CADTH

CADTH RAPID RESPONSE REPORT:
SUMMARY WITH CRITICAL APPRAISAL

Fibreoptic Endoscope Evaluation versus Video Fluoroscopic Swallowing Exams for Patients with Dysphagia: A Review of Diagnostic Accuracy and Cost-Effectiveness

Service Line:	Rapid Response Service
Version:	1.0
Publication Date:	November 15, 2019
Report Length:	18 Pages

CADTH

Authors: Rob Edge PhD, Charlene Argáez

Cite As: Fibreoptic endoscope evaluation versus video fluoroscopic swallowing exams for patients with dysphagia: a review of diagnostic accuracy and cost-effectiveness. Ottawa: CADTH; 2019 Nov. (CADTH rapid response report: summary with critical appraisal).

ISSN: 1922-8147 (online)

About CADTH: CADTH is an independent, not-for-profit organization responsible for providing Canada's health care decision-makers with objective evidence to help make informed decisions about the optimal use of drugs, medical devices, diagnostics, and procedures in our health care system.

Funding: CADTH receives funding from Canada's federal, provincial, and territorial governments, with the exception of Quebec.

Questions or requests for information about this report can be directed to Requests@CADTH.ca

Abbreviations

FEES	Fibreoptic Endoscope Evaluation of Swallowing System
NPV	Negative predictive value
OPES	Oro-pharyngo-oesophageal scintigraphy
PA-score	Penetration-aspiration score
PPV	Positive predictive value
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-analyses
QUADAS	Quality Assessment of Diagnostic Accuracy Studies
SR	Systematic Review
VFSS	Video Fluoroscopic Swallowing Exam/Study

Context and Policy Issues

Dysphagia, or difficulty in swallowing, is a common complication of many conditions including Parkinson's disease,[1] following prolonged intubation,[2,3] head and neck cancer, stroke, tracheostomized patients, vocal cord paralysis, myotonic dystrophy, critical illness polyneuropathy, osteophytes, myasthenia gravis, and progressive supranuclear palsy.[3] Dysphagia can increase the risk of pneumonia, malnutrition, and dehydration resulting in increased mortality, morbidity, and decreased quality of life.[4] An accurate and timely diagnosis of dysphagia has implications for effective interventions for these complications.[4]

Video-fluoroscopic swallowing study (VFSS) is a widely practiced procedure used for detecting dysphagia, scoring of dysphagia parameters, and revealing etiology.[3] VFSS is often considered the gold standard of dysphagia diagnosis however it exposes patients to radiation, requires a trained radiologist, and has been regarded as expensive.[5] Fibreoptic endoscope evaluation of swallowing system (FEES) offers an alternative or complementary diagnostic modality that offers a direct visualization of the pharyngeal stage of swallowing. FEES does not expose patients to radiation and is regarded as more convenient, however it is more invasive than VFSS.[6]

This report aims to retrieve and review relevant evidence on the comparative diagnostic accuracy of VFSS compared to FEES for the diagnosis and detection of dysphagia. In addition, this report aims to retrieve and review comparative cost-effectiveness studies on these two diagnostic modalities for patients with suspected dysphagia presenting in hospital or in outpatient settings.

Research Questions

1. What is the diagnostic accuracy of fibreoptic endoscopic versus video fluoroscopic swallowing exams in patients with suspected dysphagia?

2. What is the cost-effectiveness of fibreoptic endoscopic versus video fluoroscopic swallowing exams for patients with suspected dysphagia?

Key Findings

This report identified evidence of limited quality from one systematic review that conducted a meta-analysis of six studies, and two prospective comparative studies that supported both fibreoptic endoscope evaluation of swallowing system (FEES) and video fluoroscopic swallowing exam/study (VFSS) as suitable diagnostic procedures for dysphagia. With

regard to diagnostic accuracy of dysphagia parameters a lack of evidence for consistent diagnostic accuracy differences was identified. Limitations of the identified evidence included the absence of a suitable reference standard, a lack of safety data, and a lack of patient-related outcomes making the significance of diagnostic accuracy comparisons unclear. Given the lack of evidence for significant differences in diagnostic accuracy, other factors could be considered in the decision to implement FEES or VFSS. No cost-effectiveness evidence was identified comparing VFSS to FEES.

Methods

Literature Search Methods

A limited literature search was conducted by an information specialist on key resources including PubMed, the Cochrane Library, the University of York Centre for Reviews and Dissemination (CRD) databases, the websites of Canadian and major international health technology agencies, as well as a focused Internet search. The search strategy was comprised of both controlled vocabulary, such as the National Library of Medicine's MeSH (Medical Subject Headings), and keywords. The main search concepts were dysphagia and videofluoroscopy/fibreoptic endoscopic evaluation. Filters were applied to limit the retrieval to health technology assessments, systematic reviews, and meta analyses, economic studies, randomized controlled trials, and non-randomized studies. Where possible, retrieval was limited to the human population. The search was also limited to English language documents published between January 1, 2014 and October 21, 2019.

