (2006) sought to construct a reliable and accurate clinical examination program, which aimed to assess the many SD hypothesized in the literature to be involved in asthma. Of the 20 examination procedures which were utilized in the study, nine exhibited moderate accuracy (>0.5) and acceptable levels of inter-examiner reliability (kappa=0.4-0.6) (Russell, 2006). The procedures are as follows.
1. Active side bending of the truck (assessment for dysfunction)
2. Chin tuck (assessment for discomfort)
3. Thoracic spine rotation (assessment for discomfort)
4. Thoracic spine extension (assessment for discomfort)
5. Combined lateral rotation and flexion (CLRF) (assessment for discomfort)
6. Chest expansion (assessment for dysfunction)
7. Trapezius stretch (assessment for dysfunction)
8. Suboccipital tissue texture (assessment for discomfort)
9. Sternocleidomastoid tissue texture (assessment for discomfort)

The physical examination procedures used by Russell (2006) utilize one or a combination of assessments including:

- palpation for tissue texture changes
- static palpation for asymmetry of structure, observation/palpation for location of landmarks
- motion palpation for restriction in motion, hypermobility, hypomobility or for changes in quality of articulatory end feel
- palpation to elicit tenderness, pain or discomfort.

Many disciplines of manual therapy including chiropractic and physiotherapy also use these palpatory/observational procedures to diagnose SD. Whilst different professions may place more emphasis on one procedure than another, all data gathered from these procedures guide the practitioner in their clinical decision making. However, in order for practitioners to make valid and informed decisions when using these procedures, a satisfactory level of diagnostic accuracy and reliability is necessary. Diagnostic accuracy and reliability are discussed in the next section.
CHAPTER 3: A BACKGROUND INTO UNDERSTANDING RELIABILITY AND ACCURACY

Part three investigates the reliability and accuracy of previously published studies on manual palpatory procedures, investigating what types of palpatory procedures are most reliable. This section also looks at factors which affect examiner reliability and accuracy including consensus training prior to the application of procedures and examiner experience.

3.1. DIAGNOSTIC ACCURACY

Diagnostic accuracy helps determine the extent to which a procedure (or combination of tests) can discriminate between two conditions of interest (e.g. health and disease). The discriminative ability of a procedure can be quantified by using indices of diagnostic accuracy including sensitivity, specificity, predictive values and likelihood ratios.

In order to calculate the indices listed above a contingency table is generated (see Table 3). The contingency table compares the presence or absence of the component of interest against a reference standard. The reference standard should be the accepted ‘best’ method for measurement of the variable of interest. The reference standard is sometimes termed the ‘gold standard’ (Fritz & Wainner, 2001). In many manual palpation studies the variable of interest cannot be measured. For example, there is no practical objective method to measure ‘tightness’ of a muscle, and as such the accuracy of a procedure that theoretically assesses for ‘tightness’ of a muscle cannot be determined. To some degree this is overcome by:

- comparing the association between a variable of interest (e.g. tightness of the suboccipital muscle) to a pathology that is hypothesized to be one of the etiological factors that gives rise to the particular variable examined (e.g. asthma).
- using inter-examiner reliability studies to verify repeatability of results (reliability is discussed below).

Table 3: Example of diagnostic accuracy 2 x 2 contingency table

<table>
<thead>
<tr>
<th>Procedure</th>
<th>Reference Standard</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Present</td>
<td>Present</td>
<td>True positives (TP)</td>
</tr>
<tr>
<td>Absent</td>
<td>False negatives (FN)</td>
<td></td>
</tr>
<tr>
<td>Present</td>
<td>Absent</td>
<td>False positives (FP)</td>
</tr>
<tr>
<td>Absent</td>
<td>True negatives (TN)</td>
<td></td>
</tr>
</tbody>
</table>
3.2. RELIABILITY

Reliability is defined as the extent to which repeated measurements will agree (e.g. whether examiner A will get the same results as examiner B) (Rothstein et al., 1991, p. 3). A reliable procedure should provide a consistent measure of the variable of interest, and reliability must be determined in order to determine the validity and utility of the procedure. Reliability of categorical variables is commonly measured using the kappa coefficient (κ) statistic (Sim & Wright, 2005).

To determine the inter-examiner reliability of a palpatory procedure multiple examiners are used. These examiners should be blinded from each others evaluation and findings in order to determine the extent of agreement between their findings. Reliability of palpation procedures to detect SD is fundamental in validating the concept of SD in diagnosis of musculoskeletal disorders.

3.2.1. THE KAPPA STATISTIC

Results of categorical data (e.g. “present” or “absent”) from a procedure judged by two examiners are commonly entered into a 2 x 2 contingency table (see Table 4). From this table the kappa coefficient can be calculated. The kappa statistic takes into account agreement beyond change and can be considered a “true” agreement, indicating “achieved-beyond-change” agreement (Sim & Wright, 2005). In manual medicine, the kappa coefficient (κ) is the most widely used statistic to indicate the extent of agreement between two or more examiners.

Table 4: Example of reliability 2 x 2 contingency table

<table>
<thead>
<tr>
<th></th>
<th>Examiner B</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Present</td>
</tr>
<tr>
<td>Examiner A</td>
<td></td>
</tr>
<tr>
<td>Present</td>
<td>Agreement A</td>
</tr>
<tr>
<td>Absent</td>
<td>Disagreement C</td>
</tr>
</tbody>
</table>


Interpreting kappa values
The values of kappa range from -1 to 1, where +1 indicates perfect agreement, whilst 0 indicates agreement no better than chance, and -1 indicates perfect disagreement.

Whilst there is no set cut-off point for ‘acceptable’ reliability, many studies of manual therapy have arbitrarily set the threshold for ‘acceptable’ reliability to kappa=0.4 (Dixon, Munro, & Silcocks, 1998; Seffinger et al., 2004; Stochkendahl et al., 2006). Like most statistics, the kappa coefficient has weaknesses and assumptions in its calculation and use. The magnitude of the kappa can be influenced by factors such as prevalence and bias. Prevalence and bias are explained in the following section.

3.2.2. PREVALENCE AND BIAS INDICES
The kappa coefficient is appropriate when the marginal totals of the contingency tables are balanced. However, problems exist when the proportion of agreements (cells A and D of Table 4) are imbalanced and/or when the proportion of disagreements are imbalanced (cells B and C of Table 4). These imbalances in concordant and discordant findings can be measured using the Prevalence and Bias indices respectively (Cunningham, 2009; Sim & Wright, 2005).

When there is a large bias, the magnitude of the kappa is higher than when bias is low or absent. When there is a very high or very low prevalence of the index being measured, the kappa is artificially reduced. Because of this paradox, the use and reporting of the prevalence and bias adjusted kappa values as additional agreement measures are important, especially when prevalence cannot be easily achieved, such as occurs when the disorder of interest has no gold standard. Studies that report low kappa values without reporting overall agreement and adjusted kappa indexes (PABAK) tend to underestimate the strength of agreement (Huijbregts., 2002).

3.3. AN EXAMINATION OF THE LITERATURE ON RELIABILITY OF MANUAL PALPATION PROCEDURES
Numerous studies exploring the level of reliability of various manual palpation methods have been published over the last decade, including several reviews (Haneline, Cooperstein, Young, & Birkeland, 2008; Haneline & Young, 2009; Hollerwöger, 2006; Huijbregts., 2002; Hestboek & Leboeuf-Yde, 2000). However, these reviews are not comprehensive in their analysis of assessment procedures and/or anatomical regions. Some only assess one or two palpation procedures (Haneline, Cooperstein, Young, & Birkeland, 2008; Haneline & Young, 2009;
Huijbregts., 2002), whilst others focus on only one anatomical region (e.g. cervical spine) (Hestboek & Leboeuf-Yde, 2000; Hollerwöger, 2006).

There appears to be high level of variation in the type of methodological scoring systems that exist, as well as variation in cut-off limits to determine ‘high’ or ‘low’ quality studies. Studies by Haneline et al., (2008); Haneline & Young, (2009); & Stochkendahl et al., (2006) use a 6-point quality scoring system to assess for methodological quality. The authors define ‘high’/‘acceptable’ quality studies are those that score ≥ 50%. The 6-point quality criteria is based on the following: order of observers conducting the test(s) were randomized; the study population contained a mix of symptomatic and asymptomatic subjects; observer blinding to other observers findings; observer blinding to confounding information; subject blinding to observer findings; kappa or inter correlation coefficients was used for statistical analysis.

In contrast, studies by Hesboek & Leboeuf-Yde, 2000; & Gemell & Miller, 2005 use a 10-point quality scoring system to assess for methodological quality and rate high/acceptable quality studies as those that score ≥ 75 %. The 10-point quality criteria is based on the following: study population is well described; subject inclusion/exclusion criteria is clearly described; symptomatic subjects were used; subjects naïve/unable to affect result; examiners described clearly, including experience; examiners were blinded to subjects clinical information; examiners were blinded to each others findings; order of examiners was randomized; protocol is clearly described; appropriate reliability statistics were used. Those with inferior methodological scoring systems, or with low cut-off limits, include studies classified as ‘poor quality’ by other reviewers.

Notwithstanding differences in how different authors employ quality scoring methods, it appears that there is evidence for the use of palpation to elicit tenderness/discomfort or pain whilst there is a lack of evidence for the use of motion palpation, static palpation and palpation for tissue texture changes.
3.3.1. The Reliability of Manual Assessment Procedures

Palpation for tenderness/pain/discomfort

Out of all palpation procedures (palpation for tenderness, asymmetry of structure, restriction of motion or tissue texture changes) it appears that higher levels of inter-examiner reliability were found when palpation is used to elicit pain and response to pain is included. This is consistent with findings from the following four reviews.

- Hesboek & Leboeuf-Yde (2000) investigated chiropractic procedures used on the lumbo-pelvic spine. Acceptable reliability was found in studies that elicited pain on palpation. However motion palpation had mixed findings, and visual inspection of landmarks had poor agreement (Hestboek & Leboeuf-Yde, 2000).

- Stochkendahl et al (2006) assessed the reliability of manual examination procedures of the spine. The authors concluded that strong support existed for acceptable inter-examiner reliability of palpation for osseous and soft tissue pain. Strong support for clinically unacceptable levels of reproducibility was found for motion palpation and soft tissue changes (Stochkendahl et al., 2006).

- Seffinger et al (2004) reviewed the reliability of spinal palpatory procedures. The author concluded that (from highest to lowest) pain provocation, motion palpation, then landmark location demonstrated acceptable reliability. The procedures used to examine the soft tissue changes lacked reliability (Seffinger et al., 2004).

- Haneline & Young (2009) reviewed the reliability of static palpation of the spine and sacroiliac joints. Of the 10 studies on palpation procedures used to elicit tenderness/pain, 40% reported acceptable inter-examiner reliability ($K \geq 0.4$). Of the seven studies on landmark location only 14.3% reported acceptable reliability. Overall, higher levels of inter-examiner reliability were found when palpation was used to elicit pain (Haneline & Young, 2009).

Motion palpation

Two reviews have been conducted on manual palpation procedures of the spine (including sacroiliac joints (SI)) or related structures (Hollerwoger, 2006; Van Trijffel, Anderegg, Bossuyt, & Lucas 2005). Although the lumbar spine and SI joints are areas not commonly examined in patients with asthma, studies which examine these areas are included in this literature review. This is due to the similarity that exists between the procedures used to assess for SD of the lumbar and SI spine and procedures used to assess for SD of the thoracic and cervical spine.
Hollerwöger (2006) assessed the methodological quality and outcome of manual cervical spine examination procedures. The author concluded that motion palpation for segmental cervical dysfunction is questionable, and that more practically orientated study designs which include incorporating subjective data on pain and using uniform reference standards may increase reliability (Hollerwöger, 2006).

Van Trijffel et al (2005) reviewed the reliability of passive and segmental motion assessment of the cervical and lumber spine. The authors concluded that no definitive conclusions could be drawn regarding the clinical usefulness of motion palpation, as most studies did not fulfill criteria for internal and external validity. The authors also concluded that the reporting of statistical data was inadequate. The designs of the studies included in this review were so poor that no conclusions could be drawn from the data (Van Trijffel et al. 2005).

There are two further studies that deal with reliability of motion palpation procedures related to the lumber spine and SI joints.

Schneider (2008) investigated the reliability of three palpation procedures of the Lumber and SI joints: 1) pain provocation testing; 2) springing segmental mobility testing and 3) prone lumber instability test. Thirty-nine patients with low back pain (LBP) (aged 18-65yrs) were assessed by two qualified chiropractors. Results showed pain provocation procedures had the highest reliability ($K$=0.21-0.73) (PABAK= 0.34-0.74) followed by the prone instability test which showed ‘moderate’ reliability ($K$=-0.46-0.54) (PABAK=0.58-0.58), then springing segmental mobility palpation procedures which showed ‘poor’ to ‘fair’ reliability ($K$=0.17-0.17).

Amir Massoud Arab, Abdollahi, Joghataei, Golafshani, & Kazemnejad (2009) investigated the inter-examiner reliability of single motion palpation procedures for the SI joint, pain provocation procedures for the SI joint, and combinations of either pain provocation or motion palpation tests. Twenty-five subjects all with LBP were examined by two examiners. The study showed ‘moderate’ ($K$=0.44) to ‘substantial’ ($K$=0.78) reliability for the individual motion palpation and pain provocation tests. Reliability increased to between ‘moderate’ ($K$=0.52) and ‘almost perfect’ ($K$=0.92) when clusters of motion palpation or provocation procedures were used (Amir Massoud Arab et al., 2009). This study suggests that when clusters of both motion palpation and pain provocation procedures for the clinical assessment of the SI joint are used, higher reliability is generated than that of individual procedures alone. Many studies evaluate the reliability and validity of single tests. However, in clinical practice, clinicians often use combinations of
procedures; therefore it makes sense to study combinations of procedures, especially when the literature suggests that clustering procedures together may enhance reliability.

**Tissue texture changes**

Reliability of palpation to detect tissue texture changes appears to be less reliable than palpation to elicit pain and is less researched than other manual palpation procedures. Paulet & Fryer (2009) conducted a study to determine inter-examiner reliability of palpatory identification of abnormal tissue texture changes in the thoracic paraspinal region. This study used 10 final year osteopathic students who each examined 10 subjects. ‘Fair’ levels of reliability were reported for palpation of tissue texture changes \( (K=0.26) \). The authors concluded that the “assessment of texture change is complex and not highly reproducible between examiners” (Paulet & Fryer 2009, p. 96).

Fryer, Morris, & Gibbons (2004) presented a narrative-style review on literature pertaining to the detection of altered paraspinal tissue texture. They also examined the literature for explanations for altered tissue texture and evidence for the plausibility of paraspinal muscle spasm, and muscle dysfunction associated with LBP. Fryer et al. (2004) concluded that palpation for tenderness was more reliable than for tissue texture changes. However, only kappa coefficients for reliability studies on tenderness were reported, therefore we do not know how much more reliable palpation for tenderness is when compared to palpation for tissue texture changes. Also no rigorous inclusion criteria were used, therefore literature was included regardless of the methodological quality. Again it appears that than when palpation to identify tissue texture changes is used as an assessment component, the reliability generated is less than palpation for pain.

3.4. **FACTORS THAT AFFECT RELIABILITY**

Whether consensus training reduces error and thus improves reliability is still under debate. Fryer, McPherson, & O'Keefe (2005) investigated the effect of examiner training on: 1) static palpation procedures, which assess the symmetry of pelvic landmarks and 2) motion palpation procedures to detect SI dysfunction. Overall, the five osteopathic student examiners that had attended two consensus training sessions to standardize procedures achieved marginal increases in inter-examiner reliability than the five student examiners who did not undergo training (Fryer et al. 2005). Similar findings were found in a study by Degenhardt, Snider, Snider, & Johnson, (2005) who investigated the reliability of common osteopathic palpatory procedures used to
evaluate the lumbar spine. Pre-consensus training, reliability ranged from $K=-0.02-0.34$. Post-consensus training, reliability increased, particularly in procedures used to determine tissue texture changes ($K=0.45$), and procedures used to elicit tenderness ($K=0.68$), demonstrating that consensus training may improve reliability (Degenhardt et al., 2005). However, the increases in reliability observed after consensus training may not be entirely attributable to the consensus training. The second time you repeat a procedure you may be better at it purely because you have had previous practice, therefore you can’t rule out that the increases in reliability observed in Degenhardt’s et al (2005) study are due to practice.

In contrast to Degenhardt et al (2005), a systematic review on the reliability of spinal palpatory procedures concluded that procedural consensus training did not improve reliability (Seffinger et al., 2004). In another systematic review, which investigated the reliability of manual examination procedures of the spine, Stochkendahl et al (2006) also concluded that consensus training had little influence on the level of agreement achieved by examiners (Stochkendahl et al., 2006).

It is suggested by Chaitow (2008) that examiner experience may influence reliability, with the expectation that increased experience is associated with more refined palpation skills, and thus better reliability. However this expectation is not well supported in the reliability literature. A few studies show that experience has little to no effect on reliability (Kmita & Lucas, 2008; Moran & Ljubotenski, 2006; Seffinger et al., 2004; Stochkendahl et al., 2006).

Kmita & Lucas (2008) investigated the reliability of physical static palpation procedures used to detect asymmetry of pelvic landmarks, using both experienced osteopaths and final year osteopathic students. No meaningful differences in reliability existed between the experienced osteopaths versus the student osteopaths.

In a systemic review published by Stochhkendahl et al (2006), which investigated the reliability of manual examination procedures of the spine, the authors conclude that examiner experience had little influence on the studies results. The same conclusion was drawn in a systemic review on the reliability of spinal palpatory procedures by Seffinger et al (2004).

Moran & Ljubotenski (2006) conducted a study to compare the reliability of four groups of clinicians who visually assessed the depth of 60 subjects’ lumbar curves by watching video clips. Curves were assessed by placing a mark on a 100mm scale with end points labeled
“maximum lordosis” and “minimum lordosis” and the midpoint labeled “normal”. Reliability was acceptable for all groups. However, fifth year osteopathic students and clinicians with five years clinical experience had slightly better reliability than both clinicians with more than 15 years experience and first year osteopathic students. The differences between groups were small and reliability was comparable across all four groups (Moran & Ljubotenski, 2006). This study demonstrates that acceptable levels of reliability for visual assessment of the lumbar curve shape can be achieved when using experienced osteopaths or student osteopaths.

Billis, Foster, & Wright (2003) conducted a study on the reliability of a static palpation procedure which assessed for the location of spinal levels C5, T6 and L5. Three different types of examiners assessed the same five subjects. The examiner groups consisted of: 13 final year student clinicians (with no clinical experience); ten moderately experienced clinicians (with two years experience); and seven very experienced clinicians (with five years clinical experience). The results demonstrated that there were no differences between the moderately and very experienced clinicians, who had better agreement than student clinicians (Billis, Foster, & Wright, 2003). However all groups demonstrated poor reliability at all spinal levels.

In conclusion, it appears that clinical experience has little effect on reliability. The notion suggested by Chaitow (2008) that examiner experience equates to greater reliability is not well supported by the literature. However, the small number of studies on the topic of examiner experience presents a limitation in drawing conclusions about the effect of experience on examiner reliability. Also, the argument for the use of consensus training in manual palpation studies at present does not appear strong, due to a lack of available evidence. However, all practitioners regardless of experience have different ways of interpreting information based on their perceptions, beliefs and previous experience. Consensus training therefore appears to be a reasonable way to of standardizing cut-off limits between ‘dysfunctional’ findings and ‘normal’ findings, and in enhancing examiners interpretation of these cut-offs. This has the potential to improve reliability. Further studies are needed to fully investigate the matter.
SUMMARY

Asthma not only causes inflammation and constriction of the airways; but it is also hypothesized to cause biomechanical changes to the ‘respiratory pump’ resulting from the greater work required to overcome airway obstruction. The respiratory pump consists of an array of muscles, joints and neural reflexes that act in concert to create an effective pumping mechanism (lengthening and shortening the thoracic cavity) allowing lung expansion and contraction, and movement of air in and out of the lungs respectively (Breslin, 1996; Chaitow et al., 2002).

Pharmacotherapy is the mainstay of asthma treatment and has been demonstrated to be effective in decreasing asthma symptoms (Bateman, 2008). Although pharmacotherapy may not directly alter the musculoskeletal component of respiration, it may indirectly affect it by decreasing airway inflammation and obstruction (constriction), and thus also decrease the musculoskeletal work required to overcome the obstruction. However, even if asthma exacerbations are well controlled by pharmacotherapy, if the work of breathing to overcome the initial obstruction has been large or prolonged, the coordinated activity of muscles, joints and neural reflexes may be lost and a faulty ‘upper chest’ breathing pattern may remain (Courtney, 2009). Structural modifications within the respiratory pump are hypothesized to occur as a result of this faulty breathing pattern. These structural modifications are described as SD in osteopathic literature. Osteopathic theory says that by treating (eliminating or reducing) SD, functional improvements in the respiratory pump may result, which in turn will optimize pulmonary function (DiGiovanna & Schiowitz, 1991; Kuchera & Kuchera, 1994; Sammut & Searle-Barnes, 1998).

Although many researchers have investigated whether manual therapy has a role in increasing lung function and reducing asthma symptoms, none of these effects have been demonstrated in robust and systematic studies. There appears to be an absence of comprehensive examination procedures to detect SD related to asthma. Without an adequate examination procedure to evaluate all SD related to asthma clinicians are unable to make informed decisions on the type of manual therapy treatment techniques to match their findings. Currently, no asthma manual intervention study has been undertaken to investigate the treatment of SD found on examination because no reliable and valid examination protocol has been developed. Applying non-specific blanket treatment interventions as opposed to targeted treatment approaches to specific SD found upon examination needs to be developed.
The types of SD one would expect to find upon examination of people with asthma have been hypothesized in osteopathic literature, however, there is a lack of research to inform this hypothesis. One unpublished thesis by Russell (2006) sought to determine if there is an association between asthma and SD that could be identified by physical examination. Two qualified osteopaths conducted 20 physical examination procedures on a population of 17 asthmatics and 22 non-asthmatics. Examiners were blinded to the patients’ clinical histories as well as each others results. Nine of the 20 procedures had acceptable reliability (>0.4 kappa) and accuracy (>0.5). This is the first study to show an association between SD and people with asthma. It is also the first study to identify a protocol of procedures to use on people with asthma. However, Russell’s (2006) study was small and only two examiners were used. As such, research using more examiners is needed to verify the inter-examiner reliability of the examination procedures and determine their diagnostic utilization.

