Reliability of Ultrasound Measurement Using Image Enhancement Software

Sruthy Vijayan

Student ID: 49499941

Thesis Submission in partial fulfilment of the requirements for the Degree of Master of Speech and Language Science

Department of Communication Disorders

Rose Centre for Stroke Recovery and Research

University of Canterbury

February 2021

Abstract

Background: Ultrasound has not progressed to standard clinical practice despite the fact that it offers a radiation-free and non-invasive procedure for swallowing assessment and promising validity and reliability for ultrasound measurements. The cost, which may be one of the significant obstacles to restrict integration into clinical practice, can be overcome by using new pocket-sized portable ultrasound devices. However, early research suggests that poor image quality may compromise measurement reliability. To enhance the interpretation of swallowing ultrasound images, an image enhancement software is recently developed which enhances the contrast and features of acquired ultrasound images.

Aim: The focus of the present research is to determine if this image enhancement software increases intra-and inter-rater reliability of ultrasound measurement compared to the original ultrasound images.

Method: Archived data of 40 ultrasound images were measured by the primary investigator and three co-investigators to derive intra-and inter-rater reliability. Training for co-investigators was provided one week prior to ultrasound image measurement under two conditions: with and without the image enhancement software programme. A total of 80 images were measured by each researcher, which includes 40 original images and 40 enhanced images. Trials were randomized within and across the raters. Image analysis occurred across a four-day period, with 20 trials per day to minimise the effects of fatigue. One week following the completion of inter-rater reliability measurement, all investigators re-evaluated the same images, in random sequence, to evaluate intra-rater reliability. Therefore, a total of 160 measures were made by each investigator.

Results: Intra-rater reliability for enhanced images was good (> 0.75) for hyoid excursion and moderate (0.50 – 0.75) to good (> 0.75) for floor of mouth muscles, while it ranged moderate (0.50 – 0.75) to good (> 0.75) for hyoid excursion and poor (< 0.50) to good (> 0.75) for floor of mouth muscles for normal images. For thyrohyoid approximation, intra-rater reliability was found to be similar for both enhanced and normal images, ranging from poor (< 0.50) to good (>0.75). Inter-rater reliability of enhanced images ranged from moderate (0.50 -0.75) to good (> 0.75), whereas poor (< 0.50) to moderate (0.50 – 0.75) for normal images.

Conclusion: The results from this study indicate a slight increase in reliability for enhanced images compared to normal images and this difference is not potentially large enough to be clinically significant.

Acknowledgement

Firstly and foremost, I am deeply indebted to my supervisor, Professor Maggie Lee Huckabee for giving me the fantastic opportunity to finish my thesis under your supervision, honestly, it's an honour. I am truly grateful for your constructive advice and understanding over the past year. Your wealth of knowledge, enthusiasm, and dedication for people with dysphagia is genuinely inspiring. I should acknowledge Dr. Phoebe Macrae for your valuable suggestions and discussion.

My special thanks to Alan Brooks for developing the software, which is the crucial part of the thesis. Thank you for your timely response and patience while providing technical support whenever needed. Your effort in this project is sincerely appreciated.

I want to pay my special regards to Katharina Winiker, who helps me with data analysis. Thanks to your excellent knowledge of statistics. I am indebted for your constant willingness to explain and discuss. Your valuable suggestion has greatly contributed to the improvement of the thesis.

I am immensely grateful for all the participants of this study. Great thanks for your willingness to take part in this study and for making this research possible. I deeply appreciate your valuable time and contribution. Also, I am fortunate to work along with a wonderful team at Rose centre. I wish to thank each fellow students for their generous support and friendship and a great thanks to Fiona for your positive spirit.

Finally, I own my deepest gratitude to my father, C.V Vijayan, my mother Sathimol, and my brother, Ajith, for their constant support throughout this journey. A big hug to you guys. My gratitude extents to my dear friends Soorya, Sandra, Neethu, Ranna, and Anija for your positive support and friendship.

Contents

List of Abbreviations

ALS	amyotrophic lateral sclerosis
CI	confidence interval
CT	computed tomography
ICC	intraclass correlation coefficient
LAB	left anterior belly of digastric muscles
MRI	magnetic resonance imaging
Q-Q-plot	quantile-quantile plot
RAB	right anterior belly of digastric muscles
SD	standard deviation
SEM	standard error of measurement
VFSS	videofluoroscopic swallowing study

List of Figures

List of Tables

Introduction

Ultrasound technology offers a non-invasive and relatively inexpensive method to study the structures and soft tissues of the oro-pharynx and the biomechanics of the oropharyngeal system during swallowing (Sonies et al., 2003). The absence of radiation exposure, practicability, and portability are some other advantages of ultrasound over videofluoroscopy (Hsiao et al., 2012; Watkin, 1999). Ultrasound imaging has been utilized in swallowing research since the late 1970s (Skolnick et al., 1975; Stevens, 1978). Although past research documents good reliability for ultrasound measurements, it has not progressed to standard clinical practice.

The cost, which may be one of the significant factors that have restricted integration into clinical practice, can be overcome by using new small-sized portable ultrasound devices, which may increase the potential for clinical translation of ultrasound devices. Utilizing such devices could be useful for swallowing assessment at bedside or remote facilities. However, the reliability and validity of this newer technology has been investigated in recent literature and found to be insufficient for clinical application (Winiker et al., 2019). It was suggested that the image quality of this newer technology in swallowing assessment did not meet the standards of more advanced ultrasound equipment used in previous reliability studies (Hsiao et al., 2012; Huang et al., 2009; Macrae et al., 2012). Acceptable image quality is important for reliable measurements; hence, it is essential to ensure that the portable ultrasound produces good quality images for reliable measurements to ensure the integrity of diagnostic results.

Image enhancement software may facilitate the interpretation of measurements. To enhance the interpretation of swallowing ultrasound images, image enhancement software was developed that enhances the contrast and features of acquired ultrasound images[1]. The primary goal of the software was to develop algorithms for enhancing the contrast and features of ultrasound images based on intensity and texture characteristics. The focus of the present research is to determine if this image enhancement software increases the inter- and intra-rater reliability of ultrasound measurement when compared to the original ultrasound images.

[1] Alan Brooks, 2020

Literature Review

Brief History of Development of Medical Ultrasound

Medical ultrasound technology was pioneered not more than a century ago, although its origin is much earlier. The earliest work of Lazaro Spallanzani on discovering ultrasonography formed the basis for ultrasound physics when he studied the echolocation of bats in 1794 (Kane et al., 2004). The discovery of the piezoelectric effect by the Curie brothers in the 18th century led to the development of modern-day ultrasound transducer technology (Bressmann et al., 2005; Mann, 2019; Merton, 1997; Newman et al., 1998; Szabo, 2004; Woo, 2008).

Karl Dussik, a neurologist (1942) was credited with being the first to detect brain tumours using ultrasound as a medical diagnostic tool (Mann, 2008; Merton, 1997; Kane et al., 2004). Later, clinicians started to use Amplitude mode (A- mode) ultrasound and brightness mode (B-mode) ultrasound for diagnostic purposes across various settings around the world by 1950s (Merton, 1997). Inge Elder, a cardiologist in 1953, marked the beginning of echocardiography with the development of M- mode ultrasound (Newman et al., 1998). From the late 1950s, pulsed wave and continuous wave doppler ultrasound also emerged. This technological progress eventually opened the way for real-time ultrasound imaging in 1970s, which is one of the most revolutionary developments in the medical field.

Ultrasound has secured a position as a key diagnostic test with its recent technological advances, including the introduction of colour flow ultrasound imaging, 3D (three-dimensional) and 4D (four-dimensional) ultrasound imaging, availability of non-toxic contrast agents, tissue displacement imaging, high-frequency transducers, and pocket-sized portable ultrasound (Merton, 1997; Shung, 2010).

Ultrasound Working Principle

Ultrasound is an imaging technique for generating images using inaudible high-frequency sound waves (Stone, 2005; Watkin, 1999). Most common sound wave frequencies used in medical diagnosis range between 2 to 15 MHz (Jenson, 2007; Martin, 2010). Ultrasound is based on the transmission and reflection of high-frequency sound waves at 20 kHz and above that reflect at the interfaces of tissues with different acoustic impedances (Sonies et al., 2003). To correctly interpret the results while using ultrasound clinically, it is essential to have an understanding of the basic physics behind ultrasound imaging, including image formation and resolution (Martin, 2010).

Image Formation

Ultrasound waves are generated within the transducer by briefly passing an electric current through a piezoelectric crystal (Martin, 2015). The resulting pulse of ultrasound waves is delivered to the tissues (Martin et al., 2015; Stone, 2005; Watkin, 1999). When the ultrasound waves enter the tissue, some penetrate deeper while the transducer reflects some as echo signals.The reflection of these acoustic waves forms the basis for image formation (Sonies et al., 2003). Sound waves will reflect at the interfaces between tissues having different acoustic impedances as an ultrasound beam travels through the tissues (Aldrich, 2007; Kundra et al., 2011; Venables, 2011). The denser the material, such as soft tissue and bony structures, the greater the acoustic impedance ; the more echogenic it is, the more white it appears on the screen (Kristensen, 2011). Less dense structures with lower acoustic impedance (fat tissue and water) result in weak echoes and show poorly differentiated shades of grey (Aldrich, 2007; Epstein & Stone, 2005).

B-mode ultrasound provides a two-dimensional image display where many single B-mode dots are displayed and collected together to create an image on the screen (Martin, 2010). At each point, the brightness of the image is formed by the amplitude of the returned echo signal at the tissue boundaries, thereby results in B-mode (Martin, 2010). However, the intensity of ultrasound pulses is reduced or attenuated as they travel through the tissues through the process of refraction, scattering, reflection, absorption, and diffraction (Chan, 2010; Martin, 2015). Thus, only a small amount of energy is received by the ultrasound transducer from returning echoes. This is compensated by time gain compensation, a process in which a greater amplification is provided to more distant echoes from the transducer (Martin, 2015).

Image Resolution

A detailed understanding of the factors associated with the image resolution of ultrasound is essential for both the quality and interpretation of data included within the images. The factors that determine the resolution of ultrasound imaging systems include frequency, bandwidth, and the focussing properties of the transducer (Ploquin et al., 2015). The ability to detect objects of different sizes and echogenicities of an imaging system depends upon the spatial resolution. Spatial resolution can detect and display the structures that lie close together and are further categorized into axial and spatial resolution (Filoux et al., 2011).

Axial resolution refers to the minimum distance between two reflectors along the axis of the beam, which is positioned near each other (Hedrick et al., 1995; Martin, 2015; Stone et al., 2005). It is influenced by the frequency and wavelength of the transducer. The higher the frequency, the smaller the wavelength and the better the axial resolution is (Ploquin, 2015). The minimum distance between two reflectors positioned perpendicular to the ultrasound beam’s direction is referred to as the lateral resolution (Stone, 2005). It depends upon the transducer’s bandwidth (Shung, 2006; Stone, 2005). Although higher frequencies provide better image resolution, attenuation happens more rapidly and restricts penetration into deep tissues (Stone, 2005). However, the resolution of lower frequencies is poorer but reaches deeper depths.

The ultrasound beam refers to the sound waves originating from a single piezoelectric crystal (Stone, 2005). The quality of the ultrasound images and how tissues and structures are displayed on these images are affected by the ultrasound beam pattern (Kossoff, 2000). It can be either focused or unfocused and is produced with a far and near field, as the sound propagates out of the transducer (Heggie et al., 2001; Kossoff, 2000). The ultrasound beam diverges in the far field and converges in the near field. Increasing the frequency results in a longer near field and less far field divergence.

Ultrasound resolution is affected by several factors considered above. The most significant goal in ultrasound image formation is to achieve optimum image quality which may contribute to improved diagnostic outcomes.

