Intra-rater reliability of measuring diaphragm thickness utilising ultrasound imaging by a non-sonographer practitioner

Ben Giles

A 90-credit research thesis submitted in partial fulfilment of the requirements for the degree of Master of Osteopathy at Unitec Institute of Technology, 2018

Declaration

Name of candidate: Ben Giles

This thesis entitled "*Intra-rater reliability of measuring diaphragm thickness utilising ultrasound imaging by a non-sonographer practitioner*" is submitted in partial fulfillment for the requirements for the Unitec degree of Master of Osteopathy

Candidate's declaration

I confirm that:

- This thesis 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 Policies and Procedures, and has fulfilled any requirements set for this project by the Unitec Research Ethics Committee

Research Ethics Committee Approval Number: 2018-1111

Candidate Signature: …………………………………………………..Date: 09/10/2018

Student number: 1414088

Acknowledgments

First and foremost, I would like to thank my family and friends for all the love and support I received throughout my years of study. Thank you all for your words of encouragement. A special thanks to Courtney for your unwavering support and patience!

I am grateful to my supervisor Rob Moran. It has been inspiring working with you. I am so thankful to have hade your wisdom and guidance throughout the course of this project.

Thank you to all the participants and helping hands involved in this research. Particularly Scott Allen, this study would not have been possible without you.

Last but not least I would like to give a big thanks to Jacquelyn Schirmer. You have been an immense source of support and determination in helping me complete this research. It has always been a huge inspiration seeing you work. Finally, I can join you!

Introduction to Thesis

The diaphragm is well known for its primary role as the principal muscle in normal ventilation or quiet breathing. Ventilation is the physical movement of a volume of gas into and out of the lungs, whereas respiration is specific to the gas exchange of oxygen and carbon dioxide across a membrane at a cellular level. When the diaphragm is not functioning effectively or efficiently, ventilation is always compromised. This is due to accessory muscles carrying out the role of ventilation with much less efficiency (Harper et al., 2013). It is thought that dysfunction of the diaphragm is an underappreciated cause of respiratory issues and bears substantial impact not just on an individual level but also carries socioeconomic burden (Courtney & Greenwood, 2009). Additionally, it is widely accepted that diaphragm dysfunction is a catalyst for a range of health complications in several other body systems (Courtney, 2009). Due to the important role the diaphragm plays in health, there have been many methods to try and measure the diaphragm in living humans, including manual assessment (Ludwig, 2013), magnetic resonance imaging (Kolář et al., 2009; Kotani et al., 2004), and sonography (Boon et al., 2013, Harper et al., 2013). This study reported in this thesis involves the use of sonography as a measure of diaphragm structure (muscle thickness), and function via consideration of contractility.

The first stage in assessing any structure is establishing sufficient reliability of the measuring tool, in this case ultrasound. The use of ultrasound imaging (USI) as a measurement tool has been extensively studied from the view of an experienced operator (Harper et al., 2013). However, to our knowledge reliability data has not been gathered for non-sonographers or less experienced operators in the context of measuring diaphragm thickness. As the efficacy of diagnostic ultrasound is very user dependent, the same high reliability statistics may not be extrapolated to this group of operators.

Overview of the different modes of ultrasound and the gap in research

There are two different modes of ultrasound that can be used for assessing the diaphragm, these are M-mode (DiNino et al., 2014) and B-mode (Boon et al., 2013). M-mode is a time motion display of the ultrasound wave, while B-mode or 'brightness mode', uses a linear array of transducers to scan through the body and is the most common mode of ultrasound for taking tissue thickness measurements. Several previous studies have used M-mode ultrasound to quantify diaphragm excursion and movement during tidal breathing (Brown, Tseng,

Mitchell & Ridley, 2018; Baria et al., 2014). There have also been several prior studies using B-mode ultrasound that have shown varying measurements of diaphragm muscle thickness at the zone of apposition. The zone of apposition is referred to as the point at which the diaphragm connects to the ribcage. The magnitude of muscle thickness at the zone of apposition is in the range of 1.7-2.2 mm (Gerscovich et al., 2001; McCool & Tzelepis, 2012). However, these studies had several limitations including insufficient sample size, unclear methodology, no 'healthy' participant identification criteria, and lack of stratification for age and gender. There have been two studies using B-mode ultrasound with large sample sizes that were used as the basis for the methods reported in the study reported in this thesis (Boon et al., 2013; Harper et al., 2013).

A growing field within ultrasound is rehabilitative ultrasound imaging (RUSI). RUSI is designed to aid the rehabilitation of musculoskeletal disorders. It is evident, however, that there is no research that investigates the reliability of diaphragm measurements from the view of a non-sonographer. To explain further, there is an ever-growing number of healthcare practitioners wanting to add sonography as an adjunct to their already existing practice, as is seen in the increase in demand for RUSI (Kiesel et al., 2007; Teyhen & Koppenhaver, 2011).

There is an evident gap in the literature, investigating reliability of diaphragm thickness measurements by a less experienced USI operator such as rehabilitation providers and manual therapist who wish to use these machines in their practice. To address this, the primary aim of the study reported in Section II of this thesis was, 'To investigate the intra-rater reliability of diaphragm thickness measurements utilising a USI when examined by a non-sonographer. A secondary aim was to establish whether diaphragm thickness in particular 'contractility', could be a measure of self-perceived breathing quality.

The layout of this thesis is comprised of two main sections. Section 1, consists of a review of the literature, a detailed explanation of the anatomy of the diaphragm and its fascial connection. Further, different types of ultrasound will be reviewed, including an overview of the quality appraisal tools that form the methodology of a reliability study. Section 2, is the manuscript, where the methodology of this study is described. This section will provide the results of this study including a discussion of how they may be interpreted and applied in the real-world.

SECTION I – Literature Review

Literature Review

Introduction to Literature Review: Overview

In this review of the literature the relevant anatomy and physiology of the diaphragm and surrounding structures will be outlined, followed by an overview of the effects of ventilation on other body systems. Additionally, other relevant background information pertaining to this research will be discussed, such as, sonography applications including different variations and modes of ultrasound. This literature review will also cover what is already known about diaphragm contractility and how the thickness of diaphragm muscle changes between inspiration and expiration.

The later part of the review will include recent literature on key studies reporting reliability of diaphragm thickness measurements and how they formed their methodology. This will then serve as a basis for the design for this study.

Later, a review of ultrasound methodology and utility for the measurement of diaphragm thickness will be undertaken. Lastly, this literature review will include a discussion of quality appraisal tools that can be utilised to make reliability studies more robust. A review and appraisal of reliability research for the use of ultrasound methodology for both measurement of diaphragm thickness and blinding procedures, provides justification for the methods undertaken within reported study.

Part A: Background

The aim of this section is to give detailed descriptions of the anatomy of the diaphragm and the surrounding musculature including the relevant fascial systems. Additionally, we aim to give insight into the functions carried out by the diaphragm and its importance in maintaining homeostasis within multiple body systems. Subsequently, complications that arise when the

diaphragm is compromised will also be discussed. Finally, different methods of measuring the diaphragm, specifically in the field of USI will be reviewed.

1.2 Anatomy of the diaphragm muscle and its fascial connections

1.2.1 Diaphragm

The diaphragm is a dome-shaped musculotendinous structure which separates the thoracic and the abdominal cavities (Downey, 2011). The attachments of the diaphragm can be conveniently considered in three main parts; the sternal, costal, and lumbar portions.
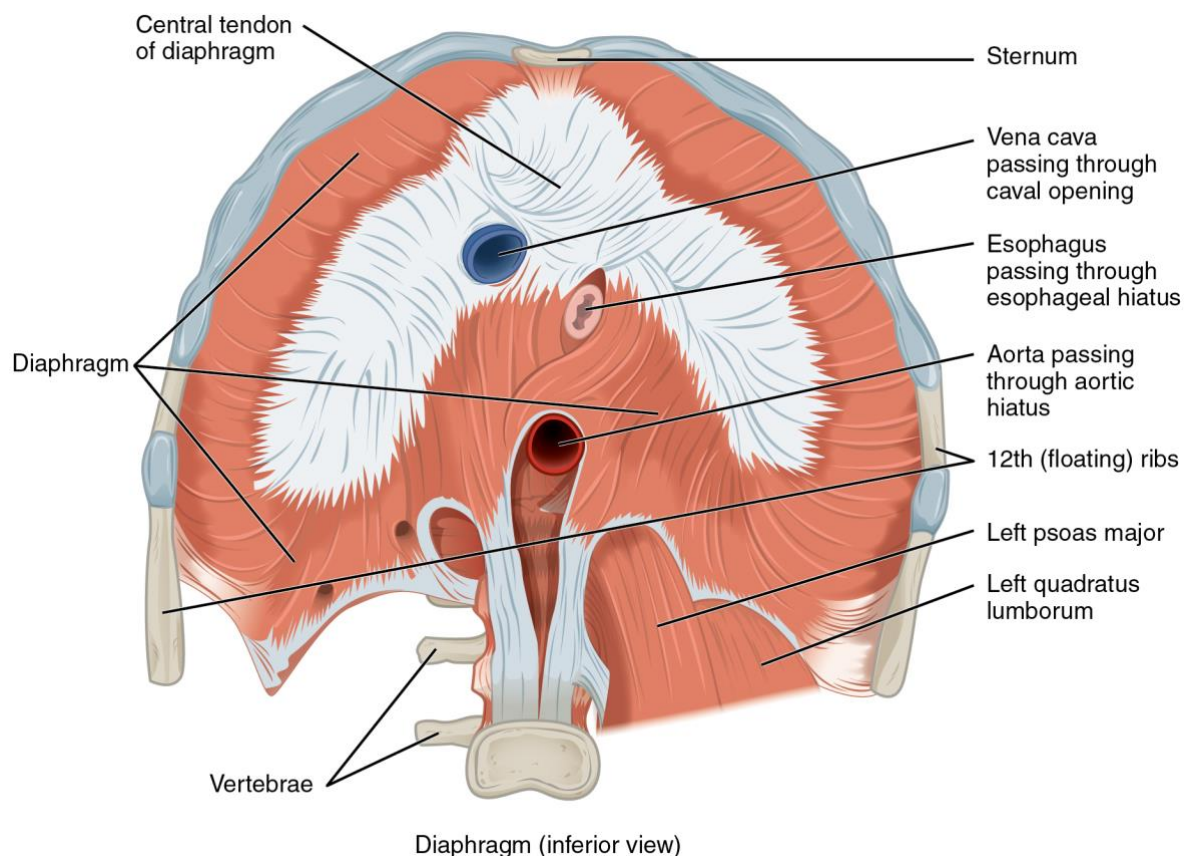

Figure 1: Inferior view of the diaphragm, showing its attachments and the structures that pass through it (Source: Wikimedia Commons https://commons.wikimedia.org/wiki/File:1113_The_Diaphragm.jpg).

1.2.2 Diaphragm – sternal portion

The sternal portion is made up of two small muscle bundles, which insert at the back of the xiphoid process near the apex creating two irregular openings in the retrosternal space known as, "the hiatus of Morgagni and Larrey"(Arráez-Aybar, González-Gómez, & Torres-García, 2009). These openings are significant not only because this is an area where the diaphragm allows communication between the thoracic and abdominal cavities, in the sense that it

creates a window in which a transfer of forces and pressure can be transferred between the two cavities, but also where anteromedial and retrocostoxiphoid hernias can occur (Bordoni & Zanier, 2013).