**The research question addressed in this thesis is therefore:** Can somatic dysfunction associated with asthma be reliably and accurately found on examination using the procedures used in Russell’s (2006) unpublished thesis by both student osteopaths and experienced osteopaths?

If a valid and reliable examination procedures to detect SD associated to asthma could be determined, then further research into the efficacy of treatment interventions for the SD can proceed on a more robust basis.
SECTION 2: EXPERIMENTAL INVESTIGATION
# Experimental Investigation

## Chapter 4: Introduction

4.1. Asthma, somatic dysfunctions and current research on manual therapy for asthma ................................ 51

4.2. Previous studies that have investigated physical examination procedures to detect somatic dysfunction in people with asthma ................................................................. 53

4.3. Aims and objectives ............................................................................................................... 54

## Chapter 5: Methods

5.1. Study design ....................................................................................................................... 56

5.2. Recruitment of participants ............................................................................................. 56

5.3. Participant inclusion / exclusion criteria .......................................................................... 57

5.4. Recruitment of examiners ................................................................................................. 58

5.5. Research ethics approval / Participant and examiner consent ......................................... 59

5.6. Experimental design ......................................................................................................... 59

5.6.1. Consensus training ....................................................................................................... 59

5.6.2. Data collection ............................................................................................................. 62

5.6.3. Physical examination procedures ............................................................................... 67

5.6.4. Recording findings ....................................................................................................... 67

5.7. Data analysis .................................................................................................................... 67

5.7.1. Data reduction .............................................................................................................. 67

5.7.2. Assessment of the performance of a diagnostic procedures ......................................... 68

5.7.3. Data analysis of inter-examiner reliability .................................................................. 71

5.7.4. Normality of the asthmatic and asymptomatic group .................................................. 75

5.8. Subgroup analysis ............................................................................................................. 75

5.8.1. Subgroup 1 and 2 ...................................................................................................... 75

5.8.2. Subgroup 3 ............................................................................................................... 75

5.8.3. Subgroup 4 ............................................................................................................... 76

## Chapter 6: Results

6.1. Participant characteristics ................................................................................................. 77

6.2. Comparison at baseline of the asthmatic and asymptomatic groups ................................ 78

6.3. Characteristics of asthmatic group .................................................................................... 78

6.4. Validity of physical examination procedures .................................................................. 80

6.5. Inter-examiner reliability ................................................................................................. 86

6.6. Subgroup analysis for examination procedures with acceptable validity and reliability ................................................................................................................................. 93

6.6.1. Subgroup 1 and 2 analysis ......................................................................................... 95

6.6.2. Subgroup 3 analysis ................................................................................................. 97

6.6.3. Subgroup 4 analysis ................................................................................................. 98
Chapter 7: Discussion .............................................................................................................. 100

7.1. Statement of principal findings ................................................................................... 100

7.2. Reliability ..................................................................................................................... 103

7.2.1. Comparing inter-examiner reliability of experienced osteopaths versus student osteopaths ....................................................................................................... 104

7.2.2. Comparing inter-examiner reliability of dysfunction components versus discomfort components of all examination procedures ........................................... 106

7.3. Validity .......................................................................................................................... 108

7.3.1. Accuracy .................................................................................................................. 108

7.3.2. Sensitivity and Specificity ....................................................................................... 111

7.3.3. Likelihood Ratios .................................................................................................... 112

7.3.4. Positive and Negative Predictive Values ............................................................... 114

7.4. Strengths, limitations and weaknesses .......................................................................... 115

7.4.1. Internal validity ....................................................................................................... 115

7.4.2. Sample size ............................................................................................................ 116

7.4.3. Training period ...................................................................................................... 116

7.4.4. Standardizing areas of palpation ......................................................................... 118

7.4.5. Reducing examiner fatigue .................................................................................... 119

7.4.6. Pre-warm up .......................................................................................................... 119

7.4.7. Randomization of examiners ............................................................................... 119

Chapter 8: Recommendations ................................................................................................. 120

Chapter 9: Conclusion ............................................................................................................. 121
ABSTRACT

Reliability and validity of manual examination procedures for the detection of musculoskeletal dysfunction in people with bronchial asthma

OBJECTIVE: The primary purpose of this study was to evaluate the validity, reliability and potential utility of ten physical examination procedures for the detection of somatic dysfunction in bronchial asthmatics using both student and experienced osteopaths.

METHODS: This study utilized a prospective blind comparison between the presence of somatic dysfunction as identified by examination and the diagnosis of asthma. Forty-one participants aged from 18-50 years were recruited from the general public and Unitec, New Zealand. Two experienced osteopaths and two post-graduate osteopathic students were recruited to be the examiners. Participants were examined for somatic dysfunction by each examiner using the ten procedures. All examination procedures (except examination procedure 10) assessed for two components: dysfunction and discomfort. The dysfunction component assessed for a combination of one or more observational and/or palpatory findings of tissue texture changes; asymmetry of structure/location of landmarks; restriction in motion; hypermobility, hypomobility or for changes in end feel (named dysfunction component). The discomfort component assessed for participant discomfort provoked during the examination procedure (named discomfort component). The examination procedures were broken up this way in order to compare whether the assessment for dysfunction generated greater reliability and accuracy than the assessment for discomfort.

Examiners were blinded to participant’s history and to each other’s results. Inter-examiner reliability was assessed using the unadjusted (Cohen’s) kappa (K) and adjusted (PABAK) kappa statistic. Validity for each group of examiners was assessed comparing the results from each component of the procedure against the presence or absence of asthma as reported by the participant. Validity was assessed using calculations of accuracy (A), sensitivity (Sn), specificity (Sp), positive likelihood ratios (+LR) and negative likelihood ratios (-LR).

RESULTS: Reliability of dysfunction components ranged from $K=-0.03$ to 0.75 (-0.27 to 0.76 PABAK) for student osteopaths, and $K=0.08$ to 0.16 (-0.71 to 0.17 PABAK) for experienced osteopaths. Reliability of discomfort components ranged from -0.04 to 1.00 (0.61 to 1.00).
PABAK) for student osteopaths and $K=-0.11$ to $0.09$ ($0.32$ to $0.76$ PABAK) for experienced osteopaths.

Three discomfort components of three procedures performed by the experienced osteopaths detected with acceptable reliability (between $K=0.46$ and 0.76) and with small but sometimes important shifts in positive likelihood ratios (between 2.2 to 3.8+LR) discomfort in subjects with asthma. The procedures were: Chest expansion for discomfort ($K=0.76$; +LR 3.81, CI 0.44 to 32.64); Trapezius stretch for discomfort ($K=0.46$; +LR=2.22, CI 0.62 to 8.00); and Sternocleidomastoid tissue texture palpation for discomfort ($K=0.66$, +LR=2.38, CI 0.49 to 11.58). However, the same results were not recreated when using the student osteopaths.

The overall mean accuracy of all ten procedures was similar for both the student and experienced osteopathic groups (0.49 and 0.52A respectively). However, when including only those asthmatics with objective lung findings of lung obstruction (Forced Expiratory Volume in one second <80% predicted), accuracy of the student and experienced osteopathic groups increased from 0.49 to 0.65A and 0.52 to 0.67A respectively. Positive Likelihood Ratio for the student and experienced osteopathic groups increased from 0.91 to 1.46+LR and 1.48 to 2.07+LR respectively. There were minimal changes in Cohen’s kappa for both the student and experienced osteopathic groups; however the adjusted (PABAK) kappa for the student and experienced osteopathic groups increased from 0.48 to 0.65 and 0.26 to 0.67 respectively.

CONCLUSION: The current study failed to show a clinically useful association between asthma and all ten procedures collectively. Our findings are contradictory to previous claims described in textbooks (Chaitow et al., 2002; DiGiovanna & Schiowitz, 1991; Kuchera & Kuchera, 1994; Sammut & Searle-Barnes, 1998) of an association between SD and asthma. This study has however identified three discomfort components of three procedures with acceptable reliability (adjusted kappa) and with small positive likelihood ratios. If when using these three procedures one finding of discomfort is found, the probability of an association between asthma and the SD found increases by 20%. However, when using the same procedures by student osteopaths in this study, the association was not present. Further research is required to confirm the validity of these procedures.

Key words: reliability, palpation, manual examination tests, respiration, ventilation, asthma, somatic dysfunction
CHAPTER 4: INTRODUCTION

4.1. ASTHMA, SOMATIC DYSFUNCTION AND CURRENT RESEARCH ON MANUAL THERAPY FOR ASTHMA

Asthma is estimated to effect one in every six New Zealanders, which is one of the highest prevalence rates of asthma in the world (Hold & Beasley, 2001). With such a high number of citizens affected, research on treatments for asthma could have significant benefits to the country, assisting those affected to better manage their disability and consequently reduce government expenditure on health. The conventional pharmacological approach to asthma has been very effective at targeting the inflammatory response; however the downstream musculoskeletal dysfunctions (tensions and restrictions in the joints and muscles of the thorax) that are hypothesized by osteopathic literature to result from asthmatic exacerbations are not addressed by this form of therapy. The musculoskeletal system is an integral component to respiratory function. Dysfunction of respiratory muscles may require increased respiratory effort. If treatment can address dysfunction, improved respiratory function may result.

Asthma is a chronic inflammatory disease of the airways involving recurrent exacerbations of airway obstruction (NAC, 2006). It is associated with symptoms of chest tightness, shortness of breath, wheezing, and coughing (NAC, 2006). During periods of asthmatic exacerbations the work of breathing increases and greater effort is required from the respiratory muscles (Chaitow et al., 2002). When this effort is large or prolonged, the coordinated activity of neural reflexes, muscular contractions and the joints involved in respiration may be lost, and a change in breathing pattern may occur, with increased and prolonged recruitment of the accessory respiration muscles (upper chest muscles) resulting in a faulty ‘upper chest’ breathing pattern (Courtney, 2009).

This faulty ‘upper chest’ breathing pattern may be present even when the inflammation and bronchial constriction has been controlled by pharmacological therapy and places greater stress on the musculoskeletal elements involved in respiration (Courtney, 2009). When these stresses extend beyond the normal capabilities of these musculoskeletal elements, resultant structural modifications occur to the tissues that produce the mechanical forces of respiration (Chaitow et al., 2002; Denslow, 1964; Glossary of osteopathic terminology, 2006; Sammut & Searle-Barnes, 1998). These structural modifications are what osteopaths term ‘Somatic Dysfunction’ (SD).
The criteria for diagnosing SD include one or more of: alterations in tissue texture, asymmetry of structure, restriction of motion, and/or tenderness on palpation (Ward, 2003). There are many SD hypothesized by osteopathic literature to be related to asthma including palpable hypertonicity of the sternocleidomastoid and scalene muscle (upper chest breathing muscles), thoracic immobility (DiGiovanna & Schiowitz, 1991), thoracic paraspinal muscle rigidity, and tension and tenderness (Ward, 2003). Authors of osteopathic textbooks claim that by treating (eliminating or reducing) these SD, functional improvements in the musculoskeletal components involved in respiration may result, which in turn will optimize pulmonary function (DiGiovanna & Schiowitz, 1991; Kuchera & Kuchera, 1994; Sammut & Searle-Barnes, 1998).

Whilst pharmacotherapy has been shown to be effective in decreasing asthma symptoms by targeting bronchi inflammation and constriction (Bateman, 2008), the musculoskeletal component of respiration is not addressed by pharmacologic therapy. As such, many studies have investigated whether manual treatment of the musculoskeletal components involved in respiration will increase pulmonary function\(^3\) (Balon et al., 1998; Bockenhauer et al., 2002; Field et al., 1998; Guiney et al., 2005; Nielsen et al., 1995; Noll et al., 2008; Wheatley et al., 2000). However, rigorous supporting evidence for the use of manual therapy remains scarce. There appears to be a lack of randomized controlled trials (RTC) with adequate methodological quality. Hondras et al (2005) undertook a systematic review of 52 studies and found that only three of the 52 studies met the inclusion criteria; they consequently concluded that there is insufficient evidence to support or refute the use of manual therapy for the treatment of asthma (Hondras et al, 2005).

The prevalence of SD in the asthmatic population is unknown. Both the number and severity of SD may be influenced by the chronicity of asthma, the severity of asthma, how well a patient’s asthma is managed by pharmacotherapy, and whether or not the person with asthma has objective findings of lung obstruction (FEV\(_1\)<80% predicted).

An essential factor in asthmatic manual intervention studies that is currently missing is an evaluation procedure to assess for all musculoskeletal dysfunction hypothesized to occur in people with asthma.

---

\(^3\) Pulmonary function is commonly assessed by measuring Forced Expiratory Volume in one second, and/or Peak Expiratory Flow and/or thoracic excursion pre- and post-treatment.
To date only one study, (Russell 2006), investigated whether SD was greater in asthmatics with objective findings of lung obstruction. Russell (2006) compared the combined validity and reliability of nine physical examination procedures performed by two examiners on 39 subjects (17 people with asthma and 22 people without asthma) to a subgroup consisting of 22 subjects (4 asthmatics with objective findings of lung obstruction and 22 non-asthmatics). Minimal changes were observed between the two groups, with no change in reliability, a minimal changes in accuracy (0.55 increased to 0.58) and a slightly stronger increase in NPV (0.58 increased to 0.87). This study however was underpowered to draw any definitive conclusions.

In many asthmatic manual treatment studies there appears to be an absence of an extensive assessment procedure to diagnose SD related to asthma prior to the application of a treatment intervention. It is possible that not all asthmatics exhibit SD, therefore applying treatment interventions in a non-specific way may explain why no positive results have developed. Despite its increasing importance, evidence-based diagnosis is missing in asthmatic manual intervention studies.

4.2. PREVIOUS STUDIES THAT HAVE INVESTIGATED PHYSICAL EXAMINATION PROCEDURES TO DETECT SOMATIC DYSFUNCTION IN PEOPLE WITH ASTHMA

Although the types of SD one would expect to find upon examination on asthmatics have been hypothesized in osteopathic literature, there is a clear lack of research-based evidence to back this up, with only one unpublished thesis by Russell (2006) demonstrating an association between SD and asthma. Twenty examination procedures were utilized in Russell’s study, nine exhibited moderate accuracy (>0.5) and acceptable levels of inter-examiner reliability ($K=0.4-0.6$) (Russell, 2006). Six of the nine procedures assessed the presence of discomfort on palpation (named assessment for discomfort), whilst the other three assessed for the presence of dysfunctional findings$^4$ (named assessment for dysfunction). The examination procedures used in Russell’s (2006) study included:

1. Active side bending of the truck (assessment for dysfunction)
2. Chin tuck (assessment for discomfort)
3. Thoracic spine rotation (assessment for discomfort)
4. Thoracic spine extension (assessment for discomfort)

$^4$ Dysfunctional findings were classified as per the instructions of the procedure. They include one or more of: 1) palpatory findings of tissue texture changes; 2) static palpation/observation for asymmetry of structure/location of landmarks; 3) motion palpation for restriction in motion, hypermobility, hypomobility or for changes in end feel.
5. Combined lateral rotation and flexion (assessment for discomfort)
6. Chest expansion (assessment for dysfunction)
7. Trapezius tension (assessment for dysfunction)
8. Suboccipital tissue texture (assessment for discomfort)
9. Sternocleidomastoid tissue texture (assessment for discomfort)

In Russell’s (2006) thesis a greater number of procedures (six of the nine procedures) that assessed for the presence of discomfort achieved acceptable levels of reliability, compared to procedures that assessed for the presence of dysfunctional findings⁴ (three of the nine procedures). Numerous studies and reviews have been published over the last decade on the level of reliability of manual palpation procedures. Assessment for discomfort, pain provocation, or tenderness appears to generate greater levels of reliability when compared to other procedures. These other procedures include: palpation for tissue texture changes, static palpation/observation for asymmetry of structure/location of landmarks, and motion palpation for restriction in motion, hypermobility, hypomobility or for changes in end feel (Stochkendahl et al., 2006; Seffinger et al., 2004).

In view of Russell’s (2006) results, the validity and inter-examiner reliability of the examination procedures (especially procedures that assess for both dysfunction and discomfort) need to be verified by additional research. If so, manual treatment interventions for SD associated to asthma may form part of a long-term management plan.

4.3. AIMS AND OBJECTIVES

The aim of this study was to evaluate the validity, reliability and potential utility of selected physical examination procedures for the detection of somatic dysfunction in bronchial asthmatics using student and experienced osteopaths.

The objectives of this study were to:
1) assess the reliability and validity of ten physical examination procedures, comparing whether the assessment for dysfunction generated greater reliability and accuracy than the assessment for discomfort
2) examine the relationship between examiner experience and the ability to perform these examination procedures by assessing the validity and reliability of procedures performed by the student osteopaths against experienced osteopaths
3) determine the degree of association between somatic dysfunction and 1) participants with asthma and objective findings of lung obstruction and 2) participants with no asthma and no objective findings of lung obstruction

The comparison of student and experienced osteopaths is important as there is a general presumption that experienced osteopaths have more advanced palpatory skills, and thus greater reliability and accuracy, but there is insufficient research to verify such an assumption.
CHAPTER 5: METHODS

5.1. STUDY DESIGN

A prospective blind comparison between the examination procedure and a reference standard was utilized. The reference standard should be the accepted ‘best’ method for measurement of the variable of interest (Fritz & Wainner, 2001). For all ten examination procedures utilized in this study no reference standard exists to compare each variable (dysfunction or discomfort) against. Therefore, in this study physical examination variables (the presence or absence of dysfunction and/or discomfort) were compared against the presence or absence of diagnosed asthma as reported by the participant to determine whether there was an association between the two variables. In no point in time does this study attempt to diagnose asthma.

All participants with asthma who were recruited into this study were diagnosed by a general practitioner and may have been prescribed drugs such as inhaled steroids to control their asthma.

5.2. RECRUITMENT OF PARTICIPANTS

Participants were recruited from the general public in Auckland and the student population at Unitec, Mount Albert, Auckland (New Zealand). Recruiting took place between August 2009 and October 2009 through:

- Posters (pertaining only to asthmatic participants), which were placed around the Unitec campus, in the Unitec health clinic, and in local community centres and shops
- emails sent to friends and osteopathic students
- flyers distributed in the Mount Albert area
- an advertisement in Unitec’s fortnightly student Magazine
- Facebook posts to associates
- word of mouth.
Potential respondents were supplied with an information sheet. The information sheet briefly outlined:

- general information about the study
- participant inclusion and exclusion criteria
- details about the actual consultation (including participant consent, health questionnaire, spirometry and physical examination procedures)
- details relating to participant confidentiality.

Forty-one participants (21 people with asthma and 20 people without asthma) volunteered to be participants in the study.

5.3. PARTICIPANT INCLUSION / EXCLUSION CRITERIA

Eligibility criteria required that participants were aged between 18 and 50 years and were either: 1) symptomatic, diagnosed with asthma by a general practitioner as reported by the participant; or 2) asymptomatic, with no history of asthma. ‘Diagnosed with asthma’ was clarified to all participants as asthma that was diagnosed by a medical practitioner.

Exclusion criteria were constructed to minimize factors that could alter lung function. Volunteers were ineligible to participate if they currently smoked, were pregnant, or had a history of serious respiratory diseases including:

- chronic bronchitis
- emphysema
- tuberculosis
- pneumonia that required hospitalization within the last two years
- cancer of the lung
- pulmonary hypertension
- pulmonary emboli
- collapsed lung
- thoracic surgery
- a history of serious injury to the spine or thorax (chest) which could alter thorax/thoracic function including rib or spinal fractures (within the last two years).
5.4. RECRUITMENT OF EXAMINERS

Two experienced osteopaths and two students were recruited from the clinical osteopathic teaching faculty and postgraduate osteopathic programme at Unitec, New Zealand.

Recruiting took place between July 2009 and August 2009. Recruiting was commenced with an email sent to both tutors at the clinical osteopathic teaching faculty and students at the postgraduate osteopathic programme at Unitec, New Zealand. The email contained an information sheet that briefly outlined what the study was about and what examiners would be required to do. This included:

- examiner consent
- assessment of participants’ thorax and neck using ten physical examination procedures
- consensus training
- time required
- examiner confidentiality.

Four examiners, two students and two experienced osteopathic clinic tutors, volunteered to be examiners in the study. The experienced osteopaths (examiner A and C) completed their training at Unitec, Auckland, New Zealand, and the British School of Osteopathy (BSO), London, United Kingdom, respectively. Both were currently practicing and had five years and 20-years of clinical experience respectively. Both students (examiner B and D) were in their first year of the Master of Osteopathy programme at Unitec, Auckland, New Zealand.

5.5. RESEARCH ETHICS APPROVAL / PARTICIPANT AND EXAMINER CONSENT

The project was approved by the Unitec Research Ethics Committee (UREC) (see appendix A). Oral and written consent was given by all participants and examiners.
5.6. EXPERIMENTAL DESIGN

5.6.1. CONSENSUS TRAINING
All examiners participated in one 90-minute consensus training session, which was conducted three weeks prior to the first data collection day. The training session consisted of:

- practical revision of examination procedures
- a small amount of feedback about the application of procedures and palpation force (to enhance consistency)
- discussion of the criteria for classifying dysfunctional or normal findings and the classification of the presence or absence of discomfort

Examiners also received verbal instructions on how to record the outcome of examination procedures onto the data collection form. Procedures and palpation strength were performed on the researcher and supervisor (an experienced osteopath), who gave verbal feedback to each examiner about the amount of manual pressure the examiners were applying, with the intended purpose of standardizing the procedure and palpation force.

5.6.2. DATA COLLECTION
Four data collection sessions (one 2-hour session and three 3-hour sessions) took place at the Unitec osteopathic clinic over a three week period. Between 3 and 5 participants were examined in every 60-minute time slot, therefore between 6 and 12 participants were examined in each data collection session. The proportion of asthmatics varied between 25 and 75% in each data collection session.