Ultrasound in Swallowing Research

Ultrasound technology offers a non-invasive and relatively inexpensive method to study the structures and soft tissues of the oro-pharynx and the biomechanics of the oropharyngeal system during swallowing (Sonies et al., 2003). The absence of radiation exposure, practicability, and portability are some other advantages of ultrasound over videofluoroscopy (Hsiao et al., 2012; Watkin, 1999). It can provide biofeedback especially for vulnerable group of patients and could be useful for the repeated use in diagnostics (Logemann, 1998; Watkin, 1999). Participants can be examined for an extended period regardless of any known adverse effects (Logemann, 1988; Sonies et al., 2003). For swallowing examinations, ultrasound uses real food rather than using contrast materials such as barium (Sonies et al., 2003). The usage of these materials may impact swallowing function

(Sonies et al., 2003; Steele, 2005). Further, videofluoroscopic examinations may be challenging for patients who cannot mobilise out of bed or those living in rural areas (Rugiu, 2007).

MRI (Magnetic Resonance Imaging) and CT (Computed Tomography) are other imaging techniques utilised to visualise muscle morphometry (Schedel et al., 1992; Sipila & Suominen, 1993) and swallowing (Hart et al., 2010; Panebianco et al., 2010). When compared to CT and MRI, ultrasound is more portable and less expensive (Chi-Fishman et al., 2004). Other limitations of these imaging methods includes the long acquisition time of MRI and ionizing nature of CT (Hartl et al., 2010; Panebianco et al., 2010). In contrast, ultrasound involves no radiation exposure and is well tolerated by individuals of different ages, from infancy to senescence (Abramovich et al.,1979; Bowie & Clair, 1982; Chi-Fishman, 2005; Mann, 2008; Marsal, 1983; Shawker et al., 1982; Shawker, Sonies et al., 1983; Shawker et al., 1984; Smith et al., 1985; Stevens, 1978; Walker, 2004; Weber et al.,1986). Additionally, CT and MRI tests are conducted in a supine position with the participants restricting feeding studies (Hiiemae & Palmer, 2003). Research comparing the measurement of muscle morphometry between imaging techniques such as ultrasound, CT and MRI has shown good agreement with electromyography and histopathology in myopathies (Chi-Fishman, 2004). This indicates that ultrasound is a valid method for evaluating cross-sectional areas of muscles (Alanen et al., 1994; Chi-Fishman, 2004; Macrae et al., 2013).

Assessment of Hyoid Excursion

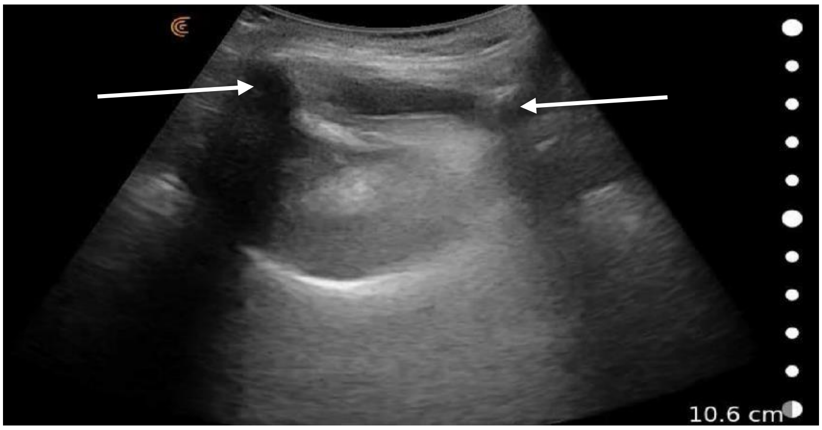
Hyoid bone movement is a major biomechanical marker of swallowing (Dodds et al., 1990; Robbins et al., 1992). It cannot be visualised directly through the ultrasound since it is a bony structure (Chi-Fisman, 2005; Cordaro & Sonies, 1993). Yet, during swallowing, the hyoid can be imaged as a dark band of hypoechoic acoustic shadow (Chi-Fishman, 2005; Chi-Fishman & Sonies, 2002). And this is utilised as a reference for the measurement of hyoid excursion (Singh et al., 2010; Walker et al., 2004). The hyoid moves towards the mentalis of the mandible during swallowing, and the hyoid shadow remains at its maximal anterior- superior displacement at the peak of swallowing (Chi-Fishman, 2005).

Different ultrasound measurement techniques for hyolaryngeal excursion have been explored in the literature. Some studies quantify hyoid excursion based on percentage change (Macrae et al., 2012) while others use absolute measurements (Chen et al., 2017; Hsiao et al., 2012; Macrae et al., 2012). Hsiao et al. (2012) used ultrasound to evaluate 30 dysphagic stroke patients, 30 non-dysphagic stroke patients, and 30 healthy individuals for assessing hyoid bone movement. Results revealed a significantly lower mean value of hyoid bone displacement in dysphagic stroke patients (1.3 cm) than in the non-dysphagic stroke patients (1.6 cm) and healthy individuals (1.7 cm). Hyoid bone displacement less than 1.5 cm was identified as the cut-off value, with a sensitivity and specificity of 73.3% and 66.7 %, respectively, for tube-feeding-dependent dysphagia.

Many studies have employed various transducer/head stabilization techniques to make sure that participant's head remained stable at the time of scanning (Chi-Fishman & Sonies, 2002a, 2002b; Peng et al., 1996; Stone & Davis, 1995). Chi-Fishman and Sonnies (2002) reported the effects of systematic bolus viscosity and volume changes on hyoid kinematics using a custom-made transducer holder. The same researchers examined discrete and rapid sequential swallowing tasks in healthy individuals using the same custom-made transducer holder device. The need for the transducer/head stabilization techniques does pose challenges for the clinical translation of this assessment method and this been eliminated to some extent by utilizing an anatomic reference point, maintaining good reliability (Chen et al., 2017; Hsiao et al., 2012; Macrae et al., 2012; Perry et al., 2016).

Research conducted by Perry et al. (2016) revealed no convincing evidence between the ultrasound measures of hyolaryngeal excursion using a fixed transducer placement and hand-held transducer placement. Twenty-four healthy individuals participated in the study, where the mandible was used as a stable reference point for both conditions. These findings suggest that utilisation of a stabilisation technique may not provide better measurement accuracy compared to a hand-held transducer.

Figure 1 *Illustration of Hyoid excursion on ultrasound*



Note. Depicts the hyoid shadow at the right side and the mandible shadow at the left side of the image.

Assessment of Thyrohyoid Approximation

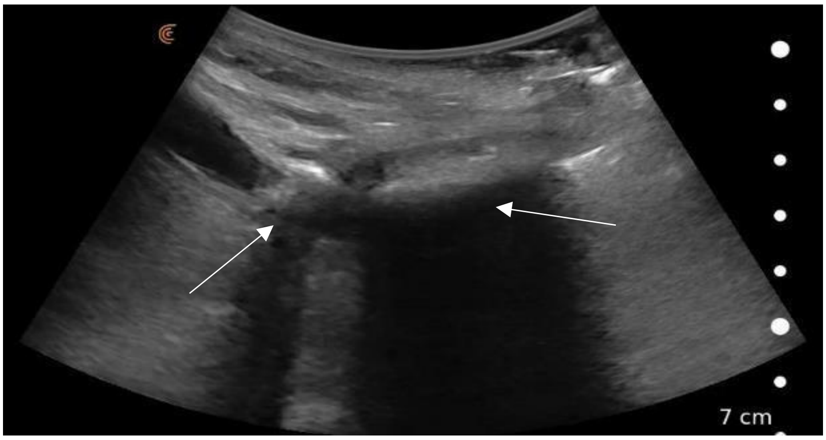
Thyrohyoid approximation is described as the shortest distance between the thyroid cartilage (top end) and the hyoid bone (lower end) during swallowing (Ahn et al., 2015). During swallowing, the approximation of hyoid and larynx leads to laryngeal vestibular closure. This results in compression of the median thyrohyoid tissue, closing the laryngeal entrance, hence protecting the airway (Fink et al., 1979; Logemann et al., 1992). Therefore it plays an essential role in swallowing.

Hyung et al. (2009) analysed the reliability of ultrasound measurement of thyrohyoid approximation in healthy participants and patients following stroke with or without dysphagia. They compared ultrasound measurements with VFSS examinations for validation purposes and found similar percentage change measures of thyrohyoid approximation between ultrasound and VFSS in stroke patients. Also, they have documented reduced thyrohyoid approximation in stroke patients with dysphagia than the non-dysphagic group.

Kuhl et al. (2003) compared thyrohyoid approximation during swallowing between 42 healthy participants and 18 dysphagic patients using ultrasound. And they have measured the shortest distance and mean distance at rest during swallowing for calculating the relative reduction of distance during swallowing. Results suggested that the patient group had significantly reduced thyrohyoid approximation (mean relative laryngeal reduction of 42%) than healthy participants (mean relative laryngeal reduction of 61%). However, the reliability or validity of measures was not examined in the study.

Ahn et al. (2015) investigated the influence of positional change using ultrasound on 20 healthy participants for measuring thyrohyoid approximation. They evaluated both sitting and supine positions with the same examiner conducting the measurements three times and averaged the findings. No significant difference was reported in swallowing between the supine and sitting positions in healthy individuals. A significant negative correlation was observed between body mass index and resting and approximation distances. However, standard reliability calculations were also lacking in this study.

Figure 2 *Illustration of Thyrohyoid approximation on ultrasound*



Note. Depicts the shadow cast by the thyroid cartilage at the right side and the shadow cast by the hyoid bone at the left side of the image.

Assessment of Floor of Mouth Muscles

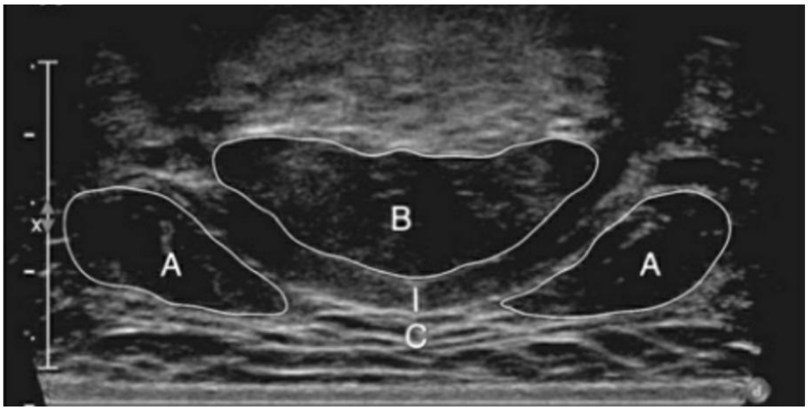
The floor of mouth muscles involved in swallowing, such as the geniohyoid, mylohyoid and anterior belly of digastric muscles can be viewed on ultrasound (Emshoff et al., 1999; Morrish et al., 1984; Shawker et al., 1984; Shimizu et al., 2016; Watkin et al., 2001). Contraction of thyrohyoid muscles and suprahyoid muscles (anterior and posterior belly of digastric muscles, mylohyoid, stylohyoid, geniohyoid) contributes to the hyolaryngeal elevation during swallowing by opening the upper esophageal sphincter (Dodds, 1989; Matsuo et al., 2008). Contraction of the floor of mouth muscles facilitates hyolaryngeal elevation by pulling the hyoid bone up and forward (Gawryszuk et al., 2018). Also, contraction of thyrohyoid and suprahyoid muscles contributes to supraglottic shortening, allows for the closing of the laryngeal vestibule and thyrohyoid approximation (Daniels et al., 2019).

A study done by Macrae et al. (2013) compared coronal ultrasound images of submental muscles to MRI in 11 healthy individuals. Their findings showed significant and relatively high correlations (left: r = 0.909, p = 0.0001; right = 0.776, p = 0.005) between the two imaging methods, although CSA measurements of MRI were higher than ultrasound measurements. They concluded that ultrasound can be utilised to measure the CSA of anterior digastric muscle accurately because of the superior clarity of muscle borders of the floor of mouth muscles compared to MRI. However, the reliability of measures was not examined in the study.