Selection Criteria and Methods

One reviewer screened citations and selected studies. In the first level of screening, titles and abstracts were reviewed and potentially relevant articles were retrieved and assessed for inclusion. The final selection of full-text articles was based on the inclusion criteria presented in Table 1.

Table 1: Selection Criteria

Population	Patients with suspected dysphagia in the hospital or outpatient setting
Intervention	Fibreoptic endoscopic evaluation of swallowing study or system (all types); also known as nasendoscopy
Comparator	Video fluoroscopic swallowing exam or study (VFSE or VFSS)
Outcomes	Q1: Diagnostic accuracy: detection of pathology of dysphagia, sensitivity, specificity, accuracy Q2: Cost-effectiveness
Study Designs	Health technology assessments, systematic reviews, meta-analyses, randomized controlled trials, non-randomized studies, and economic evaluations

FEES = flexible endoscopic evaluation of swallowing; VFSS = videofluoroscopic swallowing study.

Exclusion Criteria

Articles were excluded if they did not meet the selection criteria outlined in Table 1.

Critical Appraisal of Individual Studies

One reviewer critically appraised the included diagnostic accuracy studies using the QUADAS-2 tool,[7] and the systematic review with the AMSTAR 2 tool.[8] Summary scores

were not calculated for the included studies; rather, a review of the strengths and limitations of each included study was described narratively.

Summary of Evidence

Quantity of Research Available

A total of 777 citations were identified in the literature search. Following screening of titles and abstracts, 763 citations were excluded and 14 potentially relevant studies from the electronic search were retrieved for full-text review. No additional potentially relevant publication was retrieved from the grey literature search for full-text review. Of these 14 potentially relevant articles, five were excluded for examining an irrelevant intervention, three were excluded for examining an irrelevant comparator, one was already included in the selected systematic review, and two were excluded for being non-systematic review articles. Three publications met the inclusion criteria and were included in this report and were comprised of one systematic review (SR) with meta-analysis (MA),[4] and two non-randomized clinical studies[5,9] that compared the diagnostic accuracy of fibreoptic endoscope evaluation of swallowing system (FEES) to video-fluoroscopic swallowing study (VFSS). No cost-effectiveness studies were identified that met the inclusion criteria. Appendix 1 presents the PRISMA[10] flowchart of the study selection.

Summary of Study Characteristics

Additional details regarding the characteristics of the included studies are provided in Appendix 2.

Study Design and Country of Origin

This report identified three studies that met the selection criteria in Table 1. One SR with a meta-analysis originated from Columbia.[4] This SR searched for all relevant comparative studies published between January 1, 1988 and March 31, 2016.

Two clinical studies that examined the diagnostic accuracy of FEES versus VFSS in a prospective comparative design were also identified; Scharitzer et al. was conducted in Austria in 2019,[5] and Fattori et al. was conducted in Italy in 2016.[9]

Patient Population

The SR included studies that examined adults with suspected functional oropharyngeal dysphagia, and excluded studies that included patients with mechanical dysphagia. In total, 198 patients were examined in studies included in the SR.[4]

The diagnostic study from Austria, by Scharitzer et al., looked at adults following operative interventions for pharyngeal or laryngeal carcinoma, either ongoing or completed radiotherapy, who had postoperative dysphagia with a high risk for aspiration yet had a stable clinical condition. This study reported results on 29 enrolled patients following the exclusion of two patients.[5]

The most inclusive diagnostic study from Italy, by Fattori et al., examined patients that were referred for dysphagia studies and therefore most closely met the population of interest for this report in that patients had suspected dysphagia. This study reported results on all 60 enrolled patients.[9]

Additional details of enrolled patients in the two comparative studies is tabulated in Appendix 2.

Interventions and Comparators

All studies compared fibreoptic endoscopic evaluation of swallowing study or system (FEES) with video fluoroscopic swallowing exam or study (VFSE or VFSS).[4,5,9] The SR did not specify that included studies examined any specific instrumentation manufacturer or design to conduct either VFSS or FEES.[4]

Scharitzer et al. conducted FEES and VFSS simultaneously using a fluoroscopy unit, a flexible endoscope connected to an EndoCompact Mobile Unit, and digital storage of VFSS images on a picture archiving and communication system (PACS).[5] This study also reported that diagnostic studies were conducted by an otorhinolaryngologist or phoniatrician with 15 years of experience and a radiologist with 17 years of experience.[5] The investigators used a water-soluble, non-ionic contrast medium mixed with blue food colouring for the endoscopic examination. Bolus volumes examined were 3 mL, 5 mL, 10 mL, and 20 mL, and consistencies were modified using a cornstarch product. Scharitzer et al. did not compare the accuracy of FEES and VFSS to diagnose dysphagia, only their comparative accuracy in diagnosing parameters of dysphagia.[5]