1.2.3 Diaphragm – costal portion
The costal portion covers the most lateral aspect of the diaphragm, which originates on the inner and superior border of the lower six ribs including their cartilages, via individual interdigitations between the transversus abdominis muscle (Bordoni, Marelli, Morabito, & Sacconi, 2016).

1.2.4 Diaphragm – lumbar portion
The lumbar portion of the diaphragm derives from two crura which attach the diaphragm posteriorly to the lower thoracic and upper lumbar spine. The right and left crural arches are made up of three ligaments known as arcuate ligaments; medial, intermediate and lateral ligaments (Bordoni et al., 2016). The medial ligament bifurcates at the level of T11 creating the esophageal hiatus, carrying both the esophagus and vagus nerves, this arch is also called the median arcuate ligament. The thicker and longer right medial ligament, terminates on the anterior side of L2-L3 and occasionally to L4 as a broad tendon. Just lateral to this ligament is another smaller ligament known as the accessory medial ligament, which inserts between L1-L2 (Bordoni et al., 2016).

1.2.5 Ligaments of the diaphragm and surrounding neurovascular structures
The greater thoracic splanchnic nerve, and at times the lesser thoracic splanchnic nerve, pass through the small triangular opening created between the right medial and accessory medial ligaments, where they bend sharply to join the celiac ganglion (Gest & Hildebrandt, 2009). There is a similar situation on the other side of the abdominal aorta, where the left medial ligament terminates on the anterior bodies of L2-L3, it again has an accessory medial ligament which creates the passage for the greater and lesser thoracic splanchnic nerves on the left (Gest & Hildebrandt, 2009). The tendinous arc of the median arcuate ligament lies in front of T11 and is crossed anteriorly by the thoracic duct and aortic artery (Bordoni & Zanier, 2013). The lateral two ligaments also start here, before bifurcating into two separate tendons. The first being the medial arcuate ligament which arches over the psoas muscle and connects the body of the L1 vertebra to its transverse process. The lateral arcuate ligament inserts at the same transverse process before arching over the quadratus lumborum muscle

and terminating at the apex of rib twelve (Bordoni et al., 2016). All of the diaphragmatic musculature merges medially into a fibrous central tendon. The central tendon is a thin but strong aponeurotic sheet, which is typically described by most authors to resemble a 'cloverleaf' or 'trefoil' shape with the three leaves decreasing in size from right to left (Anraku & Shargall, 2009; Downey, 2011), and more recently described by du Plessis et al., (2015) as being more 'V-shaped' in most people.

1.2.6 Innervation of the diaphragm

The innervation of the diaphragm is provided predominantly by the phrenic nerve which provides both sensory and motor function, however, the diaphragm has additional contributions from the vagus nerve (cranial nerve X) (Bordoni et al., 2016; Nason et al., 2012). The phrenic nerve originates from the nerve roots of C3-C5 and is located posteriorly in the lateral compartment of the neck then transverses anteriorly into the thorax before running along the anterior surface of the pericardium until it reaches the diaphragm (Nason et al., 2012).

1.3 Fascia definition

Fascia is a sheet-like connective tissue that covers almost the whole body (Willard et al., 2014). It provides many functions and provides significant peripheral information including, transferring load, encapsulating structures and providing both a nociceptive and proprioceptive role (Willard et al., 2014). There are many fascial systems within the body which provide a connection between two or more structures, creating a network between them. The fascial systems directly relevant to the diaphragm are the interfascial plane, the transversalis fascia, the thoracolumbar fascia and finally the lateral raphe.

1.3.1 Fascial connections of the diaphragm; the interfascial plane & transversalis fascia

The fascial system known as the interfascial plane starts at a retroperitoneal level, it connects the inferior vena cava, aortic system, psoas muscles, quadratus lumborum muscle, phrenic-esophageal ligaments, liver and kidneys (Lee, Ku, & Rha, 2010). Laterally, another important fascial system worth considering is the transversalis fascia (Peiper, Junge, Prescher, Stumpf, & Schumpelick, 2004). This fascial system is a continuation of the endothoracic fascia and spans from the deep cervical fascia of the neck down to the pubis. Along its pathway it covers the transverse abdominis muscle, the edge of rectus abdominis, external oblique and

the inguinal canal but is also related to the pleura, pericardium and diaphragm (Bordoni & Zanier, 2013).

1.3.2 Fascial connections of the diaphragm; thoracolumbar fascia & lateral raphe

Another important fascial system is the thoracolumbar fascia, whose fascial planes unite together around the paraspinal muscles and include the trapezius muscle, latissimus dorsi, external oblique and gluteus maximus (Willard, Vleeming, Schuenke, Danneels, & Schleip, 2012). The main function of this thick sheet of fascia which extends from the cervical region through to the sacral region is to stabilise the lumbosacral spine, posteriorly. Again, diaphragm dysfunction will negatively affect this fascia by altering the length of it through the contracting parts of the diaphragm, which in turn alters the force developed during contraction as well as the force transference across the fascial plane. This is where a change in the tension will ultimately affect its length and vice-versa and therefore also has an impact on the column of muscle being supported by it (Bordoni & Zanier, 2013). This has the potential to lead to both central and peripheral symptoms, such is seen in diaphragmatic causes of cervical pain (Bain & Harrington, 1983). Lastly, another part of the fascial system that interacts with the diaphragm is the lateral raphe. This fascial sling extends from the iliac crest up to the 12th rib and has two vital functions, it dissipates the tension created by the abdominal myofascial girdle across the thoracolumbar fascia and supports the contact of the quadratus lumborum muscle onto the transverse process of L2 (Schuenke, Vleeming, Van Hoof, & Willard, 2012).

1.3.3 Pathological states of fascia

Although the role fascia might play in creating a pathological environment is still unclear, it is assumed that when muscles are not able to slide freely within the encapsulating fascia, problems with the contractile pressures between the diaphragm and other trunk and perineal muscles arise, thereby, playing an indirect role in creating an abnormal physiological state (Bordoni & Zanier, 2013).

1.4 Function of the diaphragm

The diaphragm has multiple and far reaching functions, with its primary role being the main muscle of ventilation. When the diaphragm contracts, its dome flattens and central tendon distends, which in turn draws the thoracic cavity inferiorly causing an increase in thoracic volume and decrease in intra-thoracic pressure (Merrell & Kardon, 2013). In accordance with Boyle's law, which states that the pressure of an ideal gas is inversely proportional to its volume ($P_1V_1 = P_2V_2$) where P is the gas pressure and V the volume (Quanjer et al., 1993), the decrease in intra-thoracic pressure causes air to move into the lungs from the surrounding atmosphere, allowing gas exchange to occur. The pressure difference that the contracting diaphragm creates between the thoracic and abdominal cavities also provides a pump-like function that facilitates blood flow back to the heart and creates a negative pressure immediately inferior to it which provides support for abdominal organs and helps them stay in place (Hodges et al., 2005).

1.4.1 Secondary roles of the diaphragm

Apart from being the primary muscle of ventilation, the diaphragm also has other important roles. Specifically, the diaphragm has a role in posture by providing spinal stability through increasing intra-abdominal pressure (Bordoni et al., 2016). When the diaphragm contracts and descends, it exerts pressure down upon the pelvic floor and abdominals, increasing the pressure in the abdominal cavity and therefore opposing lumbar spine extension as well as intervertebral rotation and translation (Hodges et al., 2005).

Not only does the diaphragm have a significant role in posture but also provides an anti-reflux barrier by briefly ceasing contraction in order to allow the bolus to enter the stomach during swallowing but resisting stomach contents from going the other way, similar role to a one way valve (Pickering & Jones, 2002).

In addition, the diaphragm is also directly involved in the mechanics of emesis by increasing gastric pressure in the retching phase and diverging around the oesophageal sphincter in the expulsive phase of vomiting (Pickering & Jones, 2002).

1.5 Breathing Dysfunction and Diaphragm Thickness

Breathing is suggested to be dysfunctional when it is insufficient for adapting to environmental conditions and changing requirements of the individual (Courtney, 2009). There are many possible causes of dysfunctional breathing which include musculoskeletal dysfunction, presence of co-morbidity, chronic psychological stress and other factors that affect respiratory drive and control (Courtney, 2009). Two of the most common dysfunctional breathing patterns are hyperventilation and paradoxical breathing.

1.5.1 Hyperventilation

Hyperventilation is defined as respiration that exceeds metabolic demands, resulting in a decrease in arterial partial pressure of carbon dioxide (pCO_2) and an increase in pH of the body fluids (respiratory alkalosis) (Hornsveld et al., 1996). If sustained for long enough, symptoms such as paraesthesis, trembling and dizziness ensue. This is often classified as hyperventilation syndrome (Lewis & Howell, 1986). There are certain diagnostic criteria developed to test for the presence of hyperventilation such as the hyperventilation provocation test (Hornsveld et al., 1996). This test works on the basis of voluntarily hyperventilating for several minutes and is considered positive if the induced symptoms are similar to those experienced in daily-life (Hornsveld & Garssen, 1997). Regardless of the efficacy of the test to diagnose hyperventilation syndrome, hyperventilation is a widespread health complication and has potential direct effects on the diaphragm muscle due to misuse.

1.5.2 Paradoxical breathing

Paradoxical breathing pattern is a reversal of a normal breathing pattern such that during inspiration the chest contracts and during exhalation the chest expands. Therefore, paradoxical breathing is a feature of impaired co-ordination between the chest wall and diaphragm. These complications are often a result of environmental factors but can also be a result of chest or rib deformities.

1.5.3 Ventilated patients

Mechanical ventilation is the act of assisting or replacing a person's spontaneous breathing through artificial ventilation. For obvious reasons, an individual's diaphragm becomes instantly less active as its role is now replaced by a machine. There is much debate on the

lasting impact a ventilator has on diaphragm muscle atrophy and thus thickness. One study by Goligher and colleagues (2015), tried to measure the evolution of the diaphragm thickness during mechanical ventilation. Their results showed that diaphragm thickness decreases rapidly even in the first few days after mechanical ventilation in more than 40% of participants (Goligher et al., 2015).

1.5.4 Normal thickness

There is no general consensus as to what the 'normal thickness' of the diaphragm is, or should be, within certain demographics. One study took diaphragm thickness measurements of 80 healthy subjects and found that the resting end expiration was 0.193 ± 0.044cm on the right side compared with 0.187 ± 0.039 cm on the left (Seok et al., 2017).

1.5.5 Difference in diaphragm thickness and link to dysfunctional breathing

A change in diaphragm thickness has been used as a measure of respiratory weakness and a sign of pathological change (Goligher et al., 2015; Ottenheijm, Heunks, & Dekhuijzen, 2008). A thinning or atrophy of the diaphragm can be seen in patients with congenital illnesses such as muscular dystrophy, or neurological deficits such as phrenic nerve palsies, and even chronic disuse of the diaphragm as seen in mechanically ventilated patients (Goligher et al., 2015). In the case of ventilated patients, much of this decrease in diaphragm thickness is predicted to be a result of lower levels of inspiratory effort and therefore, changes in muscle configuration and loss of density (Goligher et al., 2015). When the diaphragm contracts it thickens as does any muscle when activated, therefore, the diaphragm is at its thickest state at full inhalation and thinnest during the relaxation of expiration. This thickening ratio between full inhalation and full exhalation is therefore a measure of diaphragm activity in breathing, and may provide another means of establishing dysfunctional breathing due to underuse of the diaphragm. A recent study has also described how body position (posture), influences diaphragm thickness (Hellyer et al., 2017). Using B-mode ultrasound, Hellyer and colleagues (2017) measured diaphragm thickness in three different position; standing, sitting and supine. They found that the diaphragm is thicker when the body is more upright (standing and sitting versus supine). It is hypothesized this difference may be due to greater vertical gravitational load on the diaphragm and therefore, results in changes in the resting length of the muscle fibres (Hellyer et al., 2017). There is also extensive literature that a link between breathing dysfunction and diaphragm thickness

exists. A systematic review of 875 critically ill patients investigated the effectiveness of assessing diaphragmatic dysfunction utilising USI in these patients (Zambon et al., 2017). This systematic review also tries to establish, 'optimal cutoffs' for a healthy functional diaphragm in terms of diaphragm thickness measurements and contractility. These so called optimal cutoffs ranged from 10 to 14mm for excursion thickness and 30-36% for thickening percentage (contractility) (Zambon et al., 2017).