Using the B-mode ultrasound, Watkin et al. (2001) examined the cross-sectional area of geniohyoid muscles. They queried if ultrasound could be used to measure the tissue composition of the geniohyoid muscle and muscle changes in cross-sectional muscle area due to age and radiotherapy. The study involved 12 participants including four normal young adults (aged 25 – 28 years), four normal older adults (aged 64 – 75 years), and four head and neck cancer patients following radiation therapy. They measured the mid-point of the geniohyoid muscle as the half-way point in the total number of still frames. A significantly greater cross- sectional area was reported in the patient group than normal young adults during rest and effortful speech tasks. Also, they observed greater muscle tissue variability in the younger group than the older group. These data indicate that ultrasound might provide a method to evaluate muscle changes resulting from age, radiation treatment, and rehabilitation in head and neck patients.

The findings of an early study by Emshoff et al. (1999) suggested that ultrasound may be a reliable diagnostic procedure for examining cross-sectional areas of anterior digastric muscles. A recent study analysed the inter-, intra-, and test-retest reliability of the geniohyoid muscle using ultrasound in 10 healthy adults and indicated good reliability (Shimizu et al., 2016). However, this study failed to verify the measurement method.

Figure 3 *Submental muscle group in coronal plane*



Note. A- left and right anterior digastric muscles, B- mylohyoid muscles, C- geniohyoid muscles (Macrae et al., 2013)

Assessment of Tongue Thickness

Ultrasound is broadly used for the functional analysis of dysphagia and has vast potential in clinical research for visualizing the tongue (Hsiao et al., 2012; Tamburrini et al., 2010; Tamura et al., 2012). The earliest work on using ultrasound for observing tongue movement was done by Shawker et al. (1983). Eight normal individuals and one patient with hypoglossal palsy participated in the study and they reported decreased tongue movement in the patient with dysphagia. The researchers also determined the timing and sequence of tongue activity and hyoid elevation during swallowing in healthy individuals using ultrasound.Further, Stone and Shawker (1986) recorded vertical and horizontal tongue movements by fixing a pellet to six normal individual's tongue surface. The results showed a good correlationbetween tongue movements and propulsion of bolus into pharynx.

Hsiao et al. (2012) attempted to identify the change in tongue thickness and laryngeal elevation cut-off value to predict the severity of dysphagia in stroke patients using submental ultrasonography. These researchers studied the maximum change in tongue thickness by determining the difference between the maximum and minimum tongue thickness during swallowing. They identified that, these measures could detect the need for tube feeding in dysphagic patients along with hyoid excursion measures. Any value less than 1.5 cm for hyoid bone displacement and less than 1.0 cm for tongue thickness change correlates with poor swallowing function and the need for tube feeding. With these observations, the authors suggested that ultrasound might provide a reliable method for predicting the necessity of tube feeding in dysphagic patients.

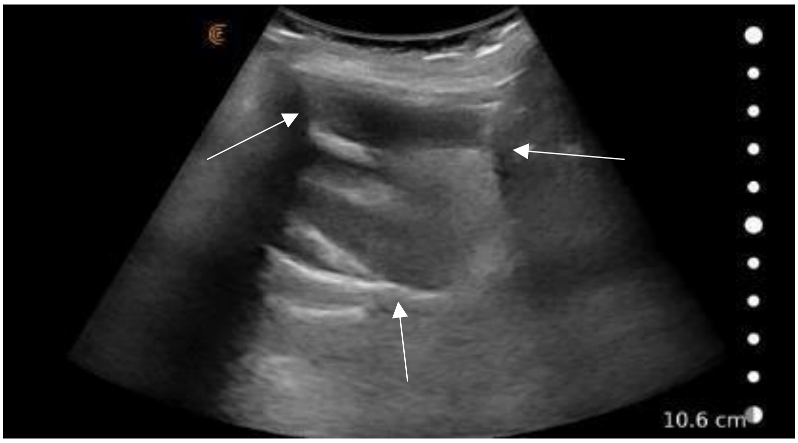
Tamura et al. (2012) employed ultrasonography to assess tongue thickness in the elderly population. The reliability of tongue thickness measurement using ultrasound has also been assessed. The study involved 104 healthy subjects between 70 & 90 years. Researchers explored the relationship between nutritional status and tongue thickness. They considered the measures of muscle area of midarm, triceps skinfold thickness, body weight and height to assess the nutritional status. A Similar correlation was obtained between tongue thickness and nutritional status measures such as body weight and midarm muscle area ($r = 0.424$; $p = 0$).

Another study by Tamburrini et al. (2010) assessed the role of video ultrasonography in nine patients with amyotrophic lateral sclerosis. The patients underwent static and dynamic ultrasound evaluation and videofluoroscopic examinations for both functional tongue evaluation and tongue morphometry. The researchers found that ultrasound could provide information about the initial alteration of static and dynamic factors associated with the oral phase of swallowing. The presence or absence of tongue atrophy was identified through the static evaluation of tongue thickness. Also, they identified patients with abnormal bolus position below the tongue in static evaluation using ultrasound. The authors suggested that ultrasound can be incorporated into diagnostic protocol for swallowing assessment in ALS patients as it has higher sensitivity than VFS in examining dynamic factors which represent the early signs of dysphagia. However, reliability analysis was not performed in this study.

Another study by Nakamori et al. (2016) utilised ultrasound for examining tongue thickness in ALS patients with a comparatively larger population than the previous research.

This study assessed tongue atrophy and tongue fasciculations in eighteen patients with ALS and eighteen age-matched healthy individuals using ultrasound and visual examination. Results suggested a significantly smaller tongue thickness in the ALS group (40.9 ± 1.0 mm, p = 0.016) than the healthy controls (44.6 ± 0.7 mm, p = 0.004). They also observed a notable progressive reduction in tongue thickness (p = 0.002) with the progression of disease (-0.315 mm/month). However, this study did not explore the reliability.

Figure 4 *Measurement of tongue thickness*



Note. Depicts the acoustic shadow cast by the hyoid bone on the right side and shadow of mandible on the left, and the surface of the tongue on the centre of the image.

Reliability of Ultrasound Measures of Swallowing

Ultrasound has been shown to have a potentially good application for the evaluation of swallowing biomechanics and reasonable correlation with videofluoroscopy (Ahn et al., 2015; Kuhl et al., 2003; Lee et al., 2016; Macrae et al., 2013; Tamburrini et al., 2010). However, reliability data is limited for the ultrasound assessment of swallowing. Prior to the implementation of medical instrumentation into standard clinical practice, reliability tests are necessary to avoid measurement errors (Portney and Watkins, 2009). Reliability gives us information about how consistent measurements can be acquired from a given method. An emphasis on the various reliability components is important for translating ultrasound technology into standard clinical practice. The main components that should be considered for the reliability of ultrasound assessment of swallowing include image acquisition, image selection and measurement of the selected image.

Reliability data of published literature are summarised in Table 1. Huang and colleagues (2009) reported high intra-and inter-rater reliability for ultrasound assessment of hyoid-larynx approximation in healthy individuals. Encouraging reliability data has also been reported for the measurement of the hyoid excursion. Hsiao et al. (2012) and Macrae et al. (2012) reported good reliability for the measurement of hyoid excursion in healthy individuals and another study (Chen et al., 2017) documented good reliability in dysphagic patients. While some studies have evaluated the reliability of the entire process (scanning procedure, image selection, and measurement) (Hsiao et al., 2012; Huang et al., 2009), other studies have excluded the scanning procedure (Macrae et al., 2012), or the scanning procedure and image selection (Chen et al., 2017) and reported only the reliability of the measurement technique in isolation. However, reliability data of the entire process of data acquisition is required for clinical translation.

Ultrasound assessment of muscle morphometry is also explored in several studies and is gaining evidence. Acceptable intra-rater reliability of CSA of the anterior belly of digastric muscles has been documented in patients with temporomandibular disorder (Emshoff et al., 1999) and in healthy individuals (Shimizu et al., 2016). The findings of these studies revealed that ultrasound may be a reliable diagnostic procedure for examining anterior digastric muscles. However, these studies lacked verification of measurement methods and therefore required further investigation. Further, moderate to good reliability has been reported for tongue thickness in patients with dysphagia (Hsiao et al., 2012) and in normal individuals (Tamura et al., 2012). These data suggest that, in a clinical setting, ultrasound assessment of tongue thickness may have the potential to predict the degree of dysphagia.

Table 1 *Summary of published reliability studies*

Authors	Measures	Participants	Inter-rater ICC	Intra rater ICC
Huang et al. (2009)	Thyrohyoid Approximation (percent displacement)	5 healthy	0.98	0.97 0.99
Chen et al. (2017)	Hyoid Excursion (absolute displacement)	10 dysphagic	0.89	0.99 0.96
Macrae et al. (2012)	Hyoid Excursion (absolute and percent displacement)	5 healthy	0.64 0.70	0.90 0.93
Hsiao et al. (2012)	Hyoid Excursion (absolute displacement)	10 healthy	0.80	0.93 0.84
Emshoff et al (1999)	Floor of mouth muscles	46 Temporomandibular disorder	-	0.91
Shimizu et al. (2016)	Floor of moth muscles	10 healthy	0.66 0.87	≥ 0.9
Hsiao et al. (2012)	Tongue thickness	60 dysphagic	0.69	0.76 0.66
Tamura et al. (2012)	Tongue thickness	104 healthy	-	0.86

These studies provide some promising data regarding the reliability of measures including hyoid excursion, thyrohyoid approximation, floor of mouth muscles, and tongue thickness using ultrasound. Though published reliability data of sophisticated ultrasound seems promising, it is not yet translated into standard clinical practice. One significant factor that hinders incorporation into clinical practice may be its cost. This may be overcome by smaller portable instrumentation. When considering the translation of this small device into clinical practice, analysis of reliability data would be of great value because good reliability is necessary to ensure the integrity of diagnostic results.

Validity and Reliability and of New Portable Ultrasound

Small portable ultrasound systems have been developed with technical progress, which may increase the potential for clinical translation of ultrasound devices. Utilizing such devices could be useful in swallowing rehabilitation for people with challenging access to videofluoroscopic evaluation such as those confined to bed or in remote facilities, and for repeated evaluation of hyolaryngeal displacement in patients (Rugiu, 2007). Such pocket-sized systems may overcome the cost obstacle compared to the large ultrasound device and may support clinical translation, but image quality requires evaluation (Winiker et al., 2019).

Winiker et al. (2019) evaluated the validity and reliability of a newly developed pocket-sized ultrasound (Clarius, Burnaby CA) in assessing hyolaryngeal excursion and thyrohyoid approximation in healthy individuals during swallowing. Researchers evaluated the validity of ultrasound with similar measures acquired from videofluoroscopic study using correlation analysis. ICC (intra-class correlation coefficient) and SEM (standard error of measurement) were measured within and across raters to evaluate reliability. Reliability of the entire process of data acquisition including the scanning, online image selection and measurement was examined to replicate the requirements of clinical application. Further, to evaluate the impact of different components of data acquisition on reliability, offline video and image measurement reliability was conducted.

The results suggested that the validity of ultrasound measurements using the Clarius™ system compared to videofluoroscopic measurement is poor for hyoid excursion and thyrohyoid approximation measures, except for hyoid excursion during dry swallowing ($r = 0.79$, $p \leq 0.001$). A moderate correlation ($r = 0.67$, $p \leq 0.002$) was found for hyoid excursion during liquid swallowing and no significant correlation during puree swallowing. Thyrohyoid approximation measure was found to have no significant association between modalities for any bolus types.

Findings of reliability study indicate insufficient reliability data for the assessment of swallowing measures using Clarius ™ system. Intra-rater reliability of offline measurement (image) of hyoid excursion range from poor to moderate (ICC: 0.49 – 0.64; CI: 0.23-0.79) and inter-rater reliability was moderate (ICC: 0.50; CI: 0.26 – 0.67). For thyrohyoid approximation, poor intra- and inter- rater reliability was reported (ICC > 0.50). However, good intra-rater reliability (ICC: 0.80 – 0.97; CI: 0.60 – 0.99) and moderate inter-rater reliability (ICC: 0.68; CI: 0.38 – 0.87) was revealed for floor of mouth muscles.

