Cochrane Library

Cochrane Database of Systematic Reviews

Screening for aspiration risk associated with dysphagia in acute stroke (Review)

Boaden E, Burnell J, Hives L, Dey P, Clegg A, Lyons MW, Lightbody CE, Hurley MA, Roddam H, McInnes E, Alexandrov A, Watkins CL

Boaden E, Burnell J, Hives L, Dey P, Clegg A, Lyons MW, Lightbody CE, Hurley MA, Roddam H, McInnes E, Alexandrov A, Watkins CL.
Screening for aspiration risk associated with dysphagia in acute stroke.
Cochrane Database of Systematic Reviews 2021, Issue 10. Art. No.: CD012679.
DOI: 10.1002/14651858.CD012679.pub2.

www.cochranelibrary.com

WILEY

TABLE OF CONTENTS

HEADER 1
ABSTRACT 1
PLAIN LANGUAGE SUMMARY 2
SUMMARY OF FINDINGS 4
BACKGROUND 8
OBJECTIVES 9
METHODS 9
Figure 1. 11
RESULTS 13
Figure 2. 15
Figure 3. 17
Figure 4. 18
Figure 5. 19
Figure 6. 20
Figure 7. 21
Figure 8. 22
Figure 9. 23
Figure 10. 24
DISCUSSION 24
AUTHORS' CONCLUSIONS 27
ACKNOWLEDGEMENTS 27
REFERENCES 28
CHARACTERISTICS OF STUDIES 35
DATA 78
Test 1. Registered Dietitian (RD) Dysphagia Screening tool - Huhmann (2004) 81
Test 2. Bedside aspiration - Combined WST & Oxygen Saturation - Lim (2001) 81
Test 3. Gugging Swallowing Screen (GUSS) - Group2 - Trappl (2007) 81
Test 4. Toronto Bedside Swallowing Screening Test (TOR-BSST) - Martino (2009) 81
Test 5. Standardized Swallowing Assessment tool (SSA) - Test2 - Perry (2001) Test2 81
Test 6. Nursing Bedside Dysphagia Screen (NBDS) - Campbell (2016) 81
Test 7. Emergency Department (ED) dysphagia screen - Turner-Lawrence (2009) 82
Test 8. Acute Stroke Dysphagia Screening (ASDS) - Aspiration - Edmiaston (2010) 82
Test 9. Barnes-Jewish Hospital-Stroke Dysphagia Screen (BJH-SDS) Aspiration - Edmiaston (2014) 82
Test 10. Edith-Huhn-Matesic Bedside Aspiration Screen (EHMBAS) plus water swallow test - Huhn-Matesic (2015) 82
Test 11. Standardized Swallowing Assessment tool (SSA) - Test1 - Perry (2001) 82
Test 12. Barnes-Jewish Hospital-Stroke Dysphagia Screen (BJH-SDS) Dysphagia - Edmiaston (2014) 82
Test 13. Modified MASA (MMASA) Neurologist 1 - Antonios (2010) 83
Test 14. Rapid Aspiration Screening for Suspected Stroke (RAS3) - Daniels (2016) 83
Test 15. Clinical examination - Daniels (1997) 83
Test 16. Acute Stroke Dysphagia Screening (ASDS) Dysphagia - Edmiaston (2010) 83
Test 17. Nurse Dysphagia Screen - Cummings (2015) 83
Test 18. Modified MASA (MMASA) Neurologist 2 - Antonios (2010) 83
Test 19. Oxygen saturation ≥ 2% - Test2 for Aspiration - Smith (2000) 84
Test 20. Bedside swallow test - WST only - Lim (2001) 84
Test 21. Chinese version of the modified SSA original 8 items - Jiang (2019) 84
Test 22. Chinese version of the modified SSA reduced 6 items - Jiang (2019) 84
Test 23. Stroke Severity using National Institutes of Health Stroke Scale (NIHSS) - Bravata (2009) 84
Test 24. Nursing Bedside Swallowing Screen (NBSS) - Ellis (2013) 84
Test 25. 2-step Swallowing Provocation Test (SPT) - step 1 - 0.4 mL - Warneke (2008) 85
Test 26. Rapid Aspiration Screening for Suspected Stroke (RAS3) - WST only - Daniels (2016) 85
Test 27. Gag function - Test3 - Perry (2001) 85

Test 28. DePaul Hospital Swallow Screener (DHSS) for Aspiration Risk - Behera (2018) 85
Test 29. DePaul Hospital Swallow Screener (DHSS) for Dysphagia - Behera (2018) 85
Test 30. Gag function - Test4 - Perry (2001) 85
Test 31. 2-step swallowing provocation test (SPT) step 2 - 2.0 mL - Warneke (2008) 86
Test 32. Johns Hopkins Hospital Brain Rescue Unit Modified 3 oz Swallow Screen - Mulheren (2019) 86
Test 33. Nursing Screening Tool - Bravata (2009) 86
Test 34. Barnes-Jewish Hospital Stroke Dysphagia Screen (BJH-SDS) – Turkish version (T-BJH) - Eren (2019) 86
Test 35. Clinical Predicative Scale of Aspiration (CPSA) - Zhou (2011) 86
Test 36. TOR-BSST water swallow item - Martino (2014) 86
Test 37. Clinical swallowing tests - 6 oromotor examinations - Nishiwaki (2005) 87
Test 38. Index taxonomy - water only 87
Test 39. Index taxonomy - water plus other consistencies 87
Test 40. Index taxonomy - other 87
Test 41. Outcome - aspiration 88
Test 42. Outcome - dysphagia 88
Test 43. Reference test - Expert Assessment and MASA 89
Test 44. Reference test - FEES 89
Test 45. Reference test - VF 89
Test 46. HCP - nurse 90
Test 47. HCP - other 90
ADDITIONAL TABLES 90
APPENDICES 93
WHAT'S NEW 101
HISTORY 101
CONTRIBUTIONS OF AUTHORS 102
DECLARATIONS OF INTEREST 102
SOURCES OF SUPPORT 102
DIFFERENCES BETWEEN PROTOCOL AND REVIEW 102

[Diagnostic Test Accuracy Review]

Screening for aspiration risk associated with dysphagia in acute stroke

Elizabeth Boaden[1], Jane Burnell[2], Lucy Hives[2], Paola Dey[3], Andrew Clegg[2], Mary W Lyons[4], C Elizabeth Lightbody[5], Margaret A Hurley[2], Hazel Roddam[6], Elizabeth McInnes[7], Anne Alexandrov[8], Caroline L Watkins[5]

[1]Faculty of Health and Care , University of Central Lancashire , Preston, UK. [2]Faculty of Health and Care, University of Central Lancashire , Preston, UK. [3]Faculty of Health and Social Care, Edge Hill University, Ormskirk, UK. [4]Liverpool School of Tropical Medicine, Liverpool, UK. [5]Faculty of Health and Care, University of Central Lancashire, Preston, UK. [6]Faculty of Allied Health and Well-being, University of Central Lancashire, Preston, UK. [7]Nursing Research Institute, Australian Catholic University, Sydney, Australia. [8]University of Tennessee Health Science Center (UTHSC), Memphis, Tennessee, USA

Contact address: Elizabeth Boaden, eboaden1@uclan.ac.uk.

Editorial group: Cochrane Stroke Group.
Publication status and date: New, published in Issue 10, 2021.

Citation: Boaden E, Burnell J, Hives L, Dey P, Clegg A, Lyons MW, Lightbody CE, Hurley MA, Roddam H, McInnes E, Alexandrov A, Watkins CL. Screening for aspiration risk associated with dysphagia in acute stroke. *Cochrane Database of Systematic Reviews* 2021, Issue 10. Art. No.: CD012679. DOI: 10.1002/14651858.CD012679.pub2.

ABSTRACT

Background

Stroke can affect people's ability to swallow, resulting in passage of some food and drink into the airway. This can cause choking, chest infection, malnutrition and dehydration, reduced rehabilitation, increased risk of anxiety and depression, longer hospital stay, increased likelihood of discharge to a care home, and increased risk of death. Early identification and management of disordered swallowing reduces risk of these difficulties.

Objectives

Primary objective

• To determine the diagnostic accuracy and the sensitivity and specificity of bedside screening tests for detecting risk of aspiration associated with dysphagia in people with acute stroke

Secondary objectives

• To assess the influence of the following sources of heterogeneity on the diagnostic accuracy of bedside screening tools for dysphagia

- Patient demographics (e.g. age, gender)

- Time post stroke that the study was conducted (from admission to 48 hours) to ensure only hyperacute and acute stroke swallow screening tools are identified

- Definition of dysphagia used by the study

- Level of training of nursing staff (both grade and training in the screening tool)

- Low-quality studies identified from the methodological quality checklist

- Type and threshold of index test

- Type of reference test

Search methods

In June 2017 and December 2019, we searched CENTRAL, MEDLINE, Embase, CINAHL, and the Health Technology Assessment (HTA) database via the Centre for Reviews and Dissemination; the reference lists of included studies; and grey literature sources. We contacted experts in the field to identify any ongoing studies and those potentially missed by the search strategy.