1.6 Sonography Applications

1.6.1 Ultrasound overview

Ultrasound imaging is considered a safe and cost-effective method of medical imaging (Reißig & Kroegel, 2005). The advantage of ultrasound as an imaging technique is that it provides real-time in vivo feedback. Observing tissue actions and movements in real-time, gives the USI operator direct feedback which can aid diagnosis, therapeutic management plan and track rehabilitation.

1.6.2 Ultrasound adverse effects

Ultrasound does not involve any exposure to ionizing radiation and is therefore generally considered as a safe form of imaging. Historically, there has been some discussion of the effect of USI cells at a biological level, but this has been difficult to assess due in *vitro* testing methods (ter Haar, 2015). Additionally, normal clinical levels of USI exposure are considerably less than exposure levels utilized in research to determine whether sonography has any effect on a cellular level. Another potential adverse effect that can arise with USI is excess heat exposure which is known as the beam's thermal bioeffects. However, modern ultrasound machines all display a thermal index (TI) when scanning, which notify the operator well before any potential tissue damage occurs (Nelson et al., 2009). USI exposure within a clinical medical setting is widely accepted as being safe (Reißig & Kroegel, 2005; Nomura & Nagdev, 2018).

1.6.3 M-mode vs B-mode ultrasound

Diagnostic ultrasound (also known as diagnostic sonography or ultrasonography) is a safe and cost effective method to see through many body layers in vivo and visualise what is occurring beneath the skin (Normura & Nagdev, 2018). The additional advantage of sonography over radiology is the ability to see soft tissue structures in real-time (Normura &

Nagdev, 2018). This is crucially important when measuring such a dynamic structure as the diaphragm. Sonographic evaluation of the diaphragm allows the user to objectively measure diaphragm thickness and the thickness ratio between inhalation and exhalation with a high degree of inter- and intra-rater reliability. There are two different methods of ultrasound available which are useful in different scenarios; B-mode (brightness mode) and M-mode (motion mode). M-mode ultrasound is used to pick up the amplitude and velocity of a specific organ or structures by taking several images in quick succession, comparable to creating a video in ultrasound. This imaging type has been used for imaging the diaphragm in the past but it has been primarily used in cardiac examination (Menegatti et al., 2014). With B-mode ultrasound an array of transducers simultaneously scan a body part which is displayed as a two-dimensional object. This allows better visualisation of the target organ and the ability to measure thickness of it accurately and therefore is the sonography method chosen in this study.

1.6.4 Previous sonographic evaluation of the diaphragm

Harper et al. (2013) set out to establish normal values of diaphragm contractility as assessed with B-mode ultrasound. Two examiners took diaphragm measurements from 150 healthy participants. Out of the 150 participants, 12 were used for inter-rater reliability and 10 for intra-rater reliability. One of the two examiners was an experienced sonographer with several years of measuring both normal and abnormal diaphragms. The experienced sonographer trained the other examiner for several weeks prior to the study. How many hours was put into training was not reported. The examiners took separate sets of images from participants on two separate days to establish both intra-rater and inter-rater reliability, however, the amount of time between re-test measurements was not reported. The diaphragm was identified by its typical 3-layered appearance, three images were then taken at the end of quiet expiration and quiet inspiration. The three measurements in each position were then used to calculate mean thickness measurements. A thickening ratio was then derived by dividing mean inspiration thickness measurements by mean expiration thickness measurements. Diaphragm thickness measurements ranged from 0.12 to 1.18 cm, with men having a slightly higher mean resting thickness. Inter-rater and intra-rater reliability were found to be very high. Inter-rater reliability ICCs were 0.97 (95% CI: 0.91-0.99) for inspiratory thickness measurements and 0.98 (95% CI: 0.94-0.99) for expiratory thickness. Intra-rater reliability ICCs were 0.94 (95%

CI: 0.79-0.98) for inspiration thickness and 0.98 (95% CI: 0.94-0.99) for thickness at expiration.

The sonography procedures used to measure diaphragm thickness in this study will closely follow the same procedure used by Boon et al (2013). Firstly, using real time B-mode ultrasound the intercostal space where the diaphragm was most easily visible was determined, the transducer covering two ribs. The diaphragm was identified by its characteristic 3-layered appearance deep to the intercostal muscles (the diaphragm is seen as a hypoechoic (dark) structure between two hyperechoic (bright) lines of pleural and peritoneal fascia). On observation, the hypoechoic muscle increased in size during the inspiratory phase. Once the most appropriate intercostal space to see the diaphragm was located on each individual, the subject was instructed to breathe quietly while three images were taken at end of quiet expiration. The subject was then instructed to breath slow and deeply as three more images where taken when the diaphragm was at maximal thickness (as identified visually by the operator or at the point at which visualisation of the diaphragm became obscured by the lung). Electronic calipers (a function of the ultrasound scanner) were used to measure the diaphragm thickness at the point where the two hyperechoic lines outlining the diaphragm were parallel. The three images for each position were then averaged to give a diaphragm thickness at end of quiet expiration and a full maximal inspiration. Boon et al (2013) measured inter-rater reliability by having two different examiners take measurements of the same subject and intra-rater reliability by having the same examiner measure the same subject on two different days.

1.7.5 Ultrasound of diaphragm: reliability.
The investigation of reliability is important within quantitative research and indicates a sufficient degree of replicability of research results using the same measurement procedures. Reliability is commonly assessed using correlation indices, indicating the 'noise' or error in the measurement (Davidson et al., 2014). Kirk et al (1986) first subcategorized reliability into three different types; (1) the degree to which any given measurement remains the same when repeated; (2) how stable a particular measurement is over time; (3) the likeness of measurements within a particular time frame.

Boon et al. (2013) investigated quantitative diaphragm thickness values in healthy subjects obtained from USI, including inter-rater and intra-rater reliability for 12 and 10 subjects,

respectively. Subjects were examined in a supine position; real time ultrasound was used to identify the intercostal space at which the diaphragm was most visible.

Baldwin et al. (2011) investigated the reliability of diaphragm and peripheral muscle thickness utilising USI, where 13(6 men and 7 women) healthy volunteers participated. Diaphragm thickness was measured on the right hemi-diaphragm in the zone of apposition found at the mid-axillary line at the level of the 9th intercostal space. Diaphragm thickness was measured from three images captured during one stable tidal breath. These three images were taken from; (1) breath hold with open glottis during expiration, (2) at 25% of maximal inspiratory capacity and (3) at 50% of maximal inspiratory capacity. To mitigate re-test differences of an individual's inspiratory capacity, spirometry measurements were used to assess lung volume of participants and calculated percentage of each participants' inspiratory capacity by the average volume displacement of a single breath hold. In addition, participants were also asked to rate the level of difficulty associated with each breath maneuver as measured using a 100-mm visual analogue scale. Participant underwent re-test measurement session between 2 hours and 2 days after the initial session. Intra-rater reliability ICCs were incredibly high with expiration thickness measurement ICCs at 0.990 (95% CI: 0.918-0.998) and measurements taken at 25% and 50% inspiratory capacity, was ICC 0.959 (95% CI: 0.870-0.988) and ICC 0.994 (95% CI: 0.980-0.998) respectively.

Designing Reliability Studies using Quality Appraisal Tools

When designing the methodology for a reliability study, there are several quality appraisal tools and reporting checklists that provide guidance. The quality appraisal tool for studies of diagnostic reliability (QAREL) framework (Lucas et al, 2010) identifies 11 items that should be considered in order to control for bias (see Table 1).

Table 1. Quality Appraisal of Diagnostic Reliability (QAREL) Checklist (Lucas et al, 2010)	
Item	Description
1	Was the test evaluated in a sample of subjects who were representative of those to whom the authors intended the results to be applied?
2	Was the test performed by raters who were representative of those to whom the authors intended the results to be applied?
3	Were raters blinded to the findings of other raters during the study?
4	Were raters blinded to their own prior findings of the test under evaluation?
5	Were raters blinded to the results of the accepted reference standard or disease status for the target disorder (or variable) being evaluated?
6	Were raters blinded to clinical information that was not intended to be provided as part of the testing procedure or study design?
7	Were raters blinded to additional cues that were not part of the test?
8	Was the order of examination varied?
9	Was the stability (or theoretical stability) of the variable being measured taken into account when determining the suitability of the time-interval between repeated measures?
10	Was the test applied correctly and interpreted appropriately?
11	Were appropriate statistical measures of agreement used?

The QAREL items in the table above will be discussed in the context of designing the reliability study that is reported in Section II. Some of the QAREL items are purposefully omitted from further discussion as they are not deemed relevant for this specific study (e.g. Item 3, as there will be only one rater in this study).

Was the test evaluated in a sample of subjects who were representative of those to whom the authors intended the results to be applied? (QAREL Item 1)

Firstly, when considering the sample of participants in a reliability study, investigators should ensure that they establish eligibility criteria which generate a sample that is representative of the target population to whom the investigators intend the results to be applied. In order to generalise the findings widely, it is necessary to design eligibility criteria that selects for a wide spread of age, gender, and breathing quality. As a child's body morphology is considerably different and changing more frequently compared to adults it does not make sense to combine both in the same sample. When designing study methodology, it will be important to identify which age range would be preferred, as it directly determines how participants are recruited. The use of both the Self Evaluation of Breathing Questionnaire (Courtney & Greenwood, 2009) and Nijmegen Questionnaire (van Dixhoorn & Duivenvoorden, 1985) could act as a proxy measure of breathing quality within a sample.

Was the test performed by raters who were representative of those to whom the authors intended the results to be applied? (QAREL Item 2)

As this study hopes to inform manual therapy practitioners who may wish to measure diaphragm thickness in the context of therapy for breathing related disorders, the rater chosen for the study should also be a manual therapist with limited experience using diagnostic ultrasound. To try and mimic the amount of training a practitioner would likely encounter before acquiring an ultrasound machine, the level of training that would be appropriate would be around two days as it would be comparable to the amount of training a rehabilitation provider might receive after purchasing an ultrasound machine

Were raters blinded to their own prior findings of the test under evaluation? (QAREL Item 4)

In order to blind a rater to their own findings throughout data collection, a portion of the sonography screen showing the measurement needs to be concealed. This is often done by sticking a shield on a portion of the screen to hide the measurement being displayed. A screen shot function on the sonography machine allows recording of measurements while still being

blinded to them. The saved images containing the measurements can then be downloaded after each session for later analyses at the completion of data collection. Having a standardized protocol like this ensures proper blinding which if not in place could falsely increase reliability.

Were raters blinded to additional cues that were not part of the test? (QAREL Item 7)
In addition, a rater should be blinded to cues associated with subjects which might be able to link as an aide to memory between measurement sessions. As the skin needs to be exposed for ultrasound imaging, identifying cues such as scars and tattoos will be unavoidable. To try and minimize the effects of this, all participants should be encouraged to just expose the area over their lower ribs, without taking their whole top off.