Although the authors did not directly compare reliability using this device to reliability using more traditional ultrasound technology, the authors hypothesised that this newer technology in swallowing assessment failed to meet the standards of large ultrasound device used in previous reliability studies (Hsiao et al., 2012; Huang et al., 2009; Macrae et al., 2012). Image quality likely effects reliability which consequently effects use in clinical practice (Steele, 2015) and this is an important factor for clear tissue interfaces which, in turn, facilitate reliable measurements (Ishida et al., 1992; Stone, 2005). Hence, to ensure the integrity of diagnostic results, it is essential to verify that the portable ultrasound produces good quality images.

Summary of Literature

The use of ultrasound imaging for swallowing is not a new development. Reliability data on ultrasound measurements in the assessment of swallowing are encouraging. However, this method has not yet translated into clinical practice, despite having multiple advantages. The cost, which may be one of the potential barrier for the translation can be overcome by using pocket-sized ultrasound devices. Recent research suggested insufficient validity and reliability for this new technology for clinical application, perhaps due to the reduced image quality of the device. To enhance the interpretation of ultrasound images, an image enhancement software is recently developed. The quality of this newly developed software needs to be examined to ensure the integrity of diagnostic results obtained from it.

Aim of Study

The aim of this research is to evaluate the influence of an image enhancement software on reliability of ultrasound measurement of swallowing images.

Hypothesis

It is hypothesized that the utilization of image enhancement software will result in increased inter- and intra- rater measurement reliability when compared to the analysis of original ultrasound images.

Method

Participants

Speech therapists with limited prior experience in ultrasound data acquisition or measurement were considered for the study, targeting four participants including the primary researcher. Colleagues in the research environment were notified of the study and given the opportunity to participate. Interested participants were required to contact the primary researcher of the study. All participants received an information sheet and a consent form and before the commencement of data collection. Once the participants familiarized themselves with their role in the project by reading the information sheet, they provided written informed consent. Ethical approval was received from the Human Ethics Committee of the University of Canterbury (HEC 2020/60).

Materials

Archived data of ultrasound images were used for the study (HDEC 17NTB214). A total of 40 images from healthy controls were analysed by the primary investigator and three co-investigators to derive intra-and inter-rater reliability, including five images each of:

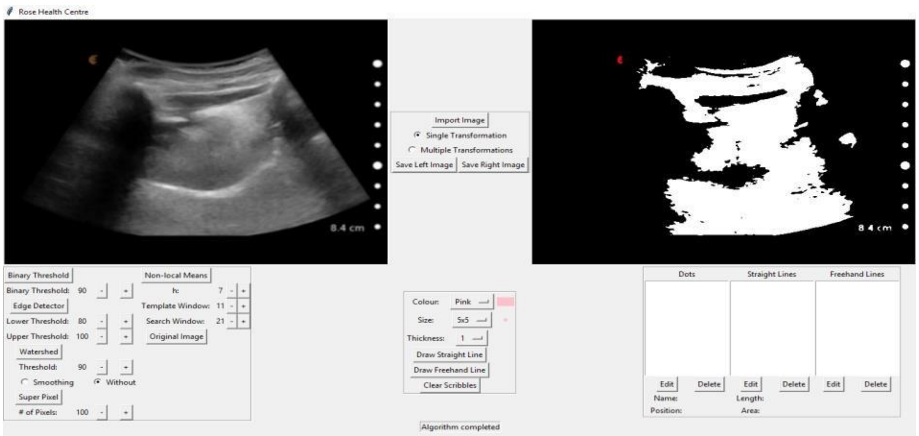
1. Distance between hyoid and mandible at rest.
2. Distance between hyoid and mandible at maximum swallowing.
3. Distance between thyroid and hyoid at rest
4. Distance between thyroid and hyoid at maximum swallowing
5. Tongue thickness
6. Cross-sectional areas
    a. Left anterior belly of digastric muscle
    b. Right anterior belly of digastric muscle and
    c. Geniohyoid[+] muscle

Instrumentation

Image Enhancement Software

Custom-designed image enhancement software[2] was used in the study to alter image quality. The software was designed to place the original image on one side and the modified adaptive image on the other side of the window (Figure 5). Based on image intensity and texture characteristics, software algorithms can be selected which automatically identify borders within the image. The regions of interest are then manually marked by the clinician. The following algorithms were included.

- ***Binary Thresholding (Figure 5)**:* This technique is a simple approach to segmenting an image by converting grayscale images into binary images, i.e., black and white. This is determined by checking each individual pixel to see if they are above or below a given threshold. For example, if the threshold was set to a numerical value 90, any pixel value in original image greater than 90 will be set to white, while pixel value less than 90 will be set to black. It can be used as a way to select areas of interest of an image, while ignoring the parts that are not of interest.
- **Figure 5** *Example of Binary Thresholding*

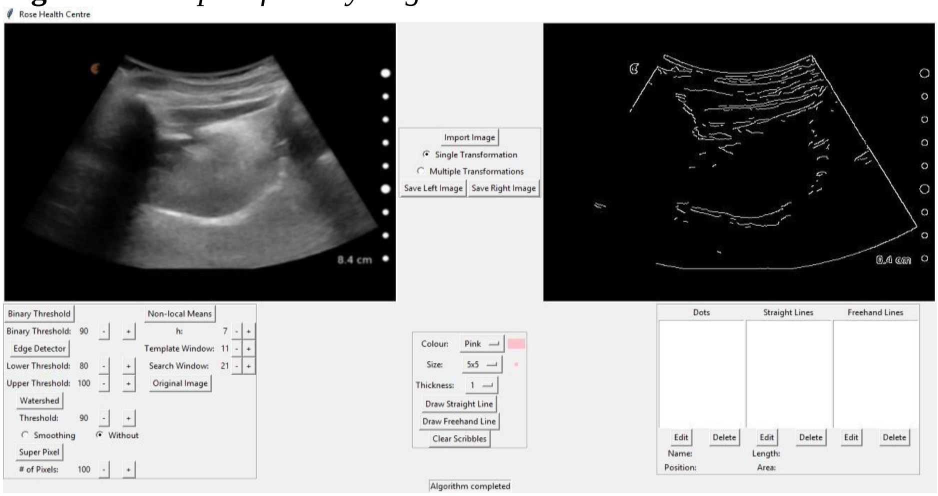


[2] Alan Brooks, 2020. University of Canterbury Summer Scholarship in Engineering

Note. Original image is shown on the left side modified on the right.

- ***Canny Edge Detection (Figure 6):*** This technique of image analysis uses a *multi-stage* algorithm to detect the edges within an image. First, it smooths the image to *eliminate noise*, then finds the *intensity gradient* and directions for each pixel in the image. The algorithm then carries out a full scan of the image and suppresses any pixel that is not at the maximum using *non-maximum suppression*. Final detection of edges is performed through *hysteresis,* which checks all leftover edges to see if they are proper edges or not. This process filters out useless information while preserving the important structural properties in an image.

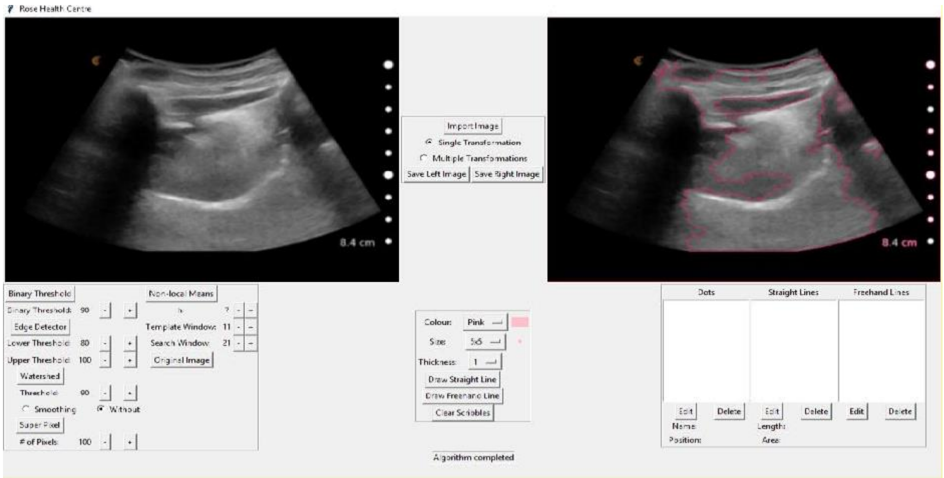
Figure 6 *Example of Canny Edge Detection*



Note. Original image is shown on the left side modified on the right.

- ***Watershed (Figure 7):*** This approach emulates a topographic map, with each point's brightness representing its height and finds the lines that run along the tops of ridges. It uses a customisable variable that identifies the boundary and finds the difference between the foreground and the background of images. It offers an additional smoothing option to smooth the created border. It can be used as a technique to provide a boundary to the images.

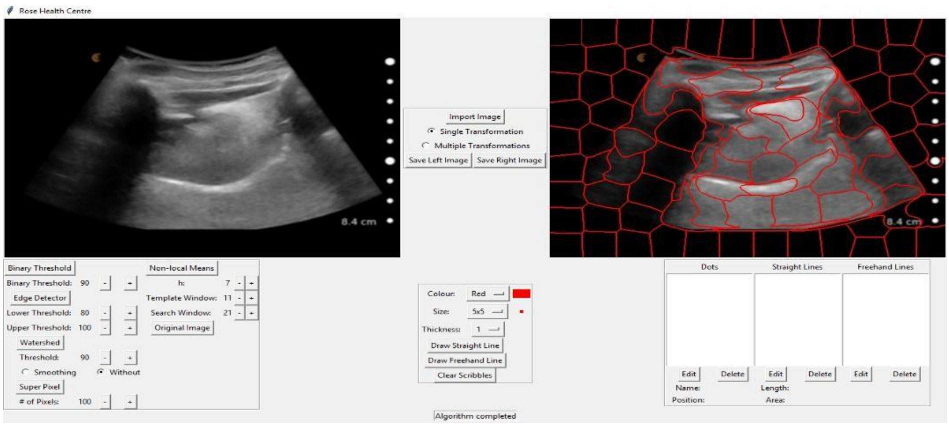
Figure 7 *Example of Watershed*



Note. Original image is shown on the left side modified on the right.

- ***Super Pixel (Figure 8)***: This technique find clusters of pixels with similar colour characteristics. And it can be described by an area as given by the result and each area has pixels with identical or similar characteristics. The customisation variable available for this algorithm is the number of super pixels the given image must be divided into. For example, 100 super pixel segmentation will be generated if the size is set to 100, while 200 segmentations will be found if it is increased to 200, thereby similar regions of image are grouped in similar super pixels.
-

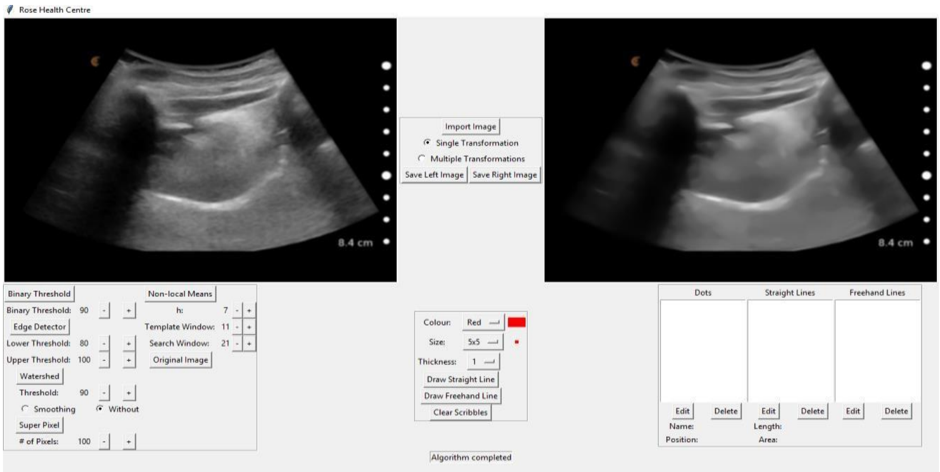
Figure 8 *Example of Super Pixel*



Note. Original image is shown on the left side modified on the right.

- ***Non-Local Means Denoising (Figure 9):*** This approach looks for similar patches of colour around the image and similar patches are then averaged and replaced with the average pixel and results in improved clarity of images.

