Anterior-posterior distension of maximal upper esophageal sphincter opening is correlated with high-resolution cervical auscultation signal features

Kechen Shu[1], James L. Coyle[2], Subashan Perera[3], Yassin Khalifa[1], Aliaa Sabry[4] and Ervin Sejdić[5]

[1] Department of Electrical and Computer Engineering, Swanson School of Engineering, University of Pittsburgh, Pittsburgh, PA, 15261, USA

[2] Department of Communication Science and Disorders, School of Health and Rehabilitation Sciences, Department of Otolaryngology, School of Medicine, University of Pittsburgh, PA, 15260, USA

[3] Division of Geriatrics, Department of Medecine, University of Pittsburgh, Pittsburgh, PA, 15261, USA

[4] Department of Communication Science and Disorders, School of Health and Rehabilitation Sciences, University of Pittsburgh, PA, 15260, USA

[5] Department of Electrical and Computer Engineering, Swanson School of Engineering, Department of Bioengineering, Swanson School of Engineering, Department of Biomedical informatics, School of Medecine, Intelligent Systems Program, School of Computing and Information, University of Pittsburgh, PA, 15260, USA

E-mail: `esejdic@ieee.org`

Abstract.

Objective: Adequate upper esophageal sphincter (UES) opening is essential during swallowing to enable clearance of material into the digestive system, and videofluoroscopy(VF) is the most commonly deployed instrumental examination for assessment of UES opening. High-resolution cervical auscultation (HRCA) has been shown to be an effective, portable and cost-efficient screening tool for dysphagia with strong capabilities in non-invasively and accurately approximating manual measurements of VF images. In this study, we aimed to examine whether the HRCA signals are correlated to the manually measured ~~AP~~ anterior-posterior (AP) distension of maximal UES opening from VF recordings, under the hypothesis that they would be strongly associated. *Approach:* We developed a standardized method to spatially measure the AP distension of maximal UES opening in 203 swallows VF recording from 27 patients referred for VF due to suspected dysphagia. Statistical analysis was conducted to compare the manually measured AP distension of maximal UES opening from lateral plane VF images and features extracted from two sets of HRCA signal segments: whole swallow segments and segments excluding all events other than the duration of UES is opening. *Main results:* HRCA signal features were ~~statistically significantly associated to~~ significantly associated with the normalized AP distension of the maximal UES opening in the longer whole-swallowing segments and the association became much stronger when analysis was performed solely during the duration of UES opening. *Significance:* This preliminary feasibility study demonstrated the potential value of HRCA signals features in approximating the objective measurements of maximal UES AP distension and paves the way of developing HRCA to non-invasively and accurately predict human spatial measurement of VF kinematic events.

Keywords: high-resolution cervical auscultation, upper esophageal sphincter, swallowing accelerometry, swallowing sounds, signal processing, deglutition

Submitted to: *Physiological Measurement*

1. Introduction

Upper esophageal sphincter (UES) opening during swallowing is an important event facilitating the passage of ingested materials into the esophagus, and its dysfunction can lead to inefficient clearance. This in turn, contributes to aspiration of hypopharyngeal residue and associated adverse pulmonary complications (Jacob, Kahrilas, Logemann, Shah & Ha 1989, Kim, Park, Oommen & McCullough 2015, Singh & Hamdy 2005, Cook, Dodds, Dantas, Massey, Kern, Lang, Brasseur & Hogan 1989). Several physiologic events contribute to opening of the UES during swallowing, and the accurate measurement of the duration and anterior-posterior (AP) distension of UES opening is essential in estimating the contribution of impaired opening to these risks (Jacob et al. 1989, Kim et al. 2015).

The videofluoroscopic (VF) swallowing study is the only imaging-based swallowing assessment technique capable of enabling visualization of UES opening in order to assess its degree of AP distension during swallowing (Sejdić, Malandraki & Coyle 2019, Ahuja & Chan 2016, Martin-Harris & Jones 2008, Logemann 1998). Although scale-based subjective judgments of UES AP distension are clinically convenient and expedient during the assessment of the swallowing function in VF, objective quantification of UES opening leads to more accurate evaluation of recovery or the effects of clinical interventions designed to mitigate impaired UES function (Martin-Harris, Brodsky, Michel, Castell, Schleicher, Sandidge, Maxwell & Blair 2008, Lee, Randall, Evangelista, Kuhn & Belafsky 2017). (Omari, Ferris, Dejaeger, Tack, Vanbeckevoort & Rommel 2012) proposed to measure UES AP distension at 15-20 mm below the level of tracheal air column and calibrate the measurements by adjacent catheter sensor distances. While ASPEKT Method suggests using line between anterior-interior corners of C2 and C4 vertebra as anatomical references, the UES position is subjectively judged (C. M. Steele & Wolkin 2019). Development of a validated non-invasive method that could objectively infer about swallow kinematic events such as UES AP distension would add a valuable measurement tool to the dysphagia diagnostic armamentarium.

Cervical auscultation (CA) is a well-known but crude non-invasive screening method that utilizes a stethoscope to listen to swallowing sound and infer about swallow physiology. Studies have provided evidence for CA's dependence on age, volume size and volume viscosity (Cichero & Murdoch 2002). And dysphagic individuals were reported to have long last and high-pitched swallow sounds through CA than normal subjects (Cichero & Murdoch 2006). The cardiac analogy hypothesis on which CA is based germanely suggests that CA acoustic signals are generated via vibrations caused by valve and pump systems within the upper aerodigestive tract during different swallowing events (Cichero & Murdoch 1998, Kurosu, Coyle, Perera & Sejdić 2019, Khalifa, Donohue, Coyle & Sejdić 2020). The nature of this acoustic information produced by the actions of these valves has yet to be completely elucidated (Sejdić et al. 2019). Although CA in clinical use is believed by devotees to provide sufficient information for assessment of swallowing physiology and biomechanics, a robust body of literature refutes its diagnostic value (Lagarde, Kamalski & Engel-Hoek 2016). This can be referred to the narrow frequency response of the stethoscope which makes it incapable of transmitting the entire spectrum of acoustic and most of vibratory information emanating from the pharynx during swallowing, thus limiting the human judgment to the audible differences only. CA might show fair sensitivity (ranges from 23% to 94%) and specificity (ranges from 50% to 74%) in dysphagia diagnosis; however, studies have argued that intra-rater reliability of CA varied widely and different raters shared poor agreement while assessing the audio of the same swallows (P. Leslie & Wilson 2004, A. E. Stroud &

Wiles 2002, Lagarde et al. 2016).

High-resolution cervical auscultation (HRCA) is an equally non-invasive, promising and advanced alternative to CA that employs electronic sensors (i.e., high-resolution accelerometers and microphones) to transmit the entire range of acoustic and vibratory information produced by the kinematic and bolus-flow events occurring during swallowing (Takahashi, Groher & Michi 1994, Dudik, Jestrović, Luan, Coyle & Sejdić 2015, Movahedi, Kurosu, Coyle, Perera & Sejdić 2017, Khalifa, Coyle & Sejdić 2020). Unlike CA which relies solely on traditional, nonstandardized human interpretation of the swallowing sounds through stethoscope, HRCA signal feature analyses are not prone to human judgment and present unbiased and more reliable interpretations than is possible with conventional CA.

