PHARYNGEAL RESIDUE MEASUREMENTS IN PATIENTS TREATED WITH CHEMORADIATION THERAPY FOR HEAD AND NECK CANCER

by

Lindsay Rozynek

BS Speech, Language, and Hearing Science, Purdue University, 2016

MS Speech-Language Pathology, University of Pittsburgh, 2019

Submitted to the Graduate Faculty of

University of Pittsburgh in partial fulfillment

of the requirements for the degree of

Master of Science Degree in Communication Science and Disorders

University of Pittsburgh

2019

UNIVERSITY OF PITTSBURGH

SCHOOL OF HEALTH AND REHABILITATION SCIENCES

This thesis was presented

by

Lindsay Rozynek

It was defended on

March, 4, 2019

and approved by

Christopher Brown, Ph.D., Assistant Professor, Communication Science and Disorders, University of Pittsburgh

Christine Matthews, CScD., BCS-S, Adjunct Assistant Professor, University of Pittsburgh

Susan Shaiman, Ph.D., Associate Professor, Communication Science and Disorders, University of Pittsburgh

Thesis Director: James L. Coyle, Ph.D., CCC-SLP, BCS-S, Professor, Departments of Communication Science and Disorders and Otolaryngology, University of Pittsburgh

PHARYNGEAL RESIDUE MEASUREMENTS IN PATIENTS TREATED WITH RADIATION THERAPY FOR HEAD AND NECK CANCER

Lindsay Rozynek, M.S.

University of Pittsburgh, 2019

Difficulty swallowing is a common consequence of individuals who are treated with primary radiation therapy (RT) or primary radiation therapy with chemotherapy (CRT) for head and neck cancer (HNC). Exercise programs are often utilized to maintain or improve function during and following RT/CRT for HNC. One component of dysphagia in people treated with RT/CRT for HNC is the retention of pharyngeal residue after swallows due to impaired clearance caused by these structural/tissue RT/CRT changes. Pharyngeal residue is measured from imaging studies using subjective and more objective methods that seek to quantify residue and indicate the level of impairment.

We sought to characterize pharyngeal residue using the Normalized Residue Ratio Scale (NRRS) described by Pearson, Molfenter, Smith, and Steele (2013). The study was done retrospectively with secondary data, collected between January 7, 2016 and June 29, 2016. Participants were patients treated with CRT for HNC and received 8 weeks of exercise therapy using either instrument-guided or non-instrument-guided protocol. The results of the residue ratings pre- and post-therapy were compared using the NRRS and number of swallows per bolus. Based on observations during pre-data analysis, we proposed a variant of the NRRS which we deployed in a second analysis to determine whether NRRS may overestimate pharyngeal residue.

Due to use of secondary data, a significant number of pairs for pre- and post-treatment data were unavailable. We performed descriptive analyses for pre- and post-treatment NRRS scores and for pre- and post-treatment number of swallows per bolus. We are unable to offer results for

pre- and post-treatment and number of swallows per bolus that offer information that would generalize to a wider sample outside of our study. We found the frame selected for measurements based on the published NRRS methods led to significantly higher residue ratio scores in the valleculae but not in the piriform sinuses, compared to our NRRS variant method of frame selection.

The results of our study suggest evidence that the NRRS may overestimate residue in the valleculae but not the piriform sinuses. Larger scale research is warranted to determine if these results generate to the overall construct of residue measurements.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

The completion of this research project and thesis would not have been possible without the guidance of many at the University of Pittsburgh. Foremost, I am sincerely grateful to Dr. James Coyle, who has served as my thesis mentor and academic advisor. I am endlessly appreciative of his expertise and patience as I navigated through my first research endeavor, and for supporting me in further pursuing my passion for swallowing and swallowing disorders. I would also like to send many thanks to the Computational Deglutition Lab (Aliaa Sabry Elbahnasy, Amanda Mahoney, and Cara Donahue). All of which have offered me their time through support and training throughout this project.

I would also like to thank the speech-language pathologists at the University Presbyterian Hospital in Pittsburgh, Pennsylvania for conducting the videofluoroscopy and providing patient care for the participants in this study. Your practice and dedication to research have made this study possible.

Thank you to all of the committee members who have served on my committee. I am thankful for the time you have taken out of your busy schedules to support me in the approval and defense of this thesis.

Finally, I would like to thank all of my loved ones who have offered me endless support throughout this thesis. Although none truly knew about pharyngeal residue, or what a valleculae was and why we would measure it, they all offered listening ears and supportive words and I am so grateful for that.

1.0 NORMAL SWALLOWING AND DYSPHAGIA

Swallowing is an innate process that begins during fetal development, and requires few voluntary contributions in healthy individuals. It is directly related to nourishment and overall health, as well as being a marked activity pleasure throughout society and specifically within certain cultures. Despite its relatively hard-wired nature, the variability between persons, and even within the same person swallowing the same liquids or solids repeatedly, has been shown to be considerable in healthy individuals (Lof & Robbins, 1990) and even more so in disease states.

The normal variability that occurs in swallowing is only beginning to be fully understood. It has been found that various patterns exist in the ways age-matched people eat solid foods and drink liquids, and in the manner in which solid and liquid boluses (food or drink being prepared for ingestion) are prepared in the oral cavity and transported to the pharynx (Hiiemae & Palmer, 1999; Saitoh et al., 2007). There are also anatomical and physiologic changes in swallowing across the lifespan of human development. During aging, the larynx descends in the upper aerodigestive tract at rest, swallowing duration increases, and pharyngeal residue becomes more common. In a seminal study of test-retest reliability of middle- and old-aged healthy individual's swallows using Modified Barium Swallow Study (MBS), the results indicated that performance was similar from one swallow to another but swallows did exhibit significant kinematic variability (Lof & Robbins, 1990). While there is variation in swallowing in healthy individuals, likewise there is variation in swallowing function parameters amongst disordered groups. For example, in a study of

Penetration-aspiration scale scores in healthy individuals, individuals with multiple strokes and HNC were compared, both disordered groups displayed significantly greater variability than the healthy cohort, and variability within the health cohort was demonstrated (Robbins, Coyle, Rosenbek, Roecker, & Wood, 1999). This has led clinicians to adopt a repeated-trials approach to conducting instrumental swallowing studies. Although individuals receiving diagnostic evaluation, most likely, will not fall under the healthy individual category, normal variation should be understood to avoid over diagnosing or over-correcting for dysphagia. With established variation in normal swallows, the prevalence of laryngeal penetration (material entering the larynx above the level of the true vocal folds) and aspiration (material entering the trachea below the level of the true vocal folds) in healthy people has also been explored (Daggett, Logemann, Rademaker, & Pauloski, 2006; Robbins et al., 1999). In a study with normal adult volunteers, MBS revealed that aspiration is not common in healthy swallowers, however, penetration is more common than originally thought but typically does not result in aspiration (Allen, White, Leonard, & Belafsky, 2010). Penetration, to some degree, occurs within normal variability and is often significantly associated with the presence of pharyngeal residue, which may or may not lead to penetration-aspiration (Allen et al., 2010; Rosenbek, Robbins, Roecker, Coyle, & Wood, 1996). And higher levels of pharyngeal residue are known to lead to aspiration in some individuals (Eisenhuber et al., 2002). Increased awareness of the variability in normal swallows has been important in the diagnosis and tracking of progress in disordered swallowing, and in the estimation of the degree of swallowing impairments caused by disease states in order to measure progress during treatment.

1.1 ANATOMY, PHYSIOLOGY, AND BIOMECHANICS

The aerodigestive tract is designed to perform respiratory and digestive functions and alternate breathing and swallowing without voluntary intervention. The tract begins at the mouth and ends in the esophagus. In order to coordinate respiration and swallowing without biomechanical error, a sophisticated network of muscles with various innervations are at work.

The aerodigestive tract continues to develop throughout the duration of the lifespan. At birth, the aerodigestive tract is compressed vertically and the tongue occupies much of the oral cavity. Infants are able to feed more efficiently and while lying horizontally due to the approximation of the larynx to the posterior oral cavity. As an infant matures, the larynx gradually descends, increasing the complexity of airway protection during swallowing and respiratory-swallowing coordination. This, however, is mitigated by the maturation of sensorimotor subsystems responsible for the coordination of these processes.

As development throughout the entire lifespan continues, change occurs within the aerodigestive tract. Motor and neural network coordination begins to slow, thereby affecting the coordination of swallowing. Analogous to age related changes in skeletal muscles controlling the limbs, swallowing muscle mass and force of contraction diminishes in advanced aging leading to prolonged durations of individual swallow events and total swallow duration. However, in healthy states, these changes in the anatomy and physiology of breathing and swallowing do not impose significant health risks until the older individual enters a disease state, in which age-related changes interact with disease processes to elevate risks of material entering the airway during swallowing, or the patient being unable to sustain their nutrition through oral intake.

During a normal swallow, a series of events takes place to safely execute transfer of the bolus (liquid or solid to be masticated and digested) to its destination in the esophagus while bypassing the airway. During this transfer, several important biomechanical processes take place. In healthy individuals, penetration or aspiration routinely trigger an ejection response (Allen et al., 2010). This response serves as a final protective mechanism for the airway.