Fattori et al. conducted FEES, VFSS, and oro-pharyngo-oesophageal scintigraphy (OPES) on the same day. Two speech-language pathologists conducted FEES using a flexible fliberoptic rhinopharyngolaryngoscope connected to a charge-coupled device (CCD) camera and colour monitor which was recorded digitally on a Digital Swallowing Workstation. VFSS was performed with a Clinodigit Compact Xframe Italray device on a PACS system.[9] For FEES, patients were administered 5 mL of semi-solid (jellied drink) or liquid boluses (water with methylene blue). For VFSS the contrast medium used was Prontobario HD diluted with different volumes of water for different consistencies. Bolus challenges were 5mL and consisted of two different consistencies (liquid and semi-solid).[9]

Outcomes

The SR used a composite reference standard of FEES and VFSS, in addition to comparisons between FEES and VFSS using the other as a reference standard to report relative sensitivity and specificity in four parameters of dysphagia: aspiration, penetration, residues, and premature spillage. The composite reference standard was built using the "OR rule" where a finding was considered positive if either VFSS or FEES was positive and negative only if both tests were negative. The methodology of combining each comparison using the other as a reference standard and the composite reference standard was described.[4]

Scharitzer et al. reported the interrater agreement of both FEES and VFSS for penetration-aspiration score (PA-score), retentions valleculae, retentions piriform sinus, and time of triggering. The PA-score uses anatomical markers (vocal folds, larynx, and trachea) to score the depth of penetration and aspiration on a scale from one to eight representing increasing severity. The investigators also reported differences of FEES and VFSS in grouped PA-scores using different bolus volumes (3 mL, 5 mL, 10 mL, and 20 mL) and consistencies (nectar, pudding, and liquid).[5]

Fattori et al. reported the sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and validity of three parameters of dysphagia (premature spillage, hypopharyngeal residue, and aspiration) for FEES and OPES using VFSS as the reference

standard (gold standard). This study included patients that were referred for dysphagia investigations and therefore some patients were found to not have dysphagia. This allowed the investigators to report the sensitivity, specificity, PPV, NPV, and validity of dysphagia detection using FEES as compared to VFSS. The authors examined these parameters in patients challenged with semi-solid and liquid boluses of 5 mL.[9]

Summary of Critical Appraisal

A tabulated summary of the critical appraisal of the included studies is presented in Appendix 3.

Systematic Review

The SR by Giraldo-Cadavid et al. was well-conducted overall, with few limitations. Limitations of the study included a lack of safety information in addition to a lack of information on the included study characteristics and the patients included in those studies. The SR had important strengths including a defined research objective, well-reported methods for the systematic literature search screened in duplicate, data-extraction methodology, and a tabulated critical appraisal of included studies using the QUADAS tool (2003). A meta-analysis was conducted with appropriate statistical methodology; data were pooled appropriately based upon examination of statistical heterogeneity, and a sensitivity analysis was conducted. An examination of publication bias was conducted which did identify a non-statistically significant tendency towards fewer publications with low diagnostic yield regarding residue outcomes but no statistically significant asymmetry with regards to comparisons of VFSS and FEES. Instead of using one method as the gold standard the authors described a method using a composite reference standard of both FEES and VFSS and the authors reported that there might be a superior solution to directly compare these diagnostic tests. Inexplicably, the authors stated that FEES was more sensitive than VFSS with regards to the dysphagia parameters of penetration and aspiration, despite the *P* values being greater than 0.05 for these comparisons. The authors of the SR also reported no potential for conflicts of interest.

Prospective Comparative Studies

Both prospective comparative studies were well-conducted with important methodological strengths such as a prospectively enrolled consecutive sample of patients with no inappropriate patient exclusions was used, all patients received both tests of interest, blinded investigators had defined roles, all patients were included in the final analysis, both tests were conducted in a short timeframe, and both reported well-defined quantitative outcomes.[5,9] Neither study provided any information on the safety of FEES or VFSS and neither provided a statistical power calculation to justify the sample size.[5,9]

While Scharitzer et al. included information on the training level of investigators that conducted the diagnostic tests, there was no discussion of the potential for bias in the method in which the diagnostic tests were conducted simultaneously. The methodology of Scharitzer et al. was also limited in the critical appraisal criteria as the authors did not use a single reference test (gold standard), and the goal of the study was not to determine which was a superior diagnostic test. Importantly this study did not compare the accuracy of FEES and VFSS to diagnose dysphagia, only their comparative accuracy in diagnosing parameters of dysphagia.