Was the order of examination varied? (QAREL Item 8)
The order of subjects to be examined should be randomized in order to control for bias which could be introduced by order effects. If the order of examination is constant between sessions the rater has a higher chance remembering details of the first scan which could have been avoided. When booking participants into appointment times, care should be taken to seek a different time for the participant from the first session to the second if possible.

Was the stability (or theoretical stability) of the variable being measured taken into account when determining the suitability of the time-interval between repeated measures? (QAREL Item 9)
As the diaphragm is a muscular structure, the thickness measurement may potentially be influenced by the participant's activities between sessions. There appears to be no previous reliability studies measuring the diaphragm reporting any controls for this variable. There are two things to consider when planning interval time frames between sessions. Firstly, the interval should not be too long as skeletal muscle such as the diaphragm can grow with exercise and training which changes the thickness of the variable being measured. However, if the interval between session is too short, it may introduce recall bias on the raters' part. For example, in a sonography reliability study if a transducer is positioned in a certain spot in relation to a participants' identifiable feature such as a tattoo or mole, it might be possible to remember this in the second session if the interval between the sessions is too short. Secondly, in a real-life clinical setting, a practitioner would often choose one to two-week follow-up for most patients following an intervention. Because of these factors, a time of

between 7 and 14 days between sessions is often desired in reliability studies to minimize their potential effects on intra-rater reliability (Hides et al, 2007; Wilson et al, 2016).

Was the test applied correctly and interpreted appropriately? (QAREL Item 10)
In order to control that the test is applied correctly by the rater, an experienced sonographer should oversee trial scans prior to data collection. Only once the experienced sonographer is satisfied that the rater was using the ultrasound machine properly and applying the test correctly, should data collection commence.

Research purpose

The preliminary research conducted by Boon et al (2013) established quantitative values of diaphragm thickness in normal subjects, but it is not clear how operator expertise might impact intra-rater reliability of thickness measurements, this warrants further investigation. Additionally, there is a lack of research investigating the relationship between diaphragm thickness and dysfunctional breathing. It is evident that diaphragm atrophy is common amongst mechanically ventilated people, as underuse of the muscle causes thickness decline. It is a lot less clear if diaphragm thickness atrophy is prevalent within a broader range of breathing dysfunction and diaphragm 'mis-use'. Given the increasing use of RUSI by manual therapists with limited sonography experience, reliability research which represents an inexperienced operator will provide an insight into the usefulness and applicability of this imaging modality for these operators.

Considering the above, the research reported in Section II of this thesis has two aims:

1. Establish non-sonographer intra-rater reliability of diaphragm thickness measurements utilising USI.
2. Establish whether there is a correlation between measured diaphragm thickness and self-perceived breathing quality as calculated by the Self Evaluation of Breathing Questionnaire and Nijmegen Questionnaire.

References

Anraku, M., & Shargall, Y. (2009). Surgical Conditions of the Diaphragm: Anatomy and Physiology. *Thoracic Surgery Clinics*, 19(4), 419-29

Arráez-Aybar, L. a, González-Gómez, C. C., & Torres-García, a J. (2009). Morgagni-Larrey parasternal diaphragmatic hernia in the adult. *Revista Española de Enfermedades Digestivas: Organo Oficial de La Sociedad Española de Patología Digestiva*, *101*(5), 357–66.

Bain, W., & Harrington, J. (1983). Head and neck manifestations of gastroesophageal reflux. *The Laryngoscope*, 93(2), 175-179

Baria, M. R., Shahgholi, L., Sorenson, E. J., Harper, C. J., Lim, K. G., Strommen, J. A., ... & Boon, A. J. (2014). B-mode ultrasound assessment of diaphragm structure and function in patients with COPD. *Chest*, *146*(3), 680-685.

Bordoni, B., Marelli, F., Morabito, B., & Sacconi, B. (2016). Manual evaluation of the diaphragm muscle. *International Journal of Chronic Obstructive Pulmonary Disease*, *11*, 1949.

Bordoni, B., & Zanier, E. (2013). Anatomic connections of the diaphragm: Influence of respiration on the body system. *Journal of Multidisciplinary Healthcare*, *6,* 281.

Boon, A. J., Harper, C. J., Ghahfarokhi, L. S., Strommen, J. A., Watson, J. C., & Sorenson, E. J. (2013). Two-dimensional ultrasound imaging of the diaphragm: Quantitative values in normal subjects. *Muscle & Nerve*, *47*(6), 884-889.

Brown, C., Tseng, S. C., Mitchell, K., & Roddey, T. (2018). Body position affects ultrasonographic measurement of diaphragm contractility. *Cardiopulmonary Physical Therapy Journal*, *29*(4), 166.

Courtney, R. (2009). The functions of breathing and its dysfunctions and their relationship to breathing therapy. *International Journal of Osteopathic Medicine*, *12*(3), 78-85.

Courtney, R., Greenwood, K. M., & Cohen, M. (2011). Relationships between measures of dysfunctional breathing in a population with concerns about their breathing. *Journal of Bodywork and Movement Therapies*, *15*(1), 24–34.

Cretikos, M., Bellomo, R., Hillman, K., Chen, J., Finfer, S., & Flabouris, A. (2008). Respiratory rate: the neglected vital sign. Medical Journal of Australia. *Medical Journal of Australia*, 188(11), 657.

Davidson, M., & Keating, J. (2014). Patient-reported outcome measures (PROMs): how should I interpret reports of measurement properties? A practical guide for clinicians and researchers who are not biostatisticians. *British Journal Of Sports Medicine*, *48*(9), 792-796.

DiNino, E., Gartman, E. J., Sethi, J. M., & McCool, F. D. (2014). Diaphragm ultrasound as a predictor of successful extubation from mechanical ventilation. *Thorax*, *69*(5), 431-435.

Downey, R. (2011). Anatomy of the Normal Diaphragm. *Thoracic Surgery Clinics*, 21(2), 273-9.

du Plessis, M., Ramai, D., Shah, S., Holland, J. D., Tubbs, R. S., & Loukas, M. (2015). The clinical anatomy of the musculotendinous part of the diaphragm. *Surgical and Radiologic Anatomy*, *37*(9), 1013–1020.

Fieselmann, J. F., Hendryx, M. S., Helms, C. M., & Wakefield, D. S. (1993). Respiratory rate predicts cardiopulmonary arrest for internal medicine inpatients. *Journal of General Internal Medicine*, 8(7), 354-360.

Gerscovich, E. O., Cronan, M., McGahan, J. P., Jain, K., Jones, C. D., & McDonald, C. (2001). Ultrasonographic evaluation of diaphragmatic motion. *Journal of Ultrasound in Medicine*, *20*(6), 597-604.

Gest, T. R., & Hildebrandt, S. (2009). The pattern of the thoracic splanchnic nerves as they pass through the diaphragm. *Clinical Anatomy*, *22*(7), 809–814.

Goligher, E. C., Fan, E., Herridge, M. S., Murray, A., Vorona, S., Brace, D., & Bolz, S. S. (2015). Evolution of diaphragm thickness during mechanical ventilation. Impact of inspiratory effort. *American Journal of Respiratory and Critical Care Medicine*, *192*(9), 1080-1088.

Harper, C. J., Shahgholi, L., Cieslak, K., Hellyer, N. J., Strommen, J. A., & Boon, A. J. (2013). Variability in diaphragm motion during normal breathing, assessed with B-mode ultrasound. *Journal of Orthopaedic & Sports Physical Therapy*, *43*(12), 927-931.

Hellyer, N. J., Andreas, N. M., Bernstetter, A. S., Cieslak, K. R., Donahue, G. F., Steiner, E. A., & Boon, A. J. (2017). Comparison of diaphragm thickness measurements among postures via ultrasound imaging. *Physical Medicine And Rehabilitation*, *9*(1), 21-25.

Hides, J. A., Miokovic, T., Belavý, D. L., Stanton, W. R., & Richardson, C. A. (2007). Ultrasound imaging assessment of abdominal muscle function during drawing-in of the abdominal wall: an intrarater reliability study. *Journal of Orthopaedic & Sports Physical Therapy*, *37*(8), 480-486.

Hodges, P. W., Eriksson, A. E. M., Shirley, D., & Gandevia, S. C. (2005). Intra-abdominal pressure increases stiffness of the lumbar spine. *Journal of Biomechanics*, *38*(9), 1873–1880.

Hornsveld, H., & Garssen, B. (1997). Hyperventilation syndrome: an elegant but scientifically untenable concept. *The Netherlands Journal of Medicine*, *50*(1), 13-20.

Hornsveld, H. K., Garssen, B., Dop, M. F., Van Spiegel, P. I., & De Haes, J. C. J. M. (1996). Double-blind placebo-controlled study of the hyperventilation provocation test and the validity of the hyperventilation syndrome. *The Lancet*, *348*(9021), 154-158.

Kiesel, K. B., Uhl, T. L., Underwood, F. B., Rodd, D. W., & Nitz, A. J. (2007). Measurement of lumbar multifidus muscle contraction with rehabilitative ultrasound imaging. *Manual Therapy*, *12*(2), 161-166.

Kirk, J., Miller, M. L., & Miller, M. L. (1986). *Reliability and validity in qualitative research* (Vol. 1). Sage Publishers.

Kolář, P., Neuwirth, J., Šanda, J., Suchanek, V., Svata, Z., Volejnik, J., & Pivec, M. (2009). Analysis of diaphragm movement during tidal breathing and during its activation while breath holding using MRI synchronized with spirometry. *Physiological Research*, *58*(3).

Kotani, T., Minami, S., Takahashi, K., Isobe, K., Nakata, Y., Takaso, M., ... & Moriya, H. (2004). An analysis of chest wall and diaphragm motions in patients with idiopathic scoliosis using dynamic breathing MRI. *Spine*, *29*(3), 298-302.

Lee, S. L., Ku, Y. M., & Rha, S. E. (2010). Comprehensive reviews of the interfascial plane of the retroperitoneum: Normal anatomy and pathologic entities. *Emergency Radiology*, 17(1), 3.

Lewis, R. A., & Howell, J. B. (1986). Definition of the hyperventilation syndrome. *Bulletin Européen De Physiopathologie Respiratoire*, *22*(2), 201-205.

Ludwig, M. (2013). Inter and intra-rater reliability of the manual assessment of respiratory motion ('MARM' technique) in adults. Unitec Institite of Technology, Master thesis.

McCool, F. D., & Tzelepis, G. E. (2012). Dysfunction of the diaphragm. *New England Journal of Medicine*, *366*(10), 932-942.

Menegatti, E., Tessari, M., Gianesini, S., Elena Vannini, M., Sisini, F., & Zamboni, P. (2014). Human internal jugular valve M-mode ultrasound characterization. *Current Neurovascular Research*, *11*(2), 149-155.

Merrell, A. J., & Kardon, G. (2013). Development of the diaphragm–a skeletal muscle essential for mammalian respiration. *The FEBS Journal*, *280*(17), 4026-4035.

Nason, L. K., Walker, C. M., McNeeley, M. F., Burivong, W., Fligner, C. L., & Godwin, J. D. (2012). Imaging of the diaphragm: anatomy and function. *Radiographics*, *32*(2), E51-E70.

Nelson, T. R., Fowlkes, J. B., Abramowicz, J. S., & Church, C. C. (2009). Ultrasound biosafety considerations for the practicing sonographer and sonologist. *Journal of Ultrasound in Medicine*, *28*(2), 139-150.