The oral stage of swallowing includes the oral preparatory stage and the oral transit stage. These two stages comprise the time the bolus enters the oral cavity to when it crosses the ramus of the mandible for the initiation of the pharyngeal stage. The oral preparatory stage begins as soon as a solid or liquid bolus enter the oral cavity. Salivary glands play an important role in the beginnings of digestion during this phase (Matsuo & Palmer, 2008).

Dependent upon the nature of the bolus (liquid or solid), the oral preparatory stage of the swallow will differ. With a liquid bolus, the tongue acts as a reservoir to hold the bolus, while the linguavelar valve is tightly closed, until the oral transit stage begins and the bolus is pushed to the posterior oral cavity where the pharyngeal stage is initiated. For solid boluses, the oral preparatory stage has marked differences. The bolus, as it is being masticated, aggregates primarily in the valleculae, but also as deep as the piriform sinuses. This variable aggregation of the bolus is now recognized as normal variation and is not denoted as abnormal in any way for this texture of swallowed material (Saitoh et al., 2007). As mastication continues, the bolus effectively 'bobs' back and forth from the posterior oral cavity to the pharynx to complete mastication, collection of the food into an organized bolus, and mixing of bolus contents with oral secretions. Once mastication is complete, the bolus is brought back up to the posterior oral cavity where the oral transit stage of the swallow is triggered, followed by the pharyngeal stage, and the bolus is transferred to esophagus (Hiiemae & Palmer, 1999; Matsuo & Palmer, 2008).

Pressure maintenance and generation are crucial to adequate propulsion of the bolus into the esophagus. Pressure generation is driven by lingual propulsion of the bolus from the oral to pharyngeal stage, the pharyngeal constrictors squeezing the bolus through the pharynx, and the final pressure generation mechanism, UES (upper esophageal sphincter) opening. Generation of pressures moves from positive (supra-atmospheric) in the oral stage to negative (sub-atmospheric) in the late pharyngeal stage. This sub-atmospheric pressure generated by the UES opening "pulls" the bolus down into the esophagus. In order for pressure to be maintained, linguavelar closure, velopharyngeal closure, and laryngeal closure must occur. In conjunction with pressure maintenance and generation, labial closure, proper bolus containment (liquid) or efficient mastication (solid), hyolaryngeal excursion (HLE), laryngeal shortening, epiglottic inversion all contribute to swallowing in healthy individuals. The biomechanical forces driving effective swallowing are crucial pieces to examine to understand the system when it does break down and become disordered.

As the bolus is transferred from the oral cavity to the pharynx, biomechanical processes are at work to deliver the bolus to the esophagus. Pressure is maintained from the oral phase by velopharyngeal closure and laryngeal closure and is generated by pharyngeal constriction and opening of the UES. By superior approximation of the larynx contributing to base of tongue (BOT) retraction, the epiglottis is inverted to further seal the trachea (Matsuo & Palmer, 2008).

Any one failure of the biomechanical actions occurring during the pharyngeal phase can result in a compromise of the airway, aspiration and/or penetration, or residue. Residue of the bolus due to biomechanical failure can occur in the vallecular space, piriform sinuses, or both. Pharyngeal residue is particularly significant due to its correlation to secondary aspiration (Perlman, Booth, & Grayhack, 1994).

With effective relaxation of the UES, the bolus is directly transferred into the esophagus. From here, peristalsis (physiological forces effectively "squeezing" the bolus from the UES to the

LES lower esophageal sphincter) begins to take place to continue the digestive process. During the esophageal stage, potential retrograde flow of the bolus is possible. Retrograde flow can remain below the level of the UES or penetrate the tightly closed seal of the UES (Matsuo & Palmer, 2008).

Retrograde flow that penetrates the UES is potentially harmful. This backflow of the bolus can remain as residue in the piriform sinus and potentially transfer anteriorly to an unprotected airway, causing penetration or aspiration.

1.2 DISORDERED BIOMECHANICS OF SWALLOWING

It is rare for a healthy individual to experience aspiration. And, if aspiration does occur, it presents minimal risk of aspiration pneumonia. Disordered swallowing, or dysphagia, increases an individual's risk of penetration or aspiration. The penetration, potentially co-occurring with pharyngeal residue which is later aspirated, or frank aspiration results as failure(s) of the biomechanical functioning of the oropharyngeal swallowing mechanism and are both prime points of focus (Allen et al., 2010).

When considering the normal biomechanics of swallowing, it is clear that every component plays a vital role in executing a safe swallow. When examining aspiration risk, penetration and residue are a key surrogate indicators of biomechanical impairment (Molfenter & Steele, 2013). Pharyngeal retention is significantly correlated to swallowing impairment, especially in regards to transfer of the bolus (Eisenhuber et al., 2002). Residue present in the vallecular space and piriform sinuses after swallowing poses a strong risk for aspiration (Eisenhuber et al., 2002; Molfenter & Steele, 2013; Perlman et al., 1994).

1.3 METHODS OF TREATING HEAD AND NECK CANCER

Patients with HNC have several methods of intervention available to them. Choice of treatment revolves around the patient's wishes and expectations regarding their plan of care. Patients with HNC can be treated as aggressively as possible if treatment provides a reasonable prognosis for a cure, or if the disease has progressed to an incurable state, choose palliative management to ensure they are comfortable throughout their disease progression. More involved treatment options include surgical resection of the tumor with adjuvant radiotherapy or chemotherapy and, a newer treatment option, organ-preservation with primary RT/CRT. While it is more likely that patients that choose organ-preservation often believe that once their cancer is cured they will return back to normal without the effects of disfiguring surgery, these aggressive treatments yield devastating effects to the swallowing and speech mechanism. Swallowing function with primary surgery generally improves over time, whereas organ-preservation treatment causes a progressive decline in functionality for many years following treatment (Arrese & Lazarus, 2013).

In the early 1990's, the emphasis of treatment for HNC switched from primary surgery followed by radiotherapy to organ-preservation methods of treatment (Logemann, 1999). Organ-preservation is a method of treatment for HNC that involves primary, intense radiotherapy, and is often accompanied by adjuvant chemotherapy or excisional surgery of a small anatomic field. This shift in treatment was prompted by a study conducted by the Department of Veterans Affairs

Laryngeal Cancer Study Group in 1991 which found traditional and organ-preservation methods of treatment to be equivocal in their effectiveness of rates of survival (Logemann, 1999).

The organ-preservation model of treatment aims to preserve the structural and neural integrity of the aerodigestive tract that is often drastically compromised with primary surgical intervention. Although structure and neural networks remain intact during and soon after organ-preservation treatment, the intense radiotherapy patients with HNC receive to effectively cure their cancer results in debilitative changes to the aerodigestive tract many years after their treatment. Organ-preservation surely preserves structure, but it does not preserve long-term functional outcomes (Dworkin, Hill, Stachler, Meleca, & Kewson, 2006).

Organ-preservation methods are attractive to patients due to its promise to cure their cancer and preserve their structure – aesthetically and for swallowing and communicative functioning. And in general, when the probability of cure is equal with surgery or organ-preservation therapy, most patients select organ preservation. The attractiveness of this method of treatment fades during the acute effects of radiotherapy which begins to take place during the 3rd to 4th week of routinely 7 weeks of treatment, during which acute pain, mucositis, and dysphagia can predominate, and over the years of survival as the long-term toxicity of RT/CRT accumulate due to progressive microvascular damage leading to sensorimotor impairments, loss of tissue flexibility and muscle contractile force, and diminished salivary flow.

1.4 DYSPHAGIA AFTER TREATMENT OF HEAD AND NECK CANCER

A diagnosis of HNC can be the primary cause of dysphagia, due to the tumor location and cranial nerve involvement. However, since HNC is treated nearly immediately following its diagnosis, the dysphagia arising from a HNC diagnosis is typically related to the treatment of the cancer. The increase in HNC over the past decade is driven by rising incidence of human papilloma virus (HPV) (Arrese & Lazarus, 2013). Research has shown that the outcomes of non-HPV and HPV positive cases vary. In those who are HPV positive, there is a greater response to treatment and rate of survival (Arrese & Lazarus, 2013).

The most widely diagnosed type of HNC is squamous cell carcinoma and commonly arises from cancerous epithelial tissue (Carrau, 2017). The location and aggressiveness of the cancerous lesion varies by primary cause of the cancer. In the past 15 to 20 years, histological studies have identified the human papilloma virus (HPV) as a cause of aggressive HNC in patients who do not have the common risk factors see historically in HNC: cigarette smoking, tobacco, and heavy alcohol use. In HPV positive cases, the BOT and tonsils are the most commonly affected areas. In those non-HPV positive cases, the cancer can appear anywhere in the oral and pharyngeal cavities and also specifically affect the tongue, mouth, and larynx (Arrese & Lazarus, 2013). The location, size, adjacent structures, and vasculature of the tumor affect how the malignancy is treated.