Figure 9 *Example of Non- Local Means of Denoising*



Note. Original image is shown on the left side modified on the right.

Procedure

The proposed research was carried out in two stages.

Stage 1: Training for Investigators

Stage 2: Original image analysis and data analysis using image enhancement software

Training

To standardize performance and assure similar interpretation of data acquisition, clinical researchers were provided with an ultrasound guidebook for review in advance, which then followed by face-to-face training with hands on experience in identifying the measurement points of different original ultrasound images. The training was provided by a clinical researcher with prior experience in ultrasound measurement and software operation. The primary researcher demonstrated each parameter of the software and provided hands-on supervised practice using the image enhancement software. The clinical researchers familiarised themselves with the following features of software programme before the image analysis including:

- **Import console**: This button helps the user to import a selected file into the application and only allows the image file (png, jpg, and ppm) to import.
- **Single/Multiple transformations**: This button represents what image is getting transformed on the window. Single transformation allows the left hand image to be transformed by the selected algorithm and it is the commonly used button. Multiple transformation allows the right hand image to be transformed and it is used rarely since it starts editing the picture beyond recognition and functionality.
- **Algorithm Window (5 algorithms)**: The algorithm window allows the user to select different algorithms which creates new characteristics in images that were not present in the original image. The five algorithms include: binary thresholding, canny edge detection, watershed, super pixel and non-local means denoising.
- **Colour console:** This dropdown menu enables the user to customize the colour of their scribbles on the right-hand image. The selected colour will be applied to dot/line. In addition to this, size and thickness of the dot/line can also be changed accordingly.
- **List console:** This console displays the list of all currently drawn dots/lines on the images in their respective boxes. There are 3 list boxes i.e., dot, straight line and freehand line. In order to view information regarding a dot or a line, simply click the entry in the list, it will results in showing the length, position and area of that particular dot/line.

After familiarising themselves with the parameters of the software, each researcher was supported to practice interpretation with the software using different algorithms.

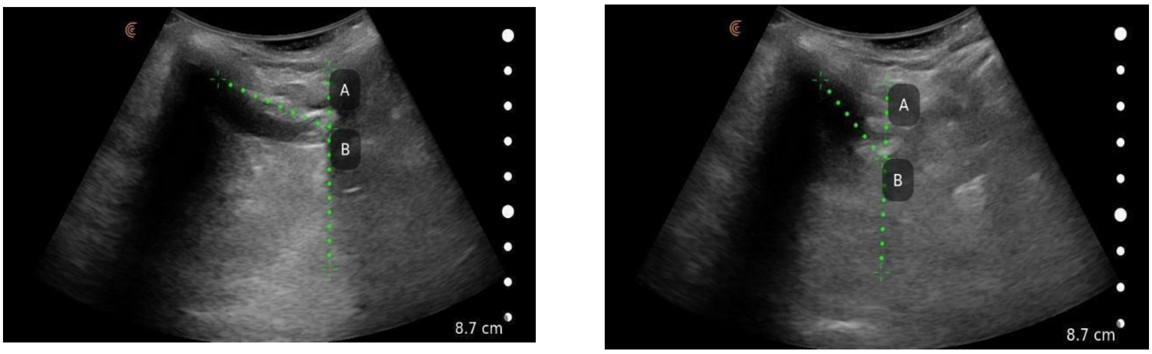
Data extraction

The primary researcher and three co-investigator then measured the archived images independently to derive a measure of inter-rater reliability. Each image was evaluated under two conditions: the original image and the image viewed through image enhancement software. When measuring the images using the image enhancement software, researchers were able to explore each image using all algorithms and base their measurements on any or all of the algorithms that best suit the image. They were asked to report the algorithms they selected for each image. Image analysis occurred across a four day period, with 20 trials per day, to minimise fatigue. In order to avoid the bias of familiarity of trials, all images were randomized within and across researchers. To avoid recall bias in the intra-rater reliability study, the four investigators completed a second round of analysis of the same images no sooner than one week after finishing the first round of analysis. The second analysis were performed using the same methods as the first analysis round. Therefore, a total of 160 images (40 images X 2 conditions X 2 trials) were measured by each researcher.

Ultrasound Measurement of Hyoid Excursion

For hyoid excursion, the measurement was done by drawing two straight lines where Line A represented a line of best fit along the anterior border of the hyoid shadow. Line B was drawn by placing one end of the calliper at the posterior border of the mandibular shadow's onset and the other end at the intersection point with the best fit line at the onset of the shadow of hyoid.

Figure 10 *Hyoid displacement at rest and maximal displacement*

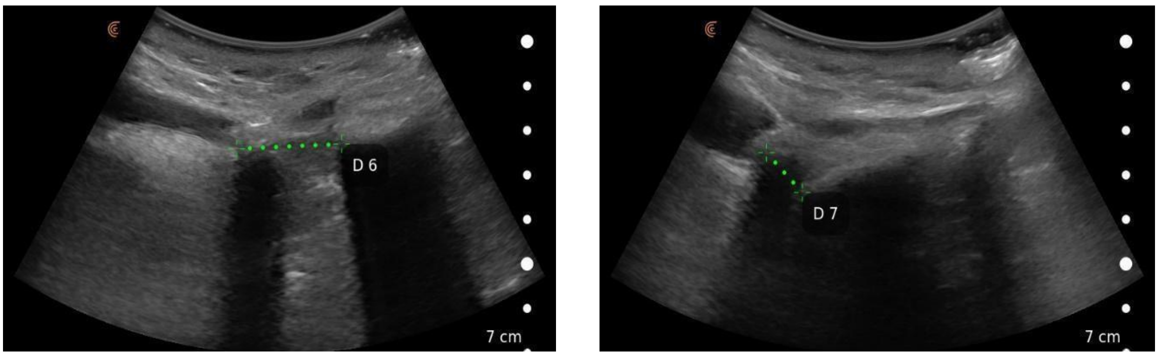


Note. Hyoid shadow is on the right side and mandible shadow is on the left

Ultrasound Measurement of Thyrohyoid approximation

Measurement was made by taking the distance between the upper border of the thyroid and hyoid by drawing a single straight line (Line D). For drawing the straight line, one calliper was placed at the beginning of the anterior border of the acoustic shadow of hyoid if consistently visible at rest and maximum or at the opacity depicting the hyoid. The other calliper was placed at either the bright opacity at the superior border of the thyroid cartilage or at the onset of the thyroid cartilage shadow if consistently visible at rest and maximum.

Figure 11 *Thyrohyoid approximation at rest and peak*



Ultrasound Measurement of Floor of Mouth Muscles

Floor of mouth muscles were measured by tracing the images representing the clearest muscle boundary using a freehand tool. Measurement of the paired geniohyoid muscles was made by tracing the region around the muscle and including the superior border of mylohyoid as border distinction is typically unclear. This is consequently referred to a geniohyoid^{+}. Left and right anterior belly of digastric muscles were calculated separately, and excluded by excluding the connective tissue taking only the region around outside of each muscle.

Figure 12 *Cross-sectional area of Geniohyoid muscle*

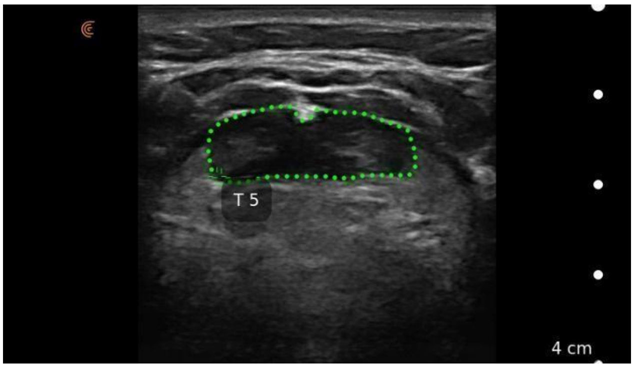
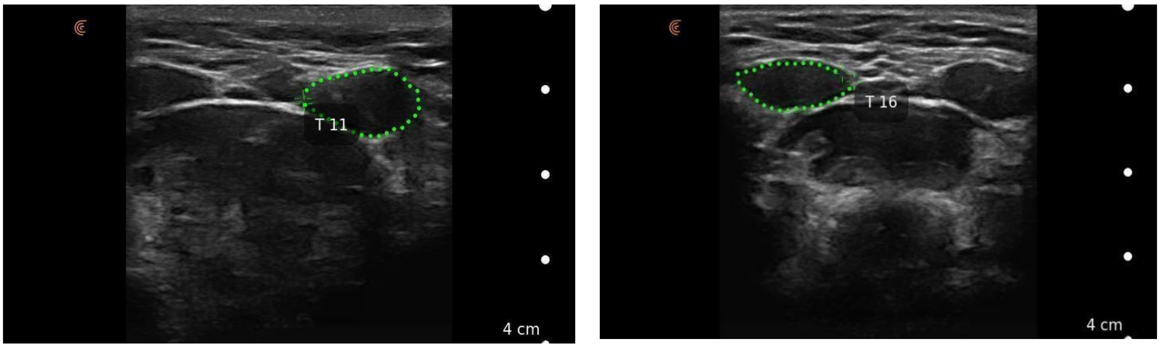


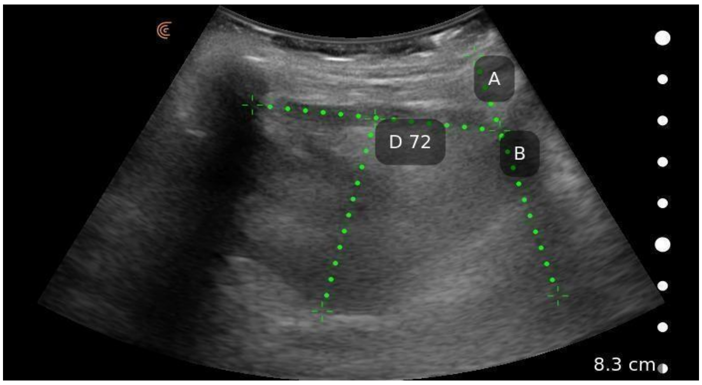
Figure 13 *Cross-sectional area of right and left anterior belly of digastric muscles*



Ultrasound Measurement of Tongue Thickness

Tongue thickness was measured by using three single straight lines bisecting the distance between mandible and hyoid shadow. A best-fit line (Line A) was drawn along the anterior border of the acoustic shadow of hyoid. For the other two lines, one end of the calliper was placed at the posterior aspect of the mandible shadow's onset whereas the other end was placed at the intersection point with the best fit line A at the onset of the shadow. And for Line D, one end of the single line was placed at the midpoint of the line between the hyoid and mandible acoustic shadow and the other end was placed at the posterior edge of the bolus (approximation of tongue to palate).

Figure 14 *Sonogram of tongue thickness measurement*



Data Analysis

Statistical analyses were carried out using R software (R Core Team, 2016) and lme4 (Bates et al., 2015) and the data were plotted using sactter and box plots. Mean, Standard deviation and Median were calculated for each ultrasound measure (Hyoid rest, Hyoid max, Thyrohyoid rest, Thyrohyoid max, FOM) across participants. ICCs with confidence intervals were calculated for intra-and inter-rater reliability and this was based on linear mixed effect analysis. Confidence intervals indicated the extend of uncertainty in the estimation of measures.

Archived data of 40 ultrasound images were intended for this reliability study and analysis were completed for the same. For the purpose of this Master thesis, statistical analysis focused only on 30 images, giving the intra-and inter-rater reliability of ultrasound measurement comparing enhanced and original ultrasound images. For intra-and inter-rater reliability assessment, following ICC analysis were completed:

1. Distance between hyoid & mandible at rest.
2. Distance between hyoid & mandible at peak.
3. Distance between thyroid & hyoid at rest
4. Distance between thyroid & hyoid at peak.
5. CSA of the anterior belly of digastric muscles
    - Left anterior belly
    - Right anterior belly

Linear mixed effect model analysis were utilised for both intra and inter-rater reliability. Measurement trial was entered as fixed and subject as random effect for intra-rater reliability and both subject and rater entered as random effects for inter-rater reliability. Data from the second measurement trial were considered for the calculation of inter-rater reliability.