The data collection sessions were arranged in three stages:

Stage 1
On arrival participants completed an information sheet, consent form and health questionnaire. The questionnaire provided the researcher with information on participants’ respiratory health, current medication history and eligibility, which was included in statistical analysis. Some participants brought their medication with them to assist in completing the medication section of the questionnaire. Those that did so were asked to keep all medication/inhalers concealed when going into the examination room in order to maintain blinding of their group allocation (i.e., asthmatic or non-asthmatic group).
Stage 2
After gaining written consent, each participant was individually taken to a separate room for measurement of weight, height and lung function (spirometry). Spirometry procedures were performed by the primary researcher using a spirometer (Model Spirolab II, Medical International Research (MIR), Via del Maggiolino, Roma, Italy). The protocol for spirometry measurements were those recommended by Miller et al (2005). Lung function parameters Forced vital capacity (FVC), Forced Expiratory Volume in one second (FEV₁), Peak Expiratory Flow Rate (PEFR) and Vital Capacity (VC) were measured using the European Respiratory Standards (ERS) (March 1993) predicted lung values. The results from the spirometry measures and from the respiratory health questionnaire enabled asthmatics to be grouped into intermittent, mild persistent, moderate persistent and severe persistent asthma based on criteria defined by the National Asthma Council Australia (NAC, 2006).

To standardize anatomical reference points used during examination procedures, the research assistant marked (with an ink marker) a cross at spinal levels T4, T6 and L2. Ink marks were also made overlying both sternocleidomastoid muscles at the widest aspect of the belly, superior to the division of the two insertions. These markings were used to ensure consistency of reference points between examiners. Upon completion of all participant markings, the primary researcher coached a five minute pre-warm up exercise class aimed at stretching muscular areas that would be evaluated during the examination procedure. The following exercises were performed:

- seated chin tuck (repeated six times)
- rotating the head to the left (repeated six times), and then the right (repeated six times)
- twisting the body to the left whilst seated (thoracic rotation) (repeated six times), and then to the right (repeated six times)
- standing side bending of the trunk to the left (repeated six times), and to the right (repeated six times)
- standing left pectoralis stretch against a wall (holding for 30 seconds), then repeated on the right (holding for 30 seconds).

Participants were given a short briefing about the examination stage (stage three). Participants were told they would be taken to an assigned room and asked to remove their upper body outer layer/s of clothing (leaving only underwear). Drapes were used to ensure the modesty of participants when necessary. Participants were told they would be examined by four different examiners, who would knock on the door, introduce themselves, conduct the examination
procedures, and leave. It was explained that there may be an interval of a few minutes between examiners; participants were briefed to remain in their rooms during this time. Participants were told that during each examination procedure the examiner would ask them to report back on the presence or absence of discomfort. Discomfort was defined to participants as any form of pain such as achy pain, tenderness or excessive stretch; however a ‘normal’ stretching feeling (“like bending your finger backwards”) was not included as discomfort. The ‘feeling’ of discomfort was demonstrated to each participant by the primary researcher or research assistant by moderately squeezing into the belly of the participants wrist extensor muscles. Each participant was randomly allocated a room number.

Stage 3
Examiners were allocated a holding room until all participants were positioned in their rooms. Once positioned, examiners (A,C,B and D) were collected and positioned at their starting room (1,2,3 and 4 respectively). The order in which the examiners examined each set of four subjects was kept the same to minimize examiner confusion. Examiner A always followed C; C followed B; B followed D and D followed A in a clockwise direction.

Once participants were positioned, examiners entered the room and proceeded to perform the examination procedures. Examiners were allowed a maximum of 15-minutes per participant to perform the 10 procedures and complete the data collection form (see Appendix G). The procedures were performed in consecutive order from procedure 1 to 10. Examiners were instructed to refrain from talking to participants during the examination procedure and on completing the examination to leave the room and wait outside the room. Once all four examiners had completed their first examination they rotated in a clockwise direction to the next participant. This meant the four participants were evaluated in the following sequence: participant one was evaluated by examiner A,C,B then D. Participant two was evaluated by examiner C,B,D,A; participant three was evaluated by B,D,A,C; and participant four was evaluated by D,A,C,B. Upon completion of the last (fourth) participants’ examination, examiners made their way back to the seminar room where they had a 20-30 minute break. This 20-30 minute break was included to minimize examiner fatigue. Following the break, examiners were collected and positioned at their starting rooms for the next group of four participants. Examiners were blinded to participants’ medical history as well as to each others results. Examiners were not permitted to discuss participant findings during data collection session or when in the seminar room. General conversation not related to the project was however permitted.
5.6.3. PHYSICAL EXAMINATION PROCEDURES

Examination procedures that were followed are described in Figures 1 to 10.

Figure 3: Active side bending of the trunk.

In order to assess the participants lateral curvature of their spine between T1 and L2, the participant was instructed to stand with their legs hip width apart and run their left hand down their left leg – bending their body over sideways without twisting. A normal curve was classified as a smooth curve formed by the spinal processes between T1-L2. A dysfunctional curve was classified as having hinging movement between T1-L2 at one or more levels. The participant was also asked to report back on any discomfort isolated to the mid-line of T1-L2 upon bending over. This procedure was repeated on the right side.

Figure 4: Chin tuck

In order to assess the participants cervical lordosis, the participant was instructed to stand with their back against the wall, their legs hip width apart and their heels approximately one foot away from the wall. The participant was then instructed to maintain contact with the back of their head on the wall, to look firstly up at the ceiling, then to their chest, then with their head in a neutral position, retract their chin toward their chest and hold that position for 10 seconds. The examiner placed a flexible ruler (approximately 3.5cm width) between the apex of the participant’s cervical spine lordosis and the wall to determine whether the distance of the participant’s lordosis was under or over the width of the ruler. A normal distance was arbitrarily defined as 3.5cm or less. A dysfunctional curve was classified as having a distance greater than 3.5cm. The participant was also asked to report on the presence or absence of discomfort in their neck.
Examination of the participant’s thoracic spine rotation was conducted with the participant seated on a plinth with his/her legs over the edge. The participant was asked to place each hand on their opposite shoulder. The examiner placed one hand on the participant’s elbows and with the other hand made a broad contact, centering the heel of their hand over T6. The examiner then induced gross left thoracic rotation. A normal result was classified as having no identifiable movement restrictions palpated under the hand, or decreased gross range of motion. Dysfunction was classified as having any of the following: movement restriction/s under the hand identified; decreased gross range of motion compared to the other side; and/or decreased range of motion compared to examiners perception of ‘reasonable’. ‘Reasonable’ was defined to examiners in the consensus training as a judgment based on their clinical knowledge of what they felt was a normal range of motion. The participant was asked to report back on the presence or absence of discomfort in their mid-back. This procedure was repeated on the right side.

In order to assess the participants thoracic spine extension, the participant was positioned in a seated position on a plinth with his/her legs over the edge. The participant was asked to place each hand on their opposite shoulder. The examiner placed one hand on the participant’s elbows, and used their other hand to make broad contact by centering the heel of their hand over T6. The examiner then induced a gross extension movement through the participant’s thoracic spine. Normal movement was classified as no movement restrictions palpated under the hand, or decreased gross range of motion identified. Dysfunction was classified as restriction under the hand palpated and/or a generalized decrease in the gross range of motion of the upper thoracic spine. The participant was asked to report back on the presence or absence of discomfort in their mid-back.
Examination of the first rib for elevation firstly involved positioning the participant in a seated position. The examiner stood behind the participant and with their left hand they passively rotated the cervical spine 45 degrees to the left. Contralateral side bending was added by moving the participants head towards their chest. Normal was classified as reasonable range of free movement in side bending and an elastic end feel. Dysfunction was classified as: limited range of motion in side bending (what the examiner perceives as ‘unreasonable’ or different to the right side); and/or a bony end feel; and/or a block to movement. The participant was asked to report back on the presence or absence of discomfort in their neck or upper ribs. This procedure was repeated on the right side.

Examination of pectoralis length was conducted with the participant positioned supine. The examiner removed the pillow under the participant’s head. The participant was instructed to interface their hands, place them behind their head and let their arms drop down to the side of the plinth. The left arm was assessed first. Normal was classified if the participants arm (whole arm and elbow) touched the table, which indicated good functional muscle length. Dysfunction was classified when the participants arm did not touch the table (indicating shortening of the pectoralis muscle). The examiner then assessed the right arm. The examiner then asked the participant to report back on the presence or absence of discomfort in their left and/or right side of their chest.
Assessment of the length of the upper fibers of the trapezius muscle was conducted with the participant positioned supine with a thin pillow under their head. The examiner stood to the right side of the participant. The examiner fixed the participant’s left shoulder with their caudal hand and then guided the participant’s neck to towards them (side bending the participants neck) using their free hand until they felt the fixed shoulder coming up into their fixed hand. Normal was classified as: reasonable range of side bending and symmetrical range of side bending (left to right); and normal end feel of movement. Dysfunction was classified as decreased side bending compared to the contralateral side, different end feel relative to the other side, and/or decreased side bending compared to examiners perception of normal. The participant was asked to report back on the presence or absence of discomfort in their neck during the procedure. This procedure was repeated on the other side (right).

Examination of the suboccipital muscles was conducted whilst the participant was placed in a supine position. The examiner palpated the left suboccipital muscle by pressing their finger pads up into the muscle (being careful not to extend the neck). Normal was classified as tissue texture state that examiners considered ideal or insignificant in a clinical setting. Dysfunction was classified as tissue texture that is described as tight, hypertonic, ropy, flaccid, loose, or boggy. The participant was asked to report back on the presence or absence of discomfort at the base of their neck on the side being palpated. This procedure was repeated on the right.
Assessment of the sternocleidomastoid muscle was conducted with the participant in a supine position. The examiner palpated the left sternocleidomastoid muscle at the point marked prior to examination. The examiner gently squeezed the muscle between the thumb and index finger. Normal was classified as tissue texture state that may be considered ideal or insignificant in a clinical setting. Dysfunction was classified as tissue texture that is described as tight, hypertonic, ropy, flaccid, loose, or boggy. The participant was asked to report back on the presence or absence of discomfort upon palpation. This procedure was repeated to the right.

The examiner was asked to gauge and record their general impressions of the participant’s breathing pattern and accessory muscle activation. Normal breathing pattern was classified as abdominal breathing. Dysfunction was classified as upper chest breathing, mouth breathing, sighing or yawning frequently. Normal accessory muscle activation was classified as observation of relaxed accessory muscles. Dysfunction was classified as observation of increased use of accessory muscles including sternocleidomastoid, scalene group, and pectoralis (major and minor).
5.6.4. RECORDING FINDINGS

After each examination procedure the examiner would record their observational and/or palpatory findings as “normal” or “dysfunctional” by marking the appropriate box on the record sheet. During each examination procedure (except examination procedure 10 which was an observational procedure) the examiner would also ask the participant to notify them of any ‘discomfort’ that was provoked during the examination. Each participant’s response was recorded as either “absent” or “present” on the record sheet. For examination procedures that assessed bilateral findings, the record sheet was broken into left and right for both “normal” or “dysfunctional” findings and the “presence” or “absence” of discomfort (See Appendix G for the physical examination record form).

5.7. DATA ANALYSIS

5.7.1. DATA REDUCTION

Seven of the 10 examination procedures assessed for bilateral (left and right) findings of dysfunction and discomfort.

In order to reduce the data, examiner responses for examination procedures that assessed for both right and left dysfunctional findings were dichotomized:
1. normal, if normal findings were found bilaterally, or
2. dysfunctional, where there were findings of either bilateral or unilateral dysfunction.

Examiner responses for examination procedures that assessed for both left and right discomfort findings were dichotomized:
1. absent, if no discomfort was elicited bilaterally or
2. present, where discomfort was elicited on one or both sides.

Examiner responses collected on observations of breathing pattern and accessory muscle activation (procedure 10) were dichotomized:
1. normal, where normal findings were recorded in all parameters and
2. dysfunctional, when dysfunction was recorded in one or both parameters.
5.7.2. ASSESSMENT OF THE PERFORMANCE OF A DIAGNOSTIC PROCEDURES

All assessment findings from each component (dysfunction or discomfort) of each examination procedure for each group of examiners (student osteopaths and experienced osteopaths) was tabulated in 2 x 2 contingency tables. These tables were used to compare the component of interest (e.g. the presence/or absence of dysfunction against the presence or absence of asthma respectively). Nineteen tables were generated for each group of examiners: 10 tables contained results relating to the presence or absence of dysfunction (dysfunction component of examination procedures 1-10); 9 tables contained results relating to the presence or absence of discomfort findings (discomfort component of examination procedures 1-9). A total of 38 contingency tables were generated in total. Examples of the contingency tables for both dysfunction findings and discomfort findings are presented in Table 5 and Table 6 respectively.

Table 5:
Example of contingency table summarizing dysfunction findings

<table>
<thead>
<tr>
<th>Asthma</th>
<th>Present</th>
<th>Absent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Present</td>
<td>A</td>
<td>B (\rightarrow G_1 = A + B)</td>
</tr>
<tr>
<td>Dysfunc</td>
<td>No. of dysfunctional findings in asthmatic participants (True positive)</td>
<td>No. of dysfunctional findings in asymptomatic participants (False positive)</td>
</tr>
<tr>
<td>Absent</td>
<td>C (\downarrow F_1 = A + C)</td>
<td>D (\downarrow F_2 = B + D)</td>
</tr>
<tr>
<td>No. of normal findings in asthmatic participants (False negative)</td>
<td>No. of normal findings in asymptomatic participants (True negative)</td>
<td></td>
</tr>
</tbody>
</table>

Table 6:
Example of contingency table summarizing discomfort findings

<table>
<thead>
<tr>
<th>Asthma</th>
<th>Present</th>
<th>Absent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Present</td>
<td>A (\rightarrow G_1 = A + B)</td>
<td>B</td>
</tr>
<tr>
<td>Discomf</td>
<td>No. of discomfort findings in asthmatic participants (True positive)</td>
<td>No. of discomfort findings in asymptomatic participants (False positive)</td>
</tr>
<tr>
<td>Absent</td>
<td>C (\downarrow F_1 = A + C)</td>
<td>D (\downarrow F_2 = B + D)</td>
</tr>
<tr>
<td>No. of normal findings in asthmatic participants (False negative)</td>
<td>No. of normal findings in asymptomatic participants (True negative)</td>
<td></td>
</tr>
</tbody>
</table>
From each contingency table the following statistics were calculated:

- sensitivity and specificity
- positive and negative predictive values
- positive and negative likelihood ratios
- accuracy

Collectively, these statistics provide information on the usefulness of the physical examination procedures (Fritz & Wainer, 2001). Confidence intervals (95% CI) for sensitivity, specificity, and positive and negative likelihood ratios were also calculated using the method described by Simel, Samsa & Matchar (1991). The formula used in this study and a brief description of the statistics generated are presented in Table 7.

The cut-off for levels for ‘adequate’ accuracy were arbitrarily set to >0.5. ‘Adequate’ levels of sensitivity and specificity were also arbitrarily set to >0.8. These cut-off levels were also used by Russell (2006).
Table 7: Formulae and descriptions of the statistics used to assess validity of physical examination procedures

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Formula</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensitivity (SN)</td>
<td>$S_n = \frac{A}{A + C}$</td>
<td>Sensitivity reflects the proportion of participants with asthma who have a dysfunctional finding</td>
</tr>
<tr>
<td>Specificity (SP)</td>
<td>$S_p = \frac{D}{B + D}$</td>
<td>Specificity reflects the proportion of participants without asthma who do not have a dysfunctional finding</td>
</tr>
<tr>
<td>Positive Predictive Value (PPV)</td>
<td>$PPV = \frac{A}{A + B}$</td>
<td>Positive predictive value indicates the probability that an individual has asthma given the observation of a dysfunctional finding</td>
</tr>
<tr>
<td>Negative Predictive Value (NPV)</td>
<td>$NPV = \frac{D}{C + D}$</td>
<td>Negative predictive value indicates the probability that an individual does not have asthma given the observation of the absence of a dysfunctional finding</td>
</tr>
<tr>
<td>Positive Likelihood Ratio (+LR)</td>
<td>$PLR = \frac{S_n}{1 - S_n}$</td>
<td>Positive Likelihood Ratio describes the change in odds favoring the presence of asthma given the observation of the presence of dysfunction</td>
</tr>
<tr>
<td>Negative Likelihood Ratio (-LR)</td>
<td>$NLR = \frac{1 - S_n}{S_n}$</td>
<td>Negative Likelihood Ratio describes the change in odds favoring the presence of asthma given the observation of the absence of dysfunction</td>
</tr>
<tr>
<td>Accuracy</td>
<td>$\frac{A + D}{A - B + C + D}$</td>
<td>Accuracy reflects the ratio of correct procedure results to the total number of procedures conducted</td>
</tr>
</tbody>
</table>

Confidence Intervals for SN

$95\% CI(S_n) = S_n \pm 1.96 \sqrt{\frac{S_n(1 - S_n)}{N_n}}$

$N_n =$ number of participants in sample with asthma

Confidence Intervals for SP

$95\% CI(S_p) = S_p \pm 1.96 \sqrt{\frac{S_p(1 - S_p)}{N_p}}$

$N_p =$ number of participants in sample without asthma

Confidence Intervals for +LR

$95\% CI(PLR) = PLR \pm 1.96 \exp \left[ \frac{1 - S_n}{A} \frac{S_n}{B} \right]$

Confidence Intervals for -LR

$95\% CI(NLR) = NLR \pm 1.96 \exp \left[ \frac{1 - S_n}{C} \frac{S_n}{D} \right]$

5.7.3. DATA ANALYSIS OF INTER-EXAMINER RELIABILITY

Findings from all procedures for the two groups of examiners (student osteopaths (examiner A and C) and experienced osteopaths (examiner B and D)) were tabulated in 2 x 2 contingency tables. The contingency tables were used to compare the agreement between the two examiners in each group. Nineteen tables were generated for each group of examiners: 10 tables contained findings relating to the presence or absence of dysfunction (dysfunction component of examination procedures 1-10); 9 tables contained findings relating to the presence or absence of discomfort (discomfort component of examination procedures 1-9). A total of 38 contingency tables were generated in total. Examples of the contingency table for dysfunction component of examination procedure 1 for the student osteopaths and experienced osteopaths are presented in Table 8 and Table 9 respectively.

Table 8: Example of contingency table summarizing Examiner A and C’s (student osteopaths) findings for dysfunction component of examination procedure 1

<table>
<thead>
<tr>
<th></th>
<th>Examiner C (student 1)</th>
<th>Normal findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dysfunctional findings</td>
<td></td>
<td>Normal findings</td>
</tr>
<tr>
<td>Examiner A (student 2)</td>
<td></td>
<td>Normal findings</td>
</tr>
<tr>
<td>Dysfunctional findings</td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>No. of dysfunctional findings in which both examiners were in agreement (Agreement)</td>
<td></td>
<td>No. of times where Examiner D observed normal findings and Examiner B observed dysfunctional findings (Disagreement)</td>
</tr>
<tr>
<td>C</td>
<td></td>
<td>D</td>
</tr>
<tr>
<td>No. of times where Examiner D observed dysfunctional findings and Examiner B observed normal findings (Disagreement)</td>
<td></td>
<td>No. of normal findings in which both examiners were in agreement (Agreement)</td>
</tr>
</tbody>
</table>

\[ G_1 = A + B \]
\[ G_2 = C + D \]
\[ F_1 = A + C \]
\[ F_2 = B + D \]
\[ N = F_1 + F_2 \]
Table 9: Example of contingency table summarizing examiner B and D’s (experienced osteopaths) findings for dysfunction component of examination procedure 1

<table>
<thead>
<tr>
<th>Examiner B (experienced osteopath 2)</th>
<th>Examiner D (experienced osteopath 1)</th>
<th>( G_1 ) = A + B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dysfunctional findings</td>
<td>A No. of dysfunctional findings in which both examiners were in agreement (Agreement)</td>
<td>( G_1 ) = A + B</td>
</tr>
<tr>
<td></td>
<td>B No. of times where Examiner C observed normal findings and Examiner A observed dysfunctional findings (Disagreement)</td>
<td>( G_1 ) = A + B</td>
</tr>
<tr>
<td>Normal findings</td>
<td>C No. of times where Examiner C observed dysfunctional findings and Examiner A observed normal findings (Disagreement)</td>
<td>( G_2 ) = C + D</td>
</tr>
<tr>
<td></td>
<td>D No. of normal findings in which both examiners were in agreement (Agreement)</td>
<td>( G_2 ) = C + D</td>
</tr>
<tr>
<td>( \downarrow ) ( F_1 ) \ = A + C</td>
<td>( \downarrow ) ( F_2 ) \ = B + D</td>
<td>( \Rightarrow N ) = ( F_1 ) + ( F_2 )</td>
</tr>
</tbody>
</table>

From each contingency table the following statistics were calculated:

- proportion of observed agreement (\( P_o \))
- proportion of expected agreement (\( P_e \))
- proportion of positive agreement (\( P_{pos} \))
- proportion of negative agreement (\( P_{neg} \)).

Analysis of reliability was determined using Cohen’s kappa (\( K \)) as recommended by Sim & Wright (2005). The kappa coefficient (\( K \)) is a statistic widely used to measure the degree of reliability between two examiners in manual medicine (Sim & Wright, 2005).

Landis and Koch (1977) proposed a categorization for the strength of agreement. It is widely referenced and is presented below in Table 10. Whilst there is no standardized level of acceptable reliability for manual therapy studies, many have arbitrarily set the level of acceptable/adequate agreement to \( K > 0.4 \) (Dixon et al., 1998; Seffinger et al., 2004; Stochkendahl et al., 2006). The kappa cut-off limit of \( K = 0.4 \) was used in this study to select examination procedures of acceptable agreement. The confidence intervals (CI) for the kappa statistic were also calculated using the method described by Sheskin (2004).
Table 10: Kappa Categorization

<table>
<thead>
<tr>
<th>Kappa statistic</th>
<th>Strength of agreement</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; 0</td>
<td>None</td>
</tr>
<tr>
<td>0-0.2</td>
<td>Slight</td>
</tr>
<tr>
<td>0.2-0.4</td>
<td>Fair</td>
</tr>
<tr>
<td>0.4-0.6</td>
<td>Moderate</td>
</tr>
<tr>
<td>0.6-0.8</td>
<td>Substantial</td>
</tr>
<tr>
<td>0.8-1.0</td>
<td>Almost perfect</td>
</tr>
</tbody>
</table>

Source: Landis and Koch (1977)

The kappa statistic alone is appropriate if the marginal totals in the 2 x 2 contingency table are relatively balanced, but if the prevalence of a response in a population is either extremely high or extremely low, the kappa may indicate low reliability, even when the observed proportion of agreement is high. Because of this kappa statistic paradox, researchers recommend reporting other values in addition to the kappa statistic (Cunningham, 2009; Sim & Wright, 2005). These include:

- the bias index (BI)
- prevalence index (PI)
- prevalence-adjusted bias-adjusted kappa (PABAK)
- maximal kappa ($K_{\text{max}}$).