Fattori et al. included patients with suspected dysphagia and was therefore able to compare the accuracy of FEES and VFSS with regard to dysphagia diagnosis. However this study

did not report the training level or experience of the investigators conducting the diagnostic interventions making the applicability to specific healthcare settings less clear.[9]

Summary of Findings

Diagnostic Accuracy of FEES versus VFSS

The SR included extracted data from six diagnostic comparative studies which were limited by small sample sizes (≤ 50 patients each), however the risk of bias in the primary studies was determined to be low based on QUADAS criteria.[4] For the evaluation of diagnostic sensitivity, the results from four studies were pooled since two studies were determined to introduce significant statistical heterogeneity, while for the evaluation of diagnostic specificity the results of all six studies were pooled. The combined results found no statistically significant differences between FEES and VFSS in diagnostic specificity of the dysphagia parameters of aspiration, penetration, residue, or premature spillage. In terms of diagnostic sensitivity FEES was statistically superior for the dysphagia parameter of residue, but there were no significant differences between FEES and VFSS for the other tested parameters (i.e., sensitivity for aspiration, penetration, and premature spillage, in addition to specificity for aspiration, penetration, residue, and premature spillage). The authors concluded that FEES had a slight advantage with regard to detection of aspiration, penetration, and residues, however these differences were not statistically significant.[4]

The interrater agreement for VFSS was excellent for PA-score, residues, and substantial for the location of swallow trigger, while the interrater agreement for FEES was excellent for the assessment of PA-score, location of swallow trigger, and substantial for residues as reported by Scharitzer et al.[5] The authors found statistically significant differences between VFSS and FEES with regard to PA-score, even when scores were grouped, depending on the raters (P = 0.045) and consistencies (P = 0.027) of contrast material. Specifically, statistically significant higher PA-scores were observed by VFSS examination than by FEES. The importance to diagnostic accuracy of this difference is unclear without a reference standard or patient-related outcomes. The authors concluded that VFSS and FEES should not be considered as interchangeable procedures.[5]

The study by Fattori et al. revealed that FEES performed with both liquid and semi-solid consistencies had good sensitivity (> 80%) and overall validity (≥ 80%) when compared to VFSS as the gold standard. The authors concluded that VFSS and FEES are both capable of detecting oro-pharyngeal dysphagia regardless of etiology. However the authors also concluded that VFSS exhibited superior sensitivity for premature spillage and aspiration, while the advantage of direct visualization of residues by FEES was an advantage over VFSS for simple quantification of even negligible volumes.[9]

Detailed accuracy outcomes are tabulated in Appendix 4.

Cost-Effectiveness of FEES vs VFSS

No relevant evidence for the cost-effectiveness of FEES as compared to VFSS was identified; therefore, no summary can be provided.

Limitations

A lack of a suitable reference standard for the comparison between FEES and VFSS was pointed out by two included studies. None of the identified studies reported any patient-related outcomes or safety data resulting in the significance of diagnostic accuracy

comparisons between VFSS and FEES unclear. Additionally, no cost-effectiveness evidence comparing VFSS with FEES was identified.

Conclusions and Implications for Decision or Policy Making

One SR, that included six comparative studies, and two prospective comparative studies were identified and included in this report. While all included studies were relatively small (< 100 patients) and had the potential for type II error (i.e., making an incorrect conclusion of "no difference"), the data suggested that both VFSS and FEES were suitable for dysphagia diagnosis.[4,5,9] When examining different dysphagia pathologies the SR identified a statistically significant difference where FEES demonstrated greater diagnostic sensitivity for the evaluation of residue. The SR did not identify any differences in diagnostic specificity for residue or the dysphagia parameters of aspiration, penetration, and premature spillage.[4] One prospective comparative study found that FEES had good overall sensitivity and validity (>80%) but lower specificity (66.7%) for the identification of dysphagia when VFSS was used as the reference standard.[9] The other identified prospective comparative study found statistically significant higher PA-scores using VFSS as compared to FEES. This study did not use a reference standard and therefore the significance to diagnostic accuracy is unclear.[5] The choice of reference standard used and comparisons made in the included studies was not correlated with any patient-related outcomes making the clinical significance of these few observed differences between FEES and VFSS unclear.[5] Despite the subjective interpretation and visual judgement of swallowing sequences required for both diagnostic methodologies, overall good to excellent interrater agreement for within VFSS and FEES was found, however the observed difference in PA-scores using VFSS depended on the rater.[5]