Assumption of normality and assumption of homoscedasticity of the residuals were checked for all the measures. A bootstrap distribution was obtained to indicate the uncertainty of the estimates. This was calculated based on the ICC and their confidence interval, taking the CIs for each ICC. Variance pattern of the residuals were detected using residual versus fitted plots. Normality of the data was examined through the visual inspection of residual Q-Q plots (quantile-quantile). Statistical analysis using Shapio-Wilk"s test was also performed. A p value of ≤ 0.05 was considered significant. Guidelines by Portney and Watkins (2009) were considered for the interpretation of the results:

ICC < 0.50 - poor reliability

ICC 0.50-0.75 - moderate reliability

ICC > 0.75 - good reliability.

Results

Descriptive Statistics

The descriptive statistics of ultrasound measures included mean, standard deviation and median for all the ultrasound measures are displayed in Table 2. There is greater variability in the mean measures of LAB and RAB (normal and enhanced) as compared to the mean measures of hyoid excursion and thyrohyoid excursion at rest and maximum (normal and enhanced). From the table, it is clear that measures of thyrohyoid approximation at rest and maximum (normal and enhanced) is showing greater uniformity as compared to other measures.

Table 2 *Ultrasound measures: Mean, Standard deviation, and median*

Measure	Mean (SD)	Median
Hyoid rest normal	41.77 (6.17)	41.75
Hyoid rest enhanced	38.75 (7.50)	37.25
Hyoid max normal	34.13 (5.68)	34.15
Hyoid max enhanced	35.10 (3.32)	35.00
TH rest normal	28.09 (5.57)	28.30
TH rest enhanced	27.73 (6.57)	25.80
TH max normal	20.86 (8.00)	21.60
TH max enhanced	22.36 (8.94)	21.60
LAB and RAB normal	82.41 (46.15)	73.50
LAB and RAB enhanced	74.28 (22.63)	71.50

Reliability

On analysis, assumptions were not met for some measures and those are noted in the Table 2 & 3. The values which does not met the assumption of normality and homoscedasticity of the residuals are denoted in the table. These results should be interpreted with caution.

Intra-rater Reliability

Findings of intra-rater reliability are displayed in Table 3. Results reveals good intra-rater reliability for hyoid excursion of enhanced images across all raters as compared to normal ultrasound images which ranges from moderate to good reliability. However, large confidence intervals were noted for both enhanced and normal images.

ICC values for thyrohyoid approximation were found to be within the same range for both enhanced and normal images for all raters, ie, poor - good reliability. Significantly higher ICC values with larger confidence intervals were noted for these measure (ICC: 0.08-0.99; CI: 0.00 – 1.0). This indicates a high degree of variability for these measures. Moderate to good reliability was revealed for the enhanced images of floor of mouth muscles as compared to normal images which indicates poor to good reliability across the four raters. Notably, ICC and their confidence intervals are considerably high for both normal and enhanced images.

Inter-rater Reliability

Table 4 reports the findings of inter-rater reliability. ICC values for inter-rater reliability of enhanced images of hyoid excursion were reported as good while moderate for normal images. For thyrohyoid approximation, moderate inter-rater reliability was reported for both enhanced and normal images. Good reliability was revealed for the enhanced images of floor of mouth muscles, however poor reliability was reported for normal images.

Table 3 *Intra-rater reliability*

Measure	ICC (95% CI) Rater 1	ICC (95%CI) Rater 2	ICC (95%CI) Rater 3	ICC (95% CI) Rater 4
H normal	0.95[a,b] (0.83,0.99)	0.87[b] (0.58, 0.97)	0.71 (0.27, 0.92)	0.67 (0.14, 0.91)
H enhanced	0.86[a] (0.58, 0.96)	0.88 (0.62, 0.97)	0.81 (0.48, 0.96)	0.89[a] (0.65, 0.97)
TH normal	0.98 (0.94, 1.0)	0.33 (0.00, 0.79)	0.75 (0.33, 0.94)	0.72 (0.20, 0.93)
TH enhanced	0.99[a] (0.96, 1.0)	0.08 (0.00, 0.69)	0.89 (0.64, 0.97)	0.82 (0.50, 0.96)
FOM normal	0.97 (0.88, 0.99)	0.34[a] (0.00, 0.81)	0.60 (0.05, 0.89)	0.88[a] (0.63, 0.98)
FOM enhanced	0.93 (0.77, 0.98)	0.80 (0.43, 0.95)	0.74 (0.31, 0.93)	0.81 (0.42, 0.95)

Note. [a] – assumption of normality may not have been met, [b] – assumption of homoscedasticity may not have been met for the condition.

Table 4 *Inter-rater Reliability*

Measure	ICC (95% CI)
H normal	0.60 (0.20, 0.82)
H enhanced	0.77 (0.44, 0.91)[a,b]
TH normal	0.65 (0.27, 0.84)
TH enhanced	0.50 (0.16, 0.75)
FOM normal	0.39 (0.03, 0.67)[a]
FOM enhanced	0.88 (0.65, 0.95)

Note. [a] - assumption of normality may not have been met.

[b] – assumption of homoscedasticity may not have been met,

Algorithm Selection

Table 5 provides the summary of the percentage of algorithms used by four raters for different measures across two round of analysis. Out of five algorithms, non-local denoising was the preferred choice across the four raters. Watershed was the second preferred algorithm among the raters. Binary thresholding and canny edge detection were the least favoured algorithms by the raters across two round of analysis.

Table 5 *Algorithms*

Algorithms	First round	Second round	Total
Non-local denoising	36.2%	43.2%	39.7%
Watershed	30%	30%	30%
Super Pixel	26.9%	21.2%	24.1%
Binary thresholding	3.2%	1.8%	2.5%
Canny Edge Detection	3.7%	3.8%	3.7%

Discussion

This is the first study to explore the reliability of ultrasound measurement of swallowing using image enhancement software. Measures considered for this reliability study included hyoid excursion, thyrohyoid approximation, and floor of mouth muscle morphometry. The findings of this study suggest only slight benefits for using image enhancement software for measuring ultrasound images as compared to original ultrasound measures.

The findings of prior reliability data using a portable device (Winiker et al., 2019) in healthy individuals are not directly comparable to our study results. Results of her study suggested that measurement using the Clarius™ system was insufficiently reliable for the acquisition of swallowing measures, with large SEMs and low ICCs for intra- and inter-rater reliability for the mean hyoid excursion and thyrohyoid approximation measure.

Reliability data obtained from another study by Hammond (2019) using the same portable device - Clarius™ system- in patients with dysphagia indicates that the raters could achieve a high agreement level when completing offline measurement of pre-selected images. All measures were found within good to excellent range in her study. Findings of this study contrast with the reliability outcomes of our study. Reliability data derived from the original images in our study demonstrated moderate to good intra-rater reliability for hyoid excursion and poor to good intra rater-reliability for thyrohyoid approximation and floor of mouth muscles. Inter-rater reliability was moderate for dynamic measures and poor for static measures. Both the previous reliability studies show a discrepancy between the results for the offline measurement of still images, and this may be due to the difference in homogeneity of data of the two studies. However, online measurement findings of both these studies showed considerably low ICC Values. This indicates a significant reduction in image selection and online measurement of data when using portable ultrasound technology in pressured clinical environment.

The findings of our study indicated a reduction in ICC values with large confidence intervals for original ultrasound images for all the measures compared to the offline measurement of pre-selected images (Hammond, 2019). One aspect that accounts for the difference in reliability is the technical issues of the newly developed software. While performing the data analysis, the software's smooth run was interrupted by bugs. Other technical problems, including malfunctioning of algorithms performance, may also likely impact our study's reliability. Additionally, while measuring, a part of freehand line of the previous image would remain on the screen even after selecting the next image. This may likely interrupt the measurement. For the purpose of offline measurement of still images, previous reliability studies utilised ImageJ analysis software, where individual rater measure and save the still images as Jpeg file. ImageJ software can adjust the brightness and contrast of the images to obtain improved image quality (Schneider et al., 2012), whereas these options are restricted in the new software. Another drawback of this software is that any image size above the specified pixel value (458 pixels high and 800 pixels width) may cause blurriness in the image quality and may affect the accuracy and spatial resolution of images. All these factors may affect the reliability findings.

Apart from technical issues, other aspects that likely impacted the results may be evaluating only still images in our study. Newly developed image enhancement software may not be applicable to ultrasound videos. Previous reliability study (Hammond, 2019) utilised video segments for the frame selection of each image where the reviewer can identify specific landmarks needed for measurement, such as echoic shadow to use for hyoid excursion and thyrohyoid approximation. This may likely increase the reliability in their study.

The newly developed image enhancement software utilised in this study did not favourably increase the reliability. There was a slight increase in reliability noted for enhanced images compared to the original images for hyoid excursion and floor of mouth muscles. Intra- and inter-rater reliability results were found to be similar across both normal and enhanced images for thyrohyoid approximation. Notably, for most ICCs, there was evidence of large confidence intervals for both normal and enhanced images across the analysed measures, which expresses uncertainty regarding the reliability estimates.

Portable ultrasound has the potential to provide an objective assessment to vulnerable patients with affordable care, where this device eliminates the challenge of transporting these patients to hospital-based diagnostic services. However, the research is limited in this area but still encouraging which indicates its potential for clinical translation. Prior research explored the reliability and validity of this portable ultrasound and found out the need for further investigation on the image quality of the acquired ultrasound images. Our study aimed to examine whether the newly developed image enhancement software could improve the quality of ultrasound images. Thus, this study focused on investigating the reliability of the software since reliable measurements are important to establish the integrity of diagnostic results. However, only superficial benefits are identified from the data generated in this study using image enhancement software comparing original ultrasound measures. These differences in measurements are not potentially large enough to be clinically significant. Though the proposed research does not indicate favourable results, this study sets the foundation for a different method for measurement analysis of the newly developed portable ultrasound device.

Limitation

Firstly, technical issues with the software are important to acknowledge. The image size of the software is not the full size of the screen and the maximum size for the canvas is 458 pixels high and 800 pixels width. Any image size greater than this may results in blurred images after re-sizing. Also, this software does not work well with laptops because of the smaller size of the screen. It can perform on desktop systems except for mac. The smooth run of the software was interrupted most of the time because of bugs. A part of the freehand line would bide on the images even after resolving the issues finally. All these technical issues may impact the results.

The sample size used in this study was small. Only three measures were considered in this study including hyolaryngeal excursion, thyrohyoid approximation and floor of mouth muscles with overall 30 images. A larger sample size is necessary to ensure reliability since the findings of this study cannot be generalised alone because of the smaller sample size.

The image enhancement software utilised in this study cannot be used for video processing; it can only be used to analyse the images. It is unknown whether this software works with other portable ultrasound machines since we considered the measurement analysis only from Clarius^{TM} system.

Conclusion

This research documents the substantial attempt aimed at the development of an image enhancement software for measuring ultrasound images. There was evidence of a slight increase in reliability for enhanced images than normal images. This software does not have significant benefits that helps in the clinical translation of portable ultrasound, but the review highlights the need for future research for measurement analysis.

References

Abramovich, D. R., Garden, A., Jandial, L., & Page, K. R. (1979). Fetal swallowing and voiding in relation to hydramnioa. *Obstetrics and Gynecology*, *54,* 15-20.

Ahn, S. Y., Cho, K. H., Beom, J., Park, D. J., Jee, S., & Nam, J. H. (2015). Reliability of ultrasound evaluation of hyoid-larynx approximation with positional change. *Ultrasound in Medicine and Biology, 41*(5), 1221-1225. https://doi.org/10.1016/j.ultrasmedbio

Alanen, Falck, Kalimo, Komu, & Sonninen. (1994). Ultrasound, computed tomography and magnetic resonance imaging in myopathies: correlations with electromyography and histopathology. *Acta neurologica scandinavica, 89*(5), 336-346.

Aldrich. (2007). Basic physics of ultrasound imaging. *Critical care medicine*, *35*(5), S131-S137.

Bowie, J. D., & Clair, M. R. (1982). Fetal swallowing and regurgitation: observation of normal and abnormal activity. *Radiology*, *144,* 877-878.