There is a growing body of literature examining and reporting the association and predictive value of HRCA signal features analysis in approximating human measurements for a variety of swallowing kinematic events. For instance, HRCA signal features were found to be strongly correlated with human measurements of VF-based hyoid displacement during swallowing (Rebrion, Zhang, Khalifa, Ramadan, Kurosu, Coyle, Perera & Sejdić 2019, He, Perera, Khalifa, Zhang, Mahoney, Sabry, Donohue, Coyle & Sejdić 2019). Moreover, HRCA signals have been used to non-invasively detect swallows and isolate swallow events from non-swallowing activity, track the location of the hyoid bone on every frame of a VF video image sequence (Mao, Zhang, Khalifa, Donohue, Coyle & Sejdić 2019), and closely approximate human measurements of UES opening duration (Khalifa, Donohue, Coyle & Sejdić 2020). Other studies have shown promising results in aspiration detection and categorization based on the penetration-aspiration scale (PAS) (Yu, Khalifa & Sejdić 2019, Rosenbek, Robbins, Roecker, Coyle & Wood 1996), unsupervised screening of healthy vs. pathological swallows (Dudik, Coyle, El-Jaroudi, Mao, Sun & Sejdić 2018), and demonstrating the association of HRCA signals with several swallowing events (e.g., laryngeal vestibule closure) (Kurosu et al. 2019). However, to the best of our knowledge, no study has addressed the relationship between UES anterior-posterior (AP) distension and HRCA signal features. In line with the cardiac analogy hypothesis, we supposed that detachment of the anterior and posterior walls of the UES may provide the valve activity that generates swallowing sounds and vibrations which can be recorded with HRCA (Cichero & Murdoch 1998, Kurosu et al. 2019). We hypothesized that HRCA signals will show strong correlation with the UES distension measurements which could be revealed by statistical tests. Therefore, we sought to examine whether HRCA signal features, from both acoustic and tri-axial acceleration signals, are ~~statistically significant to~~ significantly associated with the VF-measured maximal AP UES distension.

2. Methods

2.1. Subjects, swallows, and data acquisition

The study was approved by the institutional review board (IRB) of the University of Pittsburgh, and all participating patients provided informed consent prior to their participation. This study was performed as apart of clinical experiment conducted in the context of a standard clinical swallowing evaluation procedure rather than controlled research procedure. Speech language pathologists (SLPs) who conducted the experiment had full control over the procedure which allowed them to alter the bolus size, consistency, maneuver, and way of administration as deemed necessary and based on the patient condition. The following consistencies were used during videofluoroscopy: thin liquid (Varibar thin, Bracco Diagnostics, Inc., < 5 cPs viscosity), mildly thick liquid (Varibar nectar, 300 cPs viscosity), puree (Varibar pudding, 5000 cPs viscosity), and Keebler Sandies Mini Simply Shortbread Cookies (Kellogg Sales Company). Boluses were either self-administered by patients via a cup or a straw or administered by the clinician through the use of a spoon (3-5 mL). Two hundred three swallows were accrued from 27 patients referred to speech-language pathology for VF evaluation of suspected dysphagia at the University of Pittsburgh Medical Center Presbyterian University Hospital (Pittsburgh, PA). Participants included 20 males [ages ranged between 41 – 86 years (mean age 64.85 ± 12.72 years)] and 7 females [ages ranged between 57 – 76 years (mean age 66 ± 7.13 years)]. Of the sample, 15 patients were diagnosed with stroke while the remaining 12 patients were diagnosed with different medical conditions unrelated to stroke. Patients' demographics and characteristics are included in table 1.

VF swallowing studies were conducted while participants were sitting laterally to a standard fluoroscopic x-ray machine system (Precision 500D system, GE Healthcare, LLC, Waukesha, WI) with their head adjusted to a neutral position to have a clear view of the oral cavity, pharynx, upper esophagus and vertebral column. The VF videos were initially recorded at 30 pulses per second (PPS) and captured by a frame grabber module (AccuStream Express HD, Foresight Imaging, Chelmsford, MA) with a sampling rate of 60 frames per second (FPS).

Table 1. Subject distribution.

Characteristics	Number of samples
Total subjects	27
Total swallows	203
Female subjects	7
Male subjects	20
Stroke patients	15
Non-stroke patients	12
Age range	41-86

HRCA signals were obtained concurrently with VF by attaching a tri-axial accelerometer (ADXL 327, Analog Devices, Norwood, Massachusetts) and a microphone (model C 411L, AKG, Vienna, Austria) to the anterior neck of participants, as shown in figure 1. The tri-axial accelerometer was powered at 3 volts (model 1504, BK Precision, Yorba Linda, California) and attached at the midline of the anterior neck of participants over the arch of the cricoid cartilage (Takahashi et al. 1994, Dudik, Coyle, Perera & Sejdić 2015). The sensor's superior-inferior (SI) or vertical axis was aligned to the vertical axis of the participant's neck in the sagittal plane, with the anterior-posterior (AP) and medial-lateral (ML) axes aligned perpendicular to the SI axis in the transverse and coronal planes respectively. The microphone was powered by a B29L power supply, set to volume level 9 (model B29L, AKG, Vienna, Austria), and positioned over the lateral side of larynx slightly below the accelerometer in a position that did not interfere with x-ray imaging of the laryngeal or tracheal airway. Tri-axial accelerometry signals were then filtered by a band-pass filter from 0.1 to 3000 Hz with 5x amplification gain (model P55, Grass Technologies, Warwick, Rhode Island). All signals extracted from the tri-axial accelerometer and microphone were fed into National Instrument 6210 DAQ and recorded by Labview Program Signal Express (National Instrument, Austin, Texas) at a sampling rate of 20 kHz. This setup has demonstrated reliable effectiveness for swallowing activity detection in previous studies (Dudik, Kurosu, Coyle & Sejdić 2016, Lee, Sejdić, Steele & Chau 2010).

Each participant completed a sequence of swallows, including swallows of large volumes ranging from $5mL$ to $15mL$ depending on the participants' comfort delivered by scaled cup, and small volumes of $3mL$ delivered in a standard teaspoon. Barium of various consistencies was provided including: thin, thick (nectar), pudding and cookie. Distribution and percentage of bolus characteristics of all swallow samples are presented in table 2 .