Once a tumor is identified and diagnosed as malignant, it is classified using the tumor-node-metastasis (TNM) classification system, which was developed by Pierre Denoix at the Institute Gustave-Roussy (Logemann, 1999). Tumors are classified by T1-T4 based on their size, involvement of adjacent structures, and whether they cross the midline; T3-T4 being considered advanced stages of tumor development. N denotes the stage of local/regional lymph node involvement and M indicating whether the disease has metastasized to distant organs not drained by the lymphatic structures that drains the primary tumor site (Logemann, 1999).

With the treatment methods for HNC shifting over the years, there has also been a shift towards implementing earlier therapy. This concept of beginning dysphagia management therapy prior to HNC treatment is being investigated in its effects to minimize overall severity of dysphagia after treatment. The practice of swallowing therapy intervention prior to and during irradiation is evolving due to the growing body of literature focused on its effects on patient outcomes and quality of life.

There are three broad treatment options for HNC: traditional treatment including primary surgery or, more recently, the organ-preservation treatment method, which uses primary radiation therapy (RT), or a combination of both (RT and chemoradiation (CRT)). A third option of treatment is palliative care. In palliative care, the patient chooses comfort measures and no aggressive treatment option. Both of surgical and organ-preservation treatment methods can produce dysphagia as the result of structural, neurological, and vascular anatomic changes and tissue trauma (Arrese & Lazarus, 2013).

Following treatment for HNC, dysphagia is a common occurrence. In those diagnosed with advanced stage cancers, the rate of persistent dysphagia is 50% (Gillespie, Brodsky, Day, Lee, & Martin-Harris, 2004). Dysphagia experienced by those with HNC greatly affects overall quality of life (QOL). Many of these individuals are plagued by their cancer diagnosis long after is it cured. These individuals have a high dependency on alternative feeding methods (Bleier et al., 2007) and remain on some form of a restricted diet for an extended period of time or for the remainder of their lives. These means of alternative feeding and dietary modifications are intended to serve primarily as means of preventing adverse effects related to swallowing dysfunction and dehydration and malnutrition.

1.5 THE DIAGNOSTIC EVALUATION OF SWALLOWING

The Fiberoptic Endoscopic Evaluation of Swallowing (FEES) and MBS continue to be supported by the literature as the 'gold standard' for evaluation of swallowing (Brady & Donzelli, 2013). These two methods of imaging are often compared when considering which of the two is most appropriate to evaluate swallowing. When compared, MBS and FEES are both adequate diagnostic tools at evaluating a person's swallow though each has specific limitations. The clinician may have both MBS and FEES at their disposal, however, there are instances where that is not the case. In clinical practice, the MBS is a more ubiquitous procedure and is the focus of this research.

Several measures of biomechanical and physiologic events are typically measured from images obtained through the MBS in swallowing research. Among these are measures of displacement of the various oropharyngeal structures, scoring of airway protection, and pharyngeal residue measurements. Pharyngeal residue is one of the most complex measures to accurately quantify because FEES and MBS imaging methods obtain only two-dimensional views of residue within three-dimensional spaces (valleculae, piriform sinuses). Currently, the literature is divided in determining the preferred scale to characterize residue and link its presence to impairment severity (Kaneoka et al., 2013). Any proposed scale to quantify residue must be valid, and strong in its inter- and intra-reliability – which is one of the most tasking demands in developing the best residue quantification method (Kaneoka et al., 2013).

The MBS utilizes motion-picture radiographic imaging and swallowing of barium-impregnated boluses to examine the oral and pharyngeal phases of swallowing (Brady & Donzelli, 2013). During MBS, the patient is seated in a supported chair or wheelchair and can be viewed from the lateral and anterior position. The MBS allows for the examiner to document about the entire swallow from the bolus entering the oral cavity to its propulsion into the pharynx and eventually the esophagus.

MBS is an ideal method for observing biomechanical failures present at any given stage of a swallow, especially in regards to the nature of pharyngeal residue at every stage of the swallow. Through the MBS, it is possible to examine the oral preparatory stage, HLE, velopharyngeal closure, tongue base elevation and retraction, pharyngeal constriction, and UES opening.

When examining test-retest variability in individuals with normal swallowing, the researchers deemed MBS the most effective method to examine the swallowing behaviors of their participants (Lof & Robbins, 1990). However, the importance and complete understanding of the two imaging standards should not be forgotten.

In an effort to improve standardization of scoring of swallowing impairment-level events, Rosenbek and colleagues developed the penetration-aspiration scale (PAS), which uses an 8-point ordinal and interval appearing scale to perceptually describe penetration and aspiration events within a single or multiple swallow study (McCullough et al., 2001). In this multidimensional scale, more than one parameter of airway compromise during swallowing is evaluated: depth of the bolus into the airway (shallow or deep larynx, trachea) and the response to the airway penetration/aspiration (clearance, no clearance, no clearance despite effort). The PAS was shown to demonstrate a high degree in intra- and inter-rater reliability among trained judges. Further efforts to adopt similar quantification methods for estimating pharyngeal residue have been similarly investigated to determine their reliability among judges and have become increasingly more objective. Since residue is supported by the literature to be a direct link to aspiration, the use of other scales alongside the PAS attempts to bridge this gap in diagnostic evaluation.

Instrumental assessment tools, whether it be MBS or FEES, are the gold standard for assessing a multitude of characteristics of swallowing. Researchers and clinicians measure kinematic and temporal occurrences using these video imaging techniques. Lof & Robbins (1990) describe temporal events of swallowing in a table format. Each event can be found from viewing the MBS video and selecting the desirable frame number and time. From the selection of such frames of events, one can use the times to calculate duration measures that occur during swallowing, such as movement of the hyoid bone throughout the swallow and the duration of opening and closing of the UES.

Historically, residue has been judged categorically within the valleculae and piriform sinuses. These methods include stopping the MBS or FEES post-swallow at the desired frame, assessing the residue that is present, and assigning its significance/severity based on location, quantity, and profuseness (Pearson et al., 2013). New ways to better quantify residue are constantly being evaluated. The specific methods in which residue is quantified will be discussed in greater detail in a later chapter.

2.0 PHARYNGEAL RESIDUE-BASED IMPAIRMENT AND MEASUREMENT

Treatment of HNC, whether it is more surgical or organ-sparing, leaves patients with undeniable deficits. The treatment of the cancer causes damage to the structures, resulting in biomechanical deficits. These deficits vary on a patient to patient basis and require differing levels of intervention and compensatory strategies to return to their new "normal" of swallowing function.

Although every patient with HNC is different, dysphagia during and after treatment is typically anticipated. One of the most common symptoms of post-RT/CRT dysphagia is the patient's perception of the need to swallow multiple times to completely clear swallowed material from the pharynx. In patients treated with RT/CRT protocol, swallowing function caused by late effects of radiotherapy has been linked to widespread pharyngeal residue and a heightened risk of aspiration, including silent aspiration (Hutcheson et al., 2012). The resulting compromise of airway protection and link to aspiration risk in surviving HNC patients is of clinical importance. The presence of pharyngeal residue that accompanies, and potentially intensifies, aspiration risk with inadequate airway protection is of clinical importance.

Post-swallow retention of oral or pharyngeal remnants of swallowing is an indication of an impairment in the swallowing mechanism. Residue is classified based on its location. Pharyngeal residue is residue that resides anywhere beyond the level of the ramus of the mandible, the radiographic landmark indicating the entrance to the oropharynx. It tends to aggregate in the two cavities in the oropharynx and hypopharynx: the valleculae, which are the spaces between the tongue base and epiglottis, and the piriform sinuses, which lie lateral and posterior to the larynx and superior to the UES, and include the posterior and lateral pharyngeal walls at the level of the middle and inferior constrictor muscles (Pearson et al., 2013).

The primary concern for the presence of residue is the increased likelihood of penetration and aspiration. Residue present in the valleculae and piriform sinus can fill the space so much so that the material overflows (Eisenhuber et al., 2002). The overflowed material can penetrate the unprotected airway, resulting in aspiration. Due to the risks posed by pharyngeal residue, several variable systems of quantifying residue to indicate level of impairment and overall risk of aspiration have been developed for clinical and research purposes (Pearson et al., 2013).

2.1 METHODS OF QUANTIFYING RESIDUE

Pharyngeal residue is observed during imaging-based swallowing studies performed to identify potential treatment methods for post-RT/CRT dysphagia. To date, the judgment of residue is performed using visual inspection and assignment of either binary or ordinal descriptors. Using the MBS images, residue appears as a two-dimensional area of retained contrast with height and width, but its depth cannot be measured. As a result, this method has limitations, and research into more objective methods of measuring residue have been investigated. Residue measurements are complex, and often, clinicians utilize scales that are simpler to understand.

2.1.1 Binary, Ordinal, and Continuous methods

Binary methods of scoring residue had historically been the primary means of residue examination. The initial approach to binary measurement was two binary judgments for each anatomic location of residue (valleculae and piriform sinuses), which captured the residue based solely on its presence or absence based on its location (Pearson et al., 2013). From this initial system, further scales were devised.