Selection criteria

We included studies that were single-gate or two-gate studies comparing a bedside screening tool administered by nurses or other healthcare professionals (HCPs) with expert or instrumental assessment for detection of aspiration associated with dysphagia in adults with acute stroke admitted to hospital.

Data collection and analysis

Two review authors independently screened each study using the eligibility criteria and then extracted data, including the sensitivity and specificity of each index test against the reference test. A third review author was available at each stage to settle disagreements. The methodological quality of each study was assessed using the Quality Assessment of Studies of Diagnostic Accuracy (QUADAS-2) tool. We identified insufficient studies for each index test, so we performed no meta-analysis. Diagnostic accuracy data were presented as sensitivities and specificities for the index tests.

Main results

Overall, we included 25 studies in the review, four of which we included as narratives (with no accuracy statistics reported). The included studies involved 3953 participants and 37 screening tests. Of these, 24 screening tests used water only, six used water and other consistencies, and seven used other methods. For index tests using water only, sensitivity and specificity ranged from 46% to 100% and from 43% to 100%, respectively; for those using water and other consistencies, sensitivity and specificity ranged from 75% to 100% and from 69% to 90%, respectively; and for those using other methods, sensitivity and specificity ranged from 29% to 100% and from 39% to 86%, respectively. Twenty screening tests used expert assessment or the Mann Assessment of Swallowing Ability (MASA) as the reference, six used fibreoptic endoscopic evaluation of swallowing (FEES), and 11 used videofluoroscopy (VF). Fifteen screening tools had an outcome of aspiration risk, 20 screening tools had an outcome of dysphagia, and two narrative papers did not report the outcome. Twenty-one screening tests were carried out by nurses, and 16 were carried out by other HCPs (not including speech and language therapists (SLTs)).

We assessed a total of six studies as low risk across all four QUADAS-2 risk of bias domains, and we rated 15 studies as low concern across all three applicability domains.

No single study demonstrated 100% sensitivity and specificity with low risk of bias for all domains. The best performing combined water swallow and instrumental tool was the Bedside Aspiration test (n = 50), the best performing water plus other consistencies tool was the Gugging Swallowing Screen (GUSS; n = 30), and the best water only swallow screening tool was the Toronto Bedside Swallowing Screening Test (TOR-BSST; n = 24). All tools demonstrated combined highest sensitivity and specificity and low risk of bias for all domains. However, clinicians should be cautious in their interpretation of these findings, as these tests are based on single studies with small sample sizes, which limits the estimates of reliability of screening tests.

Authors' conclusions

We were unable to identify a single swallow screening tool with high and precisely estimated sensitivity and specificity based on at least one trial with low risk of bias. However, we were able to offer recommendations for further high-quality studies that are needed to improve the accuracy and clinical utility of bedside screening tools.

PLAIN LANGUAGE SUMMARY

Screening for aspiration risk associated with dysphagia in acute stroke

Question

How accurate are swallow screening tools for detecting when food and drink enter the airway in people with acute stroke?

Background

Stroke often affects a person's ability to swallow, allowing food and drink into the airway. This can cause choking, chest infection, malnutrition, dehydration, and reduced rehabilitation, with increased risk of anxiety, depression, discharge to a care home, and death. Early identification and management of disordered swallowing through the most accurate testing reduces these risks. If the test fails to identify swallowing difficulties, the person will continue oral intake and may experience the difficulties identified above. If the test incorrectly identifies swallowing difficulties, the person may not be given anything to eat or drink, significantly impacting quality of life, until a more detailed assessment is undertaken (usually the next day).

Study characteristics

We identified 25 studies that used a total of 37 tools. Seven tools did not use water or other consistencies, 24 used only water, and six considered water and other consistencies.

Key results

We were unable to identify a tool that could accurately identify everyone with food and drink entering their airway, as well as detect all those who definitely did not. Many studies involved different healthcare professionals, food and fluid testing consistencies, and time between stroke onset and the screening test, so it is unclear which tool is best. We were unable to directly compare the different tools because most studies used different methods.

We were able to identify the tools most able to detect people with and without risk of swallowing difficulties from studies with good quality evidence. The best combined water swallow and instrumental test was the Bedside Aspiration test, the best water plus other consistencies tool was the Gugging Swallowing Screen, and the best water only tool was the Toronto Bedside Swallowing Screening Test. However, clinicians should be cautious in their interpretation of these findings, as these tests are based on single studies with small sample sizes.

Quality of the evidence

Most included studies were poorly conducted or were unclear in reporting what they did (i.e. unclear or high risk of bias).

Conclusion

We were unable to identify a single tool with combined high levels of accuracy and good quality evidence. However, we are able to offer recommendations for further high-quality studies that are needed to improve the accuracy and clinical utility of swallow screening tools.

SUMMARY OF FINDINGS

Summary of findings 1. Review criteria and findings

Review question	What is the diagnostic accuracy of screening tools for aspiration risk associated with dysphagia in acute stroke?
Importance	A simple and reliable screening tool for identification of aspiration associated with dysphagia would reduce diagnostic delay and the risk of developing pneumonia in acute stroke patients
Patients/population	Patients (over the age of 18) who have been admitted to an acute hospital setting, where there is a clinical diagnosis of stroke. Patients with subarachnoid haemorrhage were excluded from the study
Settings	Stroke units or hospital wards where acute stroke patients are admitted
Index tests	Bedside swallowing screening tools carried out by healthcare professionals (excluding SLTs)
Reference standards	Expert assessment, including the MASA VF FEES
Studies	25 studies were included, 4 of which are presented as narratives (with no accuracy statistics reported). The included studies reported 37 screening tests
Risk of bias	Overall judgement: poor quality for the majority of studies – only 6 studies were at low risk across all 4 risk of bias domains, and 2 studies were at low risk of bias for 3 domains
	Patient selection bias: high risk = 12 studies; unclear risk = 11 studies; low risk = 2 studies
	Index test interpretation bias: high risk = 14 studies; unclear risk = 10 studies; low risk = 1 study
	Reference standard interpretation bias: high risk = 16 studies; unclear risk = 8 studies; low risk = 1 study
	Flow and timing selection bias: high risk = 13 studies; unclear risk = 6 studies; low risk = 6 studies
Applicability concern	Concerns regarding patient selection: high risk = 16 studies; unclear risk = 4 studies; low risk = 5 studies
	Concerns regarding index test: high risk = 20 studies; unclear risk = 4 studies; low risk = 1 study
	Concerns regarding reference standard: high risk = 22 studies; unclear risk = 3 studies; low risk = 0 studies
Overall findings	The best performing swallow screening tools were the Combined WST and Oxygen Saturation Test (Lim 2001a), GUSS (Trapl 2007b), and TOR-BSST (Martino 2009). The best water only swallow screening test was the TOR-BSST (Martino 2009). All demonstrate a combined highest sensitivity and specificity and low risk of bias across all domains. These tools may be considered useful in clinical practice
	The water plus various consistencies (n = 6) performed better than the water-only tools in terms of sensitivity and specificity
	Only a few studies (e.g. GUSS) (Trapl 2007b) gave direction on what food and drink consistencies should be given to an individual following the screen

	The quality of evidence varied. Studies often failed to distinguish between dysphagia and aspiration as the primary outcome. Some did not identify the time interval between stroke onset or admission to hospital and the screen, or the time interval between index test and reference test. Training required by different healthcare professionals to implement the screening tool was not reported well. This has implications for future research

FEES: fibreoptic endoscopic evaluation of swallowing.
MASA: Mann Assessment of Swallowing Ability.
SLT: speech and language therapist.
VF: videofluoroscopy.
WST: water-swallowing test.