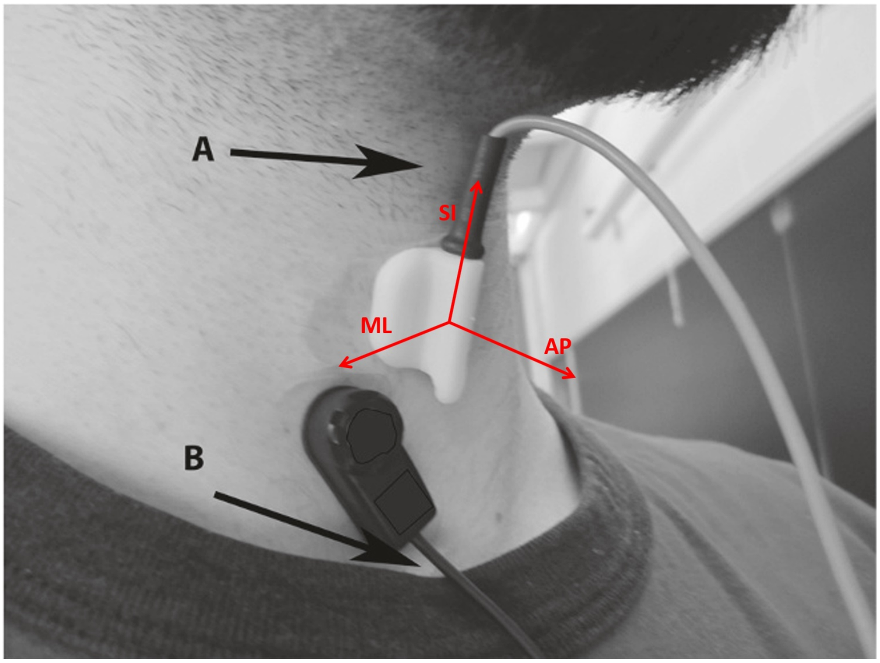

Figure 1. Positions and axes of electronic devices during the experiments. A: accelerometer attached over the arch of the cricoid cartilage on the central axis of neck; B: the microphone attached below the accelerometer and superior to the clavicle (Dudik, Jestrović, Luan, Coyle & Sejdić 2015).

Among the set of swallows considered in this study, airway penetration was observed in 7 of the 27 patients. Further, 17 patients aspirated with detailed distribution of swallows according to the PAS, included in figure 2. We excluded from analysis all swallow videos during which the bolus was not clearly visible on every frame during the entire swallow event, and in which there was a nasogastric tube in place. The set of swallows selected for this particular study varied in participant factors (e.g., age, gender, diagnosis) and bolus conditions (e.g., volume, texture, mode of administration) because this research aimed to examine whether HRCA signal features are associated with the AP distension of maximal UES opening regardless of participant variables or characteristics of swallowed materials.

Table 2. Bolus characteristics distribution.

Bolus texture and utensil	Number of samples	Percentage of samples
Thin by cup	42	20.69%
Thin by cup with straw	17	8.37%
Thin by spoon	42	20.69%
Pudding by spoon	33	16.26%
Nectar by cup	28	13.79%
Nectar in cup with straw	25	12.32%
Cookie in spoon	9	4.43%
NA	1	0.49%

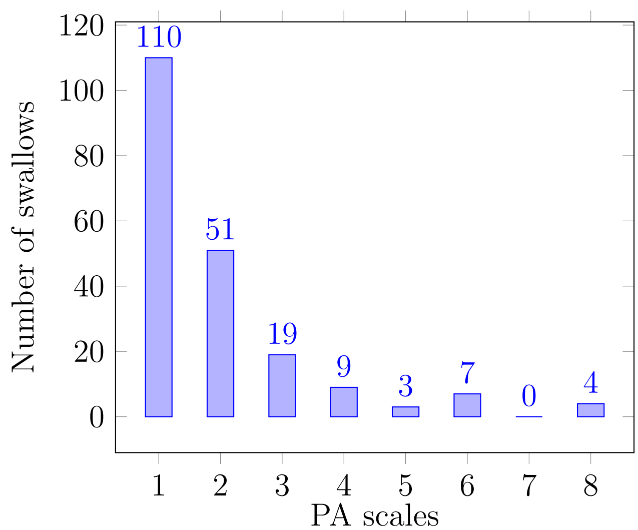

Figure 2. Distribution of tested swallow studies in PA scales

2.2. VF image analysis

All trained judges performing manual VF measurements were blinded to participant demographics, diagnosis, and bolus condition. Videos were first downsampled to 30 FPS to eliminate duplicate frames, and segmented twice into two segments of different parameters: one containing the entire swallow segment, and the other including video frames containing only the duration of UES opening. Individual swallows were segmented based on the frame in which the head of the bolus reached the ramus of the mandible (swallow onset), and the frame in which the hyoid returned to its lowest position following clearance of the bolus from the UES (swallow offset) (Dudik, Jestrović, Luan, Coyle & Sejdić 2015). The other UES opening segments were extracted from the frame of first UES opening (UES opening onset) to the frame of first UES closure (UES opening offset). The further selection of video frame containing maximal AP distension of UES along with several adjacent frames, was performed as follows.

First, the video frame in which maximal displacement of the hyoid bone was observed during the pharyngeal stage, was selected. Second, because maximal UES opening may happen during, or shortly before or after maximal hyoid displacement (Jacob et al. 1989, Cook et al. 1989), we measured the AP distension of UES opening in the 2-3 frames preceding maximal hyoid displacement, the frame with maximal hyoid displacement, and the 2-3 frames following maximal hyoid displacement (this was about 5-7 frames). The AP distension of the maximal UES opening in each video was measured using an application (UES AP distension drawing application) that was developed in Matlab to execute this function through the following steps (figure 3):

(i) Each video was uploaded into the UES AP distension drawing application.

(ii) The frames used for measuring the AP distension of UES opening were selected. The AP distension of UES opening was measured using a drawing tool for 2 or 3 frames before and after maximal hyoid displacement, as well as on the frame with maximal hyoid displacement.

(iii) In order to standardize judgments regarding the location of the superior and inferior limits of the height of the UES, we used the height of the third cervical vertebral body. The region of the proximal esophagus considered the UES has been quantified in manometric studies as coursing 1.3 cm inferiorly from the base of the plane of the true vocal folds (Cook et al. 1989). The height of the third cervical vertebra ranges from $1.11 - 1.14$ cm in adult females and $1.24 - 1.37$ cm in adult males based on midsagittal x-ray measurements (Katz, Reynolds, Foust & Baum 1975). Therefore for each selected frame, a yellow line was drawn from the anterior superior edge to the anterior inferior edge of third cervical vertebral body (C3) in figure 3a.

(iv) Next, another red line was drawn between the anterior inferior edge of the second cervical vertebrae (C2) and the anterior inferior edge of the fourth cervical vertebrae (C4) to provide a vertical axis (C2-C4) that enables the algorithms to subtract larger scale head/neck movements from the measurements (Molfenter & Steele 2014)(figure 3b). The length of the C2-C4 segment was also used as an anatomical scalar representing each subject's height.

(v) The yellow line drawn in step iii was dragged and anchored to the superior border of the posterior tracheal air column (indicating the range of UES height to limit judgments of UES opening) as shown in figure 3c.

(vi) A long blue line that is perpendicular to the C2-C4 segment was drawn and used as a referent axis to ensure alignment of the vertical and horizontal axes of measurement to participant position rather than to an arbitrary x-y coordinate system based on strict vertical and horizontal geometric axes of zero and 90 degrees. This line was then dragged superiorly and inferiorly between two ends of the dragged C3 segment to the location of maximal anterior-posterior distance of the UES opening (figure 3d).