Bressmann, T., Heng, C. L., & Irish, J. (2005). Applications of 2D and 3D ultrasound imaging in speech-language pathology. *Journal of Speech-Language Pathology and Audiology, 29*(4), 158-168.

Chen, Y.-C., Hsiao, M.-Y., Wang, Y.-C., Fu, C.-P., & Wang, T.-G. (2017). Reliability of Ultrasonography in Evaluating Hyoid Bone Movement. *Journal of Medical Ultrasound*, *25*(2), 90–95. https://doi.org/ 10.1016/j.jmu

Chi-Fishman, G., & Sonies, B. C. (2002). Effects of Systematic Bolus Viscosity and Volume Changes on Hyoid Movement Kinematics. *Dysphagia*, *17*(4), 278–287. https://doi.org/ 10.1007/s00455-002-0070-7

Chi-Fishman, G., & Sonies, B. C. (2002). Kinematic Strategies for Hyoid Movement in Rapid Sequential Swallowing. *Journal of Speech, Language, and Hearing Research*, *45*(3), 457–468. https://doi.org/10.1044/1092-4388(2002/036)

Chi-Fishman, Hicks, Cintas, Sonies, & Gerber. (2004). Ultrasound imaging distinguishes between normal and weak muscle. *Archives of physical medicine and rehabilitation, 85*(6), 980-986.

Chi-Fishman, G. (2005). Quantitative lingual, pharyngeal and laryngeal ultrasonography in swallowing research: A technical review. *Clinical Linguistics & Phonetics, 19*(6/7), 589-604.

Cordaro, M. A. & Sonies, B. C. (1993). An image processing scheme to quantitatively extract and validate hyoid bone motion based on real-time ultrasound recordings of swallowing. *IEEE Transactions on Biomedical Engineering, 40*(8), 841-844.

Dodds, Stewart, & Logemann. (1990). Physiology and radiology of the normal oral and pharyngeal phases of swallowing. AJR. American journal of roentgenology, 154(5), 953-963.

Epstein, M. A. (2005). Ultrasound and the IRB. *Clinical Linguistics & Phonetics, 19* (6/7), 567-572.

Emshoff, R., Bertram, S., & Strobl, H. (1999). Ultrasonographic cross-sectional characteristics of muscles of the head and neck. *Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology and Endodontology, 87*(1), 93-106. https://doi.org/10.1016/S1079-2104(99)70302-1

Filoux, E., Mamou, J., Aristizábal, O., & Ketterling, J. A. (2011). Characterization of the spatial resolution of different high-frequency imaging systems using a novel anechoic-sphere phantom. *IEEE transactions on ultrasonics, ferroelectrics, and frequency control,* 58(5), 994–1005. https://doi.org/10.1109/TUFFC.2011.1990

Fink, Martin, & Rohrmann. (1979). Biomechanics of the human epiglottis. *Acta oto-laryngologica, 87*(3-6), 554-559.

Gawryszuk, A., Bijl, H. P., Holwerda, M., Halmos, G. B., Wedman, J., Witjes, M. J., Dorgelo, B., & Langendijk, J. A. (2019). Functional swallowing units (FSUs) as organs-at-risk for radiotherapy. PART 1: physiology and anatomy. *Radiotherapy and Oncology*, 130, 62-67.

Hartl, D. M., Kolba, F., Bretagne, E., Bidault, F., & Sigal, R. (2010). Cine-MRI swallowing evaluation after tongue reconstruction. *European Journal of Radiology, 73,* 108-113. https://doi.org/10.1016/j.ejrad.2008.10.005.

Heggie, J. C. P., Liddell, N. A., & Maher, K. P. (2001). *Applied imaging technology* (4th ed.). St. Vincent's Hospital.

Hiiemae, K.M. & Palmer, J.B. (2003). Tongue movements in feeding and speech.

Critical Reviews in Oral Biology & Medicine, *14*(6), 413-429.

Hsiao, M.-Y., Chang, Y.-C., Chen, W.-S., Chang, H.-Y., & Wang, T.-G. (2012). Application of Ultrasonography in Assessing Oropharyngeal Dysphagia in Stroke Patients. *Ultrasound in Medicine & Biology*, *38*(9), 1522–1528. https://doi.org/10.1016/j.ultrasmedbio.2012.04.017

Huang, Y.-L., Hsieh, S.-F., Chang, Y.-C., Chen, H.-C., & Wang, T.-G. (2009). Ultrasonographic Evaluation of Hyoid–Larynx Approximation in Dysphagic Stroke Patients. *Ultrasound in Medicine & Biology*, *35*(7), 1103–1108. https://doi.org/10.1016/j.ultrasmedbio.2009.02.006

Huckabee, M.-L., Macrae, P., & Lamvik, K. (2015). Expanding Instrumental Options for Dysphagia Diagnosis and Research: Ultrasound and Manometry. *Folia Phoniatrica Et Logopaedica*, *67*(6), 269–284. https://doi.org/: 10.1159/000444636

Ishida, Y., Carroll, J., Pollock, M., Graves, J., & Legett, S. (1992). 662 Reliability of B-Mode Ultrasound in this Measurement of Body Fat and Muscle Thickness. *Medicine & Science in Sports & Exercise*, *22*(2). https://doi.org/ 10.1249/00005768-199004000-00661

Jensen, J. A. (2007). Medical ultrasound imaging. *Progress in Biophysics and Molecular Biology*, 93(1–3), 153-165. https://doi.org/10.1016/j.pbiomolbio.2006.07.025
Kane, D., Grassi, W., Sturrock, R., & Balint, P. V. (2004). A brief history of musculoskeletal ultrasound: 'From bats and ships to babies and hips.' *Rheumatology, 43,* 931-933.

Kossoff, G. K. (2000). Basic physics and imaging characteristics of ultrasound. *World Journal of Surgery, 24,*134-142.

Kristensen, M. S. (2011). Ultrasonography in the management of the airway. Acta Anaesthesiologica Scandinavica, 55(10), 1155-1173. https://doi.org/10.1111/j.1399-6576.2011.02518.x

Kuhl, V., Eicke, B. M., Dieterich, M., & Urban, P. P. (2003). Sonographic analysis of laryngeal elevation during swallowing. *Journal of Neurology*, *250*(3), 333–337. https://doi.org/10.1007/s00415-003-1007-2

Kundra, Mishra, & Ramesh. (2011). Ultrasound of the airway. *Indian journal of anaesthesia,* 55(5), 456.

Lee, Y. S., Lee, K. E., Kang, Y. K., Yi, T. I., & Kim, J. S. (2016). Usefulness of submental ultrasonographic evaluation for dysphagia patients. *Annals of Rehabilitation Medicine, 40*(2), 197-205. https://doi.org/10.5535/arm.2016.40.2.197

Leonard, R. J., Kendall, K. A., Mckenzie, S., Gonçalves, M. I., & Walker, A. (2000). Structural Displacements in Normal Swallowing: A Videofluoroscopic Study. *Dysphagia*, *15*(3), 146–152. https://doi.org/10.1007/s004550010017

Li, C., Li, J., Zhang, C., Cao, X., Li, N., Song, D., & Yu, T. (2015). Application of B+M-mode ultrasonography in assessing deglutitive tongue movements in healthy adults. *Medical Science Monitor*, 21, 1648-1655. https://doi.org/10.12659/MSM.893591

Logemann, Kahrilas, Cheng, Pauloski, Gibbons, Rademaker, & Lin. (1992). Closure mechanisms of laryngeal vestibule during swallow. *American Journal of Physiology-Gastrointestinal and Liver Physiology*, *262(*2), G338-G344.

Logemann, J. A. (1998). *Evaluation and treatment of swallowing disorders* (2nd ed.). United States of America: pro-ed.

Macrae, P. R., Doeltgen, S. H., Jones, R. D., & Huckabee, M.-L. (2011). Intra- and inter-rater reliability for analysis of hyoid displacement measured with sonography. *Journal of Clinical Ultrasound*, *40*(2), 74–78. https://doi.org/10.1002/jcu.20874

Mann, W. J. (2019). History of head and neck ultrasonography. In *Head and Neck Ultrasonography.* (pp. 1-8). Springer, Cham.

Martin, K. (2010). Properties, limitations and artefacts of B-mode images. In P. Hoskins, K. Martin, & A. Thrush (Eds.), *Diagnostic Ultrasound: Physics and Equipment*, 64-74.

Martin, K. & Ramnarine, K. (2010). Physics. In P. Hoskins, K. Martin, & A. Thrush (Eds.), *Diagnostic Ultrasound: Physics and Equipment*, 4-22. http://ebooks.cambridge.org/chapter.jsf?bid=CBO9780511750885 &cid=CBO9780511750885A011

Merton, D. A. (1997). Diagnostic medical ultrasound technology: A brief historical review. *Journal of Diagnostic Medical Sonography*, *13*, 10S-23S, https://doi.org/10.1177/875647939701300i503.

Matsuo, K., & Palmer, J. B. (2008). Anatomy and physiology of feeding and swallowing: *Normal and abnormal. Physical Medicine and Rehabilitation Clinics of North America*, *19*(4), 691. https://doi.org/10.1016/j.pmr.2008.06.001

Miller, J. L., & Watkin, K. L. (1997). Lateral pharyngeal wall motion during swallowing using real time ultrasound. Dysphagia, 12(3), 125-132. https://doi.org/10.1007/pl00009526

Minifie, Kelsey, Zagzebski, & King. (1971). Ultrasonic scans of the dorsal surface of the tongue. The Journal of The Acoustical Society of America, 49(6B), 1857-1860.

Morrish, Stone, Sonies, Kurtz, & Shawker. (1984). Characterization of tongue shape. *Ultrasonic Imaging, 6*(1), 37-47.

Nakamori, M., Hosomi, N., Takaki, S., Oda, M., Hiraoka, A., Yoshikawa, M., Matsumoto, M. (2016). Tongue thickness evaluation using ultrasonography can predict swallowing function in amyotrophic lateral sclerosis patients. *Clinical Neurophysiology, 127*(2), 1669-1674. https://doi.org/10.1016/j.clinph.2015.07.032

Newman, P. G., & Rozycki, G. S. (1998). The history of ultrasound. *Surgical clinics of north America*, *78*(2), 179-195.

Orloff, L. A., & Shindo, M. L. (2017). Normal Cervical Lymph Node Appearance and Anatomic Landmarks in Neck Ultrasound. *In Advanced Thyroid and Parathyroid Ultrasound* (pp. 207-214). Springer, Cham.

Panebianco, V., Ruoppolo, G., Pelle, G., Schettino, I., Roma, R., Bernardo, S., et al. (2010). Morpho-functional patterns of physiologic oropharyngeal swallowing evaluated with dynamic fast MRI. *European Archives of Otorhinolaryngology, 267,* 1461-1466. https://doi.org/10.1007/s00405-010-1232-0.

Pearson, W. G., Molfenter, S. M., Smith, Z. M., & Steele, C. M. (2012). Image-based Measurement of Post-Swallow Residue: The Normalized Residue Ratio Scale. Dysphagia, 28(2), 167–177. https://doi.org/ 10.1007/s00455-012-9426-9

Peng, Jost-Brinkmann, & Miethke. (1996). The cushion scanning technique: a method of dynamic tongue sonography and its comparison with the transducer-skin coupling scanning technique during swallowing. Academic radiology, 3(3), 239-244.

Perry, S. E., Winkelman, C. J., & Huckabee, M.-L. (2016). Variability in Ultrasound Measurement of Hyoid Bone Displacement and Submental Muscle Size Using 2 Methods of Data Acquisition. *Folia Phoniatrica Et Logopaedica*, *68*(5), 205–210. https://doi.org/10.1159/000473876

Ploquin, A., Baldini, C., Vuagnat, P., Makhloufi, S., Desauw, C., & Hebbar, M. (2015). Prolonged survival in a patient with a pancreatic acinar cell carcinoma. *Case reports in oncology, 8*(3), 447-450.

Robbins, Hamilton, Lof, & Kempster. (1992). Oropharyngeal swallowing in normal adults of different ages. *Gastroenterology, 103*(3), 823-829.