Summary of findings 2. Screening tests

Screening test	N participants (studies)	Reference standard	Diagnostic estimates (95% CI)	Implications
Water-only tests	2914 (13)			
Toronto Bedside Swallowing Screening Test (TOR-BSST): Martino 2009 Study 1	24 (1)	VFSS	Sens = 1.00 (0.75 to 1.00) spec = 0.64 (0.31 to 0.89)	Insufficient evidence to allow firm conclusions
Acute Stroke Dysphagia Screening (ASDS) Aspiration: Edmiaston 2010	300 (1)	Expert assessment and MASA	Sens = 0.95 (0.87 to 0.99) spec = 0.69 (0.62 to 0.75)	Large study with unclear risk of bias and low applicability concerns
Barnes-Jewish Hospital-Stroke Dysphagia Screen (BJH-SDS) Aspiration: Edmiaston 2014	223 (1)	VFSS	Sens = 0.95 (0.86 to 0.99) spec = 0.50 (0.42 to 0.58)	Large study with unclear/low risk of bias and low applicability concerns
Edith-Huhn-Matesic Bedside Aspiration Screen (EHMBAS) followed by simple water swallow test: Huhn-Matesic 2015	52 (1)	Expert assessment and MASA	Sens = 0.94 (0.71 to 1.00) spec = 0.77 (0.60 to 0.90)	Insufficient evidence to allow firm conclusions
Barnes-Jewish Hospital-Stroke Dysphagia Screen (BJH-SDS) Dysphagia: Edmiaston 2014	225 (1)	VFSS	Sens = 0.94 (0.88 to 0.98) spec = 0.66 (0.57 to 0.75)	Large study with low risk of bias and low applicability concerns.
Rapid Aspiration Screening for Suspected Stroke (RAS3): Daniels 2016	250 (1)	VFSS	Sens = 0.93 (0.77 to 0.99) spec = 0.43 (0.36 to 0.50)	Large study with low risk of bias and low applicability concerns.
Clinical examination: Daniels 1997	59 (1)	VFSS	Sens = 0.92 (0.75 to 0.99) spec = 0.64 (0.45 to 0.80)	Insufficient evidence to allow firm conclusions
Acute Stroke Dysphagia Screening (ASDS) Dysphagia: Edmiaston 2010	300 (1)	Expert assessment and MASA	Sens = 0.91 (0.83 to 0.96) spec = 0.75 (0.68 to 0.80)	Large study with unclear risk of bias and low applicability concerns
Nurse Dysphagia Screen: Cummings 2015	49 (1)	Expert assessment and MASA	Sens = 0.89 (0.65 to 0.99) spec = 0.90 (0.74 to 0.98)	Insufficient evidence to allow firm conclusions
Bedside swallow test - WST only: Lim 2001	50 (1)	FEES	Sens = 0.85 (0.65 to 0.96) spec = 0.75 (0.53 to 0.90)	Insufficient evidence to allow firm conclusions

Chinese version of the modified SSA – original 8 items: Jiang 2019	127 (1)	Expert assessment and MASA	Sens = 0.83 (0.71 to 0.91) spec = 0.59 (0.46 to 0.71)	Insufficient evidence to allow firm conclusions
Chinese version of the modified SSA – reduced 6 items: Jiang 2019	127 (1)	Expert assessment and MASA	Sens = 0.81 (0.69 to 0.90) spec = 0.64 (0.51 to 0.76)	Insufficient evidence to allow firm conclusions
2-step swallowing provocation test (SPT) - step 1 - 0.4 mL: Warnecke 2008	100 (1)	FEES	Sens = 0.74 (0.63 to 0.83) spec = 1.00 (0.82 to 1.00)	Insufficient evidence to allow firm conclusions
Rapid Aspiration Screening for Suspected Stroke (RAS3) - WST only: Daniels 2016	250 (1)	VFSS	Sens = 0.72 (0.53 to 0.87) spec = 0.60 (0.53 to 0.66)	Large study with low risk of bias and low applicability concerns.
Gag function - Test3: Perry 2001 Study 1	22 (1)	Expert assessment and MASA	Sens = 0.71 (0.42 to 0.92) spec = 0.63 (0.24 to 0.91)	Insufficient evidence to allow firm conclusions
DePaul Hospital Swallow Screener (DHSS) Aspiration: Behera 2018	226 (1)	Expert assessment and MASA	Sens = 0.70 (0.55 to 0.83) spec = 0.90 (0.85 to 0.94)	Large study with unclear risk of bias and unclear applicability concerns
DePaul Hospital Swallow Screener (DHSS) Dysphagia: Behera 2018	225 (1)	Expert assessment and MASA	Sens = 0.68 (0.55 to 0.80) spec = 0.93 (0.88 to 0.96)	Large study with unclear risk of bias and unclear applicability concerns
Gag function - Test4: Perry 2001 Study 1	157 (1)	Expert assessment and MASA	Sens = 0.67 (0.57 to 0.76) spec = 0.78 (0.65 to 0.88)	Insufficient evidence to allow firm conclusions
2-step swallowing provocation test (SPT) - step 2 - 2.0 mL: Warnecke 2008	100 (1)	FEES	Sens = 0.49 (0.38 to 0.61) spec = 1.00 (0.82 to 1.00)	Insufficient evidence to allow firm conclusions
Johns Hopkins Hospital Brain Rescue Unit Modified 3 oz Swallow Screen: Mulheren 2019	48 (1)	VFSS	Sens = 0.46 (0.28 to 0.66) spec = 1.00 (0.83 to 1.00)	Insufficient evidence to allow firm conclusions
Water plus consistencies tests	412 (5)			
Registered Dietitian (RD) Dysphagia Screening tool: Huhmann 2004	32 (1)	Expert assessment and MASA	Sens = 1.00 (0.69 to 1.00) spec = 0.86 (0.65 to 0.97)	Insufficient evidence to allow firm conclusions
Gugging Swallowing Screen (GUSS) - Group2: Trapl 2007	30 (1)	FEES	Sens = 1.00 (0.77 to 1.00) spec = 0.69 (0.41 to 0.89)	Insufficient evidence to allow firm conclusions
Standardized Swallowing Assessment tool (SSA) – Test2: Perry 2001 Study 1	68 (1)	Expert assessment and MASA	Sens = 0.97 (0.86 to 1.00) spec = 0.90 (0.74 to 0.98)	Insufficient evidence to allow firm conclusions
Nursing Bedside Dysphagia Screen (NBDS): Campbell 2016	75 (1)	Expert assessment and MASA	Sens = 0.97 (0.90 to 1.00) spec = 0.75 (0.35 to 0.97)	Insufficient evidence to allow firm conclusions
Standardized Swallowing Assessment tool (SSA) - Test1: Perry 2001 Study 1	161 (1)	Expert assessment and MASA	Sens = 0.94 (0.87 to 0.98) spec = 0.75 (0.63 to 0.84)	Insufficient evidence to allow firm conclusions

Nursing Bedside Swallowing Screen (NBSS): Ellis 2013	46 (1)	Expert assessment and MASA	Sens = 0.75 (0.35 to 0.97) spec = 0.89 (0.75 to 0.97)	Insufficient evidence to allow firm conclusions
Other tests	627 (5)			
Bedside aspiration - Combined WST & Oxygen Saturation: Lim 2001	50 (1)	FEES	Sens = 1.00 (0.87 to 1.00) spec = 0.71 (0.49 to 0.87)	Insufficient evidence to allow firm conclusions
Emergency Department (ED) dysphagia screen: Turner-Lawrence 2009	84 (1)	Expert assessment and MASA	Sens = 0.96 (0.86 to 0.99) spec = 0.56 (0.38 to 0.72)	Insufficient evidence to allow firm conclusions
Modified MASA (MMASA) Neurologist 1: Antonios 2010	150 (1)	Expert assessment and MASA	Sens = 0.93 (0.82 to 0.98) spec = 0.86 (0.78 to 0.93)	Insufficient evidence to allow firm conclusions
Modified MASA (MMASA) Neurologist 2: Antonios 2010	150 (1)	Expert assessment and MASA	Sens = 0.87 (0.75 to 0.95) spec = 0.84 (0.76 to 0.91)	Insufficient evidence to allow firm conclusions
Oxygen saturation ≥ 2% - Test2 for Aspiration: Smith 2000	53 (1)	VFSS	Sens = 0.87 (0.60 to 0.98) spec = 0.39 (0.24 to 0.57)	Insufficient evidence to allow firm conclusions
Stroke Severity using National Institutes of Health Stroke Scale (NIHSS): Bravata 2009	101 (1)	Expert assessment and MASA	Sens = 0.79 (0.63 to 0.90) spec = 0.68 (0.55 to 0.79)	Insufficient evidence to allow firm conclusions
Nursing Screening Tool: Bravata 2009	39 (1)	Expert assessment and MASA	Sens = 0.29 (0.08 to 0.58) spec = 0.84 (0.64 to 0.95)	Insufficient evidence to allow firm conclusions

FEES: fibreoptic endoscopic evaluation of swallowing.
MASA: Mann Assessment of Swallowing Ability.
sens: sensitivity.
spec: specificity.
VFSS: videofluoroscopic swallowing study.
WST: water-swallowing test.