(vii) Then, the anterior and posterior points of UES opening on the perpendicular line were marked by short blue line segments respectively on the line segment drawn in step 6 (represented by two short parallel blue lines in (figure 3e).

(viii) The coordinates of the measured length of maximal UES opening (UES opening anterior end X and Y, UES opening posterior end X and Y) were returned in the output of the application.

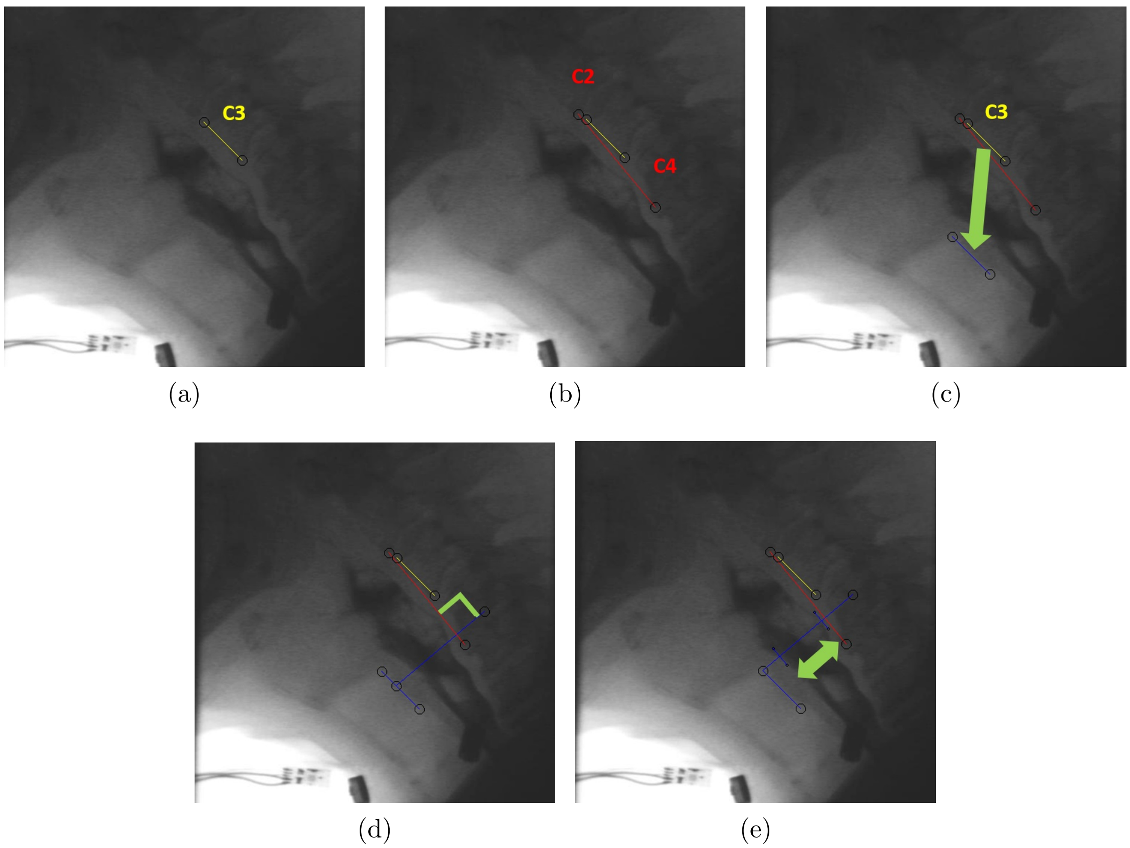

Figure 3. Illustration of steps for measuring the AP distension of maximal UES opening using the newly developed UES AP distension drawing application: (3a) The length of anterior edge of C3 is indicated by the yellow line segment; (3b) The anterior inferior edge of C2 and the anterior inferior edge of C4 were connected by the red line segment; (3c) The C3 length was dragged, following the green arrow, to the position of blue line segment with the upper ends anchored to the superior border of tracheal air column; (3d) The longer blue line segment perpendicular to the C2-C4 axis was positioned with its left ends sliding on the dragged C3 segment; (3e) When the reference line (longer blue line segment) was adjusted to across the largest width of UES opening, two short blue line segments were placed on the extremities of UES. The length of UES opening is measured between the two short segments represented by the bidirectional green arrow.

Interclass correlation coefficients (ICC) have been ubiquitous in assessing the reliability of human judgements for VF analysis in dysphagia research. In this study, the two judges' level of agreement of the manual measurements described above underwent testing of intra- and inter-judge reliability from 10% of the measured swallows using the absolute agreement ICC on an ongoing basis during measurements of study data, to mitigate judgment drift (Shrout & Fleiss 1979). Excellent (ICC > 0.98) intra- and inter-judge reliability between the judges of all VF analysis including temporal (i.e., video segmentation, identification of first UES opening and first UES closure, identification of maximal hyoid displacement), as well as spatial (i.e., UES AP distension measurements). The difference between UES distension measures during the inter-rater reliability test was illustrated by a Bland-Altman plot as shown in Figure 4. Consistency of proposed UES distension measuring method was demonstrated and the UES distension measurements were validated for further statistical analysis.

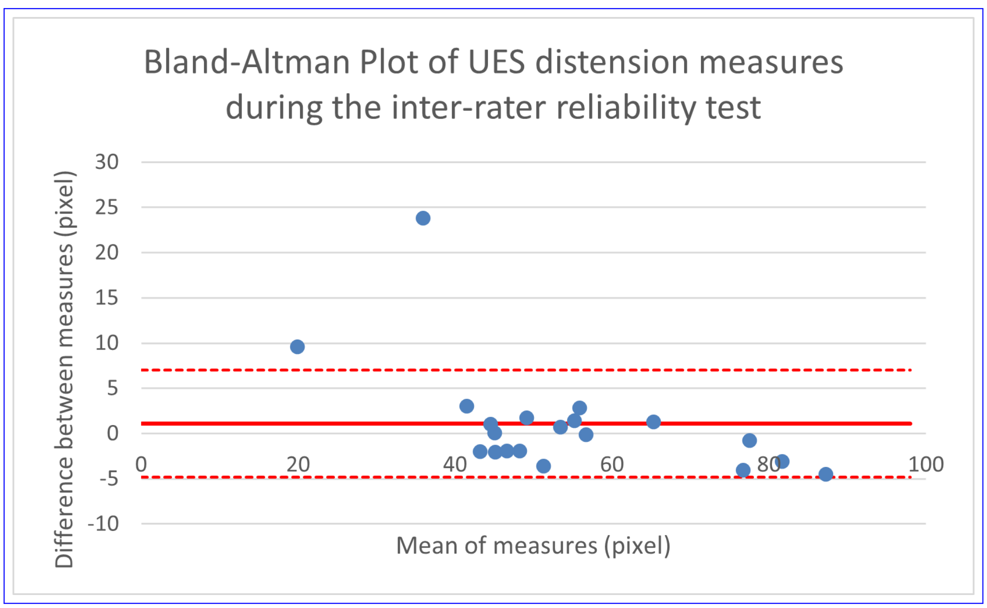

Figure 4. Bland-Altman plot of UES distension measures: the red solid line indicates the average difference between measures and red dotted lines correspond to .95 confidence interval on the average