Binary approaches that score the residue in relation to the size of the bolus were developed. This differentiation from the original 'location only' methods scale attempted to detail in percentages how much residue present in the pharyngeal space there was compared to the entire bolus swallowed (Pearson et al., 2013). However, the binary classification of residue, continually, provides no detail on the actual amount of residue present (Pearson et al., 2013). This is significant due to the fact that the amount of residue is significantly correlated with risk of aspiration (Perlman et al., 1994)

With obvious shortcomings of the binary method of classification, ordinal measures of capturing pharyngeal residue were devised. The ordinal methods created allowed for location of the residue, and also a descriptive (i.e.: mild, moderate, severe) term to acknowledge the severity of the residue, typically based on its two-dimensional appearance. Limitations in this original ordinal method were the ambiguity of the ratings. Clinicians and researchers were subject to pondering 'how much' residue a 'minimal' or 'moderate' score depicted. The vagueness of these ratings allowed for poor clinical interpretation as to risk for aspiration due to the subjective nature of the ordinal rating systems (Pearson et al., 2013).

Variations of the ordinal method have been further developed to bridge to gap between numerical score and actual risk of aspiration. These revised methods had numerical scores which then correlated to the percentage of space, vallecular or piriform sinus, that was filled with residue post-swallow (Pearson et al., 2013). One of the most well-known of the ordinal methods is the scale developed by Eisenhuber et al. (2002)

Post-swallow aspiration is frequently caused directly by pharyngeal residue. Perlman et al. (1989), rated residue based on its location (valleculae and piriform sinus) using "mild, moderate, and severe" as descriptors. Eisenhuber et al. (2002) adapted this method to an ordinal scale with scores indicating the height of the residue columns. Using this method, a score of "1" indicated that the residue column occupied less than 25% of the respective pharyngeal space. "2" indicated that the residue column occupied 25-50% of the space's height, and "3" indicated that greater than 50% of the height of each space was filled with residue. The respective score was significantly associated with risk of post-swallow aspiration (Eisenhuber et al., 2002).

Although ordinal methods using percentage-based scores have increased reliability of rating residue, limitations still exist. These methods serve as a general measurement, as opposed to one that can fixate on subtle, potentially significant, changes. Future directions have included quantifying residue in relation to actual areas, but limitations have persisted (Pearson et al., 2013).

The Normalized Residue Ratio Scale is a continuous scale and was developed to more accurately quantify residue than existing perceptual and quantitative methods and improve upon those methods and their inherent limitations. As opposed to being perceptual or subjective, the NRRS deploys actual quantification of the area of residue as a proportion of the total area of the space it occupies, producing a continuous data product on the MBS images. NRRS is a continuous method of measurement in that it utilizes "both the ratio of residue relative to the available pharyngeal space and the residue proportionate to the size of the individual" (Pearson et al., 2013).

The NRRS is divided into the $NRRS_V$ (valleculae) and $NRRS_P$ (piriform sinus). Prior to measuring the $NRRS_V$ and $NRRS_P$, and the distance from the base of anterior aspect of cervical vertebrae 2 (C2) to the inferior anterior corner of cervical vertebrae 4 (C4) is measured and logged. This measurement serves as an "internal anatomical scalar reference" to more accurately quantify the residue for the individual's size and proportions. With this scalar, an individual's residue area can be compared from test to test; however, its inclusion allows for comparisons across participants. Each pharyngeal cavity's area is measured, after which the area of reside within it is measured. The residue measure serves as the numerator while the cavity area serves as the denominator. The data are entered into a spreadsheet and the NRRS score is based on specific formulas.

NRRS measures are performed using an image processing package, ImageJ, and then a digital tablet input device attached to a personal computer to draw the areas of interest. The MBS is played and then a frame is selected after the swallow when the hyoid is at its lowest point, followed by the piriform sinus at its lowest point (Pearson et al., 2013). In theory, this method appears to be accurate without question. However, there is evidence to suggest there are limitations resulting from the method of frame selection.

2.2 OBSERVATIONS FROM USING THE NRRS

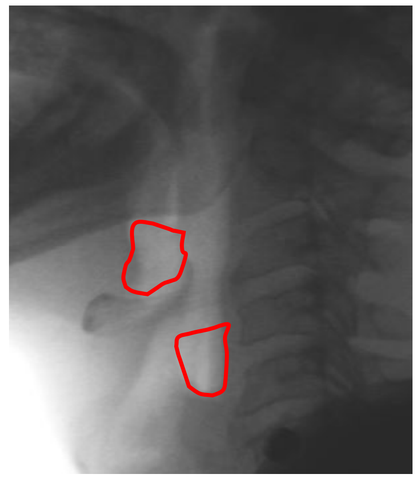
The initial premise of this study was to use pre- and post-swallow MBS data of patients with HNC who underwent CRT and received either instrument-guided (Iowa Oral Performance Instrument) tongue strengthening therapy techniques or traditional. The NRRS, as described by Pearson (2014), would then be used as a means to measure residue to indicate improvement in the participants. While training on the NRRS, a research associate in the Computational Deglutition lab and the author discovered subtle changes in the vallecular and piriform sinus spaces when the hyoid was at the lowest versus other frames recorded within a second after the hyoid was deemed at its lowest point. The NRRS clearly outlines that the measurements are to be taken using the hyoid at the lowest, and then the piriform sinuses at the lowest after that. Using this rule, however, for some suspected normal variability in some participants, could potentially lead to inflated measures of residue because the total area of the pharyngeal spaces were smaller than observed in later frames when structures had returned to rest.

As certain levels of residue are accepted throughout the literature as an indication of impairment, this leads to an issue in the current frame-selection rules for using the NRRS. To preliminarily test these findings, 20 randomly selected swallows from a different data set were measured using two different frame-selection methods. NRRS method of frame selection using the hyoid at its lowest point was initially selected and recorded. Then, we advanced the video frame by frame for up to one second (up to 30 frames) to see if there was a difference in the appearance of the area of the valleculae and piriform sinuses. The number of videos aligning with the NRRS method of frame selection and number of videos aligning with our frame selection variant were tallied. The preliminary data of 20 randomly selected swallows found 13 (65%) of the selected frames exhibited greater valleculae or piriform sinus area when using the classic NRRS frame selection rules, and 7 (35%) of the selected frames showed larger areas when an alternate frame within one second of the published rule was selected to represent "resting posture". See Figure 1 Observations Using the NRRS. This preliminary examination of swallows indicated that 35% of frames selected using NRRS methods may have inaccurate, and possibly inflated, NRRS scores and subsequently, incorrect measurement of residue to link it to impairment.

These findings led the author to investigate the NRRS residue scores versus our variant method of frame selection residue scores to preliminarily suggest inflation of residue scores using the NRRS. As described in Eisenhuber (2002), the frame selection prior to measurement of the percentage of vallecular and piriform sinus space filled is not specifically outlined and is, therefore, at the discretion of the clinician. There is no specification for frame selection, such as described by Pearson et al. (2013) as the hyoid at the lowest point. We have preliminary reason to speculate the method of frame selection is best for measurement of residue when left at the discretion of the researcher or clinician and not confined to a strict set of temporal and kinematic parameters but rather a loose set of temporal and kinematic parameters, to allow for individual variability.

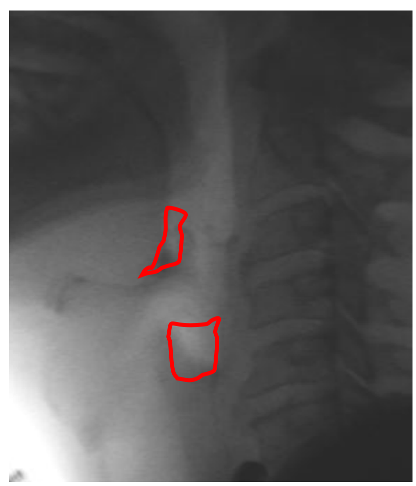
Figure 1 Observations Using the NRRS

a. Lateral view at absolute rest



b. fig a with valleculae and piriform sinuses outlined

c. Lateral view using NRRS frame selec



d. fig c with valleculae and piriform sinuses outlined

3.0 GOALS AND DESIGN OF THE STUDY

3.1 SPECIFIC AIMS

There were two aims to this study. First, we aimed to compare valleculae and piriform sinus residue before and after 8 weeks of either instrument-guided, or traditional exercise therapy, from patients treated with CRT for HNC, using NRRS, and equated the residue measures to swallows per bolus. We then repeated the NRRS with our NRRS frame-selection variant method and compared the residue measurement scores. In this pursuit, we planned to continue to investigate limitations that exist in these current residue quantification methods and if there is support for the alternate method of Pearson et al. (2013) we have described.

3.2 HYPOTHESES

Three separate hypotheses were developed in order to meet the specific aims of the study. Hypothesis 1: Patient's treated with CRT for HNC will have lower piriform sinuses and vallecular residue NRRS scores after treatment than before treatment, and that these improvements will be associated with fewer swallows per bolus. Hypothesis 2: The instrument-guided [treatment] will produce significantly greater reductions in post-swallow piriform sinuses and vallecular residue scores than the non-instrument-guided treatment. Hypothesis 3: The NRRS method will produce significantly greater residue scores in the valleculae and piriform sinuses residue ratios compared to the alternate-NRRS method.