BACKGROUND

Stroke is the second most common cause of death worldwide (Katan 2018), and it is the largest cause of complex preventable disability in adults (WHF 2016). Over 13.7 million new strokes are reported globally each year (Johnson 2019), with a projected increase in incidence due to an ageing population (Avan 2019).

The reported incidence of dysphagia (swallowing difficulties) following acute stroke varies between 37% and 78% (Martino 2005); this is related to type and severity of stroke and individual characteristics (e.g. comorbidities). It is also related to the type of diagnostic assessment used to detect dysphagia and time from admission to that assessment. Dysphagia may lead to aspiration, defined as food or fluid entering the airway below the level of the vocal cords (Blitzer 1988), then into the trachea. People with dysphagia are three times, and those with aspiration 11 times, more likely to develop pneumonia (Kumar 2010; Rofes 2011), which is a significant cause of further morbidity leading to increased hospital stay and risk of death (Intercollegiate Stroke Working Party 2012).

Stroke-associated pneumonia is more common if identification of dysphagia is delayed because of aspiration of food, drink, medications, and oral secretions (Bray 2016). Therefore, early screening for dysphagia may be an important avenue for reducing deaths from acute stroke (Bray 2016; Donovan 2013).

Internationally, different clinical guidelines are available for detecting swallowing difficulties in acute stroke. Identification of dysphagia is a criterion that is considered for obtaining Stroke Unit accreditation by the European Stroke Organisation. Within the UK, Europe, Canada, the USA, and Australia, guidelines state that on admission to hospital, people with acute stroke should have their swallowing screened promptly following admission (European Stroke Organisation 2015; Intercollegiate Stroke Working Party 2016; National Stroke Foundation 2010; Powers 2019; Teasell 2019).

Diagnostic methods such as videofluoroscopy (VF) and fibreoptic endoscopic evaluation of swallowing (FEES) are available. However, these methods have significant limitations when used for early identification of dysphagia in acute stages of stroke. Limitations may be patient specific (patients may be unable to comply with instructions due to poor posture, cognition, medical state, or fluctuating swallowing ability); organisational (speech and language therapists (SLTs) may not be available in the acute setting, and in most hospitals, instrumental examinations such as VF and FEES can be requested only after SLT evaluation; furthermore, not all staff are trained to interpret the results of these reference tests); or procedural (specialist equipment is not available in all acute settings) (Boaden 2011). Therefore, screening tools are needed to identify those likely to have dysphagia, who might be at greater risk of aspiration, so that interventions can be put in place to reduce morbidity until access to specialist equipment and staff is available for assessment and/or until the patient is fit enough to undergo the procedures. These screening tools are called bedside swallow screening tools.

Bedside swallow screening tools must be quick and easy to administer. To be clinically useful, screening tools must accurately identify those with dysphagia with its associated risk of aspiration (sensitivity), without leading to unnecessary restrictions (e.g. withholding food and drinks) for those who do not have dysphagia (specificity). The outcome of a screening test is binary (present or not), although it is acknowledged there may be different levels of severity of swallowing difficulty and subsequent management. A positive test should lead to a referral for more definitive assessment to confirm the presence of dysphagia or not, when this is available. Screening tools also need to be acceptable and feasible for use in people with a range of sequelae following stroke (e.g. different levels of consciousness, cognitive levels, postural difficulties). Screening tools must be undertaken prior to any oral intake and therefore need to be administered by staff members who are at the bedside throughout a 24-hour period.

There is wide variation in available screening tools. Some swallow screening tools rely on questionnaires, whilst others use food and drink as testing materials. Swallow screening tools vary in the types of foods and fluids tested. Water swallow tools, such as the Standardised Swallow Assessment (Perry 2001 Study 1), or the Massey Bedside Swallow Screen (Massey 2002), are very similar in that they offer people different quantities of fluids from assorted utensils; some studies use water plus a range of consistencies to prescribe food and drink consistency management plans. This group of swallow screening tests are more dissimilar from each other as they offer a variety of food and drink, give different consistencies, trial different volumes, and alter the order in which foods and drinks are offered. One concern is that some of these tools have been developed and assessed with various reference tests and different professional groups not routinely available at the bedside, which limits their clinical utility.

There is no universally accepted screening tool for identification and management of aspiration associated with dysphagia, and a wide range of screening tests are used in clinical practice throughout the world. A false-negative result is much worse than a false-positive result. If the swallow screening test fails to identify swallowing difficulties, the person will continue oral intake and may experience the difficulties identified above. If the swallow screening test erroneously identifies swallowing difficulties, the patient is placed 'nil by mouth', with significant impact on quality of life, until the SLT undertakes a more detailed assessment (usually the next day). So a bedside swallow screening tool should be chosen while taking into account this simple clinical reality. Hence, there is a need to identify the most clinically useful screening test or tools that most accurately identify the presence or absence of aspiration associated with dysphagia, and to prescribe a food and drink consistency management plan for individuals with aspiration to improve their medical, social, and psychological outcomes.

Target condition being diagnosed

Aspiration risk associated with dysphagia in people who have had an acute stroke.

Index test(s)

In this review, bedside swallowing screening tests not administered by SLTs are the index tests. A wide range of index tests (swallow screening tests) are used at the bedside by healthcare professionals for recognition or determination of whether the patient is at risk of aspirating food and fluids.

Clinical pathway

Bedside swallow screening tests are commonly used in the typical care pathway in compliance with clinical guidelines for stroke. However, there are varying degrees of compliance. Usually these

tests are implemented by allied healthcare professionals (mainly nurses) following admission to the acute care sector. Patients who are identified as at risk of swallowing difficulties are often placed 'nil by mouth' while awaiting an expert assessment or a further reference test. The care pathway for patients who have a negative test result usually allows patients to continue to eat and drink orally.

Prior test(s)

No tests to assess swallowing are conducted on patients before hospital admission.

Role of index test(s)

Index tests, that is, bedside screening tests not administered by SLTs, have the potential to improve the identification of people with risk of aspiration associated with dysphagia following stroke. Therefore, these tests may reduce the need for SLT evaluation and for more complex, invasive, and more expensive imaging methods, such as VF, FEES, or scintigraphy, which may not be available in some healthcare settings.

Alternative test(s)

Other index tests include questionnaires that rely on self-reported dysphagia symptoms. However, these tools, for example, the Sydney Swallow Questionnaire (SSQ) and the Swallowing Disturbance Questionnaire (SDQ), were used or designed and validated for use in patient populations such as those with head and neck cancer or Parkinson's disease. These are not considered in this review.

Rationale

We reviewed the diagnostic accuracy of currently available swallow screening tests. A systematic review of published evaluations of these screening tests will assist practitioners to identify those that have undergone rigorous development and testing. This review will also identify gaps in evidence for further research.

Healthcare professionals within acute stroke care settings are responsible for deciding which bedside swallow screening test they will use to detect people at risk of aspiration associated with dysphagia in adult acute stroke patients. When a bedside swallow screening test is considered, a test with high sensitivity and specificity is paramount. False-negative results may lead to serious consequences within the clinical context, as continued oral intake may precipitate aspiration pneumonia and death. High specificity is also beneficial, as false-positive results may result in patients being designated 'nil by mouth' and provided with clinically assisted nutrition and hydration unnecessarily. This adversely affects the well-being of the patient and incurs unnecessary costs. However, false-positive results are less dangerous, in that usually an SLT will see the patient within 24 hours.

A recent systematic review of bedside swallow screening tests provided a descriptive analysis of the different elements within screening tests but did not compare these with a reference standard (Almeida 2015). Other reviews have not specifically focused on a stroke population (Brodsky 2016; O'Horo 2015), have included studies in which screening was not undertaken in a timely manner (Schepp 2012), have focused on individual clinical determinants or behaviours associated with aspiration (Daniels 2012), have reviewed only multi-consistency tests (Benfield 2020), or have accepted delays longer than 24 hours between the index test and the reference standard (Daniels 2012). The results of this review will help guide policy makers and healthcare workers in acute hospital settings on selection of the most appropriate bedside swallow screening test from those currently available and might also identify the research agenda going forward.

