

Review

Not peer-reviewed version

---

# Artificial Intelligence in the Interpretation of Videofluoroscopic Swallowing Studies: Implications and Advances for Speech Pathologists

---

[Anna M Girardi](#)\*, [Elizabeth A Cardell](#), [Stephen P Bird](#)

Posted Date: 31 August 2023

doi: [10.20944/preprints202308.2166.v1](https://doi.org/10.20944/preprints202308.2166.v1)

Keywords: dysphagia; artificial intelligence; videofluoroscopic swallowing study; deep learning; machine learning; imaging; speech pathology



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Review

# Artificial Intelligence in the Interpretation of Videofluoroscopic Swallowing Studies: Implications and Advances for Speech Pathologists

Anna M. Girardi <sup>1,2,\*</sup>, Elizabeth A. Cardell <sup>1,2,3</sup> and Stephen P. Bird <sup>1,2</sup>

<sup>1</sup> School of Health and Medical Sciences, University of Southern Queensland, Ipswich QLD 4305 Australia

<sup>2</sup> Centre for Health Research, University of Southern Queensland, Toowoomba QLD 4350 Australia

<sup>3</sup> School of Medicine and Dentistry, Griffith University, Gold Coast, QLD 4222 Australia

\* Correspondence: anna.girardi@usq.edu.au

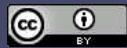
**Abstract:** Radiological imaging is an essential component of a swallowing assessment. Artificial intelligence (AI), and especially deep learning (DL) models, have enhanced the efficiency and efficacy through which imaging is interpreted, and subsequently have important implications for swallow diagnostics and intervention planning. However, the application of AI for the interpretation of videofluoroscopic swallowing studies (VFSS) is still emerging. This review showcases recent literature in the use of AI to interpret VFSS and highlights clinical implications for speech pathologists (SPs). With a surge in AI research, there have been advances made in dysphagia assessment. Several studies have demonstrated successful implementation of DL algorithms to analyze VFSS. Notably, convolutional neural networks (CNNs) have been used to detect pertinent aspects of the swallowing process with high levels of precision. DL algorithms have the potential to streamline VFSS interpretation, improve efficiency and accuracy, and enable precise interpretation of instrumental dysphagia evaluation, which is especially advantageous when access to skilled clinicians is not ubiquitous. By enhancing precision, speed, and depth of VFSS interpretation, SPs can obtain a more comprehensive understanding of swallow physiology and deliver targeted and timely intervention that is tailored towards the individual. This has practical applications for both clinical practice and dysphagia research. As this research area grows and AI technologies progress, the application of DL in the field of VFSS interpretation is clinically beneficial and has the potential to transform dysphagia assessment and management. With broader validation and inter-disciplinary collaborations, AI-augmented VFSS interpretation is likely to transform swallow evaluation and ultimately improve outcomes for individuals with dysphagia.

**Keywords:** dysphagia; artificial intelligence; videofluoroscopic swallowing study; deep learning; machine learning; imaging; speech pathology

## 1. Introduction

Dysphagia (impaired swallowing function) is difficulty moving a food or liquid bolus from the mouth to the stomach [1] and can emerge from age-related changes, neurological changes (e.g., stroke, neurodegenerative diseases), structural changes or anomalies (e.g., cancer, fistulae), and cognitive decline (e.g., dementia) [2]. Dysphagia can negatively impact quality of life and well-being, and result in social isolation, particularly as the severity of the dysphagia increases [3,4]. Further, individuals with dysphagia experience increased risk of aspiration, pneumonia, choking, malnutrition, and dehydration [5]. Therefore, the management of dysphagia is typically complex, given its multi-faceted nature and concordance with other medical conditions. Timely and accurate diagnosis is vital to ensure prompt treatment planning, and to mitigate adverse effects. While many health professionals can be involved in the assessment and management of dysphagia, in clinical practice the management of dysphagia is primarily aligned to the scope of practice of speech pathologists (SPs) [7,8].

Following referral, a speech pathologist (SP) will typically conduct a Clinical Swallow Evaluation (CSE). A CSE enables the SP to gather information on the pre-oral and oral phases of



swallowing, and form hypotheses about the pharyngeal phase of swallowing [6]. A CSE involves obtaining a thorough case history, observing the person's oral structures and oral motor movements (including lips, tongue, and jaw), and assessing the person's ability to manage varying consistencies of food and fluids through controlled trials. Swallowing screening may also include observation of cough occurrence during or following water swallows, and changes in voice quality as a potential marker for aspiration [9]. Following such assessment, the SP can make an informed decision on the persons' ability to manage oral intake and decide on intervention strategies if required. It is difficult to detect or exclude aspiration with CSE or other non-instrumental methods, with prior studies reporting sensitivity of CSE in identifying aspiration of approximately 70% [10–12]. Therefore, an important function of a CSE is to determine the need for instrumental examination of swallowing [13,14].