3.3 METHODS

3.3.1 Design

This was a retrospective study using secondary data collected throughout the course of a quasi-randomized investigation of the effects of device-mediated exercise (IOPI) in patients treated with CRT for HNC. The data were collected at the University of Pittsburgh Medical Center (UPMC) starting January 7, 2016 through June 29, 2016 and was approved by the IRB PRO12080498. All participants provided informed consent. The patients were treated with either device-mediated exercise (IOPI) and standard dysphagia therapy (individually-based exercises using lingual exercises and tongue depressors) or standard dysphagia therapy alone without device assist. The standard dysphagia therapies the patients received was individualized to each person and their needs. Twenty participants (10 in each group) pre- and post-exercise MBS data were analyzed.

For the secondary analysis performed for this thesis, the MBS was measured with the NRRS method as described by Pearson et al. (2013). The MBS data was then analyzed again using a variant of the NRRS method (“alternate-NRRS”). This method differs from the original NRRS in that we allowed up to 1 second after the NRRS resting frame rule to select the alternate NRRS frame of measurement to enable valleculae and piriform sinuses to return to, the least effaced, or a resting position.

The pre- and post-therapy MBS data was compared against one another using the NRRS measurements. We used descriptive statistics to calculate the absolute and percent differences between pre- and post-treatment measures of valleculae and piriform sinus residue. The number of swallows for each bolus swallow was also calculated for its potential relationship to the amount of residue using descriptive methods. The NRRS vs alternate-NRRS was compared using a t-test to determine if there is any significant difference in the measurement between the two methods. These methods of statistical analysis provided the most appropriate relationship or statistical difference between the two groups receiving treatment and using the different NRRS methods.

3.3.2 Participants

Twenty participants who were all diagnosed with HNC and treated with CRT made up the sample. The patients had been alternately assigned to treatment groups – 10 in the experimental group and 10 in the control group. The patients ranged from 36-81 years of age. The experimental group exhibits diversity in sex; the control group contains all male patients. The location and staging of the cancer were different across participants. The patients had a pre-and post-treatment MBSS.

Table 1 Participant Information: Experimental

Age	*Gender*	*Tumor Location*	*Tumor Classification*	*Stage*
46	Male	L. Tonsil	IVa	T2N2a
58	Female	Soft Palate	IVa	T2N2c
62	Male	Laryngeal	IVa	T3N2b
36	Female	BOT/Tonsil	IVa	T2N2c
69	Male	L. BOT	IVa	T3N2c
67	Female	NA	NA	NA
65	Male	Laryngeal	IVa	T1N2b
80	Male	Laryngeal	I	T1N0
81	Male	L. BOT	III	T3N1
56	Male	R. Tonsil	NA	NA

Table 2 Participant Information: Control

Age	*Gender*	*Tumor Location*	*Tumor Classification*	*Stage*
62	Male	R. Tonsil	T4N2b	IVa
63	Male	L. BOT	T2N3	IVb
56	Male	BOT	T4aN2cM0	IVa
52	Male	R. Tonsil	T1N2c	IVa
77	Male	Hypopharynx	T1N0	I
60	Male	L. BOT	T2N2a	IVa
63	Male	BOT	T4N0	IVa
64	Male	Hypopharynx	T2N2b	IVa
68	Male	Oropharyngeal	T3N2M0	IVa
62	Male	Laryngeal	T3N2b	IVa

3.3.3 Measurements

Frames were selected for residue measurements using the methods described by Pearson et al. (2013). The total area of the residue occupying each space (valleculae, piriform sinuses) and the total area of each entire cavity (i.e.: valleculae, piriform sinus) was outlined using the draw tool in ImageJ software using a Wacom drawing tablet and stylus. Each measurement was performed three times and the average was then copied into a calculation spreadsheet. After NRRS scores were calculated, we then repeated the measurement procedures using our alternate method of frame selection. In this method, we used the frame selection time from the NRRS and then allowed up to 1 second, or until a second swallow, to find the least effaced vallecular space. Both vallecular and piriform sinuses area residue measurements were performed on this frame. We recorded this frame time and performed the measurements each three times, as we did in the NRRS. Details for the exact procedures of frame selection and residue area measurement appear in Appendix A.

3.3.3.1 Measurements: Preparation for Analysis

In order to address hypotheses 1 and 2, after the NRRS measurements were taken for valleculae and piriform sinuses outliers were identified. The average NRRS and NRRS-alternate scores for each data set (pre- and post-treatment) were computed. Scores lying two or more standard deviations from this average were considered as outliers and omitted from the analysis.

For some participants, there were multiple trials for a single consistency in either pre- or post-treatment: For example, Participant X had 3 trials of thin liquids in the pre-MBSS and 1 trial of thin liquids for the post-MBSS. In an effort to directly compare a single NRRS residue score for pre- and post-treatment, for participants with multiple trials the average was taken for the multiple trials and used as the final NRRS residue score.

3.3.4 Reliability

The author completed the Modified Barium Swallow Impairment Profile (MBSImP) with at least 80% in the reliability zone. A CSD Visiting Scholar and PhD student, Aliaa Sabry Elbahnasy, M.D., (ASE) trained the author to use the NRRS. Informal agreement measures (percent exact agreement) were performed between ASE and the author (LMR) with 20 swallows prior to the author beginning formal data analysis using the NRRS.

ASE was provided all of the author's data of frame selection number and was blinded to the author's residue measurements. She randomly selected 10% of files, by file number, for the use in inter-rater and intra-rater reliability testing. She then performed the NRRS calculations and sent the file numbers to the author. The author performed the measurements a second time for the selected files to establish intra-rater reliability. Once the author completed her intra-rater reliability measurements, ASE sent her inter-rater reliability NRRS calculations to the author. Inter-rater reliability was then calculated by the author.

The author performed the inter-rater and intra-rater reliability measurements herself using the kappa Intraclass Correlation Coefficient described by Shrout and Fleiss (1979) using an internet tool interclass correlation coefficient (ICC) calculator (http://www.obg.cuhk.edu.hk/ResearchSupport/StatTools/IntraclassCorrelation_Pgm.php). The results are displayed below. The NRRS inter-rater reliability for valleculae was 0.9963 and for piriform sinuses was 0.759. The NRRS intra-rater reliability for valleculae was 0.9337 and for piriform sinuses was 0.9834. The alternate NRRS inter-rater reliability for valleculae was 0.8494 and for piriform sinuses was 0.8521. The alternate NRRS intra-rater reliability for valleculae was 0.8773 and for piriform sinuses 0.9202.

Koo (2016) discussed levels of ICC scores and their relative "goodness" in terms of high to low reliability judgments using ICC calculations (Koo & Li, 2016). Based on these cutoffs listed below. NRRS reliability was excellent across all categories except for inter-rater reliability for piriform sinuses which was moderate. Alternate NRRS reliability was good across all categories except intra-rater reliability piriform sinuses which was excellent.

Table 3 Reliability Data

	NRRS	Alternate NRRS
Inter-rater reliability valleculae	0.9963	0.8494
Intra-rater reliability valleculae	0.9337	0.8773
Inter-rater reliability piriform sinuses	0.7059	0.8521
Intra-rater reliability piriform sinuses	0.9834	0.9202

Table 4 Kappa ICC Reliability Cutoffs

Poor	<0.50	Good	0.75-0.90
Moderate	0.50-0.75	Excellent	>0.90

4.0 RESULTS

4.1 HYPOTHESIS 1

Hypothesis 1: Patient's treated with CRT for HNC will have lower piriform sinuses and vallecular residue NRRS scores after either exercise treatment than before treatment, and that these improvements will be associated with fewer swallows per bolus.

There were very few swallows available for comparison with statistical methods because pre-treatment MBS data did not include swallows of the same conditions (bolus volume, consistency) in each phase. Five participants produced only 12 swallows in the same conditions in the pre- and post-treatment phases. Due to this limited data set for analysis of hypothesis 1, we describe the results only. Average $NRRS_V$ pre-treatment scores were between 0.006 and 0.2254 and average post-treatment scores were 0-0.0306. Average $NRRS_P$ pre-treatment scores were between 0-0.0093 and post-treatment scores were 0-0.031. Pairwise differences, and pairwise percent differences between the pre- and post-treatment NRRS scores for these 12 matching swallows are displayed in Table 4 NRRS: Valleculae and Table 5 NRRS: Piriform Sinuses.

Nine of the swallows exhibited improved (smaller) scores after treatment while three exhibited larger post-treatment scores for Valleculae, see Table 4. For piriform sinuses, see Table 5, one swallow exhibited improved scores, 8 exhibited worse (larger) scores, and three exhibited no residue on both the pre- and post-treatment measurement (no change). Exact NRRS measurements for 4 pyriform sinuses were not possible because the pre-treatment scores were zero, and the pre-treatment score is a component of the denominator of the NRRS formula, so "Increase" is indicated in Table 5 for these swallows. Summary for swallows per bolus appears in Table 6.

Overall there was no difference between the pre- and post-treatment swallows per bolus indicating no apparent relationship between swallows per bolus and pharyngeal residue scores.