OBJECTIVES

Primary objective

- To determine the diagnostic accuracy and the sensitivity and specificity of bedside screening tests for detecting risk of aspiration associated with dysphagia in people with acute stroke

Secondary objectives

- To assess the influence of the following sources of heterogeneity on the diagnostic accuracy of bedside screening tools for dysphagia
  - Patient demographics (e.g. age, gender)
  - Time post stroke that the study was conducted (from admission to 48 hours) to ensure only hyperacute and acute stroke swallow screening tools are identified
  - Definition of dysphagia used by the study
  - Level of training of nursing staff (both grade and training in the screening tool)
  - Low-quality studies identified from the methodological quality checklist
  - Type and threshold of index test
  - Type of reference test

METHODS

Criteria for considering studies for this review

Types of studies

We considered single-gate and two-gate studies if they compared the accuracy of a bedside screening tool administered by nursing staff or other healthcare professionals (excluding SLTs) with identified reference tests (VF, FEES, scintigraphy, expert assessment). We applied no restrictions in terms of language of publication.

Participants

We included studies if they involved adults (aged 18 and over) who had been admitted to an acute hospital setting with a clinical diagnosis of stroke. We considered studies that were inclusive of people with subarachnoid haemorrhage and we excluded this sample subgroup from the analysis, when possible. We excluded studies that included only patients with subarachnoid haemorrhage. We also excluded studies that admitted patients with trauma.

Index tests

We included studies if they evaluated a bedside swallow screening test, used by nursing staff or other healthcare professionals (excluding SLTs who have expert knowledge in this area), for recognition or determination of whether patients were at risk of aspiration associated with dysphagia.

Target conditions

We included studies if they reported the accuracy of the bedside screening tool for identification of the risk of aspiration owing to dysphagia post stroke.

Reference standards

We included studies if a bedside swallow screening test, carried out by nursing staff or other healthcare professionals, was compared with VF, FEES, scintigraphy, or an expert assessment. Videofluoroscopy is an X-ray video of swallowing, allowing the swallow to be analysed in real time. FEES involves insertion of a fibreoptic flexible endoscope that is passed through the nasal passages to view the throat pre-swallows and post-swallows for secretion management, residue, and aspirated material. Scintigraphy uses radioisotopes that are swallowed, and the emitted radiation is captured by external detectors (gamma cameras) to form images of swallowed material. Expert assessments included assessments conducted by dysphagia-trained professionals such as SLTs. All reference standards are not equally valid and available (the SLT is not usually available in the acute setting and is not available 24 hours each day and seven days a week), and some procedures are more invasive than others. Expert assessment was included as a reference standard owing to the lack of immediate access to imaging or instrumental assessment in many centres, and to the inability of some patients to co-operate with these assessments. However, expert assessments have limitations in their ability to identify at bedside those patients who have swallowing difficulties owing to silent aspiration when the patient demonstrates no clinical signs of aspiration.

Search methods for identification of studies

Electronic searches

The search strategy used was developed with the help of the Cochrane Stroke Group Information Specialist. We searched relevant electronic databases for eligible diagnostic studies from inception until 9 December 2019: the Cochrane Central Register of Controlled Trials (CENTRAL; 2019, Issue 12), in the Cochrane Library; MEDLINE Ovid (from 1946); Embase Ovid (from 1980); the Cumulative Index to Nursing and Allied Health Literature (CINAHL) EBSCO (from 1937 onward), and the Health Technology Assessment (HTA) Database via the Centre for Reviews and Dissemination (University of York). The search strategy for MEDLINE is presented in Appendix 1 and was adapted for CENTRAL (Appendix 2), Embase (Appendix 3), CINAHL (Appendix 4), and HTA (Appendix 5).

Searching other resources

We checked the reference lists of all included studies, and we performed a cited reference search using Science Citation Index, to identify additional relevant studies and systematic reviews. We also contacted experts in the field to identify ongoing studies and those potentially missed by the search strategy.

We searched targeted grey literature sources, which we identified from the the Canadian Agency for Drugs and Technologies in Health (CADTH) Grey Matters document (CADTH 2018). A full list of the grey literature sources searched is displayed in Appendix 6.

Data collection and analysis

Selection of studies

Two review authors from the team (EB, JB, HR, DD) independently screened all titles and abstracts identified by the electronic database searches and excluded duplications and irrelevant records. Conflicts were resolved by involvement of a third review author (PD, AA). We then obtained full-text articles for the remaining studies, which were independently assessed for inclusion by two review authors (EB, ML, HR, PD) using the eligibility criteria described above. Disagreements were resolved by consensus meetings, with arbitration provided by a third review author (AA) who had not initially reviewed the paper. When multiple papers used the same cohort of patients, we used the paper with the most complete and up-to-date data. When we identified relevant conference abstracts, we looked for the corresponding full-text article. If we found no full-text article, we contacted the authors of the conference abstract. We excluded conference abstracts with insufficient data to calculate the 2 × 2 table. The selection process is detailed in the PRISMA flow diagram in Figure 1.

Figure 1. PRISMA flow diagram.

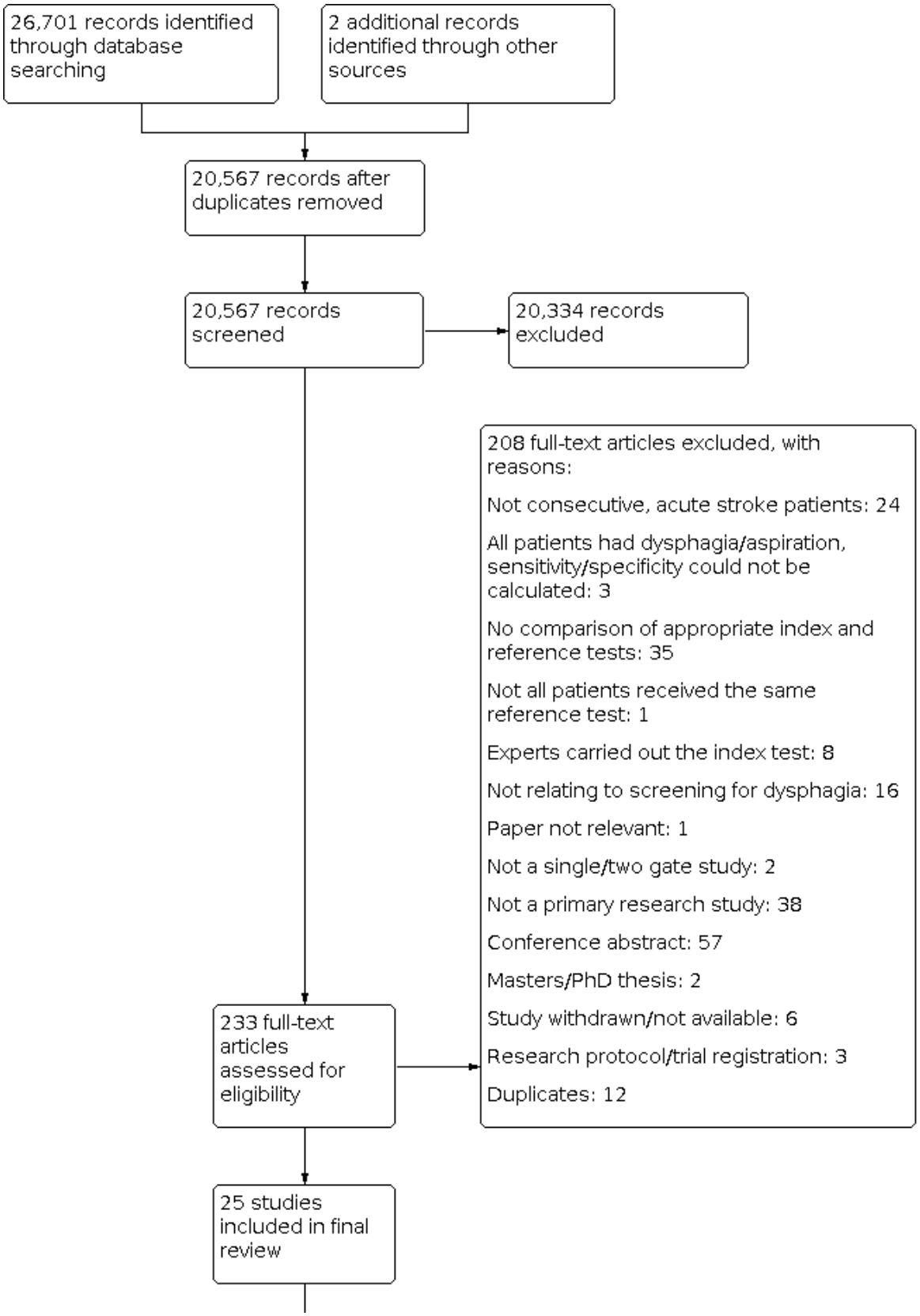
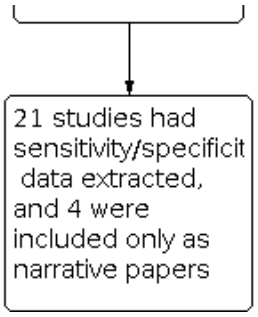


Figure 1. (Continued)



Data extraction and management

Two review authors (JB, EM) designed a bespoke data extraction form to collect details from the included studies. We piloted the form on other diagnostic accuracy studies related to acute stroke management that are beyond the scope of this review. For each study, we extracted information related to characteristics of the study, the patient population, index tests, reference standards, and any relevant outcomes.