These statistics were also calculated. Formulas for all statistics are presented in Table 11.
<table>
<thead>
<tr>
<th>Statistic</th>
<th>Formula and Description</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proportion of observed agreement (Po)</td>
<td>( P_o = \frac{A + D}{N} )</td>
<td>Reflects the proportion of observed agreement</td>
</tr>
<tr>
<td>Proportion of expected agreement (Pe)</td>
<td>( P_e = \frac{F_i \times G_i}{N} )</td>
<td>Reflects the proportion of agreement that would be expected by chance</td>
</tr>
<tr>
<td>Proportion of positive agreement (Ppos)</td>
<td>( P_{pos} = \frac{2 \times A}{N + A - D} )</td>
<td>Reflects the proportion of positive agreement</td>
</tr>
<tr>
<td>Proportion of negative agreement (Pneg)</td>
<td>( P_{neg} = \frac{2 \times D}{N - A + D} )</td>
<td>Reflects the proportion of negative agreement</td>
</tr>
<tr>
<td>Kappa coefficient (k)</td>
<td>( k = \frac{P_o - P_e}{1 - P_e} )</td>
<td>Kappa is a coefficient that measures the proportion of agreement beyond which is expected by chance. It is the ratio of observed agreement minus chance divided by the maximum possible agreement minus chance.</td>
</tr>
<tr>
<td>Observed agreement (Oij)</td>
<td>( O_{ij} = A + D )</td>
<td>Indicates the observed agreement</td>
</tr>
<tr>
<td>Expected agreement (Eij)</td>
<td>( E_{ij} = \frac{F_i \times G_i}{N} )</td>
<td>Indicates the expected agreement</td>
</tr>
<tr>
<td>Standard Error (SEk)</td>
<td>( SE_k = \frac{O_{ij}(N - O_{ij})}{N(N - N E_i)} )</td>
<td>Standard error reflects the standard deviation of the sampling distribution associated with the estimation method</td>
</tr>
<tr>
<td>Confidence Intervals for kappa (95%)</td>
<td>( 95% CI(k) = \pm 1.96 \times SE_k )</td>
<td>Confidence Intervals indicate the precision of the estimate. They indicate a range of values in the population that would lie within 95% certainty</td>
</tr>
<tr>
<td>Byrt's prevalence index (Pl)</td>
<td>( P_l = \frac{A \cdot D}{N} )</td>
<td>Prevalence reflects the proportion of examiner agreement on positive or negative findings. It can influence the magnitude of the kappa</td>
</tr>
<tr>
<td>Byrt's bias index (Bi)</td>
<td>( B_l = \frac{B \cdot C}{N} )</td>
<td>Bias reflects the proportion of examiner disagreement on positive or negative findings. It can influence the magnitude of the kappa</td>
</tr>
<tr>
<td>Prevalence-adjusted bias-adjusted kappa (PABAK)</td>
<td>( PABAK = 2 \times P_o - 1 )</td>
<td>Prevalence-adjusted bias-adjusted kappa takes into account bias and prevalence influences</td>
</tr>
<tr>
<td>Maximal kappa (Kmax)</td>
<td>Cells A, and C, of the contingency table are replaced with maximal marginal totals using the equations below: A = F1 if F1&gt;G1, else = G1 B = 0 D = F2 if F2&gt;G2, else = G2 Using A and D’s new cell totals C is calculated using the equation below: C = N(A + D) Po and Pe are calculated using the new cell values. They are then substituted into the equation below to give Kmax: ( k_{max} = \frac{P_o - P_e}{1 - P_e} )</td>
<td>Maximal kappa represents the maximal kappa value that could have been generated by the two examiners using the dataset</td>
</tr>
</tbody>
</table>

Source: Sim & Wright, 2005; Sheskin, 2004
5.7.4. NORMALITY OF THE ASTHMATIC AND ASYMPTOMATIC GROUP

Normality tests were used to explore the distribution of variables in the dataset (age, weight, height, and minimum and maximal hours/week of exercise between the asthmatic group and asymptomatic group). To investigate differences between groups, for non-normal variables Mann-Whitney test was used and for normally distributed variables independent samples t-tests was used.

5.8. SUBGROUP ANALYSIS

5.8.1. SUBGROUP 1 AND 2

Subgroup analysis was used to assess the combined validity and reliability of selected combinations of components (dysfunction and/or discomfort) of examination procedures 1-10 performed by the student osteopaths and experienced osteopaths and compared to the combined validity of all components of examination procedures 1-10 for both the student osteopaths and experienced osteopaths respectively. Criteria for each subgroup were as follows.

Subgroup 1

Components of examination procedures 1-10 that demonstrated the highest validity in terms of Accuracy (Accuracy >0.5) and reliability (Cohen’s unadjusted kappa \( \geq 0.41 \)) were selected for Subgroup 1 analysis.

Subgroup 2

Components of examination procedures 1-10 that demonstrated the highest validity in terms of Accuracy (Accuracy >0.5) and inter-examiner reliability (Adjusted PABAK kappa \( \geq 0.41 \)) were selected for Subgroup 2 analysis.

5.8.2. SUBGROUP 3

Components of examination procedures 1-10 that demonstrated the highest validity in terms of accuracy (accuracy >0.5) and inter-examiner reliability (Adjusted PABAK kappa \( \geq 0.41 \)) were selected for Subgroup 3 analysis. This was conducted to determine if the use of a cluster of components from examination procedures 1-10 resulted in greater utility. The validity of ‘one’, ‘two’ and ‘three’ positive findings out of each subgroup for both the student osteopaths and experienced osteopaths was assessed.
5.8.3. SUBGROUP 4

Participants were included in this subgroup analysis from the original sample if they either:

1) had asthma and had objective lung findings of less than 80% of their predicted Forced Expiratory Volume in one second (FEV$_1$ <80%). This level was selected from The National Asthma Council of Australia (2003) which considers FEV$_1$ <80% predicted to be moderate lung obstruction and FEV$_1$ <60% severe lung obstruction (NAC, 2006);

2) or, were in the non-asthmatic group and did not have less than 80% of their predicted Forced Expiratory Volume in one second (FEV$_1$ <80%).

The data from this subgroup was combined and presented as the validity of all physical examination procedures and compared to the original dataset.
CHAPTER 6: RESULTS

6.1. PARTICIPANT CHARACTERISTICS

Forty-one subjects (19 males and 22 females) between the ages of 18 to 50 with a mean age of 29 ± 7.4 participated in the study. Twenty-one subjects were allocated to the asthmatic group, and 20 to the asymptomatic (non-asthmatic) group. Table 12 presents the characteristics of the participants in the study.

Table 12: Characteristics of participants in the asthmatic and asymptomatic group

<table>
<thead>
<tr>
<th>Sample (SD)</th>
<th>Asthmatic Group (SD)</th>
<th>Asymptomatic Group (SD)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of participants</td>
<td>41</td>
<td>21</td>
<td>20</td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range (years)</td>
<td>18 - 50</td>
<td>21 - 50</td>
<td>18 - 47</td>
</tr>
<tr>
<td>Mean (years)</td>
<td>29.0 (7.4)</td>
<td>29.9 (8.1)</td>
<td>28.2 (6.7)</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>22.0</td>
<td>12.0</td>
<td>10.0</td>
</tr>
<tr>
<td>Male</td>
<td>19.0</td>
<td>9.0</td>
<td>10.0</td>
</tr>
<tr>
<td>Weight</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range (kg)</td>
<td>47 - 123</td>
<td>55 - 123</td>
<td>47 - 110</td>
</tr>
<tr>
<td>Mean (kg)</td>
<td>75.0 (18.7)</td>
<td>77.5 (19.6)</td>
<td>72.4 (17.9)</td>
</tr>
<tr>
<td>Height</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range (cm)</td>
<td>148 - 192</td>
<td>157 - 192</td>
<td>148 - 187</td>
</tr>
<tr>
<td>Mean (cm)</td>
<td>173.0 (9.3)</td>
<td>173.2 (9.9)</td>
<td>172.7 (8.9)</td>
</tr>
<tr>
<td>Body Mass Index (BMI)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>15.5 – 43.1</td>
<td>18.2 – 43.1</td>
<td>15.5 – 38.5</td>
</tr>
<tr>
<td>Mean</td>
<td>25.0 (5.7)</td>
<td>25.7 (6.0)</td>
<td>24.2 (5.4)</td>
</tr>
<tr>
<td>No. of participants classified as ‘overweight’ (≥25–25.99 BMI) in sample</td>
<td>13</td>
<td>7</td>
<td>6</td>
</tr>
<tr>
<td>No. of participants classified as ‘obese’ (≥30 BMI) in sample</td>
<td>5</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Weekly exercise</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range (hours/week) min</td>
<td>0 - 10</td>
<td>0 - 7</td>
<td>0 - 10</td>
</tr>
<tr>
<td>Range (hours/week) max</td>
<td>1 - 40</td>
<td>1 - 40</td>
<td>6 - 9</td>
</tr>
<tr>
<td>Mean (hours/week) min</td>
<td>2.6 (2.7)</td>
<td>2.6 (2.2)</td>
<td>2.6 (3.2)</td>
</tr>
<tr>
<td>Mean (hours/week) max</td>
<td>3.9 (6.5)</td>
<td>4.8 (8.4)</td>
<td>2.9 (3.6)</td>
</tr>
<tr>
<td>FEV1 % predicted</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range (%)</td>
<td>67 - 134</td>
<td>67 - 127</td>
<td>83 - 134</td>
</tr>
<tr>
<td>Mean (%)</td>
<td>98.5 (14.7)</td>
<td>91.4 (13.8)</td>
<td>106 (11.9)</td>
</tr>
</tbody>
</table>
Table 12: (continued)
Characteristics of participants in the asthmatic and asymptomatic group

<table>
<thead>
<tr>
<th></th>
<th>Sample (SD)</th>
<th>Asthmatic Group (SD)</th>
<th>Asymptomatic Group (SD)</th>
<th>P-value¹</th>
</tr>
</thead>
<tbody>
<tr>
<td>Handedness</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Left</td>
<td>9.0</td>
<td>8.0</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>Right</td>
<td>32.0</td>
<td>12.0</td>
<td>19.0</td>
<td></td>
</tr>
</tbody>
</table>

Note: SD = Standard Deviation; FEV1 = Forced Expiratory Volume in one second
¹ P-value for paired sample t-test
² Obesity and overweight classification are based on the International Classification for overweight and obesity according to BMI by the World Health Organization (WHO) 2006

6.2. COMPARISON AT BASELINE OF THE ASTHMATIC AND ASYMPTOMATIC GROUPS

There were no substantial differences between the groups at baseline for the variables age, weight, height, BMI, and minimum and maximal hours/week of exercise between the asthmatic group and asymptomatic group (see Table 12).

6.3. CHARACTERISTICS OF ASTHMATIC GROUP

Classification of asthma severity of the participants are summarized in Table 13. Seven of the asthmatics were classified as intermittent, 2 as mild-persistent, 6 as moderate-persistent, and 6 as severe-persistent. The duration of asthma, medication use, asthma triggers and frequency of asthma symptoms of the participants diagnosed as having asthma by a medical practitioner are summarized in Table 14. The mean duration of asthma reported was 22.4 years (± 10.9). The most common medications used to control asthma symptoms (in order of frequency) was Albuterol, followed by Budensonide, Seretide, Symbicort, Oxis, and Terbutaline. Four asthmatics reported no medication use. The three most frequently reported triggers (in order of frequency) were exercise, colds and influenza, and dust mites.

Table 13: Asthma severity

<table>
<thead>
<tr>
<th>Asthma severity</th>
<th>Number of asthmatic participants</th>
<th>Percentage in asthmatic sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intermittent</td>
<td>7</td>
<td>33%</td>
</tr>
<tr>
<td>Mild-persistent</td>
<td>2</td>
<td>10%</td>
</tr>
<tr>
<td>Moderate-persistent</td>
<td>6</td>
<td>29%</td>
</tr>
<tr>
<td>Severe-persistent</td>
<td>6</td>
<td>29%</td>
</tr>
</tbody>
</table>
Table 14:
Asthma duration, medication use, asthma triggers, and frequency of asthma symptoms (cough, wheezing, shortness of breath and/or chest tightness) as reported by participants

<table>
<thead>
<tr>
<th></th>
<th>Number of asthmatic participants</th>
<th>Percentage of asthmatic sample</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Duration of asthma</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range (years)</td>
<td>1 - 43</td>
<td></td>
</tr>
<tr>
<td>Mean (years)</td>
<td>22.4 (10.9)</td>
<td></td>
</tr>
<tr>
<td><strong>Medication use for asthma symptom control</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No asthma medication</td>
<td>4</td>
<td>19%</td>
</tr>
<tr>
<td>Inhaled corticosteroids (ICS)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Budensonide (Pulmicort)</td>
<td>3</td>
<td>14%</td>
</tr>
<tr>
<td>Fluticasone (Flixtotide)</td>
<td>2</td>
<td>10%</td>
</tr>
<tr>
<td>Long Acting Beta Agonists (LABA's)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oxis</td>
<td>1</td>
<td>5%</td>
</tr>
<tr>
<td>Combination ICS &amp; LABA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Seretide</td>
<td>3</td>
<td>14%</td>
</tr>
<tr>
<td>Symbicort</td>
<td>3</td>
<td>14%</td>
</tr>
<tr>
<td>Relievers: Short Acting Beta Agonists (SABA)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Albuterol (Salbutomol, Airomir, Asmol, Epaq, Ventolin)</td>
<td>13</td>
<td>62%</td>
</tr>
<tr>
<td>Terbutaline (Bricanyl)</td>
<td>1</td>
<td>5%</td>
</tr>
<tr>
<td><strong>Asthma triggers</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exercise</td>
<td>17</td>
<td>81%</td>
</tr>
<tr>
<td>Colds and influenza</td>
<td>16</td>
<td>76%</td>
</tr>
<tr>
<td>Dust Mites</td>
<td>14</td>
<td>67%</td>
</tr>
<tr>
<td>Pollens</td>
<td>10</td>
<td>48%</td>
</tr>
<tr>
<td>Second hand cigarette smoke</td>
<td>9</td>
<td>43%</td>
</tr>
<tr>
<td>Cold temperature</td>
<td>7</td>
<td>33%</td>
</tr>
<tr>
<td>Animals</td>
<td>5</td>
<td>24%</td>
</tr>
<tr>
<td>Stress</td>
<td>5</td>
<td>24%</td>
</tr>
<tr>
<td>Spontaneous</td>
<td>2</td>
<td>10%</td>
</tr>
<tr>
<td>Dampness/mould</td>
<td>2</td>
<td>10%</td>
</tr>
<tr>
<td>Smoke (from outdoor or indoor fires)</td>
<td>2</td>
<td>10%</td>
</tr>
<tr>
<td>Spontaneous (no specific triggers)</td>
<td>2</td>
<td>10%</td>
</tr>
<tr>
<td>Humidity</td>
<td>1</td>
<td>5%</td>
</tr>
<tr>
<td>Change in season</td>
<td>1</td>
<td>5%</td>
</tr>
<tr>
<td>Alcohol</td>
<td>1</td>
<td>5%</td>
</tr>
</tbody>
</table>
Table 14 (continued)
Asthma duration, medication use, asthma triggers, and frequency of asthma symptoms (cough, wheezing, shortness of breath and/or chest tightness) as reported by participants

<table>
<thead>
<tr>
<th>Frequency of asthma symptoms during the day</th>
<th>Number of asthmatic participants</th>
<th>Percentage of asthmatic sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>Symptoms are present less than once per week</td>
<td>13</td>
<td>62%</td>
</tr>
<tr>
<td>Symptoms are present more than once per week (but not daily)</td>
<td>5</td>
<td>24%</td>
</tr>
<tr>
<td>Symptoms are present daily, but do not generally restrict physical activity</td>
<td>3</td>
<td>14%</td>
</tr>
<tr>
<td>Symptoms are present daily and restrict physical activity</td>
<td>0</td>
<td>0%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Frequency of night-time symptoms</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Symptoms are present less than twice per month</td>
<td>7</td>
<td>33%</td>
</tr>
<tr>
<td>Symptoms are present more than twice per month, but daily</td>
<td>7</td>
<td>33%</td>
</tr>
<tr>
<td>Symptoms are present at least once per week</td>
<td>5</td>
<td>24%</td>
</tr>
<tr>
<td>Symptoms are present daily</td>
<td>2</td>
<td>10%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Frequency of asthma exacerbations</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Exacerbations are infrequent and brief</td>
<td>9</td>
<td>43%</td>
</tr>
<tr>
<td>Exacerbations occur occasionally and may affect activity or sleep</td>
<td>10</td>
<td>48%</td>
</tr>
<tr>
<td>Exacerbations are frequent and may affect activity or sleep</td>
<td>2</td>
<td>10%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Effect of asthma symptoms on physical activity</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Symptoms do not generally restrict physical activity</td>
<td>18</td>
<td>86%</td>
</tr>
<tr>
<td>Symptoms restrict physical activity</td>
<td>3</td>
<td>14%</td>
</tr>
</tbody>
</table>

6.4. VALIDITY OF PHYSICAL EXAMINATION PROCEDURES

The validity of each component (dysfunction or discomfort) of each examination procedure performed by each of the two groups (student osteopaths and experienced osteopaths) was assessed using calculations of accuracy, sensitivity, specificity, positive likelihood ratios, negative likelihood ratios, positive predictive value, and negative predictive values.

The validity of each dysfunction component of examination procedures 1-10 for both student osteopaths and experienced osteopaths are displayed in Table 15 and Table 16 respectively. The validity of each discomfort component of examination procedures 1-9 for both student osteopaths and experienced osteopaths are displayed in Table 17 and Table 18 respectively.
<table>
<thead>
<tr>
<th>Examination procedure</th>
<th>Accuracy</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Sensitivity 95% CI (SN)</th>
<th>Specificity 95% CI (SP)</th>
<th>Positive Likelihood Ratio 95% CI (+LR)</th>
<th>Negative Likelihood Ratio 95% CI (-LR)</th>
<th>PPV</th>
<th>NPV</th>
</tr>
</thead>
<tbody>
<tr>
<td>1) Active side bending of trunk</td>
<td>0.44</td>
<td>0.50</td>
<td>0.35 0.65</td>
<td>0.38</td>
<td>0.22 0.53 0.80 1.18</td>
<td>0.81 2.20 0.46 0.42</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2) Chin tuck</td>
<td>0.57</td>
<td>0.50</td>
<td>0.35 0.65</td>
<td>0.65</td>
<td>0.50 0.80 1.43 0.85</td>
<td>0.77 1.12 0.60 0.55</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3) Thoracic spine rotation</td>
<td>0.56</td>
<td>0.64</td>
<td>0.50 0.79</td>
<td>0.48</td>
<td>0.32 0.63 1.22 0.84</td>
<td>0.75 1.27 0.56 0.56</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4) Thoracic spine extension</td>
<td>0.46</td>
<td>0.31</td>
<td>0.17 0.45</td>
<td>0.63</td>
<td>0.47 0.78 0.83 0.45</td>
<td>1.10 1.51 0.46 0.46</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5) CRLF</td>
<td>0.63</td>
<td>0.81</td>
<td>0.69 0.93</td>
<td>0.45</td>
<td>0.30 0.60 1.47 1.07</td>
<td>2.02 0.42 0.21 0.86</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6) Chest expansion</td>
<td>0.50</td>
<td>0.40</td>
<td>0.26 0.55</td>
<td>0.60</td>
<td>0.45 0.75 1.01 0.60</td>
<td>1.72 0.77 0.50 0.49</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7) Trapezius stretch</td>
<td>0.66</td>
<td>0.55</td>
<td>0.40 0.70</td>
<td>0.78</td>
<td>0.65 0.90 2.43 1.29</td>
<td>4.60 0.58 0.40 0.85</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8) Suboccipital tissue texture</td>
<td>0.46</td>
<td>0.52</td>
<td>0.37 0.67</td>
<td>0.40</td>
<td>0.25 0.55 0.87 0.59</td>
<td>1.28 1.95 0.73 0.48</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9) SCM tissue texture</td>
<td>0.34</td>
<td>0.33</td>
<td>0.19 0.48</td>
<td>0.35</td>
<td>0.20 0.50 0.51 0.32</td>
<td>0.83 1.90 1.19 0.36</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10) Observation of quiet breathing</td>
<td>0.39</td>
<td>0.40</td>
<td>0.26 0.55</td>
<td>0.38</td>
<td>0.22 0.53 0.65 0.42</td>
<td>1.00 1.59 0.49 0.38</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

SN = Sensitivity; SP = Specificity; LL = Lower Limit; UL = Upper Limit; CI = Confidence Interval; +LR = Positive Likelihood Ratio; -LR = Negative Likelihood Ratio; PPV = Positive Predictive Value; NPV = Negative Predictive Value; CRLF = Combined Rotation and Lateral Flexion; SCM = Sternocleidomastoid muscle; Figures shaded grey = those that meet cut-off limits.