Table 5 NRRS: Valleculae

PARTICIPANT	Gender Age	Tumor Location	CONDITION	PRE V	Safe/ Unsafe	POST V	Safe/ Unsafe	Diff V	% Diff V
4	M, 46	T2, Tonsilar	TSS	0.0006	Safe	0.0306	Safe	0.029981	51.9031
			TSM	0.0031	Safe	0.0092	Safe	0.006087	1.95105
			TCS	0.0144	Safe	0.0034	Safe	-0.01092	-0.7603
16	F, 58	T2, Soft palate	PSM	0.2254	Unsafe	0.0247	Safe	-0.20072	-0.8906
			TSS	0.1755	Unsafe	0.0281	Safe	-0.1474	-0.8401
10	M, 65	T1, Laryngeal	TCM	0.0329	Safe	0.0027	Safe	-0.03016	-0.9169
18	M, 80	T1, Laryngeal	TSS	0.0284	Safe	0.0033	Safe	-0.02506	-0.8829
			TCS	0.0496	Safe	0.0065	Safe	-0.04308	-0.8692
			PSS	0.0635	Safe	0.0227	Safe	-0.04079	-0.6427
9	M, 56	T4a, BOT	TSS	0.0346	Safe	0.009	Safe	-0.02557	-0.7399
			TCS	0.0279	Safe	0.041	Safe	0.013088	0.46949
			PSS	0.0449	Safe	0.0053	Safe	-0.03963	-0.8828

Table 6 NRRS: Piriform Sinuses

PARTICIPANT	Gender Age	Tumor Location	CONDITION	PRE P	Safe/ Unsafe	POST P	Safe/ Unsafe	Diff P	% Diff P
4	M, 46	T2, Tonsilar	TSS	0.0004	Safe	0.0075	Safe	0.00712	19.3432
			TSM	0	Safe	0.0218	Safe	0.021765	Increase
			TCS	0	Safe	0.0034	Safe	0.003417	Increase
16	F, 58	T2, Soft palate	PSM	0	Safe	0	Safe	0	No change
			TSS	0.0001	Safe	0	Safe	-0.00014	No change
10	M, 65	T1, Laryngeal	TCM	0	Safe	0.031	Safe	0.030994	Increase
18	M, 80	T1, Laryngeal	TSS	0.002	Safe	0.0236	Safe	0.021576	10.5386
			TCS	0.0093	Safe	0.068	Unsafe	0.058698	6.34331
			PSS	0.0067	Safe	0.0176	Safe	0.010928	1.63882
9	M, 56	T4a, BOT	TSS	0	Safe	0	Safe	0	No change
			TCS	0.0028	Safe	0.0014	Safe	-0.00132	-0.4779
			PSS	0	Safe	0.0063	Safe	0.006298	Increase

Table 7 Swallows per Bolus

CONDITION	PARTICIPANT	SEX	TUMOR	PRE: # SWALLOWS	POST: # SWALLOWS
TSM	4	M	T2	2	1
TCM	22	F	T2	2	2
TCM	10	M	T1	2	2
TCM	9	M	T4a	2	3
TSM	3	M	T4	2	2

4.2 HYPOTHESIS 2

Hypothesis 2: The instrument-guided [treatment] will produce significantly greater reductions in post-swallow piriform sinuses and vallecular residue scores than the non-instrument-guided treatment.

Due to the limited pre- and post-treatment matches with our secondary data, statistical analysis could not be performed. Of our five participants who could be included, one of them received non-instrument guided treatment protocol. The single participant [9] presented with all swallow residue scores being "safe" pre- and post-treatment, compared to one participant [18] who received instrument-guided treatment and went from "safe" to "unsafe" residue scores. This data can be seen with participant 9 [non-instrument guided] and participant 18 [instrument guided].

4.3 HYPOTHESIS 3

Hypothesis 3: The NRRS method will produce significantly higher (more severe) residue ratio scores for residue in the valleculae and piriform sinuses compared to the alternate-NRRS method.

The NRRS and alternate NRRS measurements were compared using a paired t-test in Excel with an a priori p value set at $p<.05$. We performed a two-tailed t-test because it was also possible that the differences in frame selection would result in the opposite outcome, and there are no prior data to indicate a hypothesized direction of difference; however, a one-tailed t-test result was also generated. We were uncertain whether the NRRS would produce higher (more severe) or lower (less severe) scores that our alternate NRRS method.

The results of our two t-test, detailed in Table 7, revealed alternate NRRS scores for the valleculae to be significantly lower than NRRS scores (0.035 vs. 0.057, p<0001). One-tailed analysis also resulted in significance (p<.00001), see Table 7. There was not a significant difference (two-tailed) in piriform sinuses scores between the two methods (0.039, 0.031, p=0.125). The one-tailed analysis resulted in a p=0.06, this being in the opposite direction to that of the vallecular scores, indicating a probable systematic difference in these results which we believe was caused by the method of frame selection. The slightly higher mean alternate NRRS scores for piriform sinuses support our decision to use a two-tailed analysis.

Table 8 NRRS versus Alternate-NRRS: Statistical Analysis

t-Test: Paired Two Sample for Piriform Sinuses			t-Test: Paired Two Sample for Means V		
	NRRS	**Alt-NRRS**		*NRRS*	*Alt-NRRS*
Mean	0.030901	0.03955	Mean	0.05658	0.03499
Variance	0.009139	0.01503	Variance	0.00489	0.00428
Observations	81	81	Observations	81	81
Pearson Correlation	0.923889		Pearson Correlation	0.73426	
Hypothesized Mean Diff.	0		Hypothesized Mean Diff.	0	
df	80		df	80	
t Stat	-1.55235		t Stat	3.92431	
P(T<=t) one-tail	0.062262		P(T<=t) one-tail	9.15E-05	
t Critical one-tail	1.664125		t Critical one-tail	1.66412	
P(T<=t) two-tail	0.124525		P(T<=t) two-tail	0.00018	
t Critical two-tail	1.990063		t Critical two-tail	1.99006	

5.0 DISCUSSION

This study had three hypotheses addressing different questions, two of which related to the efficacy of dysphagia treatment on swallowing function, and one methodological hypothesis related to image analysis methods. Our clinical hypotheses addressed pharyngeal residue in patients with HNC after exercise programs and their association to the number of swallows a patient per bolus, instrument guided or non-instrument guided exercise programs and their effect on pharyngeal residue, and NRRS overestimating residue versus alternate-NRRS. While we encountered challenges with our secondary data set, we were able to successfully provide significant evidence for an alternate method of frame selection using the NRRS.

The results of this study in regards to pharyngeal residue before and after instrument guided versus non-instrument guided treatment are inconclusive. We are unable to draw inferences regarding whether exercise-based treatment, either with or without instrument-guided exercise, has a significant effect on reducing pharyngeal residue in patients with HNC treated with CRT. We can, however, conclude that there is evidence of a significant difference in measurements of valleculae residue based on the method of frame selection when using the NRRS versus the alternate-NRRS. Additionally, the trend toward larger alternate-NRRS scores in the piriform sinuses, which is opposite the strongly significant result for valleculae, most likely represents error caused by our method of alternate frame selection for performing the alternate-NRRS measurements. We selected the frame with the widest valleculae area at the end of each swallow and measured both valleculae and piriform sinuses on this frame. In retrospect, a better solution would have been to identify two separate frames for measurement, each of which specified the frame that appeared no more than one second after the end of the swallow as defined by Pearson et al. (2013), and that had the widest total cavity area for each of the valleculae and piriform sinus.

Our data included single swallows and multiple swallow trials. For single swallows, the conclusion of the swallow was obvious. The author experienced challenges in selecting the frames for the NRRS and alternate-NRRS when the video file was abruptly cutoff after the subject's swallow. There were cases where the frame could be selected and residue measured for the NRRS, but not enough frame time (up to 1 second after the swallow) to explore using the alternate NRRS. This led the author to exclude these swallows from her final data analysis. This is an important methodological observation that could be used to encourage investigators and clinicians to ensure adequate videofluoroscopic end-of-swallow data collection to facilitate more accurate measurements. This would add no more than one second to fluoro time for each swallow event observed during testing.

Multiple swallows presented an increased challenge compared to single swallows. For multiple swallows, there was the issue of many videos with consecutive swallows. According to the original NRRS methods by Pearson et al. (2013), for multiple swallows the frame for residue measurement is to be selected after the first swallow and before the initiation of a second swallow. This meant that all of the video files for participants who had consecutive swallows had to be excluded from the analysis. For multiple swallows, the choice to include or exclude the video was made at the discretion of the author.

Our findings that the NRRS produced significantly larger scores in the valleculae but not the piriform sinuses compared to our alternate NRRS is significant since residue is linked directly to level of impairment of a person's swallow. This is even more prominent in the population of patients with HNC, where residue is a common finding. When using the NRRS, the measurements have important meaning for a "safe" or "unsafe" swallow. A "safe" NRRS valleculae residue score is <0.082 and a "safe" piriform sinuses residue score is <0.067. While these values have established meaning, it is important to consider the process of achieving these values, and how measurement methods or error can influence clinical decision-making by over- or underestimating residue areas at the incorrect phase of the swallow. Our findings challenge the current method of frame selection using the NRRS and would potentially affect a person being "safe" or "unsafe".