The videofluoroscopic swallowing study (VFSS) is a gold standard clinical assessment for dysphagia, using x-ray imaging to examine swallowing structures including the oropharynx, pharynx, and esophagus [15]. Trained SPs observe bolus trajectory from the oral cavity to the stomach, using a radiocontrast agent (i.e., barium sulfate). Swallow kinetics and pathophysiological processes are analyzed frame by frame, which is often a time-consuming process, and, although standardized, is susceptible to human error [16]. In response to this, research has utilized computer image analysis programs to track components of VFSS [17–19]. Kellen et al. [20] used computerized image analysis to track hyoid bone movement, with high correlations with manual analysis in dysphagia and non-dysphagic patients. However, clinical usefulness has been restricted because these models often necessitate manual identification of anatomical landmarks. Recently, deep learning (DL), a component of artificial intelligence (AI) has increased the accuracy and efficiency of VFSS interpretation [21,22].

### 1.1. Artificial Intelligence

AI is the simulation of human intelligence by machines and computer systems. With rapid advances in technology, AI has established itself as a transformative force across various fields, including medical diagnostics. AI, through machine learning (ML) and DL algorithms, has the potential to analyze complex data, identify patterns, and offer diagnostic and prognostic insights that surpass human capability in both speed and accuracy [23,24]. In medical imaging, AI applications have been successful in detecting abnormalities in radiographic images and reducing manual workload [24]. Hence, AI is becoming increasingly recognized as a powerful tool in the field of radiology. While seemingly complex, AI advancements have been occurring for many years. The use of AI imaging emerged from the 1960s-1980s with the concept of computer aided diagnosis (CAD), and this developed into further diagnostic applications in the following decade. Rapid advancements have occurred in the last ten years, which has resulted in increased clinical implementation and commercialization.

### 1.2. Machine Learning and Deep Learning

Under the umbrella of AI, ML enables a program to learn and improve from exposure to data, without being explicitly programmed [25]. Put simply, ML is about prediction; feeding data to an algorithm to make predictions without a pre-defined rule [26]. The process involves training a model on a dataset, where an algorithm refines its predictions until it reaches an acceptable level of accuracy. DL, a subfield of ML, is inspired by the structure and function of the human brain and utilizes neural networks within many layers (deep neural networks) to analyze data. In essence, DL networks have self-learning ability [27], and algorithms are inspired by the structure and function of the human brain. Artificial neural networks (ANNs) consist of multiple layers, each performing a discreet task and passing its output to the next layer. There are three primary categories of DL architecture; 1) supervised, 2) unsupervised, and 3) reinforcement learning [25]. Supervised learning involves training a model on a labelled dataset, meaning each input data is paired with the correct output. An example of a supervised DL architecture are convolutional neural networks (CNNs). CNNs are commonly used for image classification tasks and have been applied to chest x-rays for detecting

diseases like pneumonia, tuberculosis and lung cancer, often out-performing human interpreters [28]. More recently, DL algorithms have been explored in analysis of aspects of VFSS. There is potential for AI to improve dysphagia management, by providing diagnostic precision and targeted, individualized treatment planning, however, there remains a paucity of literature on this emerging tool [21]. Further, given the complex nature of AI applications in imaging interpretation, there is need for studies that focus specifically on clinical implementation of AI strategies. Thus, the aim of this review is to examine key studies on the use of DL algorithms in automated VFSS analysis and explore clinical applications for dysphagia management and SP.

## 2. Literature Selection Methodology

A literature search of research published between January 2019 and July 2023 was conducted on 10<sup>th</sup> August 2023 and 23<sup>rd</sup> August 2023 utilising PubMed (MEDLINE), Web of Science (WOS), SCOPUS, and Google Scholar. The most relevant studies on VFSS analysis using ML and DL algorithms were selected. This process consisted of title filtering, abstract filtering, and full article filtering. Search terms included videofluoroscopic swallowing study; VFSS; deep learning; machine learning; artificial intelligence or AI. Article titles were reviewed, and with an abstract review conducted to select only the most relevant articles related to use of DL algorithms to analyse VFSS. Articles were then read in full to identify articles that explicitly fit the aim of this research. The initial search process returned 162 articles, with a total of 11 studies selected for final inclusion.

## 3. Results

The studies that we identified for inclusion in this review, and the clinical applications of the findings are presented below. Following a thematic reduction, three broad themes of the identified studies.

### 3.1. Detection of Aspiration

Aspiration refers to the entry of food or fluids into the trachea, and potentially, the lungs. Aspiration can lead to serious health complications, including aspiration pneumonia, which is a leading cause of morbidity and mortality in dysphagia patients [29]. An important function of VFSS is the ability to visualize whether laryngeal penetration or aspiration occurs and examine the contributing physiological factors. Studies examining the use of VFSS for swallowing diagnostics have used DL to identify presence or absence of aspiration. Kim et al. [21] used a CNN to identify the presence of aspiration in 190 participants with dysphagia with high accuracy. Similarly, Iida et al. [30] used CNNs to detect aspiration in 18, 333 images, and showed deep learning has potential to detect aspiration with precision. Lee et al. [31] used a DL model to detect airway invasion from VFSS images, without clinician input, with 97.2% accuracy in classifying image frames and 93.2% in classifying video files. Kim et al. [32] used the same DL model and found moderate to substantial inter-rater agreement between machine and human. Table 1 summarizes key studies detecting laryngeal penetration or aspiration.