2.3. Signals preprocessing and feature extraction

Independent of the temporal and spatial measurements described above, processing and extraction of signal features was performed. The preprocessing of HRCA signals consists of several steps: first, all the signals, recorded at a sampling rate of 20 kHz, were downsampled to 4 kHz to overcome the disturbing noise in the signals due to multiple environmental sources and other measurement errors. Additionally, to improve the quality of the collected signals, the device noise inherent with both the accelerometer and the microphone, was modeled by fitting an autoregressive model to the sensor signals generated with null input. Model coefficients were used to generate sensor and axis-specific finite impulse response filters to reduce the device noise (Sejdić, Kosmisar, Steele & Chau 2010). Then, the low-frequency components and motion artifacts were eliminated from accelerometer signals using fourth-order least-square splines (Sejdić, Steele & Chau 2012, Sejdić, Steele & Chau 2010*a*). Lastly, the effect of broadband noise on signals was reduced using the tenth order Meyer wavelet decomposition (Sejdić, Steele & Chau 2010*b*).

The cleaned swallowing signals, including the swallowing acoustic signal (obtained through the microphone) and the acceleration signals of 3 accelerometer axes [anterior-posterior (AP), superior-inferior (SI), and medial-lateral (ML)], passed through the feature extraction stage. The features extracted from time, frequency and time-frequency domains were later compared to the AP distension of maximal UES opening, measured from the spatial analyses using the UES AP distension drawing application. We determined which signal features are related to the AP distension of maximal UES opening and how they are related.

Regarding time-domain features, we calculated the standard deviation, skewness, and kurtosis which evaluate the extent of deviation, asymmetry, and sharpness of the peak, of the statistical distribution of signals respectively. We then determined the Lempel-Ziv complexity by using Kolmogorov complexity which measures the regularity of the signal and the entropy rate which represents the amount of information contained by the signal. Furthermore, in the frequency domain, we took the peak frequency, spectral centroid, and bandwidth into consideration. While the peak frequency and spectral centroid indicate the spectral performances of the signal, the bandwidth specifies the distribution of the power spectrum relative to the spectral centroid. Finally, we estimated the entropy rate of a tenth order discrete Meyer wavelet decomposition to provide information about the distortion of signals in the time-frequency domain. The efficiency and a detailed description of the above features (summarized in table 3) was presented in previous work (Rebrion et al. 2019, He et al. 2019, Sejdić, Kosmisar, Steele & Chau 2010, Sejdić, Steele & Chau 2010*a*).

Table 3. Definitions of extracted features.

	Feature	Definition
Time domain	Standard deviation	Variation of the signal around mean value
	Skewness	Asymetry of statistical distribution of the signal
	Kurtosis	Sharpness of the peak of signal amplitude distribution
Information-theoretic domain	Lempel-Ziv complexity	Regularity of the signal
	Entropy rate	Randomness of the signal
Frequency domain	peak frequency	Frequency that corresponds to the maximal spectral energy
	Centroid frequency	Frequency that divides the spectrum into two equal parts
	Band width	Difference between the uppermost and lowermost frequencies of the signal spectrum
Time-frequency domain	Wavelet entropy	Disordered/ordered behavior of the signal

Finally, to test whether there were differences in the strength of associations between signal features and UES opening AP distension when analyses were conducted using the entire swallow event segment versus solely the shorter sub-duration of the UES opening event, the previously described features were extracted from signals recorded from the whole swallowing segment as well as from signals recorded only during UES opening segment.

2.4. Statistical analysis

To examine the association between HRCA signal features and the AP distension of maximal UES opening, we fitted a series of linear mixed models with the UES AP distension as the response variable; each of the HRCA signal features, one at a time, as independent variables; and a participant random effect to account for multiple swallows from the same individual. To adjust statistical analysis for multiple testing, false discover rate (FDR) was applied to each testing result(Benjamini & Hochberg 1995). We used SAS® version 9.3 (SAS Institute, Inc., Cary, North Carolina) for analysis.

3. Results

3.1. Signal features analysis

3.1.1. Time-domain signal variability Tables 4 and 5 show the raw means and standard deviations of signal features extracted from the microphone and accelerometer (swallowing sound, AP, SI and ML vibrations), by using whole swallow segment data and UES opening segment data respectively. In both cases, all signal features were found to have low standard deviation, low skewness (i.e., contain both positive and negative values) and high kurtosis, indicating that the signals were distributed evenly around the mean value and exhibited peak variation at certain times. The SI movement had a higher standard deviation than AP and ML directions which implies that the dominant vibrations were caused by the superior hyolaryngeal displacement rather than other movements. This finding is consistent with previous HRCA studies on hyoid excursion (Movahedi et al. 2017, Rebrion et al. 2019, He et al. 2019, Zoratto, Chau & Steele 2010). Signals from whole swallow segment data and UES opening segment data had close values on standard deviation and skewness but differed in kurtosis. UES opening segment data showed less kurtosis thus a smoother sequence in the time-domain.

3.1.2. Information theory based features Both accelerometry and sound signals showed low Lempel-Ziv complexity and high entropy rate which corresponds to a rather low randomness and high predictability. However, among the vibration signals, the SI signal presented the greatest Lempel-Ziv complexity, and therefore less predictable behavior than ML and AP signals. Meanwhile, the sideways vibration (i.e., ML signals), showed greater entropy rate and were more regular than upward SI and forward movement (i.e., AP) signals. Compared to the tri-axial accelerometry, swallow sound signals had the most complexity and least regularity. Furthermore, UES opening segment signals showed greater Lempel-Ziv complexity and less entropy rate than the whole swallow segment signals.

3.1.3. Frequency domain distribution When considering the distribution in the frequency domain, all signals presented low peak frequency, high bandwidth, and high centroid frequency indicating that the signals had similar energy distribution over the frequency. While peak frequencies varied to a small extent, both the bandwidth and centroid frequency of HRCA signals decreased when narrowing down the data from UES opening onset to UES closure. In particular, the bandwidth of the ML signal changed tremendously. Both AP and ML signals were more widely spread over the frequency domain than SI signals and swallow sounds considering the whole swallow segment. Moreover, AP signals had the largest distribution over frequency and exhibit greater centroid frequency than others in both segments. Lastly, in the time-frequency domain, the reduction of wavelet entropy indicates that the data extracted from the UES opening segment led to more orderly behavior than the whole swallow segment in all HRCA signals.

Table 4. Mean and standard deviations of considered features from whole swallowing data (SW). (Mic) represents the recorded swallow sound; (AP) represents the accelerometer variation over Anterior-posterior axis; (SI) corresponds to stands for the acceleration over Superior-interior axis and (ML) stands for corresponds to Medial-lateral direction.