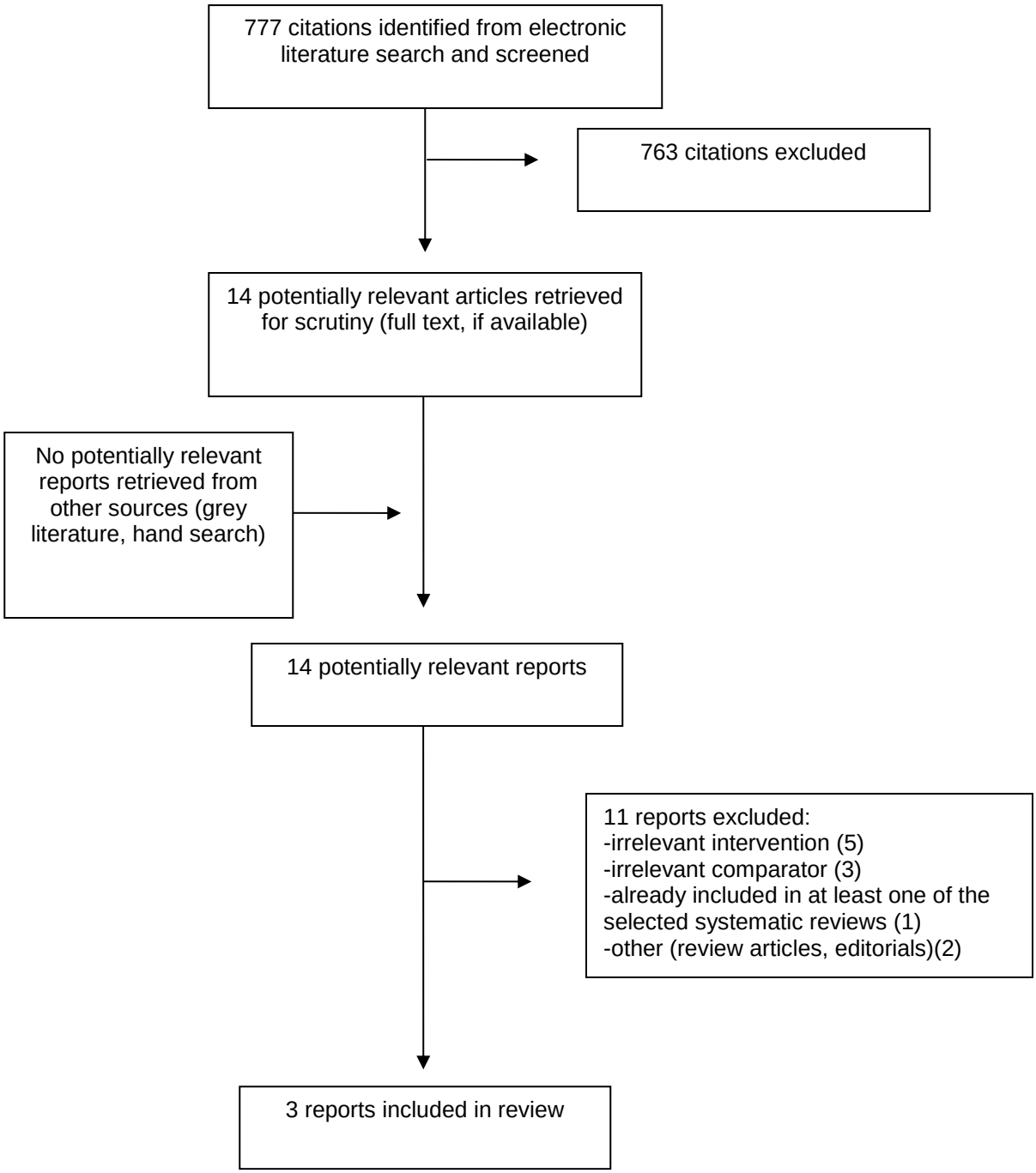
As suggested by Giraldo-Cadavid et al., large prospective studies that implement a solution for a satisfactory reference standard for comparison would be useful for comparing diagnostic accuracy of VFSS and FEES.[4] Given the lack of evidence for significant differences in diagnostic accuracy, additional research or other factors including safety and cost-effectiveness could be considered in decisions regarding VFSS or FEES implementation.

No cost-effectiveness evidence was identified that compared VFSS and FEES. One of the prospective comparative studies indicated that FEES was inexpensive, however the relative costs and outcomes of FEES and VFSS were not reported.[9]