We have found only the valleculae to be sensitive to our alternate method of frame selection in this small data set, and not the piriform sinuses. We hypothesize that the primary reason for the difference is that there may be a different method of frame selection required for the piriform sinuses to achieve more precise and correct measurement values. While the alternate-NRRS frame selection focuses in on vallecular total relaxation, it is possible that the same should be done for the piriform sinuses after the swallow.

While the focus of this endeavor was to investigate an alternate method of frame selection using the NRRS, we found reason to explore "true resting frames" of patients. A "true resting frame" is a small series of consecutive frames (1-3 seconds) obtained before any swallowing activity is elicited from the participant, and at which time the participant's upper aerodigestive mechanism is at complete rest. We speculate that this would be the most appropriate frame in which to capture the actual area of the valleculae and piriform sinuses to use as a denominator for calculating the proportion of the residue area to the total area of the anatomic spaces.

The number of participants for whom an actual resting frame had been recorded was extremely small. This meant we could not significantly explore this hypothesis for this specific project. Our hypothesis comes from the idea that the configuration of spaces between aerodigestive structures dynamically changes during any kinematic activity involving the tongue, pharyngeal constrictors, larynx and mandible, and are different from their configurations during swallowing. A specified constant amount of residue as a numerator of a residue ratio, that is measured on two different frames, each of which has a very different total area (denominator), would return very different ratios of the residue area to the total area. That is, the amount of residue within the spaces after a swallow is constant (unless it is overflowing from the space in which case it occupies 100% of the space regardless of the configuration of the space), while the space's total area can vary based on lingual, laryngeal or mandibular movements. Clinicians and researchers are concerned about the impact of residue in terms of it entering the airway typically after a swallow, or during the subsequent swallow.

Overall, we succeeded in accomplishing an analysis to test our third hypothesis while our first two hypotheses could not be tested given the limited data set. Our successful effort is based upon the availability of a set of swallows that were each measured twice with two methods, enabling pairing of the samples. This is the first study to investigate alternate methods of frame selection using the NRRS and may add to the utility of the NRRS in clinical and research domains.

5.1 LIMITATIONS

While we attempted to control extraneous variables to maintain the internal validity of our study, we encountered obstacles. The nature of this study was retrospective and the data was secondary and not collected or controlled by any of the individuals involved in conducting this project. This created challenges throughout our data analysis that were beyond our control.

We had a limited data set of 20 individuals. Of these 20 participants, there were no formal methods of data collection deployed during their MBS studies either before or after treatment. The participants were evaluated by different clinicians, with varying methods of how they organize and present bolus conditions for their MBS. This resulted in few pre-post treatment dyads of data that could be analyzed to assess the efficacy of the exercise protocol's on pharyngeal residue, as well as number of swallows per bolus. This prevented us from being able to analyze the data with inferential statistics, leaving only descriptive methods to summarize the data obtained from five participants. There is a call for prospective data collection based on these limitations.

We also only deployed the NRRS versus the alternate NRRS in a sample from a single population - patients with HNC. In order to establish sound evidence for the alternate NRRS, more research is needed and with more participants and populations. It is also necessary to acknowledge the alternate-NRRS favors the valleculae, and there may be a different frame needed to be selected in order to achieve the most accurate piriform sinuses residue scores. Our study presents with limitations, however our evidence for the NRRS overestimating residue in the valleculae is sound and worthy of further investigation outside of this pilot study.

5.2 DIRECTIONS FOR FUTURE RESEARCH IN THIS AREA

There are several methodological changes that would improve the ability of a future investigation into the clinical research questions in this study. First, prospective data collection during the pre- and post-treatment MBS studies with at least a limited set of comparable data collected at each data point, would facilitate analyses that might answer some of the research questions. Likewise, a larger sample would be ideal, though attrition of patients in similar studies of exercise in HNC has been astonishingly high.

In our methodological investigation of an alternate method of deploying a published residue measurement tool, it resulted in valleculae residue scores being significantly lower using the alternate NRRS. These findings are significant but entirely preliminary. Our proposed method of frame selection requires a larger scale study with a wider sample of participants – not just patients with HNC to determine if the methods of measurement are generalizable to all swallow residue measurements. Our alternate method of frame selection may be best suited for the valleculae only, with a separate method needed for the piriform sinuses. Only further research into the NRRS and alternate methods of frame selection will provide the necessary efficacy to firmly propose the most precise method of frame selection.

A wider scale study is required to determine the clinical significance of an alternate method of frame selection using the NRRS. We hypothesize from our results that using a "true resting frame" would yield the most precise NRRS residue measurements. In future research, we propose selecting different frames for the valleculae and piriform sinuses and capturing a resting frame of all participants pre- and post-treatment, regardless of the population, to use as the frame in which to measure the valleculae and piriform sinuses area used in the NRRS residue calculation.

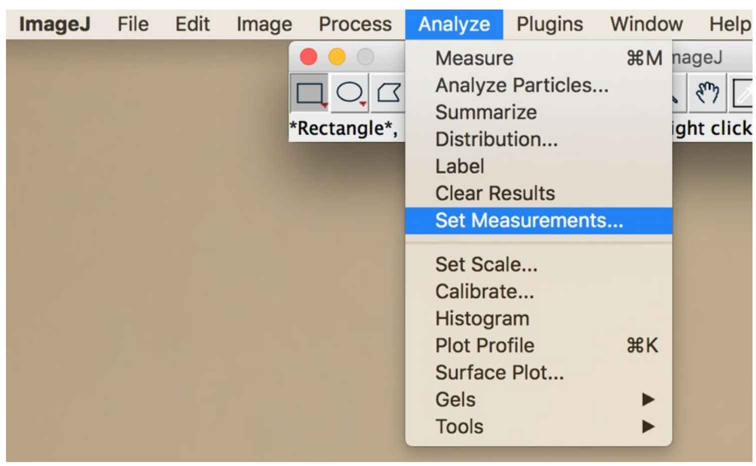
APPENDIX A

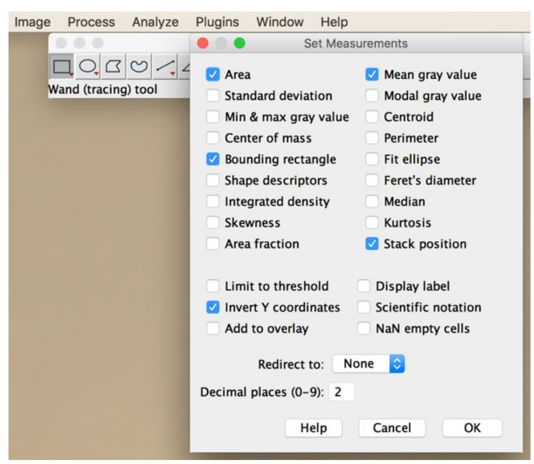
DIRECTIONS TO USE NRRS

Hypothesis 1 and 2 Methods

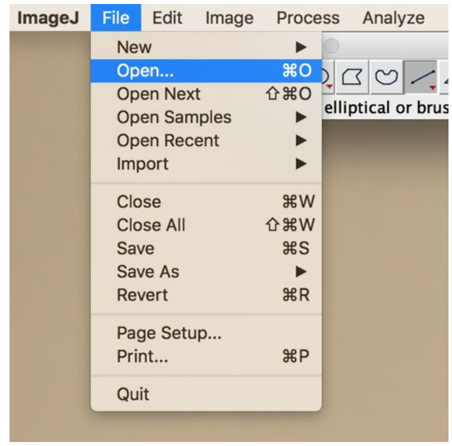
1. ImageJ software was downloaded onto a MacBook Pro. An INTUOS Pen & Touch Tablet, Wacom, Vancouver, WA 98683) was registered to the MacBook Pro. The MBS data was downloaded to the MacBook Pro from a directory containing the master recordings.

2. ImageJ was opened and the Analyze tab was clicked and then the Set Measurements tab was clicked to select: Area, Bounding rectangle, Invert Y coordinates, Mean gray value, and Stack position. Once those specifications for the measurements were selected, OK was clicked to save. An excel sheet with the formulas to calculate the NRRS measurements and an excel sheet to organize the data were opened.

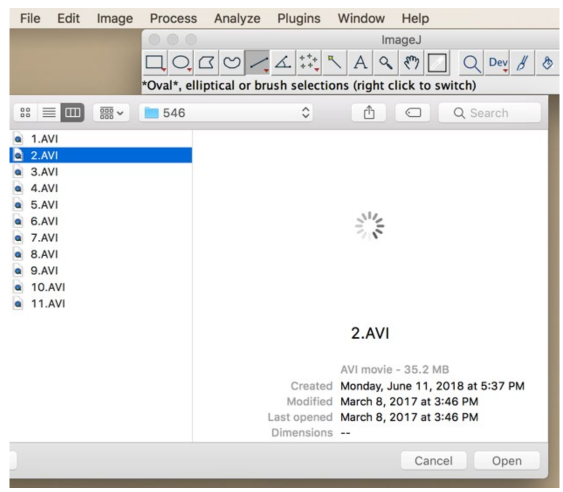




3. With ImageJ open, the File tab was clicked and then the Open tab to select the file of choice to analyze. The straight-line icon in the ImageJ bar was then selected.