Table 16:
Experienced osteopaths – validity of each dysfunction component of procedures 1-10

<table>
<thead>
<tr>
<th>Examination procedure</th>
<th>Accuracy</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Sensitivity 95% CI (SN)</th>
<th>Specificity 95% CI (SP)</th>
<th>Positive Likelihood Ratio 95% CI (+LR)</th>
<th>Negative Likelihood Ratio 95% CI (-LR)</th>
<th>PPV</th>
<th>NPV</th>
</tr>
</thead>
<tbody>
<tr>
<td>1) Active side bending of trunk</td>
<td>0.50</td>
<td>0.76</td>
<td>0.63 0.89</td>
<td>0.23</td>
<td>0.10 0.35 0.98 1.25</td>
<td>1.06 0.83 0.77 0.48</td>
<td></td>
<td>0.48 0.47</td>
<td></td>
</tr>
<tr>
<td>2) Chin tuck</td>
<td>0.50</td>
<td>0.43</td>
<td>0.28 0.58</td>
<td>0.58</td>
<td>0.42 0.73 1.01 0.99</td>
<td>0.68 1.44 0.51 0.49</td>
<td></td>
<td>0.51 0.49</td>
<td></td>
</tr>
<tr>
<td>3) Thoracic spine rotation</td>
<td>0.56</td>
<td>0.60</td>
<td>0.45 0.74</td>
<td>0.50</td>
<td>0.35 0.65 1.19 0.81</td>
<td>0.5 1.31 0.56 0.54</td>
<td></td>
<td>0.51 0.49</td>
<td></td>
</tr>
<tr>
<td>4) Thoracic spine extension</td>
<td>0.55</td>
<td>0.69</td>
<td>0.55 0.83</td>
<td>0.40</td>
<td>0.25 0.55 1.15 0.83</td>
<td>1.59 0.77 0.43 1.4</td>
<td></td>
<td>0.55 0.55</td>
<td></td>
</tr>
<tr>
<td>5) CRLF</td>
<td>0.50</td>
<td>0.57</td>
<td>0.42 0.72</td>
<td>0.43</td>
<td>0.27 0.58 0.99 0.68</td>
<td>1.44 1.01 0.61 1.67</td>
<td></td>
<td>0.51 0.49</td>
<td></td>
</tr>
<tr>
<td>6) Chest expansion</td>
<td>0.62</td>
<td>0.40</td>
<td>0.26 0.55</td>
<td>0.85</td>
<td>0.74 0.96 2.70 1.18</td>
<td>6.15 0.70 0.53 0.93</td>
<td></td>
<td>0.74 0.58</td>
<td></td>
</tr>
<tr>
<td>7) Trapezius stretch</td>
<td>0.49</td>
<td>0.33</td>
<td>0.19 0.48</td>
<td>0.65</td>
<td>0.50 0.80 0.95 0.52</td>
<td>1.74 1.03 0.75 1.4</td>
<td></td>
<td>0.50 0.48</td>
<td></td>
</tr>
<tr>
<td>8) Suboccipital tissue texture</td>
<td>0.52</td>
<td>0.57</td>
<td>0.42 0.72</td>
<td>0.48</td>
<td>0.32 0.63 1.09 0.73</td>
<td>1.61 0.90 0.56 1.45</td>
<td></td>
<td>0.53 0.51</td>
<td></td>
</tr>
<tr>
<td>9) SCM tissue texture</td>
<td>0.49</td>
<td>0.26</td>
<td>0.13 0.39</td>
<td>0.73</td>
<td>0.59 0.86 0.95 0.47</td>
<td>1.95 1.02 0.78 1.32</td>
<td></td>
<td>0.50 0.48</td>
<td></td>
</tr>
<tr>
<td>10) Observation of quiet breathing</td>
<td>0.56</td>
<td>0.60</td>
<td>0.45 0.74</td>
<td>0.53</td>
<td>0.37 0.68 1.25 0.83</td>
<td>1.89 0.77 0.48 1.23</td>
<td></td>
<td>0.57 0.55</td>
<td></td>
</tr>
</tbody>
</table>

SN = Sensitivity; SP = Specificity; LL = Lower Limit; UL = Upper Limit; CI = Confidence Interval; +LR = Positive Likelihood Ratio; -LR = Negative Likelihood Ratio; PPV = Positive Predictive Value; NPV = Negative Predictive Value; CRLF = Combined Rotation and Lateral Flexion; SCM = Sternocleidomastoid muscle; Figures shaded grey = those that meet cut-off limits.
Table 17:
Student osteopaths – validity of each discomfort component of examination procedure 1-9

<table>
<thead>
<tr>
<th>Examination procedure</th>
<th>Accuracy</th>
<th>Sensitivity</th>
<th>95% CI (SN)</th>
<th>Specificity</th>
<th>95% CI (SP)</th>
<th>95% CI (+LR)</th>
<th>95% CI (-LR)</th>
<th>PPV</th>
<th>NPV</th>
</tr>
</thead>
<tbody>
<tr>
<td>1) Active side bending of trunk (discomfort)</td>
<td>0.45</td>
<td>0.07</td>
<td>-0.01 0.15</td>
<td>0.85</td>
<td>0.74 0.96</td>
<td>0.48 1.78</td>
<td>0.94 1.28</td>
<td>0.33</td>
<td>0.47</td>
</tr>
<tr>
<td>2) Chin tuck (discomfort)</td>
<td>0.46</td>
<td>0.02</td>
<td>-0.02 0.07</td>
<td>0.93</td>
<td>0.84 1.01</td>
<td>0.32 2.93</td>
<td>0.95 1.17</td>
<td>0.25</td>
<td>0.47</td>
</tr>
<tr>
<td>3) Thoracic spine rotation (discomfort)</td>
<td>0.46</td>
<td>0.14</td>
<td>0.04 0.25</td>
<td>0.80</td>
<td>0.68 0.92</td>
<td>0.71 1.88</td>
<td>0.88 1.31</td>
<td>0.43</td>
<td>0.47</td>
</tr>
<tr>
<td>4) Thoracic spine extension (discomfort)</td>
<td>0.49</td>
<td>0.14</td>
<td>0.04 0.25</td>
<td>0.85</td>
<td>0.74 0.96</td>
<td>0.95 2.71</td>
<td>0.84 1.21</td>
<td>0.50</td>
<td>0.49</td>
</tr>
<tr>
<td>5) CRLF (discomfort)</td>
<td>0.48</td>
<td>0.07</td>
<td>-0.01 0.15</td>
<td>0.90</td>
<td>0.81 0.99</td>
<td>0.71 2.99</td>
<td>0.90 1.18</td>
<td>0.43</td>
<td>0.48</td>
</tr>
<tr>
<td>6) Chest expansion (discomfort)</td>
<td>0.49</td>
<td>0.05</td>
<td>-0.02 0.11</td>
<td>0.95</td>
<td>0.86 1.02</td>
<td>0.95 6.44</td>
<td>0.91 1.11</td>
<td>0.50</td>
<td>0.49</td>
</tr>
<tr>
<td>7) Trapezius stretch (discomfort)</td>
<td>0.48</td>
<td>0.05</td>
<td>-0.02 0.11</td>
<td>0.93</td>
<td>0.84 1.01</td>
<td>0.63 3.60</td>
<td>0.92 1.15</td>
<td>0.40</td>
<td>0.48</td>
</tr>
<tr>
<td>8) Suboccipital tissue texture (discomfort)</td>
<td>0.51</td>
<td>0.17</td>
<td>0.05 0.28</td>
<td>0.88</td>
<td>0.77 0.98</td>
<td>1.33 4.64</td>
<td>0.80 1.14</td>
<td>0.58</td>
<td>0.50</td>
</tr>
<tr>
<td>9) SCM tissue texture (discomfort)</td>
<td>0.46</td>
<td>0.00</td>
<td>0.00 0.00</td>
<td>0.95</td>
<td>0.88 1.02</td>
<td>0.00 1.05</td>
<td>0.98 1.13</td>
<td>0.00</td>
<td>0.48</td>
</tr>
</tbody>
</table>

SN = Sensitivity; SP = Specificity; LL = Lower Limit; UL = Upper Limit; CI = Confidence Interval; +LR = Positive Likelihood Ratio; -LR = Negative Likelihood Ratio; PPV = Positive Predictive Value; NPV = Negative Predictive Value; CRLF = Combined Rotation and Lateral Flexion; SCM = Sternocleidomastoid muscle; Figures shaded grey = those that meet cut-off limits.

Table 18:
Experienced osteopaths – validity of each discomfort component of examination procedure 1-9

<table>
<thead>
<tr>
<th>Examination procedure</th>
<th>Accuracy</th>
<th>Sensitivity</th>
<th>95% CI (SN)</th>
<th>Specificity</th>
<th>95% CI (SP)</th>
<th>95% CI (+LR)</th>
<th>95% CI (-LR)</th>
<th>PPV</th>
<th>NPV</th>
</tr>
</thead>
<tbody>
<tr>
<td>1) Active side bending of trunk (discomfort)</td>
<td>0.50</td>
<td>0.07</td>
<td>-0.01 0.15</td>
<td>0.85</td>
<td>0.88 1.02</td>
<td>1.25 8.11</td>
<td>0.88 1.09</td>
<td>0.60</td>
<td>0.49</td>
</tr>
<tr>
<td>2) Chin tuck (discomfort)</td>
<td>0.50</td>
<td>0.07</td>
<td>-0.01 0.15</td>
<td>0.95</td>
<td>0.88 1.02</td>
<td>1.43 8.11</td>
<td>0.98 1.09</td>
<td>0.60</td>
<td>0.49</td>
</tr>
<tr>
<td>3) Thoracic spine rotation (discomfort)</td>
<td>0.54</td>
<td>0.26</td>
<td>0.13 0.39</td>
<td>0.83</td>
<td>0.71 0.94</td>
<td>1.50 3.48</td>
<td>0.71 1.13</td>
<td>0.61</td>
<td>0.52</td>
</tr>
<tr>
<td>4) Thoracic spine extension (discomfort)</td>
<td>0.50</td>
<td>0.12</td>
<td>0.02 0.22</td>
<td>0.90</td>
<td>0.81 0.99</td>
<td>1.19 4.12</td>
<td>0.84 1.14</td>
<td>0.56</td>
<td>0.49</td>
</tr>
<tr>
<td>5) CRLF (discomfort)</td>
<td>0.48</td>
<td>0.14</td>
<td>0.04 0.25</td>
<td>0.83</td>
<td>0.71 0.94</td>
<td>0.82 2.22</td>
<td>0.86 1.25</td>
<td>0.46</td>
<td>0.48</td>
</tr>
<tr>
<td>6) Chest expansion (discomfort)</td>
<td>0.52</td>
<td>0.10</td>
<td>0.01 0.18</td>
<td>0.98</td>
<td>0.93 1.02</td>
<td>3.81 32.64</td>
<td>0.83 1.04</td>
<td>0.80</td>
<td>0.51</td>
</tr>
<tr>
<td>7) Trapezius stretch (discomfort)</td>
<td>0.54</td>
<td>0.17</td>
<td>0.05 0.28</td>
<td>0.93</td>
<td>0.84 1.01</td>
<td>2.22 8.003</td>
<td>0.77 1.06</td>
<td>0.70</td>
<td>0.51</td>
</tr>
<tr>
<td>8) Suboccipital tissue texture (discomfort)</td>
<td>0.50</td>
<td>0.17</td>
<td>0.05 0.28</td>
<td>0.85</td>
<td>0.74 0.96</td>
<td>1.11 3.02</td>
<td>0.81 1.18</td>
<td>0.54</td>
<td>0.49</td>
</tr>
<tr>
<td>9) SCM tissue texture (discomfort)</td>
<td>0.52</td>
<td>0.12</td>
<td>0.02 0.22</td>
<td>0.95</td>
<td>0.88 1.02</td>
<td>2.38 11.58</td>
<td>0.81 1.06</td>
<td>0.71</td>
<td>0.51</td>
</tr>
</tbody>
</table>

SN = Sensitivity; SP = Specificity; LL = Lower Limit; UL = Upper Limit; CI = Confidence Interval; +LR = Positive Likelihood Ratio; -LR = Negative Likelihood Ratio; PPV = Positive Predictive Value; NPV = Negative Predictive Value; CRLF = Combined Rotation and Lateral Flexion; SCM = Sternocleidomastoid muscle; Figures shaded grey = those that meet cut-off limits.
Validity of the *dysfunction* component of examination procedures 1-10

The accuracy of the dysfunction component of examination procedures 1-10 was similar for student osteopaths and experienced osteopaths. Accuracy (A) ranged from 0.34 to 0.66 for the student osteopaths and 0.49 to 0.62 for the experienced osteopaths. Four dysfunction components of four examination procedures performed by the student osteopaths achieved “adequate”\(^5\) accuracy. These examination procedures included Chin tuck (A=0.57), Thoracic spine rotation (A=0.56), CRLF (A=0.63), and Trapezius stretch (A=0.66). Five dysfunction components of 5 examination procedures performed by the experienced osteopaths achieved adequate\(^5\) accuracy. These examination procedures included: Thoracic spine rotation (A=0.55), Thoracic spine extension (A=0.55), Chest expansion (A=0.62), Suboccipital tissue texture (A=0.52), and Observation of quiet breathing (A=0.56).

Only one dysfunction component of one examination procedure performed by the student osteopaths displayed adequate\(^5\) sensitivity (Sn): CRLF (Sn=0.81, CI 0.69 to 0.93). Only one dysfunction component of one examination procedure performed by the experienced osteopaths exhibited adequate\(^5\) specificity (Sp): Chest expansion (Sp=0.85, CI 0.74 to 0.96). All other examination components performed by student and experience osteopaths and for both groups displayed inadequate sensitivity.

One dysfunction component of one examination procedure performed by the student osteopaths generated moderate changes in the Positive Likelihood Ratio (+LR): Trapezius stretch (+LR=2.43, 95% CI 1.29 to 4.60) as well as one examination component performed by the experienced osteopaths: Chest expansion (+LR=2.70, 95% CI 1.18 to 6.15). All other examination components performed by both groups generated inadequate levels of +LR’s.

One dysfunction component of one examination procedure performed by the student osteopaths generated a small but sometimes important shift in Negative Likelihood Ratio (-LR)\(^6\): CRLF (-LR=0.42, 95% CI 0.21 to 0.86). All other dysfunction components of the examination procedures performed by both groups generated inadequate levels of -LR’s.

\(^5\) Adequate accuracy cut-off was arbitrarily set to >0.5. Adequate Sensitivity and Specificity cut-off was arbitrarily set to >0.8
Validity of the discomfort component of examination procedures 1-9

All discomfort components of examination procedures 1-9 performed by both groups displayed similar levels of accuracy. Accuracy ranged from 0.45 to 0.51 for the students and 0.48 to 0.54 for the experienced osteopaths. One examination procedure performed by the student osteopaths achieved adequate\(^5\) accuracy: Suboccipital tissue texture (discomfort) (A=0.51). Four discomfort components of four examination procedures performed by the experienced osteopaths achieved adequate\(^5\) accuracy: Thoracic spine rotation (discomfort) (A=0.54), Chest expansion (discomfort) (A=0.52), Trapezius stretch (discomfort) (A=0.54), and SCM tissue texture (discomfort) (A=0.52).

All discomfort components of examination procedures 1-9 performed by both groups displayed inadequate sensitivity. Sensitivity ranged from 0.02 to 0.17 for student osteopaths and 0.07 to 0.26 for experienced osteopaths. In contrast, all discomfort components of examination procedures 1-9 performed by both groups displayed adequate\(^5\) levels of specificity. Specificity ranged from 0.80 to 0.95 for the student osteopaths and 0.81 to 0.96 for the experienced osteopaths.

Three discomfort components of three examination procedures performed by the experienced osteopaths generated small but sometimes important shifts in Positive Likelihood Ratios\(^6\). These were: Chest expansion (discomfort) (+LR=3.81, 95% CI 0.44 to 32.64), Trapezius stretch (discomfort) (+LR=2.22, 95% CI 0.62 to 8.00), SCM tissue texture (discomfort) (+LR=2.38, 95% CI 0.49 to 11.58). All other discomfort components performed by both groups generated unimportant shifts of +LR’s. All discomfort components of all examination procedures performed by either group generated unimportant shifts of Negative Likelihood Ratios.

\(^6\) +LR cut-off was set to >2.0; -LR cut-off was set to <0.5. A +LR between 2.0-5.0 or -LR between 0.2-0.5 generates small, but sometimes important shifts in probability (Fritz & Wainner, 2001).
Comparison of accuracy between the student osteopaths and experienced osteopaths for all 10 dysfunction components of examination procedures 1-10 is depicted in Figure 13 below.

Comparison of accuracy between the student osteopaths and experienced osteopaths for all 9 discomfort components of examination procedures 1-9 is depicted in Figure 14 below.

**Figure 13:** Accuracy of dysfunction components of examination procedures 1-10 (osteopathic students vs. experienced osteopaths). Note: The dotted line represents the cut-off for level for ‘acceptable’ accuracy

**Figure 14:** Accuracy of discomfort components of examination procedures 1-9 (osteopathic students vs. experienced osteopaths. Note: The dotted line represents the cut-off for level for ‘acceptable’ accuracy
The overall accuracy of each group (student osteopaths and experienced osteopaths) is similar; with no clinically important differences between the two groups.

6.5. INTER-EXAMINER RELIABILITY

The inter-examiner reliability of each dysfunction component of examination procedures 1-10 for both student osteopaths and experienced osteopaths are displayed in Table 19 and Table 20 respectively. The inter-examiner reliability of each discomfort component of examination procedures 1-9 for both student osteopaths and experienced osteopaths are displayed in Table 21 and Table 22 respectively.
Table 19: Student osteopaths – inter-examiner reliability of each dysfunction component of examination procedures 1-10

<table>
<thead>
<tr>
<th>Examination procedure</th>
<th>K</th>
<th>95% CI (K)</th>
<th>95% CI (K&lt;sub&gt;max&lt;/sub&gt;)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1) Active side bending of trunk</td>
<td>0.14</td>
<td>-0.16 - 0.44</td>
<td>0.56 - 0.49 0.61 - 0.50</td>
</tr>
<tr>
<td>2) Chin tuck</td>
<td>0.75</td>
<td>0.55 - 0.96</td>
<td>0.88 - 0.51 0.86 - 0.89</td>
</tr>
<tr>
<td>3) Thoracic spine rotation</td>
<td>0.20</td>
<td>-0.10 - 0.51</td>
<td>0.61 - 0.51 0.67 - 0.53</td>
</tr>
<tr>
<td>4) Thoracic spine extension</td>
<td>0.14</td>
<td>-0.16 - 0.43</td>
<td>0.56 - 0.49 0.36 - 0.67</td>
</tr>
<tr>
<td>5) CRLF</td>
<td>0.23</td>
<td>-0.10 - 0.56</td>
<td>0.66 - 0.56 0.75 - 0.46</td>
</tr>
<tr>
<td>6) Chest expansion</td>
<td>0.29</td>
<td>0.01 - 0.58</td>
<td>0.63 - 0.48 0.55 - 0.69</td>
</tr>
<tr>
<td>7) Trapezius stretch</td>
<td>0.39</td>
<td>0.09 - 0.68</td>
<td>0.71 - 0.52 0.63 - 0.76</td>
</tr>
<tr>
<td>8) Suboccipital tissue texture</td>
<td>-0.04</td>
<td>-0.28 - 0.20</td>
<td>0.37 - 0.39 0.43 - 0.28</td>
</tr>
<tr>
<td>9) SCM tissue texture</td>
<td>0.22</td>
<td>-0.08 - 0.52</td>
<td>0.61 - 0.50 0.60 - 0.62</td>
</tr>
<tr>
<td>10) Observation of quiet breathing</td>
<td>-0.03</td>
<td>-0.33 - 0.26</td>
<td>0.46 - 0.48 0.48 - 0.45</td>
</tr>
</tbody>
</table>

K = kappa coefficient; 95% CI = 95% confidence interval; UL = upper limit; LL = lower limit; PO = Observed agreement; PE = Chance (expected) agreement; Ppos = Proportion of positive agreement; Pneg = proportion of negative agreement; PABAK = prevalence-adjusted and bias-adjusted kappa; PI = Byrt’s prevalence index; BI = Bryt’s bias index and K<sub>max</sub> = maximum kappa coefficient; Figures shaded grey = those that met cut-off limits.