Among a group of review authors (JB, AC, LH, CEL, ML, CW), two independently extracted the data to ensure adequate reliability and quality. These two review authors then met to agree on final data extractions for each included study. When there was disagreement between the two review authors, a third review author from the group acted as an arbitrator. Once data extraction was complete, we entered the information into RevMan 5 (Review Manager 2014).

Assessment of methodological quality

Two review authors from the group (JB, AC, LH, CEL, ML, CW) independently reviewed the methodological quality of each included study, using criteria from the Quality Assessment of Studies of Diagnostic Accuracy (QUADAS-2) tool, as recommended by Cochrane (Appendix 7). As with previous diagnostic test accuracy reviews (Gupta 2016), we changed the QUADAS-2 question "Was a case-control design avoided?" to "Was a two-gate design avoided?".

We piloted the QUADAS-2 to promote agreement in interpretation within the team. We resolved disagreements with a third review author. We rated the QUADAS-2 criteria as 'yes', 'no', or 'unclear' for each included study. As stated in our protocol, guidance indicates that criteria 1, 2, 3, and 9 should receive particular attention regarding definitions used as the basis for decisions. The other criteria are more self-explanatory with regards to their interpretation.

Statistical analysis and data synthesis

From each index test, the parameters of interest were sensitivity and specificity. For a majority of index tests, these were calculated from the 2 × 2 tables. For other index tests, information from reported parameters (sensitivity, specificity, or predictive values) was used to inform the 2 × 2 table, yielding true-positive, false-positive, true-negative, and false-negative values. From available information, sensitivities, specificities, and their 95% confidence intervals were calculated by RevMan 5 and plotted in forest plots and summary receiver operating characteristic (SROC) plots. In the absence of sufficient data, to calculate the 2 × 2 tables, we contacted study authors for further information. If this information remained unavailable, we included the study in the review as a narrative study.

Data from a single study were used only once in each analysis. We checked all studies to see if all patients were included, or if the number of patients was fixed, *a priori* in the design, from a condition (from the reference test) or from the index test. If the numbers of patients were fixed, the formulae for sensitivity and specificity would be different. No included studies used such numbers fixed *a priori.*

Pooling of sensitivity and specificity results was intended for each index test. However, due to the small number of studies using the same index test (three or fewer), we did not perform any meta-analyses of diagnostic accuracy data. Pooling of screening tests by general categories was performed as part of the descriptive analysis, presented along with sensitivity and specificity.

Investigations of heterogeneity

Possible co-variates of interest were patient demographics (e.g. age, gender), time post stroke that the study was conducted, and level of training of nursing staff. However, the potential influence of these co-variates could not be investigated due to the small number of studies for each index test. The definitions of dysphagia used by the study could have been grouped, and the type and quality of the reference test, as well as the type of index test, investigated as sources of heterogeneity. For the same reason, no subgroup analyses were performed involving age, gender, time post stroke of the index test, or index test type, and whether there had been a significant change in the participant's condition between performance of index and reference tests. Index tests were pooled by broad categories (by index test type, by healthcare professional (HCP), by outcome, and by reference test), and graphical comparisons were made. We therefore performed only a narrative review running exploratory analysis in RevMan 5 to show graphically if the co-variates of interest are likely to be related to test accuracy, displayed using forest plots and SROC plots.

Sensitivity analyses

We did not carry out any sensitivity analyses due to the small number of studies for each index test. We had proposed to exclude low-quality studies and possibly reporting a delay in time between index and reference tests due to potential changes in patient consciousness and cognitive state.

Assessment of reporting bias

Since no methods are available to quantify publication bias in diagnostic test accuracy reviews, we did not conduct an assessment of reporting bias.

RESULTS

Results of the search

We conducted two searches - the first in June 2017 and the same search in December 2019. We identified 26,701 papers through database searches, and we found two additional studies by handsearching the references of existing systematic reviews, for a total of 20,567 articles once duplicates were removed. After titles and abstracts of articles had been screened, we assessed the full texts of 233 articles against the eligibility criteria. Of these, we excluded 208 articles - 35 because they did not compare appropriate index and reference tests, 24 because they did not involve consecutive acute stroke patients, 16 because they did not relate to screening for dysphagia, eight because index tests were carried out by experts, three because they involved only patients with dysphagia or aspiration and so sensitivity and specificity could not be calculated, two because they were not single-gate or two-gate studies, and one because not all patients received the same reference test (see Characteristics of excluded studies). Additionally, we excluded 57 references that were conference abstracts, 38 that were not primary research studies, 12 that were duplicate papers, six that had been withdrawn or were unavailable, three that were research protocols or trial registrations, two that were Master's or PhD theses, and one that was not relevant. Twenty-five articles went through to data extraction and were included in the final review. Four of the included articles did not include enough detail for calculating the diagnostic accuracy of the index test, thus we included these in narrative descriptions only. More detail is shown in the PRISMA diagram (Figure 1).

We checked the reference lists of included studies, but this did not reveal additional eligible papers. Details on participants, index tests, reference tests, and sensitivity and specificity of each index test are displayed in the Characteristics of included studies table.

The 25 included studies investigated a total of 37 screening tools, most of which were different tools. All 25 studies are single-gate studies. Of 37 screening tools, 21 were used by nurses and 16 by other HCPs. The other HCPs were doctor or physician (n = 6), dietician (n = 1), neurologist (n = 2), multiple HCPs (n = 4), and HCP not recorded (n = 3). Of the 37 screening tools, four could not be included in an available quantitative analysis as 2 × 2 tables could not be constructed to estimate sensitivity or specificity (Eren 2019; Martino 2014; Nishiwaki 2005; Zhou 2011). Training was given for screening tools used by 17 (81%) nurses and by four (25%) other HCPs.

Demographic information and characteristics of all 37 screening tests (including the four narrative tests) are given in Table 1. When reported, mean age was 58.6 years to 76.8 years, the percentage of males in the study was 35% to 100%, and the start date of study recruitment was between 1995 and 2016. Twenty-two (59%) tests were conducted in the USA and Canada, eight (22%) in the UK and Europe, and six (16%) in other countries; in one (3%) study, the country was not recorded. Only two (5%) studies used an established classification system to record the location of the stroke. The median sample size was 100, with an interquartile range (IQR) of 50 to 161; 16 (48%) screening tools had a sample size less than 100.

For 16 (43%) tests, admission to index test time was ≤ 24 hours, for six (16%) between 24 hours and 72 hours, and for two (5%) ≥ 72 hours; for 13 (35%), this information was not recorded. For 28 (76%) tests the index test was applied before the reference test, for one (3%) the reference test was applied before the index test, for five (14%) the order was not specified or mixed, and for three (8%) this was not recorded. The time interval between index and reference tests was ≤ 24 hours for 18 (49%) tests, was > 24 hours for 10 (27%) tests, and was not recorded for nine (24%) tests.

Twenty-four (64.9%) screening tools used water only (Behera 2018a; Behera 2018b; Cummings 2015; Daniels 1997; Daniels 2016a; Daniels 2016b; Edmiaston 2010a; Edmiaston 2010b; Edmiaston 2014a; Edmiaston 2014b; Eren 2019; Huhn-Matesic 2015; Jiang 2019a; Jiang 2019b; Lim 2001b; Martino 2009 Study 1; Martino 2014; Mulheren 2019; Nishiwaki 2005; Perry 2001 Study 1a; Perry 2001 Study 1b; Warnecke 2008a; Warnecke 2008b; Zhou 2011), six (16.2%) used water and other consistencies (Campbell 2016; Ellis 2013; Huhmann 2004; Perry 2001 Study 1c; Perry 2001 Study 1d; Trapl 2007b), and seven (18.9%) used other methods. The other methods were evaluation of patient characteristics only (Antonios 2010a; Antonios 2010b; Bravata 2009a), notes review (Bravata 2009b), combined water swallow test and oxygen saturation (Lim 2001a), oxygen saturation (Smith 2000), and evaluation of patient characteristics followed by water swallow test and oxygen saturation (Turner-Lawrence 2009). Several references reported use of more than one screening tool as detailed in Table 2.