**Table 1.** Summary of key studies detecting laryngeal penetration or aspiration.

Ref	Sample	Algorithm	Findings
[21]	190 participants with dysphagia	CNN <sup>1</sup>	The AUC* of the validation dataset of the VFSS images for the CNN model was 0.942 for normal findings, 0.878 for penetration, and 1.000 for aspiration
[30]	54 participants with aspiration, 75 participants without aspiration	Three CNNs; Simple-Layer, Multiple-Layer, and Modified LeNet	The AUC values at epoch 50 were 0.973, 0.890, and 0.950, respectively, with statistically significant differences between AUC values
[31]	106 participants with dysphagia	Deep CNN using U-Net	Detected airway invasion with overall accuracy of 97.2% in classifying image frames and 93.2% in classifying video files

[32]	49 participants with dysphagia	Deep CNN using U-Net	Kappa coefficients indicate moderate to substantial interrater agreement between AI and human raters in identifying laryngeal penetration or aspiration
------	--------------------------------	----------------------	---

<sup>1</sup> CNN = convolutional neural network, <sup>\*</sup>AUC = area under the curve.

### 3.2. Temporal Parameters of Swallowing Function

To evaluate clinical features and determine rehabilitation strategies of dysphagia, it is crucial to measure the exact response time of the pharyngeal swallowing reflex in a VFSS. Swallowing involves a sequence of precisely coordinated physiological events. Any delay in the oral, pharyngeal, or esophageal phases, or premature initiation, can cause incomplete or inefficient bolus transfer, causing adverse effects. Bandini et al. [33] examined time-points in the pharyngeal phase, namely bolus pass mandible (BPM; where the leading edge of the bolus touches or crosses the shadow of the ramus of the mandible), and upper esophageal sphincter closure (UESC; where the upper esophageal sphincter (UES) achieves closure behind the bolus tail). CCN-based approaches were able to detect these measures with high accuracy, which is congruent with research by Lee et al. [22], who automatically detected the response time for the pharyngeal swallowing reflex with high accuracy. Jeong et al. [34] measured seven temporal parameters with relatively high accuracy, however encountered difficulty measuring pharyngeal delay, and laryngeal vestibule closure, presumably due to innate variability of these phases in the swallow mechanism. Table 2 summarizes key studies using temporal measures of swallowing function.

**Table 2.** Summary of key studies measuring temporal parameters of swallowing.

Ref	Sample	Algorithm	Findings
[33]	78 healthy participants	Compared multiple CNN algorithms	Pearson's correlation coefficient of 0.951 for BPM, and 0.996 for UESC- <sup>~</sup>
[22]	27 participants with subjective dysphagia	3D CNN	Average success rate of detection 'during the pharyngeal phase' of 97.5%
[34]	547 VFSS video clips from patients with dysphagia	3D CNN	Average accuracy of 0.864 to 0.981

<sup>1</sup> CNN = convolutional neural network, <sup>\*</sup>AUC = area under the curve, <sup>~</sup>UESC = upper oesophageal sphincter, BPM = bolus pass mandible.

### 3.3. Hyoid Bone Movement

The hyoid bone is in the anterior neck, suspended by ligaments and muscles. Its localization and movement during swallowing are important for various reasons. The swallow reflex is initiated by touch receptors in the pharynx, which results in the forward and upward movement of the hyoid [35]. The upward movement of the hyoid assists in pulling open the UES, allowing the bolus to move from the pharynx into the esophagus. If the hyoid doesn't move correctly, there can be incomplete UES opening, leading to increased pharyngeal residue or increased risk of aspiration [36]. Three reviewed studies investigated hyoid bone detection and tracking through DL frameworks. Hsiao et al. [37] examined 409 videos from 233 patients using fully automated hyoid bone localization and tracking. They found excellent inter-rater reliability of hyoid bone detection between the algorithm and a group of three human annotators. Similarly, Kim et al. [38] utilized automated hyoid bone tracking and designed a network that can detect salient objects in VFSS images. Zhang et al. [39] also presented a model that could automatically detect the hyoid bone; however, inaccuracies in tracking were a limitation in this study. Lee et al. [40] proposed two types of DL networks for tracking the hyoid in VFSS images with high levels of accuracy. Table 3 summarizes key studies using hyoid bone movement.