Normura, J. T., & Nagdev, A. D. (2018). Ultrasound Safety and Infection Control. In *Ultrasound Program Management,* 243-267.

Ottenheijm, C. A. C., Heunks, L. M. A., & Dekhuijzen, R. P. N. (2008). Diaphragm adaptations in patients with COPD, *14*, 1–14.

Peiper, C., Junge, K., Prescher, A., Stumpf, M., & Schumpelick, V. (2004). Abdominal musculature and the transversalis fascia: an anatomical viewpoint. *Hernia*, *8*(4), 376-380.

Pickering, M., & Jones, J. F. (2002). The diaphragm: two physiological muscles in one. *Journal Of Anatomy*, *201*(4), 305-312.

Quanjer, P. H., Tammeling, G. J., Cotes, J. E., Pedersen, O. F., Peslin, R., & Yernault, J.-C. (1993). *Lung volumes and forced ventilatory flows*. *European Respiratory Journal* (Vol. 6).

Reißig, A., & Kroegel, C. (2005). Accuracy of transthoracic sonography in excluding post-interventional pneumothorax and hydropneumothorax: comparison to chest radiography. *European Journal of Radiology*, *53*(3), 463-470.

Schuenke, M. D., Vleeming, A., Van Hoof, T., & Willard, F. H. (2012). A description of the lumbar interfascial triangle and its relation with the lateral raphe: Anatomical constituents of load transfer through the lateral margin of the thoracolumbar fascia. *Journal of Anatomy*, 221(6), 568-576

Seok, J. I., Kim, S. Y., Walker, F. O., Kwak, S. G., & Kwon, D. H. (2017). Ultrasonographic findings of the normal diaphragm: thickness and contractility. *Annals of Clinical Neurophysiology*, *19*(2), 131-135.

ter Haar, G. (2015). Ultrasound bio-effects and safety considerations. In *Translational Neurosonology* (Vol. 36, pp. 23-30). Karger Publishers.

Teyhen, D., & Koppenhaver, S. (2011). Rehabilitative ultrasound imaging. *Journal of Physiotherapy*, *57*(3), 196.

Willard, F., Vleeming, a, Schuenke, M., Danneels, L., & Schleip, R. (2012). The thoracolumbar fascia: anatomy, function and clinical considerations. *Journal of Anatomy*, 221(6), 507-536

Wilson, A., Hides, J. A., Blizzard, L., Callisaya, M., Cooper, A., Srikanth, V. K., & Winzenberg, T. (2016). Measuring ultrasound images of abdominal and lumbar multifidus muscles in older adults: A reliability study. *Manual Therapy*, *23*, 114-119.

Zambon, M., Greco, M., Bocchino, S., Cabrini, L., Beccaria, P. F., & Zangrillo, A. (2017). Assessment of diaphragmatic dysfunction in the critically ill patient with ultrasound: a systematic review. *Intensive Care Medicine*, *43*(1), 29-38.

SECTION II – Manuscript

Reliability of diaphragm muscle thickness measurements using ultrasound imaging

Author:

Ben N Giles

Correspondence address:

Department of Community Development (Osteopathy)

Unitec Institute of Technology

Private Bag 92025

Auckland Mail Centre

Auckland 1142, NZ

Tel: +64 21 02216568

Email: beng3395@gmail.com

Table of Contents

Abbreviations

Ultrasound Imaging (USI), Self Evaluated Breathing Questionnaire (SEBQ), Nijmegen

Questionnaire (NQ)

Abstract

Reliability of diaphragm muscle thickness measurements using ultrasound imaging

Background: As the primary muscle of ventilation, the diaphragm has a pivotal role in breathing, and in maintaining homeostasis for all other body systems. There have been several previous studies assessing reliability of measuring diaphragm muscle thickness using ultrasound imaging (USI) by qualified sonographers, however, no previous study has investigated reliability of measurement by a rehabilitation practitioner. Additionally, no evaluation has been done to determine a link between diaphragm thickness and breathing measures.

Aims: (1) To investigate the reliability of diaphragm thickness measurements utilising USI by a non-sonographer ('novice operator'); and (2) To evaluate the correlation between diaphragm thickness measurements including contractility and dysfunctional breathing as measured by the Self Evaluated Breathing Questionnaire (SEBQ) and the Nijmegen Questionnaire (NQ)

Methods: High-resolution, B-mode ultrasound was utilised to execute a standardized protocol for repeated thickness measurements of diaphragm muscle. A convenience sample of 25 participants (13 males, 12 females, mean $\pm$ SD age = 27 $\pm$7.2 years, height 172.6 $\pm$ 9.3cm, body mass 79 $\pm$ 14.5 and a mean SEBQ and NQ score of 14.9 $\pm$ 11.2 and 13.3 $\pm$ 7.2, respectively) attended two sessions, separated by an interval of approximately 2 weeks. Intra-operator reliability was calculated for all thickness measurements obtained by the novice USI operator.

Results: The novice operator demonstrated 'very high' *intra*-operator reliability for diaphragm muscle thickness measurements during all stages of breathing except for maximal inspiration on the right (all other ICCs >0.8). However it was apparent that diaphragm contractility may not be an adequate measure of diaphragm function as measured by the SEBQ and NQ.

Conclusion: Within this study, the novice operator demonstrated acceptable reliability for diaphragm muscle thickness measurements using USI. The novice operator demonstrated very high *intra*-operator reliability for diaphragm measurements during quiet exhalation on both sides and at the level of maximal inspiration on the left.

Introduction

The thoracic diaphragm muscle is the primary muscle involved in the mechanics of breathing. Normal, 'quiet' breathing, otherwise known as 'diaphragmatic breathing', requires synchronized motion of the abdomen (diaphragm), lower and upper rib cages with little exertion required by the individual. The synergy between lower and upper rib cages can become disrupted, as seen in people who suffer from paradoxical breathing (i.e. in which diaphragm motion is opposite the normal direction during both inspiratory and expiratory cycles), and hyperventilation syndrome (Folgering, 1999). If the diaphragm is unable to function efficiently, ventilation issues may develop such as compromised gas exchange at the alveolar level due to insufficient lung expansion. In addition, other secondary muscles are required to contract more strongly and are therefore prone to fatigue to achieve the same level of chest wall expansion. Altered breathing mechanics can profoundly affect other body systems. For example, there appears to be an association between breathing anomalies, back pain (Smith et al., 2006), and neck pain (Kapreli, Vourazanis & Strimpakos, 2008). Due to the dual function of trunk muscles as both postural stabilisers and providing stability for breathing functions (Chaitow, 2004), it is theorised that dysfunction in either the spinal stabilisers or the diaphragm will negatively affect the other (Hodges et al., 2007), and may be related to back pain causation (Smith, Russell & Hodges, 2006).

It has been demonstrated that like any skeletal muscle, when underused the diaphragm will atrophy and decrease in mass (Schepens et al., 2015). Given the importance of the diaphragm in maintaining homeostasis of body systems, valid and reliable methods of assessment to evaluate diaphragm function are necessary. Ultrasound imaging (USI) of the diaphragm muscle is attractive as it allows real-time, point of care, dynamic visualisation of the muscle during breathing with low risk of harm and low cost. However, currently there are limited

reliability data for USI measurements of the diaphragm, and only a few studies reporting normative measures of diaphragm muscle thickness in 'normal' subjects (Baldwin, Paratz & Bersten, 2011; Boon et al., 2013; Harper et al., 2013). There is now an emergence of rehabilitation providers (eg physiotherapists, osteopaths etc) using rehabilitative ultrasound imaging (RUSI) to assess muscle function in clinical practice (Kiesel et al., 2007; Teyhen & Koppenhaver, 2011). To date, no reliability studies have investigated a rehabilitation provider, rather than a sonographer, in measurement of diaphragm thickness. Therefore, the aim of this study was to assess the intra-rater reliability of diaphragm thickness measurements when undertaken by a non-sonographer with only minimal training in USI. Additionally, a secondary aim was to explore the correlation between both diaphragm thickness and contractility, and dysfunctional breathing symptom questionnaire scores.

Methods

Design and ethics

A repeated measures, test-retest design was used to investigate the intra-operator reliability of ultrasound imaging measurements for thickness of diaphragm muscle. A novice ultrasound imaging operator completed a standardized measurement protocol. Measures were undertaken at two sessions, separated by an interval of approximately two weeks. All participants provided written informed consent and the study was approved by the institutional ethics committee (UREC 2018-1111).

Participants

Participants

A convenience sample of participants was recruited using word-of-mouth at the Unitec osteopathy teaching clinic. Inclusion criteria were: aged at least 18 years, and were able to provide informed consent. The only exclusion criterion was a history of abdominal or thoracic injuries/surgeries that made breathing either difficult or painful at the time of data collection.

Operator

The operator (B.G.) was a final year postgraduate student of osteopathy with no previous experience or formal training in sonography. The operator undertook three 2-hour tutorials with an experienced registered sonographer (S.A.) with over 20 years of clinical sonography experience and specialist expertise in musculoskeletal ultrasound imaging including ultrasound of the diaphragm.

Procedures

All images were acquired using a Philips iU22 ultrasound scanner in B-mode (Philips, Medical Systems Company, Eindhoven, NV). In order to ensure maximal superficial resolution while still attaining an acoustic window between ribs, a 12-5 MHz linear transducer was used. The imaging protocol was developed over three sessions scheduled in a two-week period prior to the study. The operator received a total of 6 hours of practical training including supervised scanning of 5 participants during these sessions. This was intended to be broadly comparable to the level of training a manual therapist using Rehabilitative Ultrasound Imaging in clinical practice might typically have undertaken (Jedrzejczak & Chipchase, 2008). Data collection proceeded only after the supervising sonographer was satisfied with the basic scanning technique and recording of measurements.

Questionnaires

All participants were required to fill out two breathing function questionnaires; the Self Evaluated Breathing Questionnaire (SEBQ) and the Nijmegen Questionnaire (NQ) (Courtney & Greenwood, 2009; Van Dixhoorn & Duivenvoorden, 1985).

The 25 item SEBQ has been developed to measure breathing related symptoms and their severity. Scoring is done on a four point scale for various breathing related symptoms; (0) never/not true at all; (1) occasionally/a bit true; (2) frequently/mostly true; and, (3) very frequently/very true. A total score of greater than 11 in the SEBQ may indicate problems with your breathing.

The 16 item NQ tests a broad range of symptoms associated with dysfunctional breathing but is mainly used to assess the presence of hyperventilation syndrome. Scoring is done on a five

point scale; (0) never, (1) Rarely, (2) sometimes, (3) often, (4) very often. A total score of over 23 in this questionnaire suggest a positive diagnosis for hyperventilation syndrome. Both questionnaires were included as one tests a broader range of breathing related symptoms (SEBQ) and the other is more specific for a diagnosis of hyperventilation syndrome (NQ).

Additionally, the Wanner et al (2014), single-item physical activity measure was used to provide a quick approximation of a participant's activity levels. The questions asked was, "how many days over the last week were you involved in physical activity". The answer was marked down as a single number between 0 and 7.