Expert assessment or the Mann Assessment of Swallowing Ability (MASA) was utilised for 20 (54.1%) screening tools as the reference test (Antonios 2010a; Antonios 2010b; Behera 2018a; Behera 2018b; Bravata 2009a; Bravata 2009b; Campbell 2016; Cummings 2015; Edmiaston 2010a; Edmiaston 2010b; Ellis 2013; Huhmann 2004; Huhn-Matesic 2015; Jiang 2019a; Jiang 2019b; Perry 2001 Study 1a; Perry 2001 Study 1b; Perry 2001 Study 1c; Perry 2001 Study 1d; Turner-Lawrence 2009); six (16.2%) used FEES (Eren 2019; Lim 2001a; Lim 2001b; Trapl 2007b; Warnecke 2008a; Warnecke 2008b), and 11 (29.7%) used VF (Daniels 1997; Daniels 2016a; Daniels 2016b; Edmiaston 2014a; Edmiaston 2014b; Martino 2009 Study 1; Martino 2014; Mulheren 2019; Nishiwaki 2005; Smith 2000; Zhou 2011).

We identified 15 (40.5%) screening tools that had an outcome of aspiration risk (Behera 2018a; Daniels 1997; Daniels 2016a; Daniels 2016b; Edmiaston 2010a; Edmiaston 2014a; Huhn-Matesic 2015; Lim 2001a; Lim 2001b; Nishiwaki 2005; Smith 2000; Trapl 2007b; Turner-Lawrence 2009; Warnecke 2008a; Warnecke 2008b), and we found 20 (54.1%) that had an outcome of dysphagia (Antonios 2010a; Antonios 2010b; Behera 2018b; Bravata 2009a; Bravata 2009b; Campbell 2016; Cummings 2015; Edmiaston 2010b; Edmiaston 2014b; Ellis 2013; Eren 2019; Huhmann 2004; Jiang 2019a; Jiang 2019b; Martino 2009 Study 1; Mulheren 2019; Perry 2001 Study 1a; Perry 2001 Study 1b; Perry 2001 Study 1c; Perry 2001 Study 1d). Only two (5.4%) narrative papers did not record the outcome (Martino 2014; Zhou 2011).

Methodological quality of included studies

Risk of bias

Results of the methodological quality assessment for each of the 25 included studies are shown in Figure 2. We considered six studies to be at low risk across all four risk of bias domains: patient selection, index test, reference standard, and flow and timing (Daniels 2016; Lim 2001; Martino 2009 Study 1; Martino 2014; Trapl 2007; Warnecke 2008). Two studies were at low risk of bias for three domains (index test, reference standard, and flow and timing) but at uncertain risk for patient selection (Edmiaston 2014; Mulheren 2019).

Figure 2. Risk of bias and applicability concerns summary: review authors' judgements about each domain for each included study.

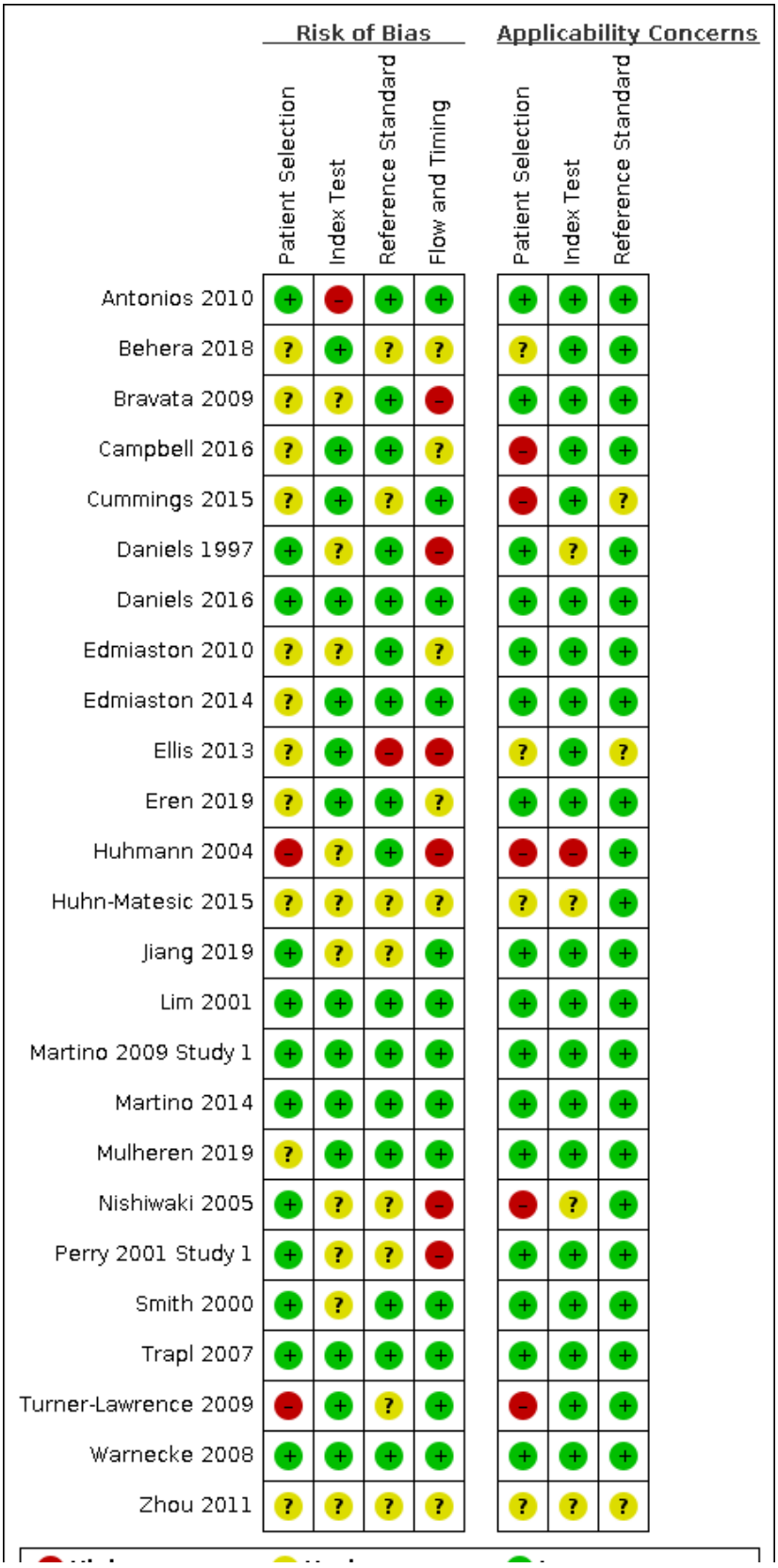
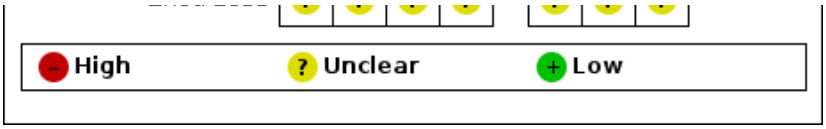


Figure 2. (Continued)



A major concern for risk of bias across the other included studies was that they had not enrolled consecutive patients. We identified 12 studies that had enrolled consecutive patients (Antonios 2010; Daniels 1997; Daniels 2016; Jiang 2019; Lim 2001; Martino 2009 Study 1; Martino 2014; Nishiwaki 2005; Perry 2001 Study 1; Smith 2000; Trapl 2007; Warnecke 2008), but for 12 studies this was unclear (Behera 2018; Bravata 2009; Campbell 2016; Cummings 2015; Edmiaston 2010; Edmiaston 2014; Ellis 2013; Eren 2019; Huhmann 2004; Huhn-Matesic 2015; Mulheren 2019; Zhou 2011), and one study employed a convenience sampling method (Turner-Lawrence 2009). For 11 studies, it is unclear whether inappropriate exclusions had been avoided when participants were recruited (Behera 2018; Bravata 2009; Cummings 2015; Edmiaston 2010; Ellis 2013; Huhmann 2004; Huhn-Matesic 2015; Martino 2014; Perry 2001 Study 1; Turner-Lawrence 2009; Zhou 2011). Finally, it is unclear in eight studies whether there had been an appropriate time interval between the index test and the reference standard (Behera 2018; Bravata 2009; Edmiaston 2010; Ellis 2013; Huhmann 2004; Huhn-Matesic 2015; Perry 2001 Study 1; Zhou 2011), and it was reported that the time interval was not appropriate in two studies (Daniels 1997; Nishiwaki 2005).