References

1. Tomita S, Oeda T, Umemura A, et al. Video-fluoroscopic swallowing study scale for predicting aspiration pneumonia in Parkinson's disease. *PLoS One.* 2018;13(6):e0197608.
2. Skoretz SA, Yau TM, Granton JT, Martino R. The feasibility of assessing swallowing physiology following prolonged intubation after cardiovascular surgery. *Pilot Feasibility Stud.* 2017;3:62.
3. Langmore SE. History of fiberoptic endoscopic evaluation of swallowing for evaluation and management of pharyngeal dysphagia: changes over the years. *Dysphagia.* 2017;32(1):27-38.
4. Giraldo-Cadavid LF, Pantoja JA, Forero YJ, Gutierrez HM, Bastidas AR. Aspiration in the fiberoptic endoscopic evaluation of swallowing associated with an increased risk of mortality in a cohort of patients suspected of oropharyngeal dysphagia. *Dysphagia.* 2019 Jul. doi: 10.1007/s00455-019-10036-7. [Epub ahead of print].
5. Scharitzer M, Roesner I, Pokieser P, Weber M, Denk-Linnert DM. Simultaneous Radiological and Fiberendoscopic Evaluation of Swallowing ("SIRFES") in patients after surgery of oropharyngeal/laryngeal cancer and postoperative dysphagia. *Dysphagia.* 2019 Dec;34(6):852-861.
6. Adachi K, Umezaki T, Kikuchi Y. Videoendoscopy worsens swallowing function: a videofluoroscopic study. A randomized controlled trial. *Eur Arch Otorhinolaryngol.* 2017;274(10):3729-3734.
7. Whiting PF, Rutjes AW, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med.* 2011;155(8):529-536.
8. Shea BJ, Reeves BC, Wells G, et al. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. *BMJ (Clinical research ed).* 2017;358:j4008.
9. Fattori B, Giusti P, Mancini V, et al. Comparison between videofluoroscopy, fiberoptic endoscopy and scintigraphy for diagnosis of oro-pharyngeal dysphagia. *Acta Otorhinolaryngol Ital.* 2016;36(5):395-402.
10. Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *J Clin Epidemiol.* 2009;62(10):e1-e34.

Appendix 1: Selection of Included Studies



Appendix 2: Characteristics of Included Publications

Table 2: Characteristics of Included Systematic Review

First Author, Publication Year, Country	Study Designs and Numbers of Primary Studies Included	Population Characteristics	Intervention and Comparator(s)	Clinical Outcomes
Giraldo-Cadavid, 2016, Columbia[4]	6 comparative studies	Adults with suspected functional oropharyngeal dysphagia (n = 198), excluding mechanical dysphagia	FEES vs. VFSS	● VFSS and FEES compared to each other with one used as reference standard ● Premature spillage ● Penetration ● Aspiration ● Pharyngeal residue

FEES = flexible endoscopic evaluation of swallowing; VFSS = videofluoroscopic swallowing study.

Table 3: Characteristics of Included Primary Clinical Studies

First Author, Publication Year, Country	Study Design	Population Characteristics	Intervention and Comparator(s)	Clinical Outcomes
Scharitzer, 2019, Austria[5]	Prospective comparative study in consecutive patients evaluated by both methods simultaneously	Adults following operation for pharyngeal or laryngeal carcinoma, either ongoing or completed radiotherapy with postoperative dysphagia with high risk for aspiration. Must be in stable clinical condition (n = 29) 82.8% male Age range (years): 48-90 <u>Site of disease</u> Nasopharynx 1/29 Oropharynx 20/29 Larynx 8/29	Compare FEES and VFSS in the same patients using a predefined exam protocol using different volumes (3 mL, 5 mL, 10 mL, 20 mL) and different viscosities (nectar, pudding, liquid)	● Interrater agreement ● PA-scores for different bolus volumes and bolus consistencies - increasing PA-score indicates increasing presence/severity, based on anatomical makers for depth of penetration and aspiration using integers from 1 to 8.
Fattori, 2016, Italy[9]	Prospective comparative study in consecutive patients evaluated by three methods on the same day	Referred for dysphagia studies (n = 60) 63.3% male Age (mean; years): 66.66 ± 16.5 SD <u>Etiology</u> Neurological 34/60	Comparing VFSS, FEES, and OPES for detection of oropharyngeal dysphagia using semisolid and liquid boluses of 5 mL. VFSS was used as reference standard	● Sensitivity ● Specificity ● PPV ● NPV ● Validity For: ● Premature spillage ● Hypoharyngeal residue ● Aspiration

First Author, Publication Year, Country	Study Design	Population Characteristics	Intervention and Comparator(s)	Clinical Outcomes
		Post-surgical head & neck cancer 15/60 Gastroenterological with reflux 7/60 Pneumological with bronchial-pulmonary disease 4/60 Mean years since onset 1.5 ± 1.2 SD		

FEES = flexible endoscopic evaluation of swallowing; OPES = oro-pharyngo-oesophageal scintigraphy; NPV = negative predictive value; PA-score = penetration aspiration score; PPV = positive predictive value; SD = standard deviation; VFSS = videofluoroscopic swallowing study.