Table 20: Experienced osteopaths – inter-examiner reliability of each dysfunction component of procedures 1-10

<table>
<thead>
<tr>
<th>Examination procedure</th>
<th>K</th>
<th>95% CI (K)</th>
<th>95% CI (K&lt;sub&gt;max&lt;/sub&gt;)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1) Active side bending of trunk</td>
<td>0.00</td>
<td>-0.33 - 0.33</td>
<td>0.54 - 0.54 0.70 - 0.00</td>
</tr>
<tr>
<td>2) Chin tuck</td>
<td>0.05</td>
<td>-0.26 - 0.37</td>
<td>0.54 - 0.51 0.46 - 0.60</td>
</tr>
<tr>
<td>3) Thoracic spine rotation</td>
<td>0.18</td>
<td>-0.10 - 0.45</td>
<td>0.54 - 0.44 0.58 - 0.49</td>
</tr>
<tr>
<td>4) Thoracic spine extension</td>
<td>0.16</td>
<td>-0.15 - 0.47</td>
<td>0.59 - 0.51 0.68 - 0.41</td>
</tr>
<tr>
<td>5) CRLF</td>
<td>0.00</td>
<td>-0.13 - 0.13</td>
<td>0.15 - 0.15 0.26 - 0.00</td>
</tr>
<tr>
<td>6) Chest expansion</td>
<td>-0.01</td>
<td>-0.38 - 0.36</td>
<td>0.59 - 0.59 0.26 - 0.71</td>
</tr>
<tr>
<td>7) Trapezius stretch</td>
<td>0.17</td>
<td>-0.12 - 0.45</td>
<td>0.56 - 0.47 0.36 - 0.67</td>
</tr>
<tr>
<td>8) Suboccipital tissue texture</td>
<td>-0.22</td>
<td>-0.46 - 0.02</td>
<td>0.29 - 0.42 0.36 - 0.22</td>
</tr>
<tr>
<td>9) SCM tissue texture</td>
<td>-0.08</td>
<td>-0.42 - 0.26</td>
<td>0.51 - 0.55 0.09 - 0.67</td>
</tr>
<tr>
<td>10) Observation of quiet breathing</td>
<td>0.03</td>
<td>-0.28 - 0.33</td>
<td>0.51 - 0.50 0.55 - 0.47</td>
</tr>
</tbody>
</table>

K = kappa coefficient; 95% CI = 95% confidence interval; UL = upper limit; LL = lower limit; PO = Observed agreement; PE = Chance (expected) agreement; Ppos = Proportion of positive agreement; Pneg = proportion of negative agreement; PABAK = prevalence-adjusted and bias-adjusted kappa; PI = Byrt’s prevalence index; BI = Bryt’s bias index and K<sub>max</sub> = maximum kappa coefficient; Figures shaded grey = those that met cut-off limits.
### Table 21:
Student osteopaths – inter-examiner reliability of each discomfort component of examination procedures 1-9

<table>
<thead>
<tr>
<th>Examination procedure</th>
<th>95% CI (K)</th>
<th>95% CI (K&lt;sub&gt;max&lt;/sub&gt;)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>K</td>
<td>LL</td>
</tr>
<tr>
<td>1) Active side bending of trunk (discomfort)</td>
<td>0.63</td>
<td>0.23</td>
</tr>
<tr>
<td>2) Chin tuck (discomfort)</td>
<td>0.47</td>
<td>-0.24</td>
</tr>
<tr>
<td>3) Thoracic spine rotation (discomfort)</td>
<td>0.32</td>
<td>-0.10</td>
</tr>
<tr>
<td>4) Thoracic spine extension (discomfort)</td>
<td>0.23</td>
<td>-0.24</td>
</tr>
<tr>
<td>5) CRLF (discomfort)</td>
<td>0.22</td>
<td>-0.42</td>
</tr>
<tr>
<td>6) Chest expansion (discomfort)</td>
<td>-0.04</td>
<td>-1.00</td>
</tr>
<tr>
<td>7) Trapezius stretch (discomfort)</td>
<td>0.36</td>
<td>-0.33</td>
</tr>
<tr>
<td>8) Suboccipital tissue texture (discomfort)</td>
<td>0.42</td>
<td>-0.01</td>
</tr>
<tr>
<td>9) SCM tissue texture (discomfort)</td>
<td>1.00</td>
<td>1.00</td>
</tr>
</tbody>
</table>

K = kappa coefficient; 95% CI = 95% confidence interval; UL = upper limit; LL = lower limit; PO = Observed agreement; PE = Chance (expected) agreement; Ppos = Proportion of positive agreement; Pneg = proportion of negative agreement; PABAK = prevalence-adjusted and bias-adjusted kappa; PI = Bryt’s prevalence index; BI = Bryt’s bias index and K<sub>max</sub> = maximum kappa coefficient; Figures shaded grey = those that met cut-off limits.

### Table 22:
Experienced osteopaths – inter-examiner reliability of each discomfort component of examination procedures 1-9

<table>
<thead>
<tr>
<th>Examination procedure</th>
<th>95% CI (K)</th>
<th>95% CI (K&lt;sub&gt;max&lt;/sub&gt;)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>K</td>
<td>LL</td>
</tr>
<tr>
<td>1) Active side bending of trunk (discomfort)</td>
<td>-0.04</td>
<td>-0.90</td>
</tr>
<tr>
<td>2) Chin tuck (discomfort)</td>
<td>-0.06</td>
<td>-0.93</td>
</tr>
<tr>
<td>3) Thoracic spine rotation (discomfort)</td>
<td>0.02</td>
<td>-0.40</td>
</tr>
<tr>
<td>4) Thoracic spine extension (discomfort)</td>
<td>-0.11</td>
<td>-0.75</td>
</tr>
<tr>
<td>5) CRLF (discomfort)</td>
<td>0.00</td>
<td>-0.51</td>
</tr>
<tr>
<td>6) Chest expansion (discomfort)</td>
<td>-0.06</td>
<td>-0.93</td>
</tr>
<tr>
<td>7) Trapezius stretch (discomfort)</td>
<td>0.09</td>
<td>-0.47</td>
</tr>
<tr>
<td>8) Suboccipital tissue texture (discomfort)</td>
<td>0.00</td>
<td>-0.51</td>
</tr>
<tr>
<td>9) SCM tissue texture (discomfort)</td>
<td>0.00</td>
<td>-0.67</td>
</tr>
</tbody>
</table>

K = kappa coefficient; 95% CI = 95% confidence interval; UL = upper limit; LL = lower limit; PO = Observed agreement; PE = Chance (expected) agreement; Ppos = Proportion of positive agreement; Pneg = proportion of negative agreement; PABAK = prevalence-adjusted and bias-adjusted kappa; PI = Bryt’s prevalence index; BI = Bryt’s bias index and K<sub>max</sub> = maximum kappa coefficient; Figures shaded grey = those that met cut-off limits.
Inter-examiner reliability of the *dysfunction* component of examination procedures 1-10

Reliability, represented by kappa ($K$) score, of the dysfunction component of examination procedures 1-10 ranged from $K=-0.03$ to 0.75 for student osteopaths and from $K=0.08$ to 0.16 for experienced osteopaths. The percentage agreement ranged from 37% to 88% for the student osteopaths and 15% to 59% for the experienced osteopaths.

Only one dysfunction component of one examination procedure performed by the student osteopaths achieved an acceptable\(^7\) level of reliability: chin tuck ($K=0.75$, 95% CI = 0.55 to 0.96) and percentage agreement (88%). No other dysfunction components of the remaining examination procedures performed by the student osteopaths or experienced osteopaths achieved acceptable levels of reliability. When taking into consideration prevalence and bias indexes using the Prevalence Adjusted Bias Adjusted Kappa (PABAK) statistic, there were decreases in reliability of both groups, more notably in the experienced osteopath group. The reliability of dysfunction components of examination procedures 1-10 using the adjusted kappa ranged from $K=-0.27$ to 0.76 for the student osteopaths and between -0.71 to 0.17 for the experienced osteopaths.

There was a low bias index (BI) present in one dysfunction component of one examination procedure performed by the student osteopaths: Suboccipital tissue texture (BI=-0.49), and in three dysfunction components of three examination procedures performed by the experienced osteopaths: CRLF (BI=-0.85), Suboccipital tissue texture (BI=-0.41), and Active side bending of trunk (BI=-0.46). When an extremely large or small bias is present the magnitude of the kappa is higher. Once the kappa is adjusted using the PABAK kappa statistic, the kappa is reduced accordingly.

---

\(^7\) Acceptable reliability was arbitrarily set to >0.4 kappa
Inter-examiner reliability of discomfort components of examination procedures 1-9
Reliability of discomfort components of examination procedures 1-9 ranged from $K = -0.04$ to 1.00 for student osteopaths and from $K = -0.11$ to 0.09 for experienced osteopaths. The percentage agreement ranged from 80% to 100% for the student osteopaths and 73% to 88% for the experienced osteopaths.

Four discomfort components performed by the student osteopaths achieved an acceptable level of reliability: Suboccipital tissue texture (discomfort) ($K = 1.00$, 95% CI 1.00 to 1.00), Active side bending of trunk (discomfort) ($K = 0.63$, 95% CI 0.23 to 1.03), Chin tuck (discomfort) ($K = 0.47$, 95% CI -0.24 to 1.18), and Suboccipital tissue texture (discomfort) ($K = 0.42$, 95% CI -0.01 to 0.85). No discomfort component performed by the experienced osteopaths achieved acceptable levels of reliability. When taking into consideration prevalence and bias indexes using the adjusted PABAK kappa statistic, there were moderate changes in the adjusted kappa range of the student osteopaths (0.61 to 1.00) and experienced osteopaths (0.32 to 0.76). All the discomfort components performed by the student osteopaths achieved acceptable reliability, and all except one discomfort component of one examination procedure (thoracic spine rotation) performed by the experienced osteopaths achieved acceptable reliability. The increases in the adjusted kappa values were due to the high prevalence in the dataset which was portrayed in the extremely low Prevalence Index, which was present in all discomfort components of all examination procedures for both groups. The prevalence index ranged from -0.66 to -0.95 for the student osteopaths, and -0.56 to -0.88 for the experienced osteopaths.
Inter-examiner reliability of both the student osteopaths and experienced osteopaths for all dysfunction components of examination procedures 1-10 and all discomfort components of examination procedures 1-9 is presented below in Figure 15 and Figure 16.

**Figure 15:** Inter-examiner reliability (Cohen's kappa) of dysfunction components of examination procedures 1-10 (osteopathic students vs. experienced osteopaths). Note: If the score for an examination procedure is zero then no bar is presented. The dotted line represents the cut-off for level for ‘acceptable’ reliability.

**Figure 16:** Inter-examiner reliability (Cohen's kappa) of discomfort components of examination procedures 1-9 (osteopathic students vs. experienced osteopaths). Note: If the score for an examination procedure is zero then no bar is presented. The dotted line represents the cut-off for level for ‘acceptable’ reliability.
The inter-examiner reliability using the adjusted kappa (PABAK) for both the student osteopaths and experienced osteopaths for all dysfunction components of examination procedures 1-10 and all discomfort components of examination procedures 1-9 is presented below in Figure 17 and Figure 18.

Figure 17: Inter-examiner reliability (adjusted kappa) of dysfunction components of examination procedures 1-10 (osteopathic students vs. experienced osteopaths). Note: The dotted line represents the cut-off for level for ‘acceptable’ reliability.

Figure 18: Inter-examiner reliability (adjusted kappa) of discomfort components of examination procedures 1-9 (osteopathic students vs. experienced osteopaths). Note: The dotted line represents the cut-off for level for ‘acceptable’ reliability.
6.6. **SUBGROUP ANALYSIS FOR EXAMINATION PROCEDURES WITH ACCEPTABLE VALIDITY AND RELIABILITY**

Accuracy versus inter-examiner reliability (Cohen’s unadjusted kappa) for each component (dysfunction or discomfort) for each examination procedure performed by the student osteopaths and experienced osteopaths is depicted in Figure 19 and Figure 20 respectively. Dysfunction or discomfort components of examination procedures that appeared in the right upper quartile, named ‘Subgroup 1’, were selected for subgroup analysis.

**Figure 19:** Student osteopaths inter-examiner reliability against accuracy of dysfunction and discomfort components for all examination procedures. Note: The shaded area contains examination procedures that had both acceptable levels of reliability and accuracy.

**Figure 20:** Experienced osteopaths inter-examiner reliability against accuracy of dysfunction and discomfort components for all examination procedures. Note: The shaded area contains examination procedures that had both acceptable levels of reliability and accuracy.
Accuracy versus inter-examiner reliability (using adjusted kappa (PABAK)) for each component (dysfunction or discomfort) for each examination procedure performed by the student osteopaths and experienced osteopaths are depicted in Figure 21 and Figure 22 respectively. Dysfunction or discomfort components of examination procedures that appeared in the right upper quartile, named ‘Subgroup 2’, were selected for subgroup analysis.

Figure 21: Student osteopath’s inter-examiner reliability (adjusted kappa) against accuracy of dysfunction and discomfort components for all examination procedures. Note: The shaded area contains examination procedures that had both acceptable levels of reliability and accuracy.

Figure 22: Experienced osteopaths inter-examiner reliability (adjusted kappa) against accuracy of dysfunction and discomfort components for all examination procedures. Note: The shaded area contains examination procedures that had both acceptable levels of reliability and accuracy.
6.6.1. SUBGROUP 1 AND 2 ANALYSIS

Subgroup 1
Dysfunction or discomfort components of examination procedures that appeared in the upper right quartile of Figure 19 and Figure 20 demonstrated adequate accuracy (>0.5) and acceptable inter-examiner reliability ($K \geq 0.41$) and were selected for further analysis. Two components of two examination procedures performed by the student osteopaths appeared in the upper right quartile of Figure 19. These included:

- Chin tuck (dysfunction)  
  Accuracy = 0.57; $K = 0.75$
- Suboccipital tissue texture (discomfort)  
  Accuracy = 0.51; $K = 0.42$

No components of examination procedures performed by the experienced osteopaths appeared in the upper right quartile of Figure 20.

Subgroup 2
Dysfunction or discomfort components of examination procedures which appeared in the upper right quartile of Figure 21 and Figure 22 that demonstrated adequate accuracy (>0.5) and acceptable inter-examiner reliability (PABAK $\geq 0.41$) were selected for further analysis. Three components of three examination procedures performed by the student osteopaths appeared in the upper right quartile of Figure 21. These included:

- 2) Chin tuck (dysfunction)  
  Accuracy = 0.57; $K = 0.76$
- 7) Trapezius stretch (dysfunction)  
  Accuracy = 0.66; $K = 0.42$
- 8) Suboccipital tissue texture (discomfort)  
  Accuracy = 0.51; $K = 0.71$

Three components of three examination procedures performed by the experienced osteopaths appeared in the upper right quartile of Figure 22. These components included:

- 6) Chest expansion (discomfort)  
  Accuracy = 0.52; $K = 0.76$
- 7) Trapezius stretch (discomfort)  
  Accuracy = 0.54; $K = 0.46$
- 9) SCM tissue texture (discomfort)  
  Accuracy = 0.52; $K = 0.66$
The data from each of these subgroups were combined and presented in comparison to the combined validity of all components of all examination procedures for both the student osteopaths and experienced osteopaths respectively.

Table 23: Combined clinical validity of all subgroups for both student osteopaths and experienced osteopaths

<table>
<thead>
<tr>
<th></th>
<th>Students osteopaths</th>
<th>Experienced osteopaths</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All 10 Procedures</td>
<td>SD</td>
</tr>
<tr>
<td>Overall accuracy</td>
<td>0.49</td>
<td>0.07</td>
</tr>
<tr>
<td>PPV</td>
<td>0.45</td>
<td>0.16</td>
</tr>
<tr>
<td>NVP</td>
<td>0.49</td>
<td>0.08</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>0.30</td>
<td>0.25</td>
</tr>
<tr>
<td>Specificity</td>
<td>0.69</td>
<td>0.22</td>
</tr>
<tr>
<td>+LR</td>
<td>0.91</td>
<td>0.54</td>
</tr>
<tr>
<td>-LR</td>
<td>1.05</td>
<td>0.31</td>
</tr>
<tr>
<td>Cohen’s Kappa</td>
<td>0.31</td>
<td>0.26</td>
</tr>
<tr>
<td>PABAK</td>
<td>0.48</td>
<td>0.35</td>
</tr>
</tbody>
</table>

SD = Standard Deviation; PPV = Positive Predictive Value; NVP = Negative Predictive Value; +LR = Positive Likelihood Ratio; -LR = Negative Likelihood Ratio; PABAK = prevalence-adjusted and bias-adjusted kappa; Figures shaded grey = those that meet cut-off limits.

Student osteopaths’ subgroup 1 and 2 were compared to results from all components of the original 10 examination procedures. The student osteopaths overall accuracy increased in both subgroups with subgroup 2 showing the largest rise (from 0.49 to 0.58). Sensitivity also increased in both subgroups; subgroup 2 showed a largest increase from 0.30 to 0.41, however, these levels are still below an acceptable level (<0.80).

Specificity increased in all subgroups performed by the student osteopaths, with subgroup 2 showing the largest increase (from 0.69 to 0.77). Subgroup 2 also showed the largest +LR and smallest -LR (1.73 and 0.77 respectively). There were increases in Cohen’s kappa across all subgroups, with subgroup one showing the largest increase from 0.31 to 0.58. Adjusted kappa (PABAK) values also increased, with subgroup 1 also showing the largest increase (from 0.48 to 0.73).

Experienced osteopaths subgroup 2 was compared to results from all components of the original 10 examination procedures. There were marginal differences in accuracy between the two groups. Specificity increased from 0.71 to 0.95, however sensitivity decreased from 0.34 to 0.13. Subgroup two showed the largest +LR and smallest -LR (2.80 and 0.92 respectively). There was no change in Cohen’s kappa, but the adjusted kappa (PABAK) increased (from 0.26 to 0.67).
6.6.2. **SUBGROUP 3 ANALYSIS**

Dysfunction or discomfort components examination procedures 1-10 which appeared in the upper right quartile of Figure 21 and Figure 22 demonstrated adequate accuracy (>0.5) and inter-examiner reliability (PABAK ≥0.41) and were selected for further analysis. This was conducted to determine if the use of a cluster of components exhibited stronger validity than the original dataset.

Three components of three examination procedures performed by the *student osteopaths* appeared in the upper right quartile of Figure 21. These included:

2) Chin tuck (dysfunction)  
Accuracy=0.57; $K=0.76$

7) Trapezius stretch (dysfunction)  
Accuracy=0.66; $K=0.42$

8) Suboccipital tissue texture (discomfort)  
Accuracy=0.51; $K=0.71$

Three components of three examination procedures performed by the *experienced osteopaths* appeared in the upper right quartile of Figure 22. These included:

6) Chest expansion (discomfort)  
Accuracy=0.52; $K=0.76$

7) Trapezius stretch (discomfort)  
Accuracy=0.54; $K=0.46$

9) SCM tissue texture (discomfort)  
Accuracy=0.52; $K=0.66$

The validity of ‘one’, ‘two’ and ‘three’ positive dysfunction and/or discomfort findings out of each subgroup of examination procedures for both the student osteopaths and experienced osteopaths was assessed and presented in Table 24 and Table 25 respectively.

### Table 24: Clinical validity of subgroup 3 performed by the student osteopaths

<table>
<thead>
<tr>
<th></th>
<th>'1 of 3' positive findings</th>
<th>'2 of 3' positive findings</th>
<th>'3 or more' positive findings</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>95% CI</td>
<td>95% CI</td>
<td>95% CI</td>
</tr>
<tr>
<td>Overall accuracy</td>
<td>0.66</td>
<td>0.57</td>
<td>0.51</td>
</tr>
<tr>
<td>Positive Predictive Value</td>
<td>0.63</td>
<td>0.67</td>
<td>1.00</td>
</tr>
<tr>
<td>Negative Predictive Value</td>
<td>0.73</td>
<td>0.54</td>
<td>0.50</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>0.83</td>
<td>0.72</td>
<td>0.95</td>
</tr>
<tr>
<td>Specificity</td>
<td>0.48</td>
<td>0.32</td>
<td>0.63</td>
</tr>
<tr>
<td>Positive Likelihood Ratio</td>
<td>1.59</td>
<td>1.15</td>
<td>2.20</td>
</tr>
<tr>
<td>Negative Likelihood Ratio</td>
<td>0.35</td>
<td>0.17</td>
<td>0.74</td>
</tr>
</tbody>
</table>

Note: N/C = Not calculable; Figures shaded grey = those that met cut-off limits.
When a positive finding was generated from ‘one of the three’ components of the three examination procedures performed by the student osteopaths this cluster of procedures generated adequate sensitivity (0.83) adequate accuracy (0.66) and small but probable changes in -LR (0.35). However this subgroup had inadequate levels of specificity (0.48) and unimportant shifts in +LR (1.59). Although the -LR value 0.35 is small, a negative result (absence of dysfunction or discomfort) generated from performing these three examination procedures may produce small (approximately 25%) (see McGee (2002) bedside estimates - Table 29) but useful changes in probability that a negative finding is not associated to asthma (McGee, 2002).

Table 25: Clinical validity of subgroup 3 performed by the experienced osteopaths

<table>
<thead>
<tr>
<th></th>
<th>‘1 of 3’ positive findings</th>
<th>‘2 of 3’ positive findings</th>
<th>‘3 or more’ positive findings</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>95% CI</td>
<td>95% CI</td>
<td>95% CI</td>
</tr>
<tr>
<td></td>
<td>LL</td>
<td>UL</td>
<td>LL</td>
</tr>
<tr>
<td>Overall accuracy</td>
<td>0.60</td>
<td>0.50</td>
<td>0.49</td>
</tr>
<tr>
<td>Positive Predictive Value</td>
<td>0.76</td>
<td>0.60</td>
<td>0.60</td>
</tr>
<tr>
<td>Negative Predictive Value</td>
<td>0.55</td>
<td>0.49</td>
<td>0.49</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>0.31</td>
<td>0.17</td>
<td>0.45</td>
</tr>
<tr>
<td>Specificity</td>
<td>0.90</td>
<td>0.81</td>
<td>0.99</td>
</tr>
<tr>
<td>Positive Likelihood Ratio</td>
<td>3.10</td>
<td>1.10</td>
<td>8.70</td>
</tr>
<tr>
<td>Negative Likelihood Ratio</td>
<td>0.77</td>
<td>0.61</td>
<td>0.96</td>
</tr>
</tbody>
</table>

Note: N/C = Not calculable; Figures shaded grey = those that met cut-off limits.

When a positive finding was generated from ‘one of the three’ components of the three examination procedures performed by the experienced osteopaths this cluster of procedures generated adequate specificity (0.90) adequate accuracy (0.60) and small but probable changes in +LR (3.10). However it had inadequate levels of sensitivity (0.31) and unimportant shifts in -LR (0.77). Although the +LR of 3.10 is small, one positive finding (the presence of discomfort) generated from one of these three examination procedures provides small (approximately 20%) (see McGee (2002) bedside estimates - Table 29) changes in the probability of the finding being related to asthma

6.6.3. SUBGROUP 4 ANALYSIS

Participants were included in subgroup 4 from the original sample if they either: 1) had asthma and had objective lung findings of less than 80% of their predicted Forced Expiratory Volume in one second (FEV₁ <80%); or, 2) if they were in the non-asthmatic group and did not have objective lung findings of FEV₁ <80%. Three subjects from the asthmatic group and 20 from the non-asthmatic group were included in Subgroup 4. The data from these 23 participants was combined and presented in comparison to the combined validity of all components of all 10
examination procedures for the 41 participants for both the student osteopaths and experienced osteopaths respectively.