Measurement protocol

The protocol design was informed by previously published methods for measurement of the diaphragm muscle (Boon et al, 2013). Each participant was instructed to assume a supine position on a standard adjustable treatment table (Aster, Metron Medical) with the headrest offset at ~30° with one cervical pillow. Once oriented to the procedures, participants were instructed to only talk if necessary in order to not disturb the image during the scan. Initially the operator observed several cycles of normal quiet respiration to establish a baseline. If this was not achieved within the first minute of observation, the operator used verbal cues to encourage the participant to relax and establish a normal breathing rhythm. The transducer was placed between the ribs, approximately midway between the mid-clavicular and mid-axillary lines. The transducer was then swept parallel to the ribs to establish a view of the diaphragm. Subtle movements of the transducer were utilised until all three layers of the diaphragm were visible such that the echogenic lines that make up the layers were the thinnest. These three layers are as seen as a hypoechoic (dark) structure between two hyperechoic (bright) fascial layers (Figure 1). To establish the correct intercostal space where

the measurement would be taken, the transducer was initially placed on an intercostal space in the vicinity of ribs 7 and 8. The participant was instructed to take a full breath in and out, and the operator observed the lung as a hyperechoic white band moving caudally. If view of the diaphragm was obscured during full breaths it indicated the transducer was placed too cephalad and was moved caudally one intercostal space. This step was reproduced until the lung could be observed on-screen during a full inspiratory phase. This area of the diaphragm just below the visible lung was purposefully used for the location of the measurement to take place. Thickness measurements were taken here as preliminary testing demonstrated this to be more reproducible compared with the alternative of instructing the participant to undertake one maximal inspiratory breath where the 'maximal' breath would often differ between cycles. Once the desired intercostal space was found, and a cine-loop of a full inspiration and expiration cycle was recorded, diaphragm thickness measurements were then taken using the on-screen caliper function. The zoom function was utilized to improve visualisation of the three fascial layers. All images were taken in the following sequence; quiet expiration on the right, maximal inspiration on the right, quiet expiration on the left and maximal inspiration on the left. This was repeated three times in each session and the mean value of the three repetitions was used for all subsequent calculations. All images were digitally captured as screenshots and saved for later offline analyses. During the whole data collection process, the operator was blinded to all measurements displayed on the ultrasound screen using several layers of self-adhesive paper over the measurement fields on the display screen. The second measurement session was undertaken using the same procedures as the first session.
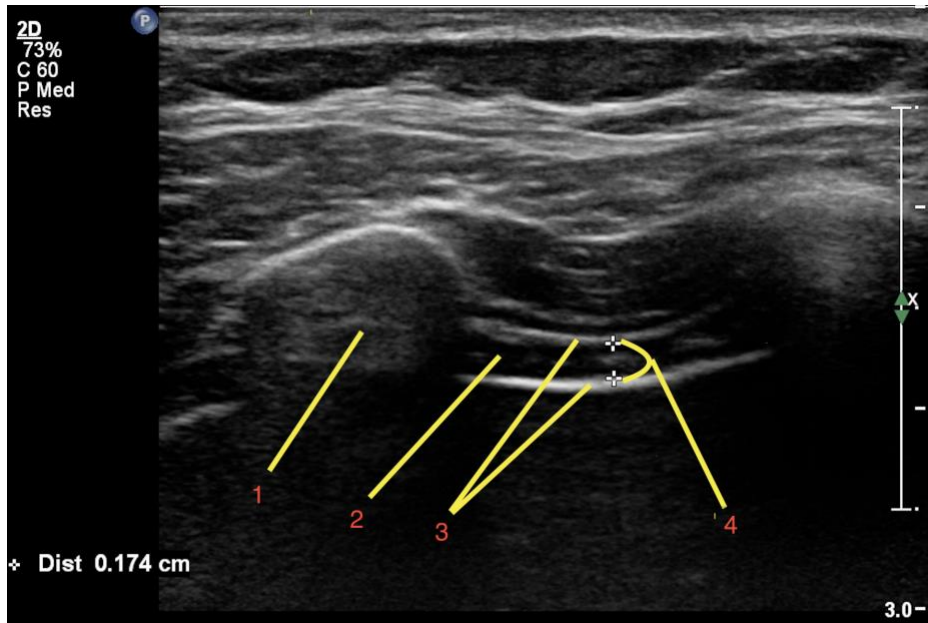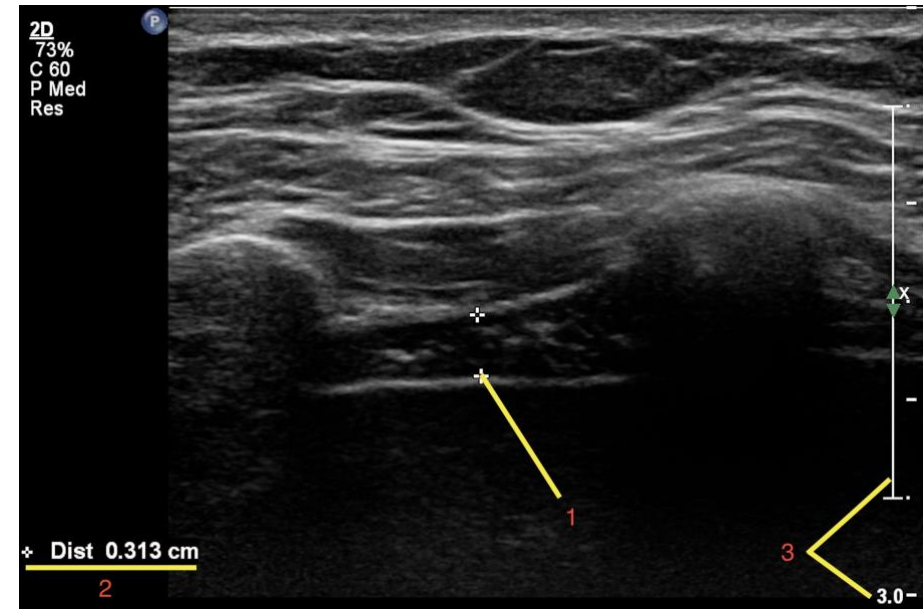
Panel A   **Panel B**

Figure 1. Illustrated screen shot of diaphragm and surrounding tissues during quiet expiration (Panel A), and maximal inspiration (Panel B).

Panel A, Notes: 1: Rib, 2: Hypoechoic (dark) middle layer of the diaphragm, 3: Hyperechoic (light) outer layers of the diaphragm, 4: Diaphragm thickness measuremen

inside of the two hyperechoic layers. **Panel B, Notes:** 1: Calipers placed on the inside of the two hyperechoic layers of the diaphragm in order to take thickness measur

Distance between the calipers in cm, 3: Scale which shows depth in cm.

Data analysis

Utilising G*Power 3.1.2 (Faul et al., 2009), a calculation was made in the planning stage of the study for sample size pertaining to correlational analyses. Using the parameters of power of 0.95, alpha of 0.05, and correlation coefficient of 0.7 a sample size of 20 was calculated to be sufficient. After completion of all data collection, raw values were extracted from screenshots and tabulated in a spreadsheet. The mean value of three measurements for each variable (ie QExp, maxInsp) were used for all subsequent analyses. All statistical analyses were undertaken using SPSS v23 (IBM Corp., Armonk, NY). Descriptive statistics were generated to describe participant characteristics. For all diaphragm measurements, exploration of normality was undertaken using visual inspection of P-P and Q-Q plots, by interpreting the Shapiro-Wilk statistic, and by consideration of kurtosis and skewness (Razali & Wah, 2011). To determine intra-rater reliability an intraclass correlation coefficient was calculated using a two-way random model (model 2,1) and 95% confidence interval calculated. Reliability coefficients were interpreted based on the qualitative descriptors recommended by Hopkins (2002). An *a priori* threshold for ICC coefficients that would be interpreted as clinically acceptable was defined as ICC > 0.6 (Chin et al., 1991). The Standard Error of Measurement (SEM) was calculated using the formula SEM = 1.96 * SQRT(1-ICC), and the Minimum Detectable Change (MDC) was calculated using the formula MDC = SEM * SQRT2 (Wu et al., 2011). In order to assess diaphragm contractility, each participant's maximal diaphragm thickness was divided by their resting thickness measurements to get a ratio, and also expressed as a percentage. To address the secondary aim, the difference between breathing questionnaire scores and relative diaphragm contractility was calculated. For this analysis participants were classified as either having dysfunctional breathing (DB(Yes)) if either their SEBQ scores were greater than 11 or if their

NQ scores were greater than 23 (Courtney & Greenwood, 2009; Van Dixhoorn &

Duivenvoorden, 1985). Participants with lower scores than these cut-offs in both

questionnaires were classified as having no dysfunctional breathing (DB (No)). Differences

in contractility between DB(Yes) and DB(No) were calculated for the left and right

diaphragm using the Mann-Whitney U test.

Results

Measurement protocol reliability

Twenty-five adults (13 males, 12 females) participated in this study and their descriptive

characteristics are shown in Table 1. Out of the twenty-five participants, twelve were

categorised as dysfunctional breathers as having scored higher than the cut-off in either of the

two questionnaires (SEBQ or NQ). All of the twelve 'dysfunctional breathers' were

categorised through their SEBQ scores. Only one person was above the threshold for the NQ,

however, that person also scored as a dysfunctional breather in the SEBQ score. The mean

thickness and range of diaphragm measures at different stages of breathing are reported in

Table 2. Intra-operator reliability ICCs, SEM and MDC are shown for both sides of the

diaphragm including during both quiet expiration and maximal inspiration (Table 3). The

operator showed 'very high' intra-operator reliability for measurements of the diaphragm in

both quiet expiration and maximal inspiration on the left but only during quiet expiration on

the right (Hopkins, 2002). The intra-operator reliability for measuring the diaphragm

thickness during maximal inspiration on the right was 'moderate'. Diaphragm contractility

measurements were recorded as a percentage and also defined as a ratio (see Table 4). There

were no significant difference in contractility measurements between the dysfunctional

breathing group (DB (Yes)) and the non-dysfunctional breathing group (DB (No)) for either

the left ($z=-0.174$, $p=0.862$) or right side of the diaphragm ($z=-0.492$, $p=0.622$).

Table 1. Participant characteristics

	Males	Females	Combined
Age (y)	26.5 (8.3)	27.8 (6.7)	27 (7.2)
Height (cm)	177.9(8.2)	167.3 (6.9)	172.6 (9.3)
Body Weight (kg)	86.7 (10.4)	70.9 (13.3)	79 (14.5)
Physical Activity[1]	3 (1-6, IQR=4)	3.5 (0-4, IQR=2)	3 (0-7, IQR=3)
Nijmegen Score[2]	11.3 (8.2)	15.3 (5.7)	13.3 (7.2)
SEBQ Score[3]	12.8 (12.1)	16.4 (10)	14.9 (11.2)

Notes: All values are mean (SD), or median (min-max, IQR). 1. Physical activity measured using the Wanner et al Single Item Physical Activity measure (Wanner et al., 2014). 2. Nijmegen score measured using the Nijmegen Questionnaire (Dixhoorn & Duivenvoorden, 1985). 3. SEBQ score measured using the Self Evaluated Breathing Questionnaire (Courtney & Greenwood, 2009)

Table 2. Diaphragm thickness measurements (cm)

	Right Side		Left Side	
	Mean (SD)	Range (min-max)	Mean (SD)	Range (min-max)
Day 1				
Quiet expiration	0.17 (0.06)	0.09 - 0.36	0.15 (0.04)	0.09 - 0.25
Max inspiration	0.30 (.099)	0.16 – 0.59	0.24 (0.07)	0.15 - 0.41
Day 2				
Quiet expiration	0.17 (.06)	0.09 – 0.34	0.16 (0.05)	0.09 - 0.26
Max inspiration	0.29 (.08)	0.17 – 0.52	0.24 (0.07)	0.15 - 0.38

Notes: ICC (95%CI) = interclass correlation coefficient with a 95% confidence interval; SEM = standard error of measurement; MDC95 = minimum detectable change