Rugiu, M. G. (2007). Role of videofluoroscopy in evaluation of neurologic dysphagia. *Acta Otorhinolaryngologica Italica,* 27(6), 306-316.

Shawker, T. H., Sonies, B., Stone, M., & Baum, B. J. (1983). Real-time ultrasound visualization of tongue movement during swallowing. *Journal of Clinical*

Ultrasound, 11, 485-490.

Schedel, H., Reimers, C. D., Nägele, M., Witt, T. N., Pongratz, D. E., & Vogl, T. (1992). Imaging techniques in myotonic dystrophy. A comparative study of ultrasound, computed tomography and magnetic resonance imaging of skeletal muscles. *European journal of radiology, 15*(3*),* 230-238.

Shawker, T. H., Sonies, B. C., & Stone, M. (1984). Soft tissue anatomy of the tongue and floor of the mouth: An ultrasound demonstration. *Brain and Language*, 21(2), 335–350.

Shawker, T. H., Sonies, B., Hall, T. E., & Baum, B. F. (1984). Ultrasound analysis of tongue, hyoid, and larynx activity during swallowing. *Investigative Radiology, 19,* 82-86.

Shawker, T. H., Doppman, J. L., Dunnick, N. R., & McCarthy, D. M. (1982). Ultrasonic investigation of pancreatic islet cell tumors. *Journal of Ultrasound in Medicine, 1*(5), 193-200.

Shimizu, S., Hanayama, K., Metani, H., Sugiyama, T., Abe, H., Seki, S., Hiraoka, T., & Tsubahara, A. (2016). Retest reliability of ultrasonic geniohyoid muscle measurement. *Japanese Journal of Comprehensive Rehabilitation Science, 7,* 55-60. https://doi.org/10.11336/jjcrs.7.55

Shung, K. K. (2006). Diagnostic ultrasound: *Imaging and blood flow measurements.* http://www.crcnetbase.com.ezproxy.canterbury.ac.nz/isbn/9780824740962.

Singh, M., Chin, K. J., Chan, V. W., Wong, D. T., Prasad, G. A., & Yu, E. (2010). Use of sonography for airway assessment: An observational study. *Journal of Ultrasound in Medicine, 29*(1), 79-85. https://doi.org/10.7863/jum.2010.29.1.79

Singh, K., Singh, S., Gupta, R., Gathwal, C., Bansal, P., & Singh, M. (2017). A feasibility study to assess vallecula and pyriform sinus using protocol-based ultrasonic evaluation of floor of mouth and upper airway. *Saudi Journal of Anaesthesia, 11*(3), 299-304. https://doi.org/10.4103/1658-354X.206799

Sipila, S. & Suominen, H. (1993). Muscle ultrasonography and computed tomography in elderly trained and untrained women. *Muscle & Nerve, 16,* 294-300.

Sonies, Shawker, Hall, Gerber, & Leighton. (1981). Ultrasonic visualization of tongue motion during speech. The Journal of the Acoustical Society of America, 70(3), 683-686.

Sonies, Parent, Morrish, & Baum. (1988). Durational aspects of the oral-pharyngeal phase of swallow in normal adults. Dysphagia, 3(1), 1-10.

Sonies, B. C., Chi-Fishman, G., & Miller, J. L. (2003). Ultrasound Imaging and Swallowing. *Normal and Abnormal Swallowing*, 119–138. https://doi.org/10.1007/978-0-387-22434-3_8

Steele, C. M. (2015). The blind scientists and the elephant of swallowing: A review of instrumental perspectives on swallowing physiology. *Journal of Texture Studies, 46*(3), 122-137. https://doi.org/10.1111/jtxs.12101

Stevens, D. (1978). "Ultrasound swallow". *British Medical Journal*, *2*(6154), 1789-1790.

Stone, M., & Davis, E. P. (1995). A head and transducer support system for making ultrasound images of tongue/jaw movement. *The Journal of the Acoustical Society of America, 98*(6), 3107-3112. https://doi.org/10.1121/1.409676

Stone, M. (2005). A guide to analysing tongue motion from ultrasound images. *Clinical Linguistics & Phonetics, 19*(6/7), 455-501.

Szabo, T. L. (2004). Diagnostic imaging: inside out. Elsevier Academic Press.

Tamburrini, Solazzo, Sagnelli, Del Vecchio, Reginelli, Monsorro, & Grassi. (2010). Amyotrophic lateral sclerosis: sonographic evaluation of dysphagia. La radiologia medica, 115(5), 784-793.

Tamura, Kikutani, Tohara, Yoshida, & Yaegaki. (2012). Tongue thickness relates to nutritional status in the elderly. Dysphagia, 27(4), 556-561.

Venables. (2011). How does ultrasound work? Ultrasound, 19(1), 44-49.

Walker, F. O. (2004). Neuromuscular ultrasound. *Neurologic Clinics, 22,* 563-590.

Watkin, K. L. (1999). Ultrasound and swallowing. *Folia Phoniatrica et Logopaedica, 51*(4-5), 183-198. https://doi.org/10.1159/000021496

Watkin, K. L., Diouf, I., Gallagher, T. M., Logemann, J. A., Rademaker, A. W., & Ettema, S. L. (2001). Ultrasonic quantification of geniohyoid cross-sectional area and tissue composition: A preliminary study of age and radiation effects. *Head & Neck*, *23*(6), 467–474. https://doi.org/10.1002/hed.1061

Winiker, K. S. (2019). Assessment and behavioural modulation of the upper oesophageal sphincter in healthy swallowing.

Woo, J. (2002). A short history of the development of ultrasound in obstetrics and gynecology. *History of Ultrasound in Obstetrics and Gynecology, 3,* 1-25.

Yabunaka, K., Sanada, H., Sanada, S., Konishi, H., Hashimoto, T., Yatake, H., et al. (2011). Sonographic assessment of hyoid bone movement during swallowing: a study of normal adults with advancing age. *Radiological Physics and Technology, 4*(1), 73-77. https://doi.org/10.1007/s12194-010-0107-9.

Appendices

Appendix A: Participant Information Sheet

Information Sheet for Participants

Reliability of Ultrasound Measurement using Image Enhancement Software

Aim of the study

This project aims to evaluate the influence of image enhancement software on reliability of ultrasound measurement. Ultrasound has not progressed to standard clinical practice despite the fact that it offers a radiation-free and non-invasive procedure for swallowing assessment and promising validity and reliability for ultrasound measurements. The cost, which may be one of the significant obstacles to restrict integration into clinical practice, can be overcome by using new pocket- sized portable ultrasound devices. However, early research suggests that poor image quality may compromise measurement reliability. To enhance the interpretation of swallowing ultrasound images, a summer student in collaboration with the Rose Centre for Stroke Recovery and Research, recently developed image enhancement software which enhances the contrast and features of acquired ultrasound images The focus of my research is to determine if this image enhancement software increases inter- intra-rater reliability of ultrasound measurement when compared to the original ultrasound images.

You have been approached to take part in this study because you have experience working with patients with swallowing impairment.

Research Procedure

If you choose to take part in this study, you will first attend a four-hour training session on the measurement of ultrasound imaging both with and without the image enhancement software. During this training, you will receive supervised training from a clinician with prior experience in ultrasound measurement and software operation. Then, you and other three clinicians will analyze a total of 80 images which includes 40 original images and 40 software enhanced images. This will take place across a four-day period, with 20 trials per day analyzed each day to minimise the effects of fatigue. One week following the completion of inter-rater reliability measurement, you will re-evaluate the same images to evaluate intra-rater reliability. The second analysis will be performed using the same methods as the first analysis round. Therefore, a total of 160 measures will be analyzed by each researcher and and the participants are asked to contribute approximately one hour each day. Participants will receive a $75 voucher for their time upon the completion of the procedure. Once you have submitted your data, no further follow-up is required.

In the performance of the tasks and application of the procedures there are no foreseeable risks of any kind.

Participation

Participation is voluntary and you have the right to withdraw at any stage without penalty. You may ask for your raw data to be returned to you or destroyed at any point. If you withdraw, I will removeinformation relating to you. However, once analysis of raw data starts on mid of August, it will become increasingly difficult to remove the influence of your data on the results.

Confidentiality

The results of the project may be published, but you may be assured of the complete confidentiality of data gathered in this investigation: your identity will not be made public without your prior consent. Participant data will be coded with a participant number. The only identifying information is on the consent form; the participant number will not be linked to the consent form thus data are anonymised. The identifying data will be destroyed after 5 years of Master thesis completion by the supervisor of the study. The de-identified raw data may be used in the further research. A thesis is a public document and will be available through the UC Library.

Questions

The project is being carried out as a requirement for Master of Science in Speech and Language Sciences by Sruthy Vijayan under the supervision of Maggie Lee Huckabee and Dr. Phoebe Macrae. If you have further questions regarding the research, please feel free to contact Sruthy Vijayan on *sruthy.vijayan@pg.canterbury.ac.nz*. Alternatively, Professor Maggie Lee Huckabee can be contacted on *maggie-lee.huckabee@canterbury.ac.nz*. She will be pleased to discuss any concerns you may have about participation in the project.

This project has been reviewed and approved by the University of Canterbury Human Ethics Committee, and participants should address any complaints to The Chair, Human Ethics Committee, University of Canterbury, Private Bag 4800, Christchurch (human-ethics@canterbury.ac.nz).

Appendix B: Consent Form

UNIVERSITY OF CANTERBURY
Te Whare Wānanga o Waitaha
CHRISTCHURCH NEW ZEALAND

Department: School of Psychology, Speech and Hearing
Telephone: +64 33695124
Email: sruthy.vijayan@pg.canterbury.ac.nz

Reliability of Ultrasound Measurement using Image Enhancement Software

Consent Form

☐ I have been given a full explanation of this project and have had the opportunity to ask questions.

☐ I understand what is required of me if I agree to take part in the research.

☐ I understand that participation is voluntary and I may withdraw at any time without penalty. Withdrawal of participation will also include the withdrawal of any information I have provided should this remain practically achievable.

☐ I understand that any information or opinions I provide will be kept confidential to the researcher and supervisors involved in the study and that any published or reported results will not identify the participants. I understand that a thesis is a public document and will be available through the UC Library.

☐ I understand that the de-identified raw data may be used in the further research.

☐ I understand that I can contact the researcher, Sruthy Vijayan, *sruthy.vijayan@pg.canterbury.ac.nz* or supervisor Maggie Lee Huckabee, *maggie-lee.huckabee@canterbury.ac.nz* for further information. If I have any complaints, I can contact the Chair of the University of Canterbury Human Ethics Committee, Private Bag 4800, Christchurch (human-ethics@canterbury.ac.nz)

☐ I would like a summary of the results of the project.

☐ By signing below, I agree to participate in this research project.

Name:____________________Signed:________________________Date:

Email address:

Appendix C: Advertisement for participants

Can You Help Us?

We are seeking Speech Language Pathologists to participate in a reliability study of ultrasound measurement.

Ultrasound has not progressed to standard clinical practice despite the fact that it offers a radiation-free and non-invasive procedure for swallowing assessment and promising validity and reliability for ultrasound measurements. The cost, which may be one of the significant obstacles to restrict integration into clinical practice, can be overcome by using new pocket- sized portable ultrasound devices. However, poor image quality may compromise measurement reliability.

This project aims to evaluate the influence of a recently developed image enhancement software on reliability of portable ultrasound measurement of swallowing images.

- Participation is required for 9 days in total including training session.
- One day training session will be provided before data analysis (four- hour programme).
- Data analysis session is approximately 1 hour each day to measure 20 images in a two week period (four days a week).
- Participants will complete two rounds of analysis to derive inter- and intra-rater reliability.
- Sessions take place at The Rose Centre for Stroke Recovery and Research, St Georges Hospital, Merivale.
- In appreciation of time, participants will receive a $75 voucher upon the completion of the procedure.

For more information, please contact:

Sruthy Vijayan, 0212162171, sruthy.vijayan@canterbury .ac.in

University of Canterbury, Rose Centre for Stroke Recovery and Research

Level 1, Leinster Chambers, St Georges Medical Centre, 249 Papanui Road, Christchurch, 65804