	MIC_{sw}	AP_{sw}	SI_{sw}	ML_{sw}
Time domain				
Standard deviation	0.014 ± 0.009	0.021 ± 0.019	0.042 ± 0.024	0.012 ± 0.008
Skewness	-0.433 ± 2.917	0.491 ± 3.868	-0.456 ± 1.992	-0.056 ± 2.432
Kurtosis	37.900 ± 89.202	54.812 ± 134.569	20.432 ± 51.416	29.705 ± 83.436
Lempel-Ziv complexity	0.240 ± 0.083	0.161 ± 0.068	0.216 ± 0.067	0.173 ± 0.068
Entropy rate	0.912 ± 0.041	0.945 ± 0.026	0.939 ± 0.023	0.952 ± 0.022
Frequency domain				
Peak frequency (Hz)	18.145 ± 28.090	10.363 ± 21.343	8.425 ± 8.692	9.357 ± 10.462
Bandwidth (Hz)	108.765 ± 28.090	167.814 ± 152.323	71.258 ± 76.123	156.279 ± 145.890
Centroid frequency (Hz)	84.494 ± 64.215	121.917 ± 170.669	42.471 ± 68.798	69.908 ± 102.359
Time-frequency domain				
Wavelet entropy	1.171 ± 0.816	0.904 ± 0.744	0.880 ± 0.744	0.785 ± 0.733

Table 5. Mean and standard deviation of same features obtained from UES opening (UESO) data segment.

	MIC_{ueso}	AP_{ueso}	SI_{ueso}	ML_{ueso}
Time domain				
Standard deviation	0.013 ± 0.010	0.019 ± 0.018	0.043 ± 0.027	0.011 ± 0.007
Skewness	-0.430 ± 2.213	0.556 ± 2.567	-0.394 ± 1.540	0.034 ± 1.602
Kurtosis	23.821 ± 47.005	31.466 ± 60.382	12.348 ± 21.019	13.891 ± 37.620
Lempel-Ziv complexity	0.287 ± 0.093	0.213 ± 0.082	0.264 ± 0.074	0.243 ± 0.071
Entropy rate	0.878 ± 0.060	0.915 ± 0.039	0.909 ± 0.034	0.919 ± 0.032
Frequency domain				
Peak frequency (Hz)	19.717 ± 28.959	14.113 ± 44.726	9.447 ± 12.904	8.944 ± 10.230
Bandwidth (Hz)	94.508 ± 81.338	121.633 ± 150.981	57.583 ± 63.966	59.911 ± 79.896
Centroid frequency (Hz)	76.397 ± 72.924	110.857 ± 198.028	36.924 ± 55.771	36.535 ± 59.638
Time-frequency domain				
Wavelet entropy	0.973 ± 0.746	0.737 ± 0.671	0.708 ± 0.603	0.716 ± 0.665

3.2. Maximal UES AP distension is associated with HRCA signal features

3.2.1. Maximal UES AP distension measurements Table 6 depicts the difference in the values of mean and standard deviation between the original AP distension of maximal UES opening in pixels, the values normalized to the length of C2-C4, and the values normalized to the anterior length of C3. Interestingly, the maximum AP distension of UES opening closely approximated the height of the anterior plane of the C3 vertebral body.

Table 6. Mean and standard deviation of maximum anterior-posterior AP distension of UES.

	Values
Maximum AP distension width (Pixel)	50.935 ± 16.513
Maximum AP distension width normalized to one C2-C4 length	0.362 ± 0.117
Maximum AP distension width normalized to one C3 length	0.937 ± 0.312

3.2.2. Statistical significicance between UES distension and HRCA signal features

Following statistical analysis, significant correlations were demonstrated between certain features extracted from the whole swallow signals and the values of the AP distension of maximal UES opening divided by C2-C4 length, as depicted in table 7. These features, with a FDR adjusted p-value $< .05$, include the following; standard deviation and Lempel-Ziv complexity of swallow sound; standard deviation and wavelet entropy of AP signal; standard deviation, skewness, Lempel-Ziv complexity and peak frequency of SI signal, and finally standard deviation of ML signal.

When comparing associations between signal features and maximum UES opening based on the entire swallow event segment versus solely the duration of UES opening, more HRCA features were found to be significantly correlated to the AP distension of the maximal UES opening divided by C2-C4 length from UES opening segment, as shown in table 8. Lempel-Ziv complexity and entropy rate of all signals (sound, AP, SI, and ML) were significantly correlated to the measured UES AP distension. Additionally, skewness of swallow sound, bandwidth of AP vibrations, kurtosis and centroid frequency of SI vibrations

Table 7. Statistical significance and regressive coefficients between the maximum UES AP distension normalized to C2-C4 length (Dmax/C2C4) with the accelerometer and the microphone signal features using whole swallow data (NS stands for not significant).

	Dmax/C2C4			
	MIC_{sw}	AP_{sw}	SI_{sw}	ML_{sw}
Standard deviation	$p = 0.0185$ $Coef = 2.2971$	$p = 0.0005$ $Coef = 1.4111$	$p = 0.0195$ $Coef = 0.8187$	$p = 0.0272$ $Coef = 2.7806$
Skewness	NS -	NS -	$p = 0.0195$ $Coef = -0.0009$	NS -
Kurtosis	NS -	NS -	NS -	NS -
Lempel-Ziv complexity	$p = 0.0434$ $Coef = -0.1980$	NS -	$p = 0.0432$ $Coef = -0.2345$	NS -
Entropy rate	NS -	NS -	NS -	NS -
Centroid frequency	NS -	NS -	NS -	NS -
Peak frequency	NS -	NS -	$p = 0.0432$ $Coef = -0.0018$	NS -
Bandwidth	NS -	NS -	NS -	NS -
Wavelet entropy	NS -	$p = 0.0432$ $Coef = 0.0201$	NS -	NS -

became significant to the normalized AP distension of maximal UES opening, while skewness and peak frequency of SI vibrations lost their importance. Far more features were found to be related to the UES opening segment data.

3.2.3. Linear regressive coefficients between UES distension and HRCA signal features

Regressive coefficients (noted as Coef) in Tables 7 and 8 refer to how the maximal UES AP distension (normalized to C2-C4 length) varies across the HRCA signal features, as suggested by the linear mixed model established previously. For instance, if the standard deviation of the sound signal in a whole swallowing segment is increased by 1 unit, then

Table 8. Statistical significance and regressive coefficients between the maximum UES AP distension normalized to C2-C4 length (Dmax/C2C4) with the accelerometer and the microphone signal features using UES opening data (NS stands for not significant).

	Dmax/C2C4			
	MIC_{ueso}	AP_{ueso}	SI_{ueso}	ML_{ueso}
Standard deviation	$p = 0.0100$ $Coef = 2.2571$	$p = 0.0069$ $Coef = 1.4044$	$p = 0.0100$ $Coef = 0.7907$	$p = 0.0185$ $Coef = 3.1927$
Skewness	$p = 0.0352$ $Coef = -0.0079$	NS -	NS -	NS -
Kurtosis	NS -	NS -	$p = 0.0100$ $Coef = 0.0010$	NS -
Lempel-Ziv complexity	$p = 0.0020$ $Coef = -0.2956$	$p = 0.0048$ $Coef = -0.2969$	$p = 0.0051$ $Coef = -0.3096$	$p = 0.0100$ $Coef = -0.3059$
Entropy rate	$p = 0.0100$ $Coef = 0.3561$	$p = 0.0288$ $Coef = 0.4549$	$p = 0.0204$ $Coef = 0.5520$	$p = 0.0185$ $Coef = 0.5889$
Centroid frequency	NS -	NS -	$p = 0.0185$ $Coef = 0.0004$	NS -
Peak frequency	NS -	NS -	NS -	NS -
Bandwidth	NS -	$p = 0.0185$ $Coef = 0.0002$	NS -	NS -
Wavelet entropy	NS -	$p = 0.0434$ $Coef = 0.0220$	NS -	NS -

the UES AP distension divided by C2-C4 length will increase by 2.2971 units. Negative values mean that the UES AP distension will decrease if the value of the considered feature increases.