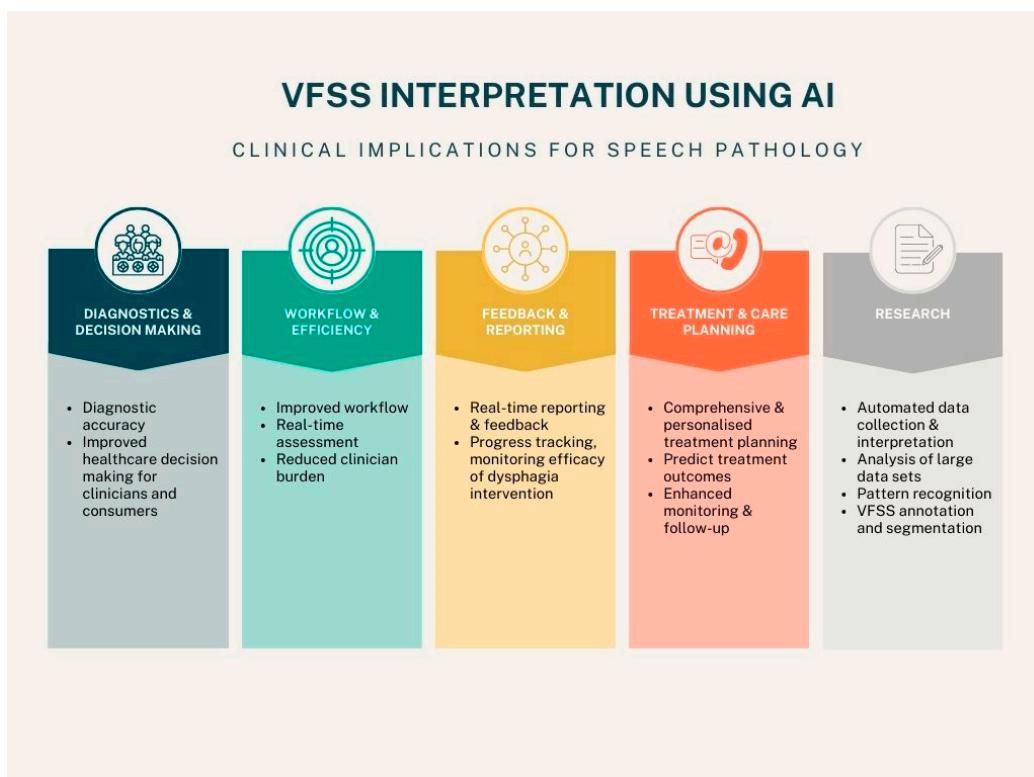
**Table 3.** Summary of key studies using AI Hyoid Bone Movement Detection or Measurement.

Ref	Sample	Algorithm	Findings
[37]	44 participants with dysphagia	CNN-based algorithm, the Cascaded Pyramid Network	Excellent inter-rater reliability for hyoid bone detection, good-to-excellent inter-rater reliability for displacement and the average velocity of the hyoid bone in horizontal or vertical directions, moderate-to-good reliability in calculating the average velocity in horizontal direction
[38]	207 participants with dysphagia	CNN; U-Net	mAP of 91% for hyoid bone detection
[39]	265 participants with dysphagia	CNN <sup>1</sup> ; SSD*	mAP <sup>~</sup> of 89.14%
[40]	77 participants; healthy individuals, and individuals with Parkinson's Disease, and stroke.	CNN; MDNet	DSC <sup>β</sup> results for the proposed method were 0.87 for healthy individuals, 0.88 for patients with Parkinson's Disease, 0.85 for patients with stroke, and a total of 0.87.

<sup>1</sup> CNN = convolutional neural network\* SSD = single-shot detector, <sup>~</sup>mAP = mean average precision, <sup>β</sup>DSC = dice similarity coefficient.

#### 4. Implications for Speech Pathology

AI has been increasingly integrated into various healthcare fields and has the potential to increase clinical efficiency, accuracy, and improve patient outcomes. AI also has the potential to capture diverse data, with enhanced precision, for research. This research can then be used to inform targeted interventions. In the field of speech pathology, VFSS are considered the gold standard for dysphagia assessment. However, VFSS are time-consuming to conduct, and can be laborious to interpret, particularly for inexperienced clinicians. Further, VFSS interpretation is prone human error (32). DL frameworks have the potential to improve the accuracy and speed with which VFSS are interpreted, thus have several clinical applications depending on the method of detection used. Figure 1 outlines five key areas of AI interpretation of VFSS in speech pathology, including diagnostics and decision making, workflow and efficiency, feedback and reporting, treatment and care planning, and research.

**Figure 1.** VFSS Interpretation Using AI.

The elements of VFSS that have been investigated in key AI studies identified in this review comprise three broad groups: laryngeal penetration and aspiration detection, temporal aspects of swallowing, and hyoid bone detection and localization. Each of these areas has clinical implications for SPs and individuals with dysphagia.