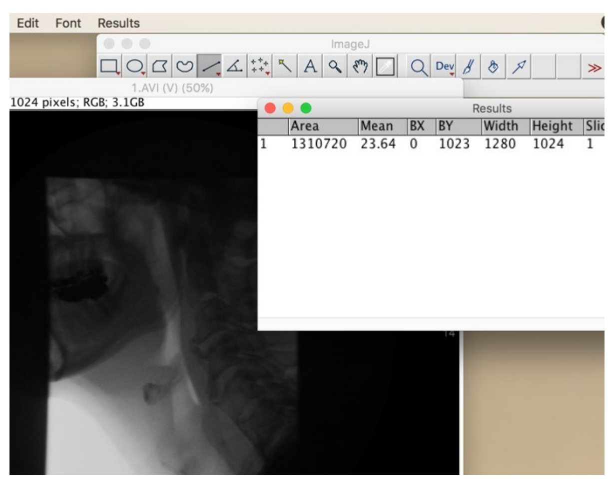


4. The MBS file video was paused on the standard NRRS frame based upon the frame selection specifications described by Pearson et al. (2014). The time and frame number were recorded. The tablet and stylus were used to measure the linear distance between base of C2 to base of C4. Control M was clicked to save the measurement in a new ImageJ window.



Note: For MBS files with multiple swallows, discretion of the author was determined whether the video was usable or not based on if there was a true rest frame fitting the specifications. In many cases, continuous tongue and hyoid movement and continuous swallows negated use of videos for analysis.

5. The ImageJ bar Freehand selections tab was then selected. The tablet and stylus were used to first measure the vallecular residue (if any) and then Control M was clicked to save the measurement in the new ImageJ window. The entire vallecular space was then measured and saved with Control M. The same was repeated for the piriform sinuses.



6. Once the five measurements were taken, the ImageJ window with the measurements was copied and pasted into the NRRS measurement excel sheet which produced a $NRRS_V$ (Normalized Residue Ratio Scale Valleculae) and $NRRS_P$ (Normalized Residue Ratio Scale Piriform Sinuses) value. The values were then copied into the excel sheet to organize the data. This entire measurement process for the vallecular and piriform sinuses was performed three times. The three measurements were then averaged.

   Note: The base of C2-C4 measurement was taken once and was a constant for all three separate measurements for consistency.

Hypothesis 3 methods of measuring residue were the same for Hypothesis 1 and 2 except for step 4. The below step is to be used in place of step 4.

During measurements, the investigator noticed in many swallows, the that standard method of frame selection rendered a frame for measurement in which the cavity containing the barium residue (valleculae, piriform sinuses) were partially relaxed due to participant motion. The videos were then assessed for possible alternate-NRRS measurements. For alternate-NRRS measurements, all previously stated methods were used except for the method of frame selection (Step 4). To determine if there was an alternate-NRRS frame, the original NRRS frame selection time was used as the starting point. From that selected resting frame, the frames were advanced for up to one second or until before the initiation of a second swallow was initiated, to find the frame in which the vallecular space was widest. Advancing frames up to one second or until a second swallow was initiated in many individuals appeared to yield a more “at rest” vallecular and piriform sinus space.

Note: For some of the individuals, this method failed to return a true “resting frame”. In those cases, a rest frame that was taken prior to any ingestion of barium used to measure the actual area of these two cavities and against which to compare the area occupied by residue. In the instance of a rest frame, the rest frame was used as the alternate NRRS frame. The rest frame file and swallow file were opened simultaneously. The base of C2-C4 measurements were taken from the rest frame. The residue was measured from the swallow file from the original NRRS rest frame selection methods and then the vallecular and piriform sinuses spaces were measured using the rest frame.

BIBLIOGRAPHY

Allen, J. E., White, C. J., Leonard, R. J., & Belafsky, P. C. (2010). Prevalence of penetration and aspiration on videofluoroscopy in normal individuals without dysphagia. *Otolaryngol Head Neck Surg, 142*(2), 208-213. doi:10.1016/j.otohns.2009.11.008

Arrese, L. C., & Lazarus, C. L. (2013). Special groups: head and neck cancer. *Otolaryngol Clin North Am, 46*(6), 1123-1136. doi:10.1016/j.otc.2013.08.009

Bleier, B. S., Levine, M. S., Mick, R., Rubesin, S. E., Sack, S. Z., McKinney, K., & Mirza, N. (2007). Dysphagia after chemoradiation: analysis by modified barium swallow. *Ann Otol Rhinol Laryngol, 116*(11), 837-841. doi:10.1177/000348940711601108

Brady, S., & Donzelli, J. (2013). The modified barium swallow and the functional endoscopic evaluation of swallowing. *Otolaryngol Clin North Am, 46*(6), 1009-1022. doi:10.1016/j.otc.2013.08.001

Daggett, A., Logemann, J., Rademaker, A., & Pauloski, B. (2006). Laryngeal penetration during deglutition in normal subjects of various ages. *Dysphagia, 21*(4), 270-274. doi:10.1007/s00455-006-9051-6

Dworkin, J. P., Hill, S. L., Stachler, R. J., Meleca, R. J., & Kewson, D. (2006). Swallowing function outcomes following nonsurgical therapy for advanced-stage laryngeal carcinoma. *Dysphagia, 21*(1), 66-74. doi:10.1007/s00455-005-9001-8

Eisenhuber, E., Schima, W., Schober, E., Pokieser, P., Stadler, A., Scharitzer, M., & Oschatz, E. (2002). Videofluoroscopic assessment of patients with dysphagia: pharyngeal retention is a predictive factor for aspiration. *AJR Am J Roentgenol, 178*(2), 393-398. doi:10.2214/ajr.178.2.1780393

Gillespie, M. B., Brodsky, M. B., Day, T. A., Lee, F. S., & Martin-Harris, B. (2004). Swallowing-related quality of life after head and neck cancer treatment. *Laryngoscope, 114*(8), 1362-1367. doi:10.1097/00005537-200408000-00008

Hiiemae, K. M., & Palmer, J. B. (1999). Food transport and bolus formation during complete feeding sequences on foods of different initial consistency. *Dysphagia, 14*(1), 31-42. doi:10.1007/pl00009582

Hutcheson, K. A., Lewin, J. S., Barringer, D. A., Lisec, A., Gunn, G. B., Moore, M. W., & Holsinger, F. C. (2012). Late dysphagia after radiotherapy-based treatment of head and neck cancer. *Cancer, 118*(23), 5793-5799. doi:10.1002/cncr.27631

Kaneoka, A. S., Langmore, S. E., Krisciunas, G. P., Field, K., Scheel, R., McNally, E., . . . Cabral, H. (2013). The Boston Residue and Clearance Scale: preliminary reliability and validity testing. *Folia Phoniatr Logop, 65*(6), 312-317. doi:10.1159/000365006

Koo, T. K., & Li, M. Y. (2016). A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *J Chiropr Med, 15*(2), 155-163. doi:10.1016/j.jcm.2016.02.012

Lof, G. L., & Robbins, J. (1990). Test-retest variability in normal swallowing. *Dysphagia, 4*(4), 236-242.

Matsuo, K., & Palmer, J. B. (2008). Anatomy and physiology of feeding and swallowing: normal and abnormal. *Phys Med Rehabil Clin N Am, 19*(4), 691-707, vii. doi:10.1016/j.pmr.2008.06.001

McCullough, G. H., Wertz, R. T., Rosenbek, J. C., Mills, R. H., Webb, W. G., & Ross, K. B. (2001). Inter- and intrajudge reliability for videofluoroscopic swallowing evaluation measures. *Dysphagia, 16*(2), 110-118. doi:10.1007/pl00021291

Molfenter, S. M., & Steele, C. M. (2013). The relationship between residue and aspiration on the subsequent swallow: an application of the normalized residue ratio scale. *Dysphagia, 28*(4), 494-500. doi:10.1007/s00455-013-9459-8

Pearson, W. G., Jr., Molfenter, S. M., Smith, Z. M., & Steele, C. M. (2013). Image-based measurement of post-swallow residue: the normalized residue ratio scale. *Dysphagia, 28*(2), 167-177.

Perlman, A. L., Booth, B. M., & Grayhack, J. P. (1994). Videofluoroscopic predictors of aspiration in patients with oropharyngeal dysphagia. *Dysphagia, 9*(2), 90-95.

Robbins, J., Coyle, J., Rosenbek, J., Roecker, E., & Wood, J. (1999). Differentiation of normal and abnormal airway protection during swallowing using the penetration-aspiration scale. *Dysphagia, 14*(4), 228-232. doi:10.1007/pl00009610

Rosenbek, J. C., Robbins, J. A., Roecker, E. B., Coyle, J. L., & Wood, J. L. (1996). A penetration-aspiration scale. *Dysphagia, 11*(2), 93-98.

Saitoh, E., Shibata, S., Matsuo, K., Baba, M., Fujii, W., & Palmer, J. B. (2007). Chewing and food consistency: effects on bolus transport and swallow initiation. *Dysphagia, 22*(2), 100-107. doi:10.1007/s00455-006-9060-5

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: uses in assessing rater reliability. *Psychol Bull, 86*(2), 420-428.