Appendix 3: Critical Appraisal of Included Publications

Table 4: Strengths and Limitations of Clinical Studies using the AMSTAR-2 tool[8]

Strengths	Limitations
Systematic Review	
Giraldo-Cadavid et al., 2016[4]	
● Defined research objective ● Literature search selection/inclusion/exclusion methodology clear ● Literature screened in duplicate ● Critical appraisal of included studies tabulated ● Statement of no conflict of interest ● Data extraction methodology described ● Statistical heterogeneity examined ● Sensitivity analysis conducted ● Statistical methodology outlined and appropriate ● Publication bias assessment	● Limited information on included study characteristics ● Limited information on patient characteristics within included studies ● No safety information reported

Table 5: Strengths and Limitations of Clinical Studies using QUADAS-2[7]

Strengths	Limitations
Prospective Comparative Studies	
Scharitzer, 2019, Austria[5]	
Risk of Bias ● Consecutive sample of patients enrolled ● Case-control design avoided ● No inappropriate patient exclusion ● All patients received same tests ● Blinded investigators conducted tests ● Role of investigators clear ● All patients included in analysis ● No loss to follow-up ● Patient flow clear ● Outcomes well defined ● Statement of no COI ● Discussion of study limitations **Applicability** ● Training level of diagnostic investigators reported	**Risk of Bias** ● No discussion on potential limitations of doing both diagnostic interventions simultaneously ● Goal was not to test superiority - no gold standard ● No statistical power calculation **Applicability** ● No safety outcomes ● All dysphagia patients - no negative controls
Fattori, 2016, Italy[9]	
Risk of Bias ● Consecutive sample of patients enrolled ● Case-control design avoided ● No inappropriate patient exclusion ● All patients received same reference standard ● Blinded investigators conducted tests ● Role of investigators clear ● All patients included in analysis ● No loss to follow-up	**Risk of Bias** ● No COI statement ● No statistical power calculation **Applicability** ● No safety outcomes ● Training level of diagnostic investigators not reported

Strengths	Limitations
● Patient flow clear ● Outcomes well defined **Applicability** ● Broad patient inclusion	

COI = conflict of interest.

Appendix 4: Main Study Findings and Authors' Conclusions

Table 6: Summary of Findings of Included Primary Clinical Studies

<table>
<tr><th>Main Study Findings</th><th>Authors' Conclusion</th></tr>
<tr><td colspan="2">Systematic Review</td></tr>
<tr><td colspan="2">Giraldo-Cadavid et al., 2016[4]</td></tr>
<tr><td><u>Pooled Accuracy Measures against a composite reference standard:</u>
<u>Sensitivities (95% CI)</u> | <u>FEES</u> | <u>VFSS</u>
Aspiration (P = 0.055) | 0.86 (0.77 to 0.93) | 0.77 (0.66 to 0.85)
Penetration (P = 0.057) | 0.96 (0.89 to 0.99) | 0.89 (0.81 to 0.95)
Residue (P = 0.015) | 0.93 (0.84 to 0.98) | 0.80 (0.68 to 0.89)
Premature spillage (P = 0.267) | 0.68 (0.50 to 0.82) | 0.77 (0.68 to 0.89)

<u>Specificities (95% CI)</u>
Aspiration (P = 1.00) | 0.98 (0.96 to 0.99) | 0.98 (0.96 to 0.99)
Penetration (P = 1.00) | 0.98 (0.94 to 0.99) | 0.98 (0.94 to 0.99)
Residue (P = 1.00) | 0.97 (0.91 to 0.99) | 0.97 (0.91 to 0.99)
Premature spillage (P = 1.00) | 0.93 (0.68 to 0.99) | 0.93 (0.68 to 0.99)</td><td>"The information available to date shows that FEES tends to be more sensitive than VFSS to penetration, aspiration, and residues, and that both tests are equally sensitive to premature spillage. A prospective study of the tests' diagnostic accuracy should be performed using a satisfactory solution for a reference standard to determine which of the tests has greater accuracy." (p2,009)</td></tr>
<tr><td colspan="2">Clinical Studies</td></tr>
<tr><td colspan="2">Scharitzer, 2019, Austria[5]</td></tr>
<tr><td><u>Interrater agreement</u>
<u>Weighted kappa (95% CI)</u> | <u>FEES</u> | <u>VFSS</u>
PA-score | 0.911 (0.864 to 0.959) | 0.979 (0.963 to 0.994)
Retentions valleculae | 0.613 (0.528 to 0.697) | 0.819 (0.748 to 0.930)
Retentions piriform sinus | 0.762 (0.686 to 0.837) | 0.857 (0.784 to 0.930)
Time of triggering | 0.828 (0.750 to 0.906) | 0.771 (0.689 to 0.853)

<u>Grouped PA-scores and number of swallows for different bolus volumes and consistencies</u>
3 mL (P = 0.003)
<u>Grouped PA-score</u> | <u>FEES</u> | <u>VFSS</u>
1 | 48 | 40
2-5 | 18 | 19
6-8 | 9 | 16