4. Discussion

4.1. Objective measurement of UES AP distension

In this study, we presented a standardized method of maximal UES anterior-posterior AP distension measurement in lateral view VF which calibrates against subject height and corrects for any possible head or neck movement by using C2-C4 normalization and supplement referent axis to indicate subject's vertical axis. The resulting measurement of UES AP distension was placed in proximity to anterior planes of C5 and C6 which is coherent to the protocol of UES opening measurement in Omari et al.'s (2012) and Kim et al.'s (2015) studies. The ratios between values of the AP distension of maximal UES opening and C3 length are close to 1, which indicates that the C3 length may be a potential anatomic reference of sufficient UES opening that enables clinicians objective evaluate UES opening during the VF study thus leads to trial interventions. We plan to investigate the feasibility and reliability of such a reference in VF analysis of images obtained from healthy participants in the future.

4.2. Features extracted from HRCA signals

Both mean and standard deviation for most of the features extracted from HRCA recordings exhibit consistency with the previous results (Movahedi et al. 2017, Rebrion et al. 2019, Dudik, Coyle, Perera & Sejdić 2015) in terms of magnitude. On the other hand, the UES opening segments of HRCA signals present smoother distribution in the time domain, narrower distribution in the frequency domain, as well as, greater complexity and randomness than the whole swallow segments. This observation reflects that complex swallowing events occur during UES opening onset.

4.3. AP distension of maximal UES opening and HRCA signal features

A robust portion of HRCA signal features extracted from whole swallow segments was significantly correlated to the normalized AP distension of maximal UES opening. However more signal features were significantly relevant to the maximal UES AP distension when we performed the analysis from the shorter-segments of UES opening duration alone. This finding is reasonable since the UES AP distension equals zero outside the UES opening segment. Still, additional work to ascertain whether UES opening onset segmentation contributes to more accurate UES distension prediction is warranted as this would represent an important and clinically useful automated measurement adjunct to VF interpretations. Several features from AP vibrations turned out to be statistically significant to significantly associated with the UES AP distension, which is expected given the fact that the UES distends in the AP direction to allow the passage of bolus during swallowing (Jacob et al. 1989, Kahrilas 1997). The swallowing sound, generated predominantly by the valve and pump activity caused by UES opening and closure, also contained more influential HRCA features from UES opening segments rather than whole swallow segments (Cichero & Murdoch 1998).

In aspects of specific features, the AP distension of maximal UES opening was more associated with Lempel-ziv complexity and entropy rate in all HRCA signals, indicating that the amount of information and the extent of complexity of UES opening segments strongly reflects the evaluation of UES AP distension. Our statistical analysis demonstrated that a significant association between HRCA signals and the AP distension of maximal UES opening exists. In addition, narrowing down the segment under investigation from HRCA signals to UES opening only helps strengthen the association by introducing a larger proportion of signal features that were influential to the estimation of UES AP distension.

5. Conclusion

This study provides evidence that a significant number of HRCA signal features are statistically correlated to the maximal UES AP distension measured from VF. This preliminary finding demonstrates the feasibility of potential HRCA signals analysis algorithms in the non-invasive prediction of maximal UES AP distension during swallowing. Further research may investigate the interpretation of detailed swallow events from HRCA signals and developing HRCA screening into a more effective and accurate dysphagia assessment tool in clinical practice.

Acknowledgments

The research reported in this publication was supported by the Eunice Kennedy Shriver National Institute of Child Health & Human Development of the National Institutes of Health under Award Number R01HD092239, while the data was collected under Award Number R01HD074819. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

References

A. E. Stroud, B. W. L. & Wiles, C. M. (2002), 'Inter and intra-rater reliability of cervical auscultation to detect aspiration in patients with dysphagia', *Clinical Rehabilitation* **16**(6), 640–645.

Ahuja, N. K. & Chan, W. W. (2016), 'Assessing upper esophageal sphincter function in clinical practice: a primer', *Current Gastroenterology Reports* **18**(2), 7.

Benjamini, Y. & Hochberg, Y. (1995), 'Controlling the false discovery rate: A practical and powerful approach to multiple testing', *Journal of the Royal Statistical Society. Series B (Methodological)* **57**(1), 289–300.

C. M. Steele, M. Peladeau-Pigeon, C. A. E. B. B. T. G. A. M. N.-M. W. V. N. S. S. M. S. T. T. J. V. A. A. W. & Wolkin, T. S. (2019), 'Reference values for healthy swallowing across the range from thin to extremely thick liquids', *Journal of Speech, Language, and Hearing Research* **64**, 1338–1363.

Cichero, J. A. & Murdoch, B. E. (1998), 'The physiologic cause of swallowing sounds: answers from heart sounds and vocal tract acoustics', *Dysphagia* **13**(1), 39–52.

Cichero, J. A. Y. & Murdoch, B. E. (2002), 'Acoustic signature of the normal swallow: Characterization by age, gender, and bolus volume', *Annals of Otology, Rhinology & Laryngology* **111**(7), 623–632.

Cichero, J. A. Y. & Murdoch, B. E. (2006), *Dysphagia: Foundation, theory and practice*, Wiley, Chichester, England.

Cook, I. J., Dodds, W. J., Dantas, R. O., Massey, B., Kern, M. K., Lang, I. M., Brasseur, J. G. & Hogan, W. J. (1989), 'Opening mechanisms of the human upper esophageal sphincter', *American Journal of Physiology - Gastrointestinal and Liver Physiology* **257**(5), G748–G759.

Dudik, J. M., Coyle, J. L., El-Jaroudi, A., Mao, Z., Sun, M. & Sejdić, E. (2018), 'Deep learning for classification of normal swallows in adults', *Neurocomputing* **285**, 1–9.

Dudik, J. M., Coyle, J. L., Perera, S. & Sejdić, E. (2015), 'Dysphagia screening: Contribution of cervical auscultation signals and modern signal-processing techniques', *IEEE Transactions on Human-Machine Systems* **45**(4), 465–477.

Dudik, J. M., Jestrović, I., Luan, B., Coyle, J. L. & Sejdić, E. (2015), 'Characteristics of dry chin-tuck swallowing vibrations and sounds', *IEEE Transactions on Biomedical Engineering* **62**(10), 2456–2464.

Dudik, J. M., Kurosu, A., Coyle, J. L. & Sejdić, E. (2016), 'A statistical analysis of cervical auscultation signals from adults with unsafe airway protection', *Journal of Neuroengineering and Rehabilitation* **13**(1), 7.