#### *4.1. Clinical Applications of AI Detection of Laryngeal Penetration and Aspiration*

Studies have shown AI can detect presence of laryngeal penetration or aspiration with accuracy [21,30–32]. Accurate detection of aspiration has the potential to improve patient outcomes. Aspiration, if undetected, can result in adverse health outcomes, and is a significant cause of mortality and morbidity in vulnerable clinical populations [41]. Accurate and timely detection of aspiration also has the potential to improve informed decision making. When healthcare providers, people with dysphagia, and their families, are fully informed about the presence and extent of aspiration, they can make reasoned decisions about their health and healthcare. This may include modification of diet or fluids, or alternative feeding or hydration methods.

#### *4.2. Clinical Applications of AI Measurement of Temporal Parameters of Swallowing*

Swallowing involves rapid, sequential physiologic components. The late oral and early pharyngeal components of swallowing are arguably the most crucial from a safety perspective. When the bolus reaches the oropharynx, the pharyngeal swallow is initiated. While the onset of the pharyngeal swallow is variable relative to bolus position [42], once initiated, there are a series of pharyngeal and laryngeal events that protect the airway and clear ingested material from the pharynx [43]. While an intricate description of the physiology of swallowing is outside the scope of this review, the reader will appreciate the alignment of pharyngeal motion, with physiological events, for successful swallow function. AI frameworks can be a clinically useful tool for estimating the absence or delayed response time of the swallowing reflex in patients with dysphagia and improving poor inter-rater reliability of evaluation of response time of pharyngeal swallowing reflex between expert and unskilled clinicians. The frameworks in these studies can be used to provide considerable clinical information for dysphagia treatment. Clinically, this DL application can also be expanded to other spatiotemporal parameters in VFSS.

#### *4.3. Clinical Applications of AI Hyoid Bone Movement Detection or Measurement*

The hyoid bone is a salient anatomical feature, commonly monitored during analysis of VFSS [44]. The anterior-superior movement of the hyoid bone plays a significant role in preventing aspiration and opening the UES to enable the food bolus to move into the esophagus [45]. Thus, evaluating hyoid bone movement during VFSS is an important factor in clinical dysphagia management. Manual tracking of hyoid movement is considered the gold standard for SPs in dysphagia management, however, of course is prone to human error and is time-consuming. Zhang et al. [39] and Lee et al. [40] used DL models to track the hyoid bone with precision. During swallowing, the hyoid bone moves upwards and forwards, at times becoming obscured by the shadow of the mandible, thus becoming difficult to detect by human reviewers. Both models could detect the hyoid bone, even when obscured, and therefore appear to be promising as a widely applicable pre-processing step for dysphagia research and, eventually, clinically [39,40].

### **5. Limitations and Future Directions**

While the benefits of using AI in VFSS interpretation are likely to outweigh the risks, there are various limitations to consider in its implementation. First, the accuracy of AI depends largely on the quality and diversity of training data. Studies require access to VFSS data to train AI models. Therefore, if certain patient demographics or groups are not represented in training data, widespread application may be limited. A second limitation is that the complexity of the human swallowing process may be difficult to quantify in AI models. The nuances and individual differences in swallowing function that a trained clinician may identify may not be elucidated by AI models.

Notably, the studies reviewed in this paper utilized heterogeneous participant groups, with varying dysphagia presentation and severity both within and between studies. A series of studies with homogeneity around dysphagia cause would be useful to train AI models to identify diagnostic variability within specific populations. The wider clinical context will need to remain at the center of AI interpretation for VFSS. A clinical evaluation of swallowing involves consideration of not only instrumental evaluation of swallowing, but the person's medical history, presentation and individual needs and goals. AI interpretation will provide another tool that must be applied within the larger context of client-centered care. As the use of AI becomes widely spread, there is risk of clinicians becoming reliant on the AI interpretation. When clinicians accept the AI interpretation without critique or use of clinical expertise, ethical and diagnostic issues may arise [46,47]. Further, identification of cases that do not fit within the defined or 'trained' contexts of AI models will be a challenge.

Despite these limitations, the integration of AI into VFSS interpretation holds promise for enhancing diagnostic accuracy, automating routine components of analysis, and assisting clinicians with what can be a time-consuming task. In areas where access to expert clinicians is limited, AI can enable rapid, accurate assessment of swallowing, and facilitate targeted intervention for individuals, regardless of the skill level of the treating SP. AI, at least in its initial implementation in the field, should be viewed as a complement to human expertise in swallow diagnostics. It is anticipated that with broader clinical validation and inter-disciplinary collaborations, AI-augmented VFSS interpretation will become the cornerstone of dysphagia management in the future.