Applicability concerns

There were no applicability concerns for 15 studies across the three applicability domains: patient selection, index test, and reference standard (Antonios 2010; Bravata 2009; Daniels 2016; Edmiaston 2010; Edmiaston 2014; Eren 2019; Jiang 2019; Lim 2001; Martino 2009 Study 1; Martino 2014; Mulheren 2019; Perry 2001 Study 1; Smith 2000; Trapl 2007; Warnecke 2008). However, concern was high in five studies (Campbell 2016; Cummings 2015; Huhmann 2004; Nishiwaki 2005; Turner-Lawrence 2009), and concern was unclear in four other studies (Behera 2018; Ellis 2013; Huhn-Matesic 2015; Zhou 2011), which included patients that did not match the review question. Concern that the index test, its conduct, or its interpretation differs from the review question was high in Huhmann 2004 and unclear in four studies (Daniels 1997; Huhn-Matesic 2015; Nishiwaki 2005; Zhou 2011).

Findings

Healthcare professionals' level of training across studies

The level of training for healthcare professionals undertaking the index test in dysphagia and/or use of the index test was underreported, with only 21 of 37 screening tests stating that training was offered. Of those that detailed the training, there was variation in both length and content. Some index tests required 10 minutes of training (n = 4), some took several hours (n = 3), and two detailed two days of training. The content of training varied between training to use the test plus element of competence (n = 5), instructions on how to use the index test (n = 4), instructions with a practice component (n = 2), training in anatomy and physiology of swallowing, identification and management of dysfunction and five successful practice assessments (n = 2), on-line training in swallowing anatomy and physiology (n = 1), clinical signs of dysphagia and aspiration (n = 1), and digitised examples of five stroke patients with a review of basic anatomy and physiology of swallowing (n = 1). Some index tests reported that training was offered but did not detail the contents (n = 5).

Synthesis of results

The high level of heterogeneity between studies as identified by the review means that statistical pooling of diagnostic accuracy data was not possible. We performed a descriptive analysis from extracted data (2 × 2 tables) and sensitivity and specificity for all but the four screening tests for which this data could not be extracted (Eren 2019; Martino 2014; Nishiwaki 2005; Zhou 2011); these four studies are not included from this point onwards.

Values for sensitivity and specificity presented here include the point estimate (95% confidence interval (CI)). The criteria of both high sensitivity and high specificity indicate that the best performing test overall was the Standardized Swallowing Assessment tool (SSA) (Perry 2001 Study 1d) (n = 68), which had sensitivity of 0.97 (95% CI 0.86 to 1.00) and specificity of 0.90 (95% CI 0.74 to 0.98). However, Perry 2001 Study 1 performed poorly on the risk of bias assessment, showing, specifically, high risk of bias in the flow and timing domain and unclear risk of bias in both index test and reference standard domains, so this result should be interpreted with caution. Several tests performed better on sensitivity but less well on specificity: the Registered Dietician (RD) Dysphagia Screening tool (n = 32) had sensitivity of 1.00 (95% CI 0.69 to 1.00) with specificity of 0.86 (95% CI 0.65 to 0.97) (Huhmann 2004); the Bedside Aspiration test (n = 50) had sensitivity of 1.00 (95% CI 0.87 to 1.00) with specificity of 0.71 (95% CI 0.49 to 0.87) (Lim 2001a); the Gugging Swallowing Screen (GUSS) (n = 30) had sensitivity of 1.00 (95% CI 0.77 to 1.00) with specificity of 0.69 (95% CI 0.41 to 0.89) (Trapl 2007b); and the Toronto Bedside Swallowing Screening Test (TOR-BSST) (n = 24) had sensitivity of 1.00 (95% CI 0.75 to 1.00) with specificity of 0.64 (95% CI 0.31 to 0.89) (Martino 2009 Study 1).

Of the best performing screening tests, we have the most confidence in the following studies: Bedside Aspiration test (Lim 2001a), GUSS (Trapl 2007b), and TOR-BSST (Martino 2009 Study 1), which we rated as having low risk of bias across all four QUADAS-2 risk of bias domains. Of those with low risk of bias, the best performing screening test using only water was TOR-BSST (Martino 2009 Study 1), and the best performing screening test using water plus other consistencies was GUSS (Trapl 2007b). A description of these tests is provided in Appendix 8. However, all of these screening tools had small sample sizes (n < 100), which limits interpretation of the estimates of reliability of these tests.

Screening tests were grouped into general categories according to taxonomy (water only tests, water plus other consistencies tests, and other tests), healthcare professionals who administered the test (nurses versus other healthcare professionals), types of

outcomes (aspiration versus dysphagia), and types of reference standards (expert assessment or MASA versus VF versus FEES), and we generated related summary receiver operating characteristic (ROC) plots using RevMan 5 (see Figure 3; Figure 4; Figure 5; Figure 6). This informed the descriptive analysis, but no meta-analysis was done because of the small number of studies identified for each individual screening test.

Figure 3. Summary ROC plot of tests: Grouped by Index taxonomy - water only, water plus bolus and other.

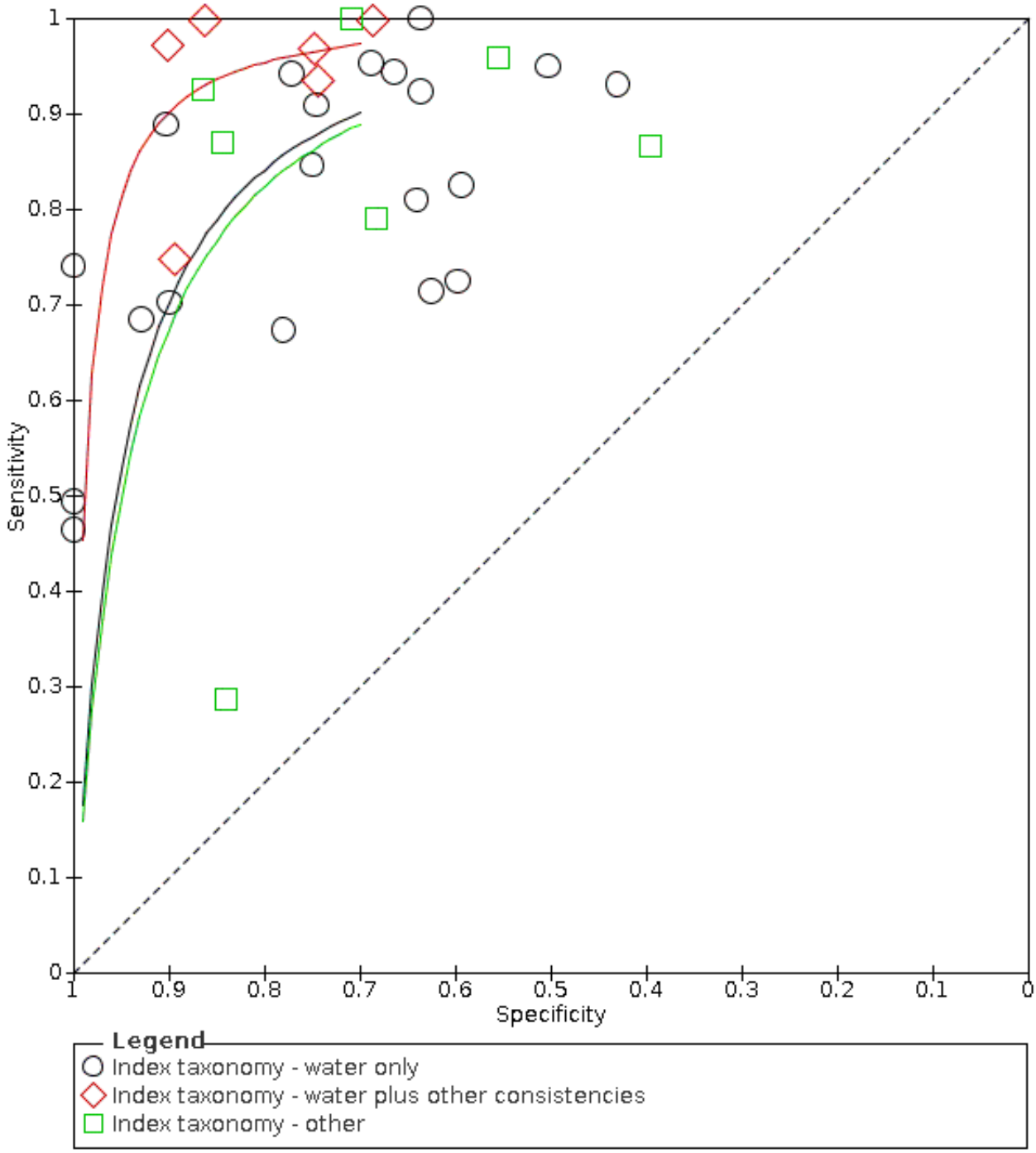


Figure 4. Summary ROC plot of tests grouped by index test healthcare professional - nurse and other.