Table 26: Comparison of combined clinical validity of all components of all 10 examination procedures compared to subgroup 4

<table>
<thead>
<tr>
<th></th>
<th>Students osteopaths</th>
<th>Experienced osteopaths</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All 10 Procedures</td>
<td>Subgroup 4</td>
</tr>
<tr>
<td>(N=41)</td>
<td>SD</td>
<td>(N=23)</td>
</tr>
<tr>
<td>Overall accuracy</td>
<td>0.49</td>
<td>0.07</td>
</tr>
<tr>
<td>PPV</td>
<td>0.45</td>
<td>0.16</td>
</tr>
<tr>
<td>NPV</td>
<td>0.49</td>
<td>0.08</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>0.30</td>
<td>0.25</td>
</tr>
<tr>
<td>Specificity</td>
<td>0.69</td>
<td>0.22</td>
</tr>
<tr>
<td>+LR</td>
<td>0.91</td>
<td>0.54</td>
</tr>
<tr>
<td>-LR</td>
<td>1.05</td>
<td>0.31</td>
</tr>
<tr>
<td>Cohen's Kappa</td>
<td>0.31</td>
<td>0.26</td>
</tr>
<tr>
<td>PABAK</td>
<td>0.48</td>
<td>0.35</td>
</tr>
</tbody>
</table>

SD = Standard Deviation, PPV = Positive Predictive Value, NPV = Negative Predictive Value, +LR = Positive Likelihood Ratio, -LR = Negative Likelihood Ratio, PABAK = prevalence-adjusted and bias-adjusted kappa; Figures shaded grey = those that met cut-off limits.

The overall mean accuracy of all examination procedures was similar for both the student and experienced osteopathic groups (0.49 and 0.52 respectively). However, when including only those asthmatics with objective lung findings of FEV1<80%, accuracy increased from 0.49 to 0.65 for the student osteopaths and from 0.52 to 0.67 for the experienced osteopaths. NPV increased from 0.49 to 0.88 for the student osteopaths and from 0.51 to 0.89 for the experienced osteopaths. Positive likelihood ratio increased from 0.91 to 1.46 for the student osteopaths and 1.48 to 2.07 for the experienced osteopaths. There were minimal changes in Cohen’s kappa, but the adjusted (PABAK) kappa increased from 0.48 to 0.65 for the student osteopaths and 0.26 to 0.67 for the experienced osteopaths.
CHAPTER 7: DISCUSSION

7.1. STATEMENT OF PRINCIPAL FINDINGS

Various authors of osteopathic literature have proposed that asthmatics exhibit Somatic Dysfunction (SD) related to breathing obstruction (Chaitow et al., 2002; DiGiovanna & Schiowitz, 1991; Green, 2000; Sammut & Searle-Barnes, 1998; Ward, 2003). Somatic Dysfunction is said to occur from stresses imposed on the neuromusculoskeletal system beyond its normal capacity (Denslow, 1964; Glossary of osteopathic terminology, 2006; Sammut & Searle-Barnes, 1998). Somatic Dysfunction is clinically identified by detection of alteration in tissue texture on palpation, asymmetry of structure, restriction of joint motion, and tenderness/discomfort on palpation (Ward, 2003). Authors of osteopathic literature have indicated there are specific areas of the cervical spine, thoracic spine, ribs and respiratory muscles where SD related to asthma are thought to commonly occur (Chaitow et al., 2002; DiGiovanna & Schiowitz, 1991; Kuchera & Kuchera, 1994; Sammut & Searle-Barnes, 1998). However, there are no published studies that support this putative relationship between SD and asthma. One unpublished study, conducted by Russell (2006), investigated the association between asthma and SD using 20 physical examination procedures and identified a group of nine examination procedures which demonstrated adequate accuracy >0.5 and acceptable reliability ≥0.4 kappa. The current study continued on from Russell’s (2006) study, and aimed to evaluate the validity, reliability and potential utility of these nine manual examination procedures, plus an additional observational procedure which assessed for breathing dysfunction using both student osteopaths and experienced osteopaths. It has been assumed that clinical experience leads to more refined palpation skills, and thus better reliability (Chaitow, 2008). To investigate this assumption the present study used both experienced and student osteopaths to assess the diagnostic usability of the examination procedures.

Each of the nine examination procedures were divided into two components:

- The first component assessed for one or more palpatory and/or observational findings of: alteration in tissue texture, asymmetry of structure and/or restriction of motion (named ‘dysfunction component’ [one to ten]);
- The second component assessed for reporting of discomfort from the subject during the examination procedure (named ‘discomfort component’ [one to nine]).
All nine examination procedures were analysed using these two components to determine if the discomfort component of the procedure generated greater reliability than the dysfunction component which assessed for findings of: alteration in tissue texture, asymmetry of structure and/or restriction of motion. An additional examination procedure (examination procedure 10) was added to the examination protocol as it assessed for a dysfunctional ‘faulty’ upper chest breathing pattern, a pattern hypothesized to occur in asthmatics (Courtney, 2009). Unlike examination procedures 1 to 9 examination which assessed for two components (dysfunction and discomfort), examination procedure ten assessed for one component (dysfunction). The complete examination protocol consisted of a total of 10 examination procedures which included 10 dysfunction components and 9 discomfort components.

The current study failed to show a clinically useful association between asthma and all components of all 10 examination procedures collectively. Our findings are contradictory to previous claims described in textbooks (Chaitow et al., 2002; DiGiovanna & Schiowitz, 1991; Kuchera & Kuchera, 1994; Sammut & Searle-Barnes, 1998) of an association between SD and asthma.

This study has however, identified three discomfort components of three examination procedures with acceptable reliability (adjusted kappa) and with small positive likelihood ratio. The examination procedures were: 1) Chest expansion for discomfort; 2) Trapezius stretch for discomfort; and 3) Sternocleidomastoid tissue texture palpation for discomfort. If discomfort is reported by the patient during any of these examination procedures, the probability of an association between asthma and the SD found increases by 20%. In manual therapy, this is a clinically important finding – as many manual palpatory procedures exhibit much lower reliability. To put this finding in context it is useful to consider some other examples of shifts in probability for other common procedures. A common orthopedic procedure for knee joint effusion from internal derangement, the ballottement test, exhibits less than 10% probability in favor of a positive finding being associated with arthroscopic changes (Kastelein et al., 2009). Additionally, nuchal rigidity and Kernig's signs are common clinical procedures that are almost universally used on patients suspected to have meningitis. Each of these procedures also exhibit less than 10% probability in favor of a positive finding being associated with meningeal irritation (Waghdhare, Kalantri, Joshi, & Kalantri, 2010). The results generated by the experienced osteopaths when using these three procedures were not reproducible by the student osteopaths used in our study, therefore further research is required to confirm the validity of these tests.
One of the objectives of this study was to compare the reliability and validity of *dysfunction* components of examination procedures 1-10 against the *discomfort* components of examination procedures 1-9. This study found that the *discomfort* components of examination procedures 1-9 (participant reports of discomfort) were more reliable (had higher Kappa scores) and had higher specificity when compared to the *dysfunction* components of examination procedures 1-10 that relied on examiner perception.

The second objective of this study was to examine the relationship between experience of the examiner and the ability to reliably identify SD. The student osteopaths showed greater reliability than their more experienced counterparts.

The third objective of this study was to determine the association (if any) between SD and asthmatics with lung obstruction. This was done by comparing the overall mean validity and reliability of all 10 examination procedures from the original dataset against overall mean validity and reliability of all 10 examination procedures from subgroup 4 – a dataset consisting of only asthmatic participants with objective findings of lung obstruction (FEV₁<80% predicted) (n=3) and non-asthmatic participants with no objective findings of lung obstruction (n=20). The sample size of subgroup 4 was underpowered to draw any definitive conclusions. However, notwithstanding the limitations of the dataset, increases in +LR, accuracy and PABAK for both experienced and student osteopaths were found when compared to the original dataset.

This research was limited by the small number of subjects (n=41) and examiners (n=4), the fact that most asthmatics who participated in this study were taking medication, and that few had objective findings of lung obstruction. The degree of lung obstruction, the duration of obstruction, and how well it is managed by pharmacotherapy may influence the extent to which SD is exhibited in people with asthma. We hypothesize that if asthma is diagnosed early, and airflow obstruction is small, or not prolonged, or well controlled by medication, then the work of breathing to overcome the initial obstruction may not exceed a threshold where anatomical or functional adaptation of the respiratory pump (thoracic spine, cervical spine and respiratory muscles) occurs. In such a case, there may be no clinically observable differences in anatomical structure or function between those with well controlled asthma and normal, non-asthmatic, people.
7.2. RELIABILITY

Investigating the reliability of physical examination procedures presents many problems. Firstly, the commonly used statistic to measure reliability of dichotomous tests, the kappa statistic, becomes unstable when there is limited spread in the findings. Secondly, palpation itself is inherently variable. The following section discusses these factors in relation to the findings in this study.

The kappa coefficient ($\kappa$) is the most widely used statistic to measure the extent of agreement between two raters in manual medicine. Kappa takes into account agreement beyond chance and can therefore be considered to represent “true” agreement – defined as agreement that is achieved beyond chance alone.

Influence of unbalanced prevalence on reliability

Imbalances in concordant and discordant findings can be measured using the Prevalence and Bias index respectively. When there is a large bias the magnitude of the kappa is higher than when bias is low or absent. When there is a very high or very low prevalence of the index being measured, the kappa is artificially reduced. This concept is illustrated in Table 27 where there is an unequal distribution of agreements for discomfort examination procedure 2. The figures are then altered it to illustrate equal distribution of agreements (see Table 28). For discomfort component of examination procedure 2, both experienced osteopaths agreed on 36 negative findings (cell D) and 0 positive findings (cell A). The kappa value indicated agreement no better than chance (-0.06) yet the percentage of examiner agreement was 88%. A high percentage agreement and low kappa can be explained by an imbalance in the prevalence findings in the dataset and reflected in the Prevalence Index. The Prevalence Index for discomfort examination procedure two was extremely low at -0.88. Table 28 illustrates an ‘ideal’ scenario where the proportion of agreements for discomfort examination procedure 2 are evenly balanced. The figure from cell D of Table 27 (n=36) was divided equally into both negative and positive findings and placed into cells A (n=18) and D (n=18) of Table 28, giving a prevalence index of zero. The percentage agreement is still 88%, however the kappa value has risen substantially and now indicates substantial agreement ($\kappa = 0.76$). This illustrates the problem that occurs in interpreting the kappa statistic when there is an extremely low or high prevalence index. Because of this paradox, the use and reporting of the prevalence and bias adjusted kappa values as additional agreement measures are important, especially when prevalence cannot be easily
achieved. Studies that report low kappa values without reporting overall agreement and adjusted kappa indexes (PABAK) can be misleading (Huijbregts, 2002).

Table 27: Discomfort component of examination procedure 2 performed by the experienced osteopaths (with extremely low prevalence)

<table>
<thead>
<tr>
<th>Examiner</th>
<th>Rater B</th>
<th>Positive Finding</th>
<th>Negative Finding</th>
<th>Marginal Totals</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td></td>
<td>Positive Finding</td>
<td>Negative Finding</td>
<td>Marginal Totals</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0 (A)</td>
<td>2 (B)</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3 (C)</td>
<td>36 (D)</td>
<td>39</td>
</tr>
<tr>
<td>Marginal Totals</td>
<td>43 (a+c)</td>
<td>38 (b+d)</td>
<td>41 (N)</td>
<td></td>
</tr>
</tbody>
</table>

Table 28: Discomfort component of examination procedure 2 performed by the experienced osteopaths (with zero prevalence)

<table>
<thead>
<tr>
<th>Examiner</th>
<th>Rater B</th>
<th>Positive Finding</th>
<th>Negative Finding</th>
<th>Marginal Totals</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td></td>
<td>Positive Finding</td>
<td>Negative Finding</td>
<td>Marginal Totals</td>
</tr>
<tr>
<td></td>
<td></td>
<td>18 (A)</td>
<td>2 (B)</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3 (C)</td>
<td>18 (D)</td>
<td>21</td>
</tr>
<tr>
<td>Marginal Totals</td>
<td>21 (a+c)</td>
<td>20 (b+d)</td>
<td>41 (N)</td>
<td></td>
</tr>
</tbody>
</table>

7.2.1. COMPARING INTER-EXAMINER RELIABILITY OF EXPERIENCED OSTEOPATHS VERSUS STUDENT OSTEOPATHS

The second objective of this study was to compare the reliability estimates of experienced practicing osteopaths with those of first year master of osteopathic students. Student osteopaths were more reliable for all components of all examination procedures compared to their experienced osteopaths. These results are surprising and contrary to the presumption that experienced osteopaths would have more advanced palpatory skills, and thus greater reliability (Chaitow, 2008).

The inter-examiner reliability of discomfort components of examination procedures 1-9, calculated with Cohen’s kappa, revealed a broad range of agreement for the student osteopaths, ranging from less than chance to perfect agreement. In contrast, the experienced osteopaths exhibited poor agreement from less than chance to no better than chance.

The level of palpatory skill depends on both the individuals capacity to sense stimuli (a ‘sensory component’) and the individuals ability to interpret the findings, based on the individuals perception of normal or dysfunctional (an ‘interpretation component’) (Beal & Burns, 1989). These two factors may be influenced by experience. The sensory component of palpation, namely tactile acuity thresholds of the fingers are largely determined by the size of the cortical representation of the fingers in somatosensory cortex (Duncan & Boynton, 2007). The size of the cortical representation of the fingers increases with tactile acuity training, but reaches a threshold
limit relatively quickly. Tactile acuity and finger representations of the somatosensory cortex also decrease quickly with immobility (Lissek et al., 2009). This means that once training reaches its threshold limit, further training will not increase tactile acuity and that if you discontinue training your tactile acuity decreases. The ‘interpretation component’ of palpation however does not have an upper limit threshold and expands and changes as the practitioner trains and gains experience. The ‘interpretation component’ will therefore vary greatly between people with difference experience and training levels. As such, attempts were made in this study to use examiners with similar training and experience to decrease the variability of the ‘interpretation component’ of palpation. The student osteopaths had undertaken the same undergraduate training, and had the same level of clinical experience. However, due to availability, the experienced practicing osteopathic examiners used in this study had different educational backgrounds (different institutions in different countries) different clinical experience levels (20 and 6-years), and had worked on different demographic population bases. This is likely to have contributed to the experienced osteopaths’ relatively low reliability kappa coefficients. This compares with the relatively high reliability of the student osteopaths, whose palpatory skills developed in response to the same undergraduate educational experience, and consequently were similar to each other. Because both students education was the same, the chances of variability are probably reduced.

Marx, Bombardier, & Wright (1999) also suggest that experience and expertise may affect the performance of an examination procedure (Marx, Bombardier, & Wright, 1999). Varying reliability results have been generated from manual palpation, and clinical observation studies that have used both student osteopaths and experienced osteopaths. Similar to our study, Moran & Ljubotenski (2006) noted greater reliability of post graduate osteopathic students compared to clinicians with 15 years or more experience (Moran & Ljubotenski, 2006). However other studies have concluded that experience causes no appreciable differences in reliability (Kmita & Lucas, 2008; Seffinger et al., 2004; Stochkendahl et al., 2006), and findings by Billis, Foster & Wright (2003) demonstrated that physiotherapy students had poorer agreement than clinicians with two years experience and manual therapists with greater than five years experience (Billis et al., 2003).

Based on participant feedback after the data collection sessions it became apparent during the study, that one examiner (an experienced osteopath) had slightly firmer palpation strength and pressure compared to the other examiners. It is unknown whether firmer pressure influences the outcome of the examination procedure (i.e. identifies more dysfunction or the opposite).
Although depth and intensity of pressure may affect subjective reports of discomfort, in this study the examiner agreement for examination procedures on the absence of discomfort was relatively high for both groups, indicating that if there were variances in palpation pressure this had marginal impact on participants’ positive responses. However, when examining the examiner findings of the dysfunction components of examination procedures 1-10, it was apparent that examiner B had a bias towards ‘dysfunctional findings’ compared to the other examiners (A, C and D).

These factors may partly explain why student osteopaths achieved a greater number of examination procedures with acceptable reliability (when comparing both adjusted and unadjusted kappa values) than their more experienced counterparts.

7.2.2. COMPARING INTER-EXAMINER RELIABILITY OF DYSFUNCTION COMPONENTS VERSUS DISCOMFORT COMPONENTS OF ALL EXAMINATION PROCEDURES

The first objective of this study was to assess the validity and reliability of dysfunction components of examination procedures 1-10 versus discomfort components of examination procedures 1-9.

Acceptable levels of reliability and percentage agreement were obtained from one dysfunction component of one examination procedure performed by the student osteopaths. All dysfunction components of examination procedures 1-10 performed by the experienced osteopaths exhibited low levels of reliability. Poor examiner agreement was also indicated by the low levels of percentage agreement (observed agreement) for the remaining dysfunction components performed by the student osteopaths which did not exceed over 71% and all dysfunction components performed by the experienced osteopaths which did not exceed over 59%.

Acceptable levels of reliability were obtained for four discomfort components of four examination procedures performed by the student osteopaths. All discomfort components of all examination procedures performed the experienced osteopaths exhibited poor levels of reliability. However the percentage agreement (observed agreement) for the discomfort components was greater than 80% for the osteopathic students and greater than 73% for the experienced osteopaths. Whilst the percentage agreement for discomfort components was high, this was not reflected in the kappa values for all of these tests. A high percentage agreement and low kappa can be explained by an imbalance in the prevalence findings in the dataset. In our
dataset there was higher agreement between student and experienced osteopaths on the absence of discomfort than agreement on the presence of discomfort. This imbalance was expressed in the extremely low prevalence indexes for discomfort examination procedures (>-.50) for both student and experienced osteopaths. As discussed earlier, if the prevalence of the attribute being measured in the sample is extremely high or low, the unadjusted kappa is artificially reduced. Once the kappa value was readjusted to take into consideration these imbalances, using the adjusted (PABAK) kappa statistic, there were notable increases in the reliability of both groups. As calculated with PABAK kappa, all discomfort components by the student osteopaths achieved acceptable levels of reliability and 8 of the 9 discomfort components performed by the experienced osteopaths achieved acceptable levels of reliability.

Overall, the discomfort component of examination procedures proved to be more reliable (when comparing both unadjusted and adjusted kappa values) than the dysfunction component of examination procedures. Similar findings were also identified by Russell (2006). Several other published studies have also concluded that palpation for discomfort, tenderness, or pain generates superior reliability than motion palpation, static palpation or palpation for tissue texture changes (Haneline & Young, 2009; Hestboek & Leboeuf-Yde, 2000; Seffinger et al., 2004; Stochkendahl et al., 2006).

One reason why the discomfort components of examination procedures may have generated higher reliability is the reduced chance of variability involved. Unlike the reliability of dysfunction components of examination procedures, which compare two examiners palpation abilities to sense stimuli and interpret findings based on their perception of ‘normal’ or ‘dysfunctional’, reliability of discomfort components compare only the subjects ability to sense, interpret and report back the pressure as either ‘discomfort’ or ‘non-discomfort’. If the palpation pressure used on the subject is consistent, then the results are likely to be more consistent with the subjects’ perception of discomfort or non-discomfort.
7.3. VALIDITY

Procedure validity measures the performance of an examination procedure in comparison to a reference standard. In this study, the reference standard was defined as the presence of self-reported diagnosed asthma. When considering the validity of dysfunction and discomfort components of all examination procedures, the following statistics were reported: accuracy; sensitivity and specificity; positive and negative predicative values; positive and negative likelihood ratios. It is important to appreciate the correct interpretation of these statistics as they have varying contributions to validity. The following sections discuss these statistics in relation to the results generated from this study.

7.3.1. ACCURACY

Accuracy reflects the proportion of positive and negative results that are correct against the reference standard. A ‘perfect test’ with 100% accuracy (indicated as the decimal 1.0) is one that every positive examination procedure result (presence of SD) is associated with the presence of diagnosed asthma as reported by the participant and that every negative examination procedure result (absence of SD) is associated with the absence of asthma as reported by the participant. In order to achieve perfect accuracy, every asthmatic must exhibit SD and every non-asthmatic must not exhibit SD.

There is no standardized level of acceptable accuracy for manual therapy studies. In order to determine cut-off limits for diagnostic procedures (determining the best value for distinguishing between disease and non-disease) receiver-operating characteristic (ROC) curves can be plotted (Goutham, 2003). Sensitivity is plotted on the y-axis and 1-specificity is plotted on the x-axis, giving a graphical representation of the balance between sensitivity and specificity. The shape of a ROC curve and the area under the curve gives an estimation of the discriminative power of a procedure. Generally speaking, diagnostic accuracy is determined by the area under the curve. Šimundić (n.d.) have proposed a categorization of the strength of accuracy based on cut-off levels values from area under the ROC curve (Šimundić, n.d). It was not within the scope of this study to construct ROC curves for each procedure and determine the best accuracy cut-off limits for each test. The cut-off level for ‘adequate’ accuracy was arbitrarily set to >0.5. This cut-off point was used in this study to select components of examination procedures of acceptable accuracy. Acceptable levels of sensitivity and specificity were also arbitrarily set to >0.8. These were the same cut-off levels previously used by Russell (2006).
The overall accuracy of all components of all 10 examination procedures ranged ‘inadequate’ to ‘adequate’ for both the student and experienced osteopathic groups, with no appreciable differences in the accuracy between both groups. The students’ accuracy for all components of all examination procedures ranged from 0.34 to 0.66, whilst the experienced osteopaths’ accuracy ranged from 0.49 to 0.62.

A limitation with the examination procedures used in this study is that they are not specific to detecting only asthmatic related SD. Many other influences could also alter the outcomes of these examination procedures. Chaitow, Bradley, & Gilbert (2002) state that the following may also cause evidence of joint and soft tissue restrictions and imbalances:

- anxiety and emotional states
- occupational and leisure influences
- proprioceptive and other neural inputs
- inborn characteristics (for example an anatomical short leg)
- habitual patterns of use (for example upper-chest breathing) (Chaitow et al., 2002).

Screening for the presence of these factors was not undertaken as part of this study, therefore any influences these factors may have had on the results could not be assessed. For example, ‘slouched posture’ has been related to a forward head posture, and it has been suggested that forward head posture is associated with shortness of the upper trapezius muscle (Janda, 1983). Shortness/hypertonicity of the upper trapezius muscles has also been implicated to occur in asthmatics (Chaitow et al., 2002; Janda, 1983). Therefore examination procedure 7, trapezius length assessment, may cause false-positives in non-asthmatics that also demonstrate slouched posture. In a clinical context, this may not be very important since most clinicians are likely to address slouched posture and forward head posture regardless of its aetiology. In a research sense, this is interesting but may not be very clinically important to address as knowing the aetiology is unlikely to greatly alter the clinical management approach.

Also, in a normal clinical setting, examiners would not be limited to the defined areas assessed in this study. Commonly, assessment protocols would include a broader area, thus clinicians may be able to make judgments about whether any dysfunction in the area examined in this research was associated (a reflection of adaptation) with a dysfunctional area, trauma or strain distal to the sites assessed in this study. For example, a participant with an anatomically short leg may have a structural scoliosis in their thoracic spine. Therefore, when an examiner performed examination procedure 3, thoracic rotation, they may find dysfunctional and/or discomfort
findings. However the actual finding may be associated with the subject’s leg length discrepancy, rather than asthma and thus the examiner could apply their clinical judgment when considering this factor. Again, such a finding may cause false-positives in the non-asthmatic group. Although an exclusion criteria was utilized to exclude injuries and surgery to areas assessed in the study, no exclusion criteria or pre-screening session was conducted for dysfunctions in distal areas to those assessed, or for anatomical abnormalities.

Another limitation in this study is that the validity of each individual examination procedure is unknown - it is unknown whether each examination procedure actually measures what it is hypothesized to measure. This is because the majority of examination procedures used in this study do not have a reference criterion for comparisons. For example, the dysfunction component of the suboccipital tissue texture procedure is thought to measure ‘tightness’ of the suboccipital muscles. However, there is no objective measurement tool that measures ‘tightness’ of these muscles, and consequently we have nothing to compare our results against.

As such, this study compares the association between SD (e.g. tightness of the suboccipital muscles) and asthma; a pathology that is hypothesized to be one of the etiological factors that gives rise to the particular SD examined. The SD assessed in this study is believed to result from the increased work of breathing that occurs in asthmatics due to their intermittent lung obstruction. However the prevalence of SD in the asthmatic population is unknown; having asthma may not always result in overt findings of SD. Instead there may be other factors involved such as the chronicity and the severity of asthma, how well a patient’s asthma is managed by medication, and whether or not the person with asthma has objective findings of lung obstruction. High levels of accuracy in this study may be impossible to achieve due to these limitations. For this reason interpretation of the study findings should be made cautiously and with awareness of these limitations.

The majority of asthmatic subjects (18 of 21) in this study were medicated. None of the asthmatics were symptomatic at the time of the study, and only three of the 21 asthmatics exhibited signs of lung obstruction (FEV<sub>1</sub> <80% predicted) which was assessed using a spirometry by the primary researcher. Most spirometry is performed in a clinical setting by a spirometry nurse or respiratory specialist. Both undertake specialised training in spirometry. Although the protocol described by Miller et al (2005) was followed, the primary researcher had limited experience and it is possible that measurements collected may be less than optimal because it has been demonstrated that practitioners who do not undergo specialised spirometry
training generate a lower number of acceptable maneuvers and have decreased reproducibility than trained practitioners (Eaton et al., 1999).

The third objective of this study was to determine the degree of association between somatic dysfunction and 1) participants with asthma and objective findings of lung obstruction; and 2) participants with no asthma and no objective findings of lung obstruction. This was done by comparing the combined validity of all components of all examination procedures from only those asthmatics with objective findings of lung obstruction, and non-asthmatics with no signs of objective lung obstruction was assessed (named subgroup 4). This subgroup was compared against the mean overall combined accuracy of all components of all examination procedures. Subgroup 4 showed higher reliability, higher positive likelihood ratios, and higher accuracy for both student osteopaths and experienced osteopaths. The accuracy of student osteopaths increased from 0.49 to 0.65, and the accuracy of experienced osteopaths increased from 0.52 to 0.67. The reliability of each group of examiners also increased when using the adjusted (PABAK) kappa. Reliability increased from $K = 0.48$ to 0.65 for the student osteopaths and from $K = 0.26$ to 0.67 for the experienced osteopaths. The rise in accuracy and reliability may be due to this subgroup of asthmatics having more obvious musculoskeletal dysfunction/palpable deviations from ‘normal’ function. A study by Fryer (2006) demonstrated greater inter-examiner reliability when the structural asymmetry of medial malleolus level was large (>4mm). The study by Fryer (2006) suggested that the amount of asymmetry has a major effect on inter-examiner reliability; however we don’t know whether this applies to other regions of the body.

### 7.3.2. Sensitivity and Specificity

In this study specificity represents the ability of an examination procedure to identify the absence of discomfort or dysfunction in non-asthmatics (Fritz & Wainner, 2001). Sensitivity represents the ability of an examination procedure to identify discomfort or dysfunction in asthmatics.

In our study, discomfort components of examination procedures 1-9 generated high specificity values. Only one discomfort component of one examination procedure performed by the student osteopaths (CRLF) exhibited high sensitivity, whilst no other discomfort components performed by the experienced osteopaths exhibited high sensitivity. All discomfort components of examination procedures 1-9 performed by student osteopaths exhibited specificity greater than 80%, and all but one of the nine discomfort components performed by the experienced osteopaths exhibited specificity greater than 80%. High specificity of discomfort components
were also reported in Russell’s (2006) thesis; however no explanation for this was included in her unpublished thesis (Russell, 2006).

High specificity and low sensitivity values occurred in discomfort components of examination procedures 1-9 because there was a high proportion of non-discomfort findings in both non-asthmatics (true-negatives) and asthmatics (false-negatives) with a subsequent low proportion of discomfort findings diagnosed in asthmatics (true-positives) and non-asthmatics (false-positives).

High specificity and low specificity either represents a situation where an examination procedure lacks discriminative power in determining a positive examination procedure result or a situation where there is an examiner bias towards diagnosing an examination procedure as negative. Caution must therefore be used when interpreting results from examination procedures with high values of specificity and low values of sensitivity as they may provide little clinical value. The real measure of validity lies in examination procedures which exhibit high values of both specificity and sensitivity.

7.3.3. LIKELIHOOD RATIOS
Likelihood ratios are a measure of diagnostic accuracy (Steven, 2002). Conventional use of likelihood ratios (LR) requires the pre-test probability (prevalence) of the finding of interest. Using the LR, the pre-test probability is then converted to the post-test probability. The prevalence of SD associated with asthma in the asthmatic population is unknown hence the traditional use of LR has limited use. However McGee (2002) advocates the use of ‘bedside LR estimates’ – reported in Table 29. These estimates are independent of pre-test probability and allow the approximate change in the probability to be calculated. McGee (2002) does however state that this method is inaccurate for pre-test probabilities less than 10% or greater than 90%, however, these polar extremes indicate diagnostic certainty in most clinical scenarios, thus it would be unnecessary to undertake further testing (McGee, 2002).
While the likelihood ratios reported in the present research are not definitive or substantial, they have at least some clinical utility. For example, the +LR obtained by the experienced osteopaths for subgroup 2 (a cluster of three examination procedures) was 3.10. If the three examination procedures from subgroup two were used on a subject with asthma and a SD was found, then that the probability of that patients SD being related to their asthma (when using McGee (2001) quick bedside estimates) increases by approximately 20% (see Table 29) (McGee, 2001).

Conversely, the -LR obtained by the student osteopaths for subgroup 2 (a cluster of three examination procedures) was 0.35. If the three examination procedures from the student osteopaths subgroup 2 were used on a subject with asthma and no SD was found, the probability that a negative finding is not associated to the patient’s asthma increases by approximately 25%.

However, in this study the -LR holds lesser clinical meaning. This is because in a normal clinical scenario, the practitioner would already know whether the patient had asthma. Therefore for examination procedures to have any usefulness they must aid in determining whether the asthmatic patient has SD associated with their asthma, rather than whether the asthmatic patient does not have SD associated to their asthma.

Although it may be argued that the positive likelihood ratios generated from this study are too low to provide any clinical usefulness, they do show a small probability of SD being associated with asthma. Somatic dysfunction may in clinical practice, along with other findings (symptoms of breathing dysfunction, objective findings of lung obstruction etc) aid the clinician in raising or lowering the probability of an association of their findings with asthma.
It is important to note that the indices: accuracy, +LR, -LR, sensitivity and specificity must be viewed within the limitations in this study, primarily the reference standard (which is the presence of asthma). This reference standard is not an absolute criterion for “asthmatic related SD” and as such, comparisons may not be entirely accurate. For this reason the findings from this study must be interpreted with caution.

Although underpowered to draw any definitive conclusions we found subgroup four$^8$ to have an overall higher mean validity (+LR and accuracy) and reliability (adjusted (PABAK) kappa) when compared to the original dataset for both groups (student osteopaths and experienced osteopaths). These results may indicate that there is a greater association between SD and asthmatics with objective findings of lung obstruction. Russell (2006) also demonstrated minor increases in accuracy when only using asthmatics with objective findings of lung obstruction; however, like our dataset it was also underpowered. This relationship should be the subject of further research. A sample that includes both 1) non-asthmatics and 2) asthmatics with confirmed lung obstruction would be useful in assessing this relationship.

### 7.3.4. POSITIVE AND NEGATIVE PREDICTIVE VALUES

Positive predictive values (PPV) are used to express the proportion of those with a positive result that truly have the specified disease, in this case, asthma. It is calculated by: True Positives/True Positives + False Positives. Negative predictive values (NPV) express the proportion of those with a negative result that do not have the specified disease. It is calculated by: True Negatives/True Negatives + False Negatives. To illustrate how PPV and NPV are calculated consider the example: examination procedure 1: active side bending of trunk performed by the two student osteopaths resulted in findings of dysfunction in 46 out of the 82 procedures performed on the 41 participants. Only 21 of these participants had asthma. The proportion of correct diagnosis was 21/46 = 0.46. Similarly, the student osteopaths reported an absence of dysfunctional findings in 36 of the 82 participants. Only 15 of these participants were non-asthmatics. The proportion of correct diagnosis was 15/36 = 0.42. However, these proportions are of only limited validity. These measures are critically dependent on the sample chosen, and the prevalence of the abnormality in the patients being tested. When there is a larger prevalence of the abnormality in the patients being tested, the positive predictive value of an examination procedure rises. Therefore the positive predictive value of an examination procedure reflects the prevalence of the abnormality in the patients being tested and not the property of the

---

$^8$ Subgroup 4 contained only asthmatics with objective findings of lung obstruction and non-asthmatics with no objective findings of lung obstruction
examination procedure itself. Because the prevalence of SD related to asthma is unknown we have not focused on these statistics (Fritz & Wainner, 2001).

7.4. STRENGTHS, LIMITATIONS AND WEAKNESSES

7.4.1. INTERNAL VALIDITY

There were a number of limitations and weaknesses in this study which limit the interpretation of these results, some of which have already been discussed. Strengths of this study include: blinding examiners from information that may have influenced their findings. Examiners were blinded to details of the patient’s status (presence or absence of asthma), each others findings, and all other patient information. Steps were also taken to ensure any objective cues, such as subject medication or inhalers, were hidden from examiners. The examination procedures were also undertaken in isolation. In order to do this, examiners were asked to refrain from talking to participants during the examination to prevent examiners from gaining information or additional cues about their subject. This is not representative of normal clinical practice, but essential to exclude any information that may have influenced examiners decision processes in order to identify the extent to which the examination procedures alone can indicate dysfunction or discomfort. The implication of examining examination procedures in isolation is that findings are more conservative and probably underestimate what can be achieved in normal clinical practice – since in practice further information about the participant can be gained by the examiner. The findings from this study therefore represent the utility of the examination procedures separate from the outcomes that may be achieved in clinical practice when the examination procedures are combined with participant history and other examination procedures (not used in this study) and verbal cues.

A weakness in this study lies in the recruitment of participants and examiners. Participants were not randomly selected and therefore the potential for sample bias is possible. Although there was reasonable diversity in age, sex and morphology in the sample, readers should be cautious when drawing conclusions from these findings because the extent of any possible sample bias is unknown. The nature and aims of the study was also disclosed to potential examiners – they were aware that asthmatic participants would be involved. This was done in order to increase potential examiners interest but also to avoid deception which would have been problematic for ethical approval of the study. This would tend to bias the study in favor of overestimating the accuracy of procedures but may better represent what would happen in usual practice, where the examiner is likely to have a working hypothesis in mind and be seeking confirmation of findings
from the physical examination process. Also, both subjects and examiners were drawn from the
general public, including student and teaching staff at Unitec NZ. It is possible that the
examiners (students and clinic tutors at Unitec NZ) may have seen potential participants reading
recruitment advertisements placed in the Unitec health clinic and around the Unitec campus. It
was also possible that some of the participants and/or examiners were known to each other.
However, those subjects that disclosed they did know one or more of the examiners
communicated that the examiners would have no knowledge of their prior medical history.
However this also introduces a potential bias of examiners guessing whether subjects do or don’t
have asthma.

7.4.2. Diagnosis of Asthma
The diagnosis for asthma as described by NZGG (2002) relies on satisfying two criteria. The
first is subjective and relies on key diagnostic indicators (history of a cough, wheeze and
shortness of breath) as provided by the patient. The second criterion is objective and relies on
evidence of reversible airflow obstruction which occurs either spontaneously over time or in
response to pharmacotherapy. The objective criterion is necessary, but not sufficient, to arrive at
the diagnosis of asthma. The subjective diagnostic indicators must also be met as objective
evidence of lung obstruction on its own could be due to a number of causes other than asthma
such as chronic obstructive pulmonary disease and central airway obstruction from a foreign
body. The subjective diagnostic indicators of asthma rely on participant self-reporting.
Therefore, there is inevitably some subjectivity about the accuracy of asthma diagnosis. This
subjectivity, however, is a common limitation when studying any disease process whose
diagnostic criteria contain by definition subjective (self-reported) indicators.

7.4.3. Sample Size
The sample size was limited to 41 participants due to the time limitations and costs involved
with this study. This means that confidence intervals are wider, and consequently the estimates
of validity are less precise than they could have been with a larger group.

7.4.4. Training Period
It has been suggested that standardization of physical examination procedures improves
reliability (Fritz & Wainner, 2001). However, there is debate on what influence consensus
training has on reproducibility studies and how much consensus training (if any) should be used.
Two systemic reviews concluded that consensus training has little influence on reliability
(Seffinger et al., 2004; Stochkendahl et al., 2006). However, in both reviews, the amount of consensus training, or type of training involved was not mentioned. It is possible that studies included in these reviews had minimal consensus training which is reflected in the apparent lack of influence on reliability. Other studies have shown improved reliability after the use of consensus training (Degenhardt et al., 2005; Fryer et al., 2005). One study in particular has shown significant improvements in reliability of palpatory diagnostic tests of the lumber spine after one two-hour training sessions each week over a four month period. The training consisting of a protocol designed to promote consensus by modifying examiner technique until a high percentage agreement was identified between examiners for each procedure (Degenhardt et al., 2005). Inter-examiner reliability was assessed pre and post consensus training. Pre consensus training, reliability of tenderness was ‘fair’ ($K = 0.34$). All other palpatory tests (tissue texture, and positional asymmetry) were ‘poor’ ($K < 0.20$). After consensus training reliability of detecting tissue texture changes of the lumber spine increased to ‘moderate’ ($K = 0.45$), reliability of detecting tenderness increased to ‘substantial’ ($K = 0.68$) and positional asymmetry increased to ‘fair’ ($K = 0.34$).

Whilst rigid training protocols may enhance internal validity, it does so at the expense of external validity by diverging from normal clinical practice. However, because of the variation in performance of psychomotor skills between practitioners in general (Streiner & Norman, 2003) a certain amount of training is needed in many manual therapy studies such as ours to clarify procedures define operational criteria for ‘dysfunctional’ and ‘normal’, standardize palpation strength for discomfort, and ensure consistency in recording results. As such, efforts were made in this study to use a 1.5-hour examiner familization session prior to the data collection. This session was thought to sufficiently familiarize examiners with examination procedures and the recording protocol. The session was more reflective of what is undertaken in normal clinical practice and we therefore consider the study to be more representative of effectiveness rather than efficacy (what reliability can be achieved with training). By conducting our familization session in this way, the effect of changes to external validity was minimized.

Based on participant feedback after the data collection sessions it became apparent during the study, that one examiner (an experienced osteopath) had slightly firmer palpation strength and pressure compared to the other examiners. The 1.5-hour familization session may have been insufficient to overcome individual palpation strength differences. Another factor that may account for the variances in palpation strength was that the time elapsed between the familization session and first data collection was three weeks. This may have been too long and
examiners may have forgotten the appropriate palpation pressure to use. More consensus training closer to the data collection date may have minimized these differences. Also, an initial check on the first day to ensure examination procedures were performed adequately and that palpation strength and depth being used was consistent between examiners may have also minimized palpation variances. The use of a pressure device, for example pressure sensors attached to the palmer surface of the fingers like that used in a study by Marcotte, Normand & Black (2005) may also have aided in improving consistency of palpation pressure.

7.4.5. STANDARDIZING AREAS OF PALPATION

Hicks, Fritz, Delitto, & Mishock (2003) suggested that errors in individual spinal segmental motion testing may be associated with inadequately localizing the certain spinal segment to be evaluated. To minimize this type of error, certain landmarks were marked prior to testing to ensure examiners were assessing the same areas. However, appraisal of the examiners consistency of placing their hands on the correct area was not conducted.
7.4.6. REDUCING EXAMINER FATIGUE

It is unlikely that fatigue was a factor in decision making. The data collection was divided into five sessions to decrease the potential of examiner fatigue. Brismée et al. (2006) conducted a study on the reliability of a passive physiological inter-vertebral motion testing in the mid-thoracic spine. They indicated that a rest period between a 9-hour testing sessions sufficiently limited tester fatigue. This was measured by looking at the correlation between percentage agreement and subject testing order which was non-significant. All examiners in our study had a 15 minute break between each set of four participants. Each data collection session lasted between 2-3 hours. Both the student osteopaths and experienced osteopaths would be use to concentrating for a 2-3 hour period in the student practice and their own clinical practices respectively.

7.4.7. PRE-WARM UP

The protocol of 10 examination procedures was performed by all four examiners on each participant; therefore each participant was examined four times. Although no literature exists on the topic on the effects of using a repetitive evaluation protocol it was proposed that changes in tissue length may occur during the initial evaluation protocol as the tissues maybe warmed up and thus inadvertently affect examiner findings. To minimize these effects the investigator considered that pre-stretching the evaluation areas using exercises that closely mimicked movements used in the examination would aid in maximizing the length of the tissues to their functional limitations, and thus maximize accuracy and reliability. The effect of warm-up on repetitive movement has not, to our knowledge been investigated to date. It may however influence the outcome of physical assessment. Its influence, if any to date is unknown.

7.4.8. RANDOMIZATION OF EXAMINERS

Although each set of four subjects were randomly allocated to either one of four rooms, the order in which the examiners examined each set of four subjects was kept the same. Examiner A always followed C; C followed B; B followed D and D followed A in a clockwise direction. This order was kept the same throughout the data collection process to simplify the logistics. If there was an order effect it would be consistent throughout the entire data collection process.
CHAPTER 8: RECOMMENDATIONS

This study has identified three discomfort components from each of three examination procedures with acceptable reliability (adjusted kappa) and with small positive likelihood ratios of a discomfort finding generated from one of these three examination procedures being related to asthma. To enable the findings of this study to be more generalisable this study would need to be replicated using a more representative sample of practitioners.

Our study was underpowered to draw any definitive conclusions on whether asthmatics with objective findings of lung obstruction exhibited a greater number of SD. Further research may investigate this relationship by using a population of asthmatics with only objective findings of lung obstruction.

Further studies should be undertaken to investigate the diagnostic validity of each physical examination procedure by confirming the musculoskeletal structures they are hypothesized to assess.

Future studies could also investigate the association between SD and the degree of severity and length of time the asthma is present (chronicity) as this may influence the extent to which SD is manifested.
CHAPTER 9: CONCLUSION

In the osteopathic literature there are several descriptions of SD being related to asthma, including but not limited to, palpable hypertonicity in the sternocleidomastoid and scalene muscles (upper chest muscles), thoracic immobility (DiGiovanna & Schiowitz, 1991), thoracic paraspinal muscle rigidity, tension and tenderness (Ward, 2003).

Somatic dysfunction associated with asthma are thought to arise from increased respiratory effort that is required during asthmatic exacerbations. When this effort is substantial or prolonged, the coordinated activity of neural reflexes, muscular contractions and the joints involved in respiration may be lost, and a change in breathing pattern may occur, with increased and prolonged recruitment of the accessory respiration muscles (upper chest muscles) resulting in a faulty ‘upper chest’ breathing pattern (Courtney, 2009). As a result of this faulty breathing pattern, greater stresses are imposed on the musculoskeletal elements involved in respiration. When these stresses extend beyond the normal capabilities of these musculoskeletal elements, resultant modifications occur to the tissues that produce the mechanical forces of respiration (Chaitow et al., 2002; Denslow, 1964; Glossary of osteopathic terminology, 2006; Sammut & Searle-Barnes, 1998). These structural modifications are what osteopaths term ‘Somatic Dysfunction’.

Evidence for the existence of SD in asthmatics is currently very limited with only one unpublished thesis by Russell (2006) reporting an association between SD and asthma.

The current study failed to show a clinically useful association between asthma and all ten examination procedures. Our findings do not support previous claims described in textbooks (Chaitow et al., 2002; DiGiovanna & Schiowitz, 1991; Kuchera & Kuchera, 1994; Sammut & Searle-Barnes, 1998) of an association between SD and asthma.
This study did however, identify three discomfort components from each of three examination procedures performed by the experienced osteopaths that detected with acceptable reliability and with small positive likelihood ratios discomfort in the pectoralis muscles, trapezius muscle and sternocleidomastoid muscles in subjects with asthma. The three examination procedures are listed below:

- Chest expansion (discomfort)
- Trapezius stretch (discomfort)
- Sternocleidomastoid tissue texture (discomfort)

However, when using the same examination procedures by student osteopaths in this study, the association was not present. Further research is required to confirm the validity of these tests.

Our study also demonstrates that examination procedures that assess for discomfort are more reliable than other physical examination procedures. This finding is consistent with other research on the reliability of physical examination procedures.

Our study also identified that osteopathic students were more reliable than their more experienced counterparts. These results are contrary to the presumption that experienced osteopaths would have more advanced palpatory skills, and thus greater reliability.

Although, underpowered to draw any definitive conclusions, this study identified a greater association between somatic dysfunction and asthmatics with objective findings of lung obstruction. This relationship should be the subject of further research.