**Author Contributions:** Conceptualization, A.G., S.P.B; methodology, A.G., E.C.; data curation, A.G.; writing—original draft preparation, A.G.; writing—review and editing, E.C., S.P.B.; project administration, A.G., E.C., S.P.B. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** No new data were created or analyzed in this study. Data sharing is not applicable to this article.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. O'Rourke, F.; Vickers, K.; Upton, C.; Chan, D. Swallowing and Oropharyngeal Dysphagia. *Clin. Med.* **2014**, *14*, 196–199, doi:10.7861/clinmedicine.14-2-196.
2. Azer, S.A.; Kanugula, A.K.; Kshirsagar, R.K. Dysphagia. In *StatPearls*; StatPearls Publishing: Treasure Island, FL, USA, 2023.
3. Ekberg, O.; Hamdy, S.; Woisard, V.; Wuttge-Hannig, A.; Ortega, P. Social and Psychological Burden of Dysphagia: Its Impact on Diagnosis and Treatment. *Dysphagia* **2002**, *17*, 139–146. doi: 10.1007/s00455-001-0113-5.
4. Smith, R.; Bryant, L.; Hemsley, B. The True Cost of Dysphagia on Quality of Life: The Views of Adults with Swallowing Disability. *Int. J. Lang. Commun. Disord.* **2023**, *58*, 451–466, doi:10.1111/1460-6984.12804.
5. Altman, K. W.; Yu, G.; Schaefer, S. D. Consequence of Dysphagia in the Hospitalized Patient: Impact on Prognosis and Hospital Resources. *Arch. Otolaryngol. Head Neck Surg.* **2010**, *136*, 784–789. doi:10.1001/archoto.2010.129
6. Barnaby-Mann, G.; Lenius, K. The Bedside Examination in Dysphagia. *Phys. Med. Rehabil. Clin. N. Am.* **2008**, *19*, 747–768, doi:10.1016/j.pmr.2008.05.008.
7. Speech Pathology Australia Clinical Guidelines. Available online: [https://www.speechpathologyaustralia.org.au/SPAweb/Members/Clinical\\_Guidelines/spaweb/Members/Clinical\\_Guidelines/Clinical\\_Guidelines.aspx?hkey=f66634e4-825a-4f1a-910d-644553f59140](https://www.speechpathologyaustralia.org.au/SPAweb/Members/Clinical_Guidelines/spaweb/Members/Clinical_Guidelines/Clinical_Guidelines.aspx?hkey=f66634e4-825a-4f1a-910d-644553f59140) (accessed on 28 August 2023).
8. Scope of Practice in Speech-Language Pathology Available online: <https://www.asha.org/policy/sp2016-00343/> (accessed on 28 August 2023).
9. Garand, K.L. F; McCullough, G.; Crary, M.; Arvedson, J.C.; Dodrill, P. Assessment Across the Life Span: The Clinical Swallow Evaluation Available online: [https://pubs.asha.org/doi/epdf/10.1044/2020\\_AJSLP-19-00063](https://pubs.asha.org/doi/epdf/10.1044/2020_AJSLP-19-00063) (accessed on 23 August 2023).