Table 3. Intra-rater Reliability of Diaphragm Thickness Measurements

	Diaphragm thickness during quiet expiration			Diaphragm thickness during max inspiration		
	ICC (95%CI)	SEM (cm)	MDC95 (cm)	ICC (95%CI)	SEM (cm)	MDC95 (cm)
Left	0.892 (0.749-0.956)	0.01	0.04	0.806 (0.596-0.913)	0.03	0.08
Right	0.82 (0.599-0.924)	0.02	0.06	0.538 (0.171-0.774)	0.06	0.16

Notes: ICC (95%CI) = interclass correlation coefficient with a 95% confidence interval; SEM = standard error of measurement; MDC95 = minimum detectable change

Table 4. Diaphragm Contractility Measurements and Relation to Dysfunctional Breathing Questionnaire Scores

	Contractility on right	Contractility on left
Mean	76.625%	58.879%
SD	22.708	15.192
Range (min-max)	39 - 115.8	34.2 - 92.2
Defined as a ratio	1.74	1.56
DB (Yes)	1.74 ± 0.22	1.59 ± 0.2
DB (No)	1.79 ± 0.25	1.59 ± 0.11

Notes: DB = Dysfunctional breathing as measured by fitting into either of the cut off criteria for the NQ or SEBQ. DB values are mean (SD)

Discussion

The main aim of this study was to investigate test re-test reliability for thickness measurement of the diaphragm muscle using ultrasound imaging by a novice operator. The main finding was that the operator was able to measure diaphragm muscle thickness with 'moderate' to 'high' reliability, during maximal inspiration and quiet expiration on the left side and during quiet expiration on the right. However, when measuring maximal inspiration on the right, intra-operator reliability was in the range 'low' to 'medium'. The secondary aim of this study was to establish diaphragm contractility measures for each participant and investigate the presence of group differences between contractility and dysfunctional breathing scores. The inclusion criteria used in this study were purposefully broad. One element of the quality appraisal tools developed by Lucas et al (2010), is that the test sample evaluated should be representative of those to whom the authors intended the results to be applied. In this study the eligibility criteria were purposefully broad to ensure a wide spectrum of participants that would be representative of real world scenarios.

Comparison with previous studies

To our knowledge, this is the first study to report the reliability of diaphragm thickness measurements utilising USI by a non-sonographer operator. There appears to be an increasing number of non-sonographer USI operators, especially in physiotherapy (Jedrzejczak & Chipchase, 2008; Ellis et al., 2018). Establishing novice operator reliability of muscle thickness measurement (and other tissues) using USI may support application in clinical practice. There are several studies that demonstrate the effectiveness of diaphragm tissue measurements acquired by experienced sonographers (Baldwin et al., 2011; Ferrari et al., 2014; Goligher et al., 2015). Boon et al (2013) had a reliability component to their study when measuring diaphragm thickness, however, the extent of operator experience undertaking the ultrasound scans was not reported. Their study reported both inter-rater and intra-rater reliability for sonography operators to be 'very high' in a study of 150 healthy subjects. Intra-rater reliability ICCs were 0.94 (95%CI 0.79-0.98) for resting diaphragm thickness and 0.89 (0.69-0.97) for diaphragm thickness at the end of maximal inspiration. The results of Boon et al (2013) are promising, in that the diaphragm could be reliably measured during both inspiration and expiration, but it is still unclear whether these results could be replicated by a non-sonographer. It was one of our aims to investigate whether a novice operator with little sonography experience could achieve an acceptable level of

reliability for clinical application of this technique. As the price of portable ultrasound machines has decreased, an increasing number of rehabilitation providers are acquiring machines to support rehabilitation practice (Koppenhaver et al., 2009; Teyhen, 2011). As this emerging field of RUSI grows, it is evident there are several applications to breathing retraining and diaphragm monitoring. It is therefore important to determine that RUSI operators can scan and reliably measure a structure such as the diaphragm. The novice rater had comparable sonography training to what could be achieved on a short weekend length course, that is, no more than 8 hours of practical scanning experience. Here, the findings did not match the almost 'perfect' ICCs achieved by Boon et al., (2013). One explanation for the difference in observed reliability between our study and Boon et al may be the experience of the operator. However, other factors in their procedures also need to be considered including: excluding participants with a history of dyspnea which might remove variability between breaths; and only scanning one side of the diaphragm as they did not report on taking measurements bilaterally (Boon et., 2013). Additionally, Boon et al., (2013) failed to report whether raters were blinded to their own findings and possible additional cues (e.g. skin marking, tattoos etc.) (Lucas et al, 2010), as well as whether examination order was varied. Collectively, these aspects represent possible sources of bias that were not reported and may have inflated the reliability reported by Boon et al (2013).

Exploration of diaphragm measures and breathing questionnaire scores

A secondary aim was to establish whether diaphragm thickness and contractility could be a measure of dysfunctional breathing. This was addressed by calculating diaphragm contractility for all participants and considering the extent to which breathing dysfunction as measured by breathing questionnaires (NQ and SEBQ) differed from diaphragm contractility ratios. To our knowledge there are no studies investigating the correlation between diaphragm thickness and especially contractility to dysfunctional breathing. There have been two different methods for calculating contractility reported in the literature. To measure diaphragm contractility as a percentage change, the following formula is used: $(T_{insp}-T_{exp})/T_{exp}*100$ (Terada et al., 2016). Boon et al (2013) use a simple formula in order to calculate diaphragm contractility ratio: T_{max}/T_{min}, where (T_{max}) is the thickness at full inspiration and (T_{min}) is the thickness at the resting end expiration. There are also different approaches to interpreting contractility findings. According to Boon et al (2013), average thickening ratios are ~1.8, with a lower limit of normal of 1.2. They also suggest that

contractility measures may be a more reliable measure of a person's diaphragm function than thickness measurements alone. The results in this study are more in-line with those of Harper et al (2013), who suggest diaphragm contractility may not be an adequate measure of diaphragm function. Harper et al argue that low contractility measures may be found in healthy active subjects, because in this group only very limited contraction of the diaphragm takes place during quiet breathing (Harper et al, 2013). This suggests that a low contractility value could indicate either a breathing problem where a subject's diaphragm has lost most of its contractile properties, or conversely, those who are very healthy and as such are very efficient and need only exert minimal diaphragm contraction at rest.

Internal validity – controls for bias

The characteristics of participants in reliability studies should be designed to be representative of the target population to whom the study findings are intended to be applied (Lucas et al., 2010). Here, the eligibility criteria were purposefully selected to include a wide variation in age, gender, and level of breathing dysfunction in order to be representative of those who might present to rehabilitation providers (i.e. physiotherapy and osteopathy practitioners). In addition, both SEBQ (Courtney & Greenwood, 2009) and Nijmegen (van Dixhoorn & Duivenvoorden, 1985) questionnaires, were completed by all participants prior to undergoing scans. This was to ensure that there was an adequate spectrum of self-perceived breathing quality across the sample. By doing this, the results of this study could be extrapolated to the general public. The classification for dysfunctional breathing in this study was based on threshold scores for the SEBQ (Courtney & Greenwood, 2009) and NQ (van Dixhoorn & Duivenvoorden, 1985). These questionnaires have been tested with rigor and have been used in many breathing studies to date to show breathing efficacy. Being over the threshold in either of those two questionnaires resulted in the participant being grouped in the dysfunctional breathing (DB) category. This was decided to have a wide range of participants with dysfunctional breathing and those who are completely asymptomatic. The operator who performed the USI measures was a postgraduate osteopathy student, with no previous sonography experience. This was intended so that the results could be directly applied to rehabilitation providers such as osteopaths and physiotherapists who may employ RUSI with minimal training in the context of managing breathing related dysfunction (Benjamin et al., 2016).

Recall bias within this study could include rater knowledge of previous diaphragm measurements during data collection, which should be controlled for when establishing intra-rater reliability (Lucas et al, 2010). To mitigate this potential bias, the operator was blinded to all measures throughout the course of data collection. This was achieved by applying self-adhesive paper shields on the area of the ultrasound screen to blind the operator from measurements. Images showing the measurements were saved for later analysis, and so the operator was not privy to the measurements between trials in the same session but was also blinded from any of the measurements throughout the data collection process until data collection was completed. In order to minimize the effects of operator bias related to knowledge of participants' scores to breathing dysfunction indicators, the operator was blinded to the results of both the SEBQ and Nijmegen questionnaires prior to undertaking each scan.

A further source of potential bias in the design of test re-test reliability studies may be the recognition of additional cues or identifiable characteristics on some participants (e.g. skin moles, scars, tattoos etc.) (Lucas et al., 2010). An example of this would be if a participant had a tattoo over his/her ribs which would guide the operator in the second session rather than using the established scanning protocol. The operator was blinded to additional cues to the extent that was reasonably achievable. As each participant was required to expose their lower ribs to allow scanning to take place, it was not considered reasonable to hide features such as tattoos or scars on their torso from the operator. However, care was taken to only expose the area of skin that needed to be scanned which minimized the potential for these additional cues to influence the reliability.

In reliability studies, the interval between measures should be designed with the routine clinical application of the test in mind (Lucas et al., 2010) The two-week interval between sessions was selected for two reasons, firstly, that it would be sufficient time for the operator to forget specific characteristics (e.g. skin markings) about any one individual; and secondly, that the two week interval would be a period over which a practitioner might undertake a repeat measurement during the course of a clinical intervention.

The order of subjects to be examined should be randomized in order to control for order bias (Lucas et al, 2010). However, this was logistically difficult to achieve due to participant availability and logistical constraints in accessing the ultrasound machine. Although not truly

random, there was some variation of the order in which participants were examined because participants were sometimes not able to attend the exact same time for both scanning sessions due to personal scheduling conflicts. Beyond this, no further attempt at varying examination order was made.

In all reliability studies, it is important that the rater carries out the test correctly to ensure good intra-rater reliability (Lucas et al., 2010). For instance, in measurement of the diaphragm using USI, adequate knowledge of surface anatomy and identification of specific intercostal spaces is required in order to place the transducer in the correct location, making the diaphragm visible. As the rater was a novice in the use of USI, it took multiple trial runs with an experienced musculoskeletal sonographer to ensure that the test was applied correctly. It was not until the experienced sonographer was satisfied that the method and technique of extracting diaphragm measurements was demonstrated over several pilot subjects that the main data collection took place.

Clinical applicability

The results of this study should be of interest to practitioners and researchers involved in measuring diaphragm thickness. It is important to establish whether the same high reliability when measuring a dynamic structure as the diaphragm can also be seen in operators with minimal training and experience. In addition, a correlation between diaphragm thickness including contractility and breathing quality had not been previously investigated. For rehabilitation purposes, a practitioner with a USI device could then foreseeably note progress in diaphragm thickness variation.

Limitations

A limitation of this study include relatively small sample size (n=25), with a generally young mean age (27 years of age). Although the sample was adequate for reliability purposes, having both a larger sample size and also older participants would improve the generalisability of the findings across the range of ages who seek care for breathing related disorders, with the exception of those under the age of 18 years. Hence, these results cannot be extrapolated to measuring children and caution must be taken in generalising to older adults. An additional limitation is that only intra-rater reliability was assessed and not *inter-*rater reliability.