5 mL (P = 0.052)
<u>Grouped PA-score</u> | <u>FEES</u> | <u>VFSS</u>
1 | 39 | 35
2-5 | 14 | 15
6-8 | 6 | 9

10 mL (P = 0.007)
<u>Grouped PA-score</u> | <u>FEES</u> | <u>VFSS</u>
1 | 33 | 27
2-5 | 4 | 7
6-8 | 2 | 5

20 mL (P = 0.011)
<u>Grouped PA-score</u> | <u>FEES</u> | <u>VFSS</u>
1 | 27 | 20
2-5 | 2 | 8
6-8 | 0 | 1</td><td>"In conclusion, our study shows that videofluoroscopy adds significant and crucial information to the findings of the FEES for this specific patient group, namely the detection of aspiration and the quantity of pharyngeal residues. Both modalities should not be considered as interchangeable procedures, and, with regard to the relative benefits of each procedure, both provide relevant information in dysphagic patients after pharyngeal or laryngeal cancer and radiotherapy." (p8)

"Significant differences between both methods were found when assessing the penetration-aspiration scale (p = 0.001, tendency of higher scores by videofluoroscopic (median = 2.59) as opposed to fiberendoscopic (median = 2.14) and the residue severity scores in the valleculae (p = 0.029) and the sinus piriformes (p = 0.002) with larger residues scored by fiberendoscopic evaluation of swallowing." (p1)</td></tr>
</table>

Main Study Findings	Authors' Conclusion
Nectar (*P* < 0.001) Grouped PA-score / FEES / VFSS 1 / 63 / 52 2-5 / 11 / 14 6-8 / 3 / 11 **Pudding (*P* = 0.257)** Grouped PA-score / FEES / VFSS 1 / 43 / 40 2-5 / 13 / 16 6-8 / 10 / 10 **Liquid (*P* = 0.001)** Grouped PA-score / FEES / VFSS 1 / 41 / 30 2-5 / 14 / 19 6-8 / 4 / 10 **Median of median scores of both raters (*P* = 0.001)** VFSS = 2.59 FEES = 2.14 This difference was dependent on the raters (*P* = 0.016), and the consistency (*P* = 0.039)	
Fattori, 2016, Italy[9]	
Accuracy Measures of FEES using VFSS as reference: Sensitivity / Semi-solid / Liquid FEES vs. VFSS (detection) / 85.2 / 80.4 Premature spillage / 60.0 / 60.0 Hypopharyngeal residue / 75.6 / 61.4 Aspiration / 33.3 / 37.0 Specificity / Semi-solid / Liquid FEES vs. VFSS (detection) / 66.7 / 77.8 Premature spillage / 84.4 / 86.7 Hypopharyngeal residue / 73.3 / 75.0 Aspiration / 87.9 / 87.9 PPV / Semi-solid / Liquid FEES vs. VFSS (detection) / 95.8 / 95.3 Premature spillage / 56.3 / 60.0 Hypopharyngeal residue / 89.5 / 87.1 Aspiration / 69.2 / 71.4 NPV / Semi-solid / Liquid FEES vs. VFSS (detection) / 33.3 / 41.2 Premature spillage / 86.4 / 86.7 Hypopharyngeal residue / 50.0 / 41.4 Aspiration / 61.7 / 63.0 Validity / Semi-solid / Liquid FEES vs. VFSS (detection) / 83.3 / 80.0	"Our study leads us to conclude that the VFS, FEES and OPES tests are all capable of detecting oro-pharyngeal dysphagia, whichever disorder is at the basis of it. Nevertheless, VFS must still be considered by speech-language pathologists as the gold standard since it supplies values that are more reliable than those obtained with the other two tests, at least as far as the swallowing parameters we took into account are concerned." (*p400*) "In addition, as reported in the literature, FEES has a great advantage over VFS in that it uses real food during the test and allows a better view of the larynx movement Therefore, on the grounds of these considerations and our results, we maintain that FEES should always be considered as a valid test for studying swallowing, particularly since it is able to replace the VFSS for investigating oropharyngeal dysphagia, and that it should be performed first of all when it is not possible to use VFS." (*p401*) Thus, all three of these tests, FEES, VFS and OPES, are capable of supplying an

Main Study Findings			Authors' Conclusion
Premature spillage	78.3	80.0	accurate diagnosis of oro-pharyngeal dysphagia." (*p400*)
Hypopharyngeal residue	75.0	65.0	
Aspiration	63.3	65.0	

CI = confidence interval; FEES = flexible endoscopic evaluation of swallowing; OPES = oro-pharyngo-oesophageal scintigraphy; NPV = negative predictive value; PA-score = penetration aspiration score; PPV = positive predictive value; VFSS/VFS = videofluoroscopic swallowing study.