10. Lynch, Y.T.; Clark, B.J.; Macht, M.; White, S.D.; Taylor, H.; Wimbish, T.; Moss, M. The Accuracy of the Bedside Swallowing Evaluation for Detecting Aspiration in Survivors of Acute Respiratory Failure. *J. Crit. Care* **2017**, *39*, 143–148, doi:10.1016/j.jcrc.2017.02.013.
11. DePippo, K.L.; Holas, M.A.; Reding, M.J. Validation of the 3-oz Water Swallow Test for Aspiration Following Stroke. *Arch. Neurol.* **1992**, *49*, 1259–1261.
12. McCullough, G.H.; Rosenbek, J.C.; Wertz, R.T.; McCoy, S.; Mann, G.; McCullough, K. Utility of Clinical Swallowing Examination Measures for Detecting Aspiration Post-Stroke. *J. Speech Lang. Hear. Res.* **2005**, *48*, 1280–1293, doi:10.1044/1092-4388(2005/089).
13. Desai, R.V. Build a Case for Instrumental Swallowing Assessments in Long-Term Care Available online: <https://leader.pubs.asha.org/doi/10.1044/leader.OTP.24032019.38> (accessed on 23 August 2023).
14. Warner, H.; Coutinho, J.M.; Young, N. Utilization of Instrumentation in Swallowing Assessment of Surgical Patients during COVID-19. *Life* **2023**, *13*, 1471, doi:10.3390/life13071471.
15. Costa, M.M.B. Videofluoroscopy: The Gold Standard Exam for Studying Swallowing and Its Dysfunction. *Arq. Gastroenterol.* **2010**, *47*, 327–328, doi:10.1590/S0004-28032010000400001.
16. Kerrison, G.; Miles, A.; Allen, J.; Heron, M. Impact of Quantitative Videofluoroscopic Swallowing Measures on Clinical Interpretation and Recommendations by Speech-Language Pathologists. *Dysphagia* **2023**, doi:10.1007/s00455-023-10580-3.
17. Hossain, I.; Roberts-South, A.; Jog, M.; El-Sakka, M.R. Semi-Automatic Assessment of Hyoid Bone Motion in Digital Videofluoroscopic Images. *Comput. Methods Biomed. Eng. Imaging Vis.* **2014**, *2*, 25–37, doi:10.1080/21681163.2013.833859.
18. Natarajan, R.; Stavness, I.; Pearson, W. Semi-Automatic Tracking of Hyolaryngeal Coordinates in Videofluoroscopic Swallowing Studies. *Comput. Methods Biomed. Eng. Imaging Vis.* **2017**, *5*, 379–389, doi:10.1080/21681163.2015.1046190.
19. Lee, W.H.; Chun, C.; Seo, H.G.; Lee, S.H.; Oh, B.-M. STAMPS: Development and Verification of Swallowing Kinematic Analysis Software. *Biomed. Eng. OnLine* **2017**, *16*, 120, doi:10.1186/s12938-017-0412-1.
20. Kellen, P.M.; Becker, D.L.; Reinhardt, J.M.; Van Daele, D.J. Computer-Assisted Assessment of Hyoid Bone Motion from Videofluoroscopic Swallow Studies. *Dysphagia* **2010**, *25*, 298–306, doi:10.1007/s00455-009-9261-9.
21. Kim, J.K.; Choo, Y.J.; Choi, G.S.; Shin, H.; Chang, M.C.; Chang, M.; Park, D. Deep Learning Analysis to Automatically Detect the Presence of Penetration or Aspiration in Videofluoroscopic Swallowing Study. *J. Korean Med. Sci.* **2021**, doi:10.21203/rs.3.rs-884186/v1.
22. Lee, J.T.; Park, E.; Hwang, J.-M.; Jung, T.-D.; Park, D. Machine Learning Analysis to Automatically Measure Response Time of Pharyngeal Swallowing Reflex in Videofluoroscopic Swallowing Study. *Sci. Rep.* **2020**, *10*, 14735, doi:10.1038/s41598-020-71713-4.
23. Jiang, F.; Jiang, Y.; Zhi, H.; Dong, Y.; Li, H.; Ma, S.; Wang, Y.; Dong, Q.; Shen, H.; Wang, Y. Artificial Intelligence in Healthcare: Past, Present and Future. *Stroke Vasc. Neurol.* **2017**, *2*, doi:10.1136/svn-2017-000101.
24. Litjens, G.; Kooi, T.; Bejnordi, B.E.; Setio, A.A.A.; Ciompi, F.; Ghafoorian, M.; van der Laak, J.A.W.M.; van Ginneken, B.; Sánchez, C.I. A Survey on Deep Learning in Medical Image Analysis. *Med. Image Anal.* **2017**, *42*, 60–88, doi:10.1016/j.media.2017.07.005.
25. Janiesch, C.; Zschech, P.; Heinrich, K. Machine Learning and Deep Learning. *Electron. Mark.* **2021**, *31*, 685–695, doi:10.1007/s12525-021-00475-2.
26. Jordan, M.I.; Mitchell, T.M. Machine Learning: Trends, Perspectives, and Prospects. *Science* **2015**, *349*, 255–260, doi:10.1126/science.aaa8415.
27. Busnatu, S.; Niculescu, A.G.; Bolocan, A.; Petrescu, G.E.D.; Păduraru, D.N.; Năstăsă, I.; Lupușoru, M.; Geantă, M.; Andronic, O.; Grumezescu, A.M.; et al. Clinical Applications of Artificial Intelligence—An Updated Overview. *J. Clin. Med.* **2022**, *11*, 2265, doi:10.3390/jcm11082265.
28. Rajpurkar, P.; Irvin, J.; Zhu, K.; Yang, B.; Mehta, H.; Duan, T.; Ding, D.; Bagul, A.; Langlotz, C.; Shpanskaya, K.; et al. CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning 2017.
29. Langmore, S.E.; Terpenning, M.S.; Schork, A.; Chen, Y.; Murray, J.T.; Lopatin, D.; Loesche, W.J. Predictors of Aspiration Pneumonia: How Important Is Dysphagia? *Dysphagia* **1998**, *13*, 69–81, doi:10.1007/PL00009559.
30. Iida, Y.; Näppi, J.; Kitano, T.; Hironaka, T.; Katsumata, A.; Yoshida, H. Detection of Aspiration from Images of a Videofluoroscopic Swallowing Study Adopting Deep Learning. *Oral Radiol.* **2023**, *39*, 553–562, doi:10.1007/s11282-023-00669-8.
31. Lee, S.J.; Ko, J.Y.; Kim, H.I.; Choi, S.I. Automatic Detection of Airway Invasion from Videofluoroscopy via Deep Learning Technology. *Appl. Sci.* **2020**, *10*, 6179, doi:10.3390/app10186179.
32. Kim, Y.; Kim, H.I.; Park, G.S.; Kim, S.Y.; Choi, S.I.; Lee, S.J. Reliability of Machine and Human Examiners for Detection of Laryngeal Penetration or Aspiration in Videofluoroscopic Swallowing Studies. *J. Clin. Med.* **2021**, *10*, 2681, doi:10.3390/jcm10122681.