One of the challenges when scanning the diaphragm is locating a good acoustic window in which to visualise the muscle thickening during inspiration. Boon et al (2013), stated they used real time ultrasound to identify the intercostal space at which the diaphragm was most easily visualised and where the encroachment of lung tissue did not obstruct view of the diaphragm. This description of locating the portion of the diaphragm they wished to measure, although clear, was somewhat ambiguous. Soon after commencing pilot work it became apparent that it would be difficult to locate the same portion of the diaphragm from one week to the next. Variance in body morphology and rib angles meant transducer placement was different from person to person and often all three layers of the diaphragm were visible over two to three intercostal spaces. This made it challenging to decide where best to make measurements as multiple locations were possible. In addition, the breath volume used by the participants when recording maximal inspiration sometimes varied between measures within the single session. It appears that this potential problem has not previously been addressed in the literature. When piloting this study, our solution to both these obstacles was to initially place the scanner on a more cephalad intercostal space and scan in each intercostal space until the desired intercostal space was identified. A strength of this study was that the operator adapted the way in which diaphragm thickness measurements were taken to use the patient's physiology as a marker instead of a subjective variable such as attempting to have the participant take the same 'maximal' inspiratory breath for each measurement. Specifically, this was achieved by commencing the scan at an intercostal space around ribs 7-8 where the diaphragm was visible. At this stage the participant would be instructed to "take a full breath in" while we observed the monitor. If the hyperechoic band of the lung came into view to obstruct the view of the diaphragm, the intercostal space was considered to be too high and the next caudal intercostal space was reviewed. This process was repeated until the lung was visualised and started to separate the diaphragm from the adjacent parietal layer of the lung but stopped before completely shielding all three layers of the diaphragm.

Future research

This study focused only on intra-rater reliability of a single novice USI operator with minimal experience. Future research should establish whether the same high reliability can be achieved between different operators, i.e. inter-rater reliability. This is especially important between operators of different skill and experience level. It would also be interesting to evaluate the effects of therapy on diaphragm thickness and contractility as measured using USI. Future research could also investigate interventions such as exercise or

a breathing retraining protocol (Benjamin et al., 2016) on possible changes in diaphragm thickness and contractility. Additionally, USI could be utilised before and after a single treatment or technique to observe possible changes within the diaphragm. Another variable that would be interesting to investigate further would be smoking status, as this could change diaphragm thickness measurements and contractility.

Conclusions

Within this study, a non-sonographer demonstrated acceptable reliability for diaphragm measurements using ultrasound. However, more research has to be conducted in order to establish whether a correlation between diaphragm thickness including contractility and dysfunctional breathing exists.

REFERENCES

Baldwin, C. E., Paratz, J. D., & Bersten, A. D. (2011). Diaphragm and peripheral muscle thickness on ultrasound: Intra-rater reliability and variability of a methodology using non-standard recumbent positions. *Respirology*, *16*(7), 1136-1143.

Benjamin, J. G., Bacon, C. J., Verhoeff, W. J., & Moran, R. W. (2016). Preliminary development of a complex intervention for osteopathic management of dysfunctional breathing. *International Journal of Osteopathic Medicine*, *21*, 29-39.

Boon, A. J., Harper, C. J., Ghahfarokhi, L. S., Strommen, J. A., Watson, J. C., & Sorenson, E. J. (2013). Two-dimensional ultrasound imaging of the diaphragm: Quantitative values in normal subjects. *Muscle & Nerve*, *47*(6), 884-889.

Chaitow, L. (2004). Breathing pattern disorders, motor control, and low back pain. *Journal of Osteopathic Medicine*, *7*(1), 33-40.

Chinn, S. (1991). Statistics in respiratory medicine. 2. Repeatability and method comparison. *Thorax*, *46*(6), 454.

Courtney, R., & Greenwood, K. (2009). Preliminary investigation of a measure of dysfunctional breathing symptoms: The Self Evaluation of Breathing Questionnaire (SEBQ). *International Journal of Osteopathic Medicine*, 12(4), 121- 127.

Ellis, R., Rohan, M., Fox, J., Hitt, J., Langevin, H., & Henry, S. (2018). Ultrasound Elastographic Measurement of Sciatic Nerve Displacement and Shear Strain During Active and Passive Knee Extension. *Journal of Ultrasound in Medicine*.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G* Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*(4), 1149-1160.

Ferrari, G., De Filippi, G., Elia, F., Panero, F., Volpicelli, G., & Aprà, F. (2014). Diaphragm ultrasound as a new index of discontinuation from mechanical ventilation. *Critical Ultrasound Journal*, *6*(1), 8.

Folgering, H. (1999). The pathophysiology of hyperventilation syndrome. *Monaldi Archives for Chest Disease= Archivio Monaldi per le Malattie del Torace*, *54*(4), 365-372.

Goligher, E. C., Laghi, F., Detsky, M. E., Farias, P., Murray, A., Brace, D., ... & Ferguson, N. D. (2015). Measuring diaphragm thickness with ultrasound in mechanically ventilated patients: feasibility, reproducibility and validity. *Intensive Care Medicine*, *41*(4), 642-649.

Harper, C. J., Shahgholi, L., Cieslak, K., Hellyer, N. J., Strommen, J. A., & Boon, A. J. (2013). Variability in diaphragm motion during normal breathing, assessed with B-mode ultrasound. *Journal of Orthopaedic & Sports Physical Therapy*, *43*(12), 927-931.

Hodges, P. W., Sapsford, R., & Pengel, L. H. M. (2007). Postural and respiratory functions of the pelvic floor muscles. *Neurourology and Urodynamics*, *26*(3), 362-371.

Hopkins, W. G. (2002). Probabilies of clinical or practical significance. *Sportscience*, *6*(201), 16.

Jedrzejczak, A., & Chipchase, L. S. (2008). The availability and usage frequency of real time ultrasound by physiotherapists in South Australia: an observational study. *Physiotherapy Research International*, *13*(4), 231-240.

Kapreli, E., Vourazanis, E., & Strimpakos, N. (2008). Neck pain causes respiratory dysfunction. *Medical Hypotheses*, *70*(5), 1009-1013.

Kiesel, K. B., Uhl, T. L., Underwood, F. B., Rodd, D. W., & Nitz, A. J. (2007). Measurement of lumbar multifidus muscle contraction with rehabilitative ultrasound imaging. *Manual Therapy*, *12*(2), 161-166.

Koppenhaver, S. L., Hebert, J. J., Parent, E. C., & Fritz, J. M. (2009). Rehabilitative ultrasound imaging is a valid measure of trunk muscle size and activation during most isometric sub-maximal contractions: a systematic review. *Australian Journal of Physiotherapy*, *55*(3), 153-169.

Kunik, M. E., Roundy, K., Veazey, C., Souchek, J., Richardson, P., Wray, N. P., & Stanley, M. A. (2005). Surprisingly high prevalence of anxiety and depression in chronic breathing disorders. *Chest*, *127*(4), 1205-1211.

Lucas, N. P., Macaskill, P., Irwig, L., & Bogduk, N. (2010). The development of a quality appraisal tool for studies of diagnostic reliability (QAREL). *Journal of Clinical Epidemiology*, *63*(8), 854-861.

Razali, N. M., & Wah, Y. B. (2011). Power comparisons of shapiro-wilk, kolmogorov-smirnov, lilliefors and anderson-darling tests. *Journal of Statistical Modeling and Analytics*, *2*(1), 21-33.

Schepens, T., Verbrugghe, W., Dams, K., Corthouts, B., Parizel, P. M., & Jorens, P. G. (2015). The course of diaphragm atrophy in ventilated patients assessed with ultrasound: a longitudinal cohort study. *Critical Care*, *19*(1), 422.

Smith, M. D., Russell, A., & Hodges, P. W. (2006). Disorders of breathing and continence have a stronger association with back pain than obesity and physical activity. *Australian Journal of Physiotherapy*, *52*(1), 11-16.

Terada, M., Kosik, K. B., Mccann, R. S., & Gribble, P. A. (2016). Diaphragm Contractility in Individuals with Chronic Ankle Instability. *Medicine and Science in Sports and Exercise*, *48*(10), 2040-2045.

Teyhen, D. S. (2011). Rehabilitative ultrasound imaging for assessment and treatment of musculoskeletal conditions. *Manual Therapy*, *16*(1), 44.

Teyhen, D., & Koppenhaver, S. (2011). Rehabilitative ultrasound imaging. *Journal of Physiotherapy*, *57*(3), 196.

Van Dixhoorn J, Duivenvoorden H. Efficacy of Nijmegen questionnaire in recognition of hyperventilation syndrome. Journal Psychosomatic Research. 1985;29(2):199-205.

Wanner, M., Probst-Hensch, N., Kriemler, S., Meier, F., Bauman, A., & Martin, B. W. (2014). What physical activity surveillance needs: validity of a single-item questionnaire. *Journal of Sports Medicine*, *48*(21), 1570-1576.

Wu, C. Y., Chuang, L. L., Lin, K. C., & Hong, W. H. (2011). Responsiveness, minimal detectable change, and minimal clinically important difference of the Nottingham Extended Activities of Daily Living Scale in patients with improved performance after stroke rehabilitation. *Archives of Physical Medicine and Rehabilitation*, *92*(8), 1281-1287.

Appendix: Ethics Approval

Ben Giles

Health Care Practice Pathway

Osteopathy

Dear Ben,

Ethics application number: 2017-1043

Thank you for completing and submitting the amendments requested. As Primary Reader of your application and under delegated authority from the Unitec Research Ethics Committee (UREC) I now authorise you to begin your research.

Please note, if you have not yet done so at the time of receiving this advice; please email one copy of your final amended ethics application and any additional documents to the UREC secretary at: ethics@unitec.ac.nz. You will receive a formal letter of approval after the next UREC meeting. Note meetings are held monthly.

The dates that must be referred to on the Information Sheet AND Consent Forms given to all participants and appear on your documents are as follows:

Start date: 28 June 2017

Finish date: 27 June 2018 (or as appropriate but not after this date)

Please note, you must inform UREC, in advance of any ethically-relevant modification in the project as this may require additional approval.

Best wishes for your project.

Signed,

Dr Nigel Adams

Deputy Chair: Unitec Research Ethics Committee

Full name of author: *Ben Giles*

ORCID number (Optional): ....

Full title of thesis/dissertation/research project ('the work'):

Intra-rater reliability of measuring diaphragm thickness utilizing ultrasound imaging by a non-sonographer practitioner

Practice Pathway: *Osteopathy*

Degree: *Masters of Osteopathy*

Year of presentation: *2018*

Principal Supervisor: *Rob Moran*

Associate Supervisor: *Catherine Bacon (resigned)*

Permission to make open access
I agree to a digital copy of my final thesis/work being uploaded to the Unitec institutional repository and being made viewable worldwide.

Copyright Rights:
Unless otherwise stated this work is protected by copyright with all rights reserved.
I provide this copy in the expectation that due acknowledgement of its use is made.

AND

Copyright Compliance:
I confirm that I either used no substantial portions of third party copyright material, including charts, diagrams, graphs, photographs or maps in my thesis/work or I have obtained permission for such material to be made accessible worldwide via the Internet.

Signature of author: ....

Date: *12 / 11 / 2018*

Declaration

Name of candidate: **Ben Giles**

This Thesis/Dissertation/Research Project entitled : **Intra-rater reliability of measuring diaphragm thickness utilizing ultrasound imaging by a non-sonographer practitioner** is submitted in partial fulfillment for the requirements for the Unitec degree of

Principal Supervisor: **Rob Moran**

Associate Supervisor/s: **Catherine Bacon (resigned)**

CANDIDATE'S DECLARATION

I confirm that:

- This Thesis/Dissertation/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: **2017 - 1043**

Candidate Signature: _____ Date: **12/11/18**

Student number: **1414088**