He, Q., Perera, S., Khalifa, Y., Zhang, Z., Mahoney, A. S., Sabry, A., Donohue, C., Coyle, J. L. & Sejdić, E. (2019), 'The association of high resolution cervical auscultation signal features with hyoid bone displacement during swallowing', *IEEE Transactions on Neural Systems and Rehabilitation Engineering* **27**(9), 1810–1816.

Jacob, P., Kahrilas, P. J., Logemann, J. A., Shah, V. & Ha, T. (1989), 'Upper esophageal sphincter opening and modulation during swallowing', *Gastroenterology* **97**, 1469–78.

Kahrilas, P. J. (1997), 'Upper esophageal sphincter function during antegrade and retrograde transit', *American Journal of Medicine* **103**(5A), 56S–60S.

Katz, P. R., Reynolds, H. M., Foust, D. R. & Baum, J. K. (1975), 'Mid-sagittal dimensions of cervical vertebral bodies', *American Journal of Physical Anthropology* **43**(3), 319–326.

Khalifa, Y., Coyle, J. L. & Sejdić, E. (2020), 'Non-invasive identification of swallows via deep learning in high resolution cervical auscultation recordings', *Scientific Reports* **10**, 8704.

Khalifa, Y., Donohue, C., Coyle, J. L. & Sejdić, E. (2020), 'Upper esophageal sphincter opening segmentation with convolutional recurrent neural networks in high resolution cervical auscultation', *IEEE Journal* *of Biomedical and Health Informatics* p. 1.

Kim, Y., Park, T., Oommen, E. & McCullough, G. (2015), 'Upper esophageal sphincter opening during swallow in stroke survivors', *American Journal of Physical Medicine & Rehabilitation* **94**(9), 734–739.

Kurosu, A., Coyle, J. L., Perera, S. & Sejdić, E. (2019), 'Detection of swallow kinematic events from acoustic high-resolution cervical auscultation signals in patients with stroke', *Archives of Physical Medicine and Rehabilitation* **100**(3), 501–508.

Lagarde, M. L., Kamalski, D. M. & Engel-Hoek, L. V. (2016), 'The reliability and validity of cervical auscultation in the diagnosis of dysphagia: a systematic review', *Clinical Rehabilitation* **30**(2), 199–207.

Lee, J., Sejdić, E., Steele, C. M. & Chau, T. (2010), 'Effects of liquid stimuli on dual-axis swallowing accelerometry signals in a healthy population', *Biomedical Engineering Online* **9**(1), 7.

Lee, J. W., Randall, D. R., Evangelista, L. M., Kuhn, M. A. & Belafsky, P. C. (2017), 'Subjective assessment of videlfluoroscopic swallow studies', *Otolaryngology – Head and Neck Surgery* **156**(5), 901–905.

Logemann, J. A. (1998), *Evaluation and Treatment of Swallowing Disorders*, Pro Ed, Austin, TX.

Mao, S., Zhang, Z., Khalifa, Y., Donohue, C., Coyle, J. L. & Sejdić, E. (2019), 'Neck sensor-supported hyoid bone movement tracking during swallowing', *Royal Society Open Science* **6**(7), 181982.

Martin-Harris, B., Brodsky, M. B., Michel, Y., Castell, D. O., Schleicher, M., Sandidge, J., Maxwell, R. & Blair, J. (2008), 'MBS measurement tool for swallow impairment — MBSImp: Establishing a standard', *Dysphagia* **23**(4), 392–405.

Martin-Harris, B. & Jones, B. (2008), 'The videofluorographic swallowing study', *Physical Medicine and Rehabilitation Clinics of North America* **19**(4), 769–785.

Molfenter, S. M. & Steele, C. M. (2014), 'Use of an anatomical scalar to control for sex–based size differences in measures of hyoid excursion during swallowing', *Journal of Speech, Language, and Hearing Research* **57**(3), 768–778.

Movahedi, F., Kurosu, A., Coyle, J. L., Perera, S. & Sejdić, E. (2017), 'Anatomical directional dissimilarities in tri-axial swallowing accelerometry signals', *IEEE Transactions on Neural Systems and Rehabilitation Engineering* **25**(5), 447–458.

Omari, T. I., Ferris, L., Dejaeger, E., Tack, J., Vanbeckevoort, D. & Rommel, N. (2012), 'Upper esophageal sphincter impedance as a marker of sphincter opening diameter', *American Journal of Physiology - Gastrointestinal and Liver Physiology* **302**, G909–G913.

P. Leslie, M. J. Drinnan, P. F. G. A. F. & Wilson, J. A. (2004), 'Reliability and validity of cervical auscultation: A controlled comparison using videofluoroscopy', *Dysphagia* **19**, 231–240.

Rebrion, C., Zhang, Z., Khalifa, Y., Ramadan, M., Kurosu, A., Coyle, J. L., Perera, S. & Sejdić, E. (2019), 'High-resolution cervical auscultation signal features reflect vertical and horizontal displacements of the hyoid bone during swallowing', *IEEE Journal of Translational Engineering in Health and Medicine* **7**, 1–9.

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

Sejdić, E., Kosmisar, V., Steele, C. M. & Chau, T. (2010), 'Baseline characteristics of dual-axis cervical accelerometry signals', *Annals of Biomedical Engineering* **38**(3), 1048–1059.

Sejdić, E., Malandraki, G. A. & Coyle, J. L. (2019), 'Computational deglutition: Using signal-and-image processing methods to understand swallowing and associated disorders', *IEEE Signal Processing Magazine* **36**(1), 138–146.

Sejdić, E., Steele, C. M. & Chau, T. (2010*a*), 'The effects of head movement on dual-axis swallowing accelerometry signals', *BMC Research Notes* **3**, 269.

Sejdić, E., Steele, C. M. & Chau, T. (2010*b*), 'A procedure for denoising dual-axis swallowing acelerometry signal', *Physiological Measurement* **31**(1), 1–9.

Sejdić, E., Steele, C. M. & Chau, T. (2012), 'A method of removal of low frequency components associated with head movement on dual-axis cervical accelerometry signals', *PLOS ONE* **7**(3), 1–8.

Shrout, P. & Fleiss, J. L. (1979), 'Intraclass correlations: Uses in assessing rater reliability', *Psychological Bulletin* **86**(2), 420–8.

Singh, S. & Hamdy, S. (2005), 'The upper oesophageal sphincter', *Neurogastroenterology & Motility* **17**(Suppl. 1), 3–12.

Takahashi, K., Groher, M. E. & Michi, K. (1994), 'Methodology for detection swallowing sounds', *Dysphagia* **9**(1), 54–62.

Yu, C., Khalifa, Y. & Sejdić, E. (2019), Silent aspiration detection in high resolution cervical auscultations, *in* '2019 IEEE Engineering in Medicine and Biology Society (EMBS) International Conference on

Biomedical Health Informatics (BHI)', pp. 1–4.

Zoratto, D. C. B., Chau, T. & Steele, C. M. (2010), 'Hyolaryngeal excursion as the physiological source of swallowing accelerometry signals', *Physiological Measurement* **31**(6), 843.