33. Bandini, A.; Steele, C.M. The Effect of Time on the Automated Detection of the Pharyngeal Phase in Videofluoroscopic Swallowing Studies. *Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.* **2021**, 3435–3438, doi:10.1109/EMBC46164.2021.9629562.
34. Jeong, S.Y.; Kim, J.M.; Park, J.E.; Baek, S.J.; Yang, S.N. Application of Deep Learning Technology for Temporal Analysis of Videofluoroscopic Swallowing Studies; *Research Square* **2002**. doi: 10.21203/rs.3.rs-2311543/v1.
35. Matsuo, K.; Palmer, J.B. Anatomy and Physiology of Feeding and Swallowing – Normal and Abnormal. *Phys. Med. Rehabil. Clin. N. Am.* **2008**, 19, 691–707, doi:10.1016/j.pmr.2008.06.001.
36. Cook, I.J.; Dodds, W.J.; Dantas, R.O.; Massey, B.; Kern, M.K.; Lang, I.M.; Brasseur, J.G.; Hogan, W.J. Opening Mechanisms of the Human Upper Esophageal Sphincter. *Am. J. Physiol.* **1989**, 257, G748–759, doi:10.1152/ajpgi.1989.257.5.G748.
37. Hsiao, M.Y.; Weng, C.H.; Wang, Y.C.; Cheng, S.H.; Wei, K.C.; Tung, P.Y.; Chen, J.Y.; Yeh, C.Y.; Wang, T.G. Deep Learning for Automatic Hyoid Tracking in Videofluoroscopic Swallow Studies. *Dysphagia* **2023**, 38, 171–180, doi:10.1007/s00455-022-10438-0.
38. Kim, H.I.; Kim, Y.; Kim, B.; Shin, D.Y.; Lee, S.J.; Choi, S.I. Hyoid Bone Tracking in a Videofluoroscopic Swallowing Study Using a Deep-Learning-Based Segmentation Network. *Diagn. Basel Switz.* **2021**, 11, 1147, doi:10.3390/diagnostics11071147.
39. Zhang, Z.; Coyle, J.L.; Sejdic, E. Automatic Hyoid Bone Detection in Fluoroscopic Images Using Deep Learning. *Sci. Rep.* **2018**, 8, 12310–12310, doi:10.1038/s41598-018-30182-6.
40. Lee, D.; Lee, W.H.; Seo, H.G.; Oh, B.-M.; Lee, J.C.; Kim, H.C. Online Learning for the Hyoid Bone Tracking During Swallowing with Neck Movement Adjustment Using Semantic Segmentation. *IEEE Access* **2020**, 8, 157451–157461, doi:10.1109/ACCESS.2020.3019532.
41. Shin, D.; Lebovic, G.; Lin, R.J. In-Hospital Mortality for Aspiration Pneumonia in a Tertiary Teaching Hospital: A Retrospective Cohort Review from 2008 to 2018. *J. Otolaryngol. Head Neck Surg.* **2023**, 52, 23, doi:10.1186/s40463-022-00617-2.
42. Martin-Harris, B.; Brodsky, M.B.; Michel, Y.; Lee, F.-S.; Walters, B. Delayed Initiation of the Pharyngeal Swallow: Normal Variability in Adult Swallows. *J. Speech Lang. Hear. Res.* **2007**, 50, 585–594, doi:10.1044/1092-4388(2007/041).
43. Martin-Harris, B.; Jones, B. The Videofluorographic Swallowing Study. *Phys. Med. Rehabil. Clin. N. Am.* **2008**, 19, 769–785, doi:10.1016/j.pmr.2008.06.004.
44. Donohue, C.; Mao, S.; Sejdic, E.; Coyle, J.L. Tracking Hyoid Bone Displacement During Swallowing Without Videofluoroscopy Using Machine Learning of Vibratory Signals. *Dysphagia* **2021**, 36, 259–269, doi:10.1007/s00455-020-10124-z.
45. Wei, K.C.; Hsiao, M.Y.; Wang, T.G. The Kinematic Features of Hyoid Bone Movement during Swallowing in Different Disease Populations: A Narrative Review. *J. Formos. Med. Assoc.* **2022**, 121, 1892–1899, doi:10.1016/j.jfma.2022.04.007.
46. Goisauf, M.; Abadía, M. Ethics of AI in Radiology: A Review of Ethical and Societal Implications. *Front. Big Data* **2022**, 5, doi:10.3389/fdata.2022.850383.
47. Naik, N.; Hameed, B.M.Z.; Shetty, D.K.; Swain, D.; Shah, M.; Paul, R.; Aggarwal, K.; Ibrahim, S.; Patil, V.; Smriti, K.; et al. Legal and Ethical Consideration in Artificial Intelligence in Healthcare: Who Takes Responsibility? *Front. Surg.* **2022**, 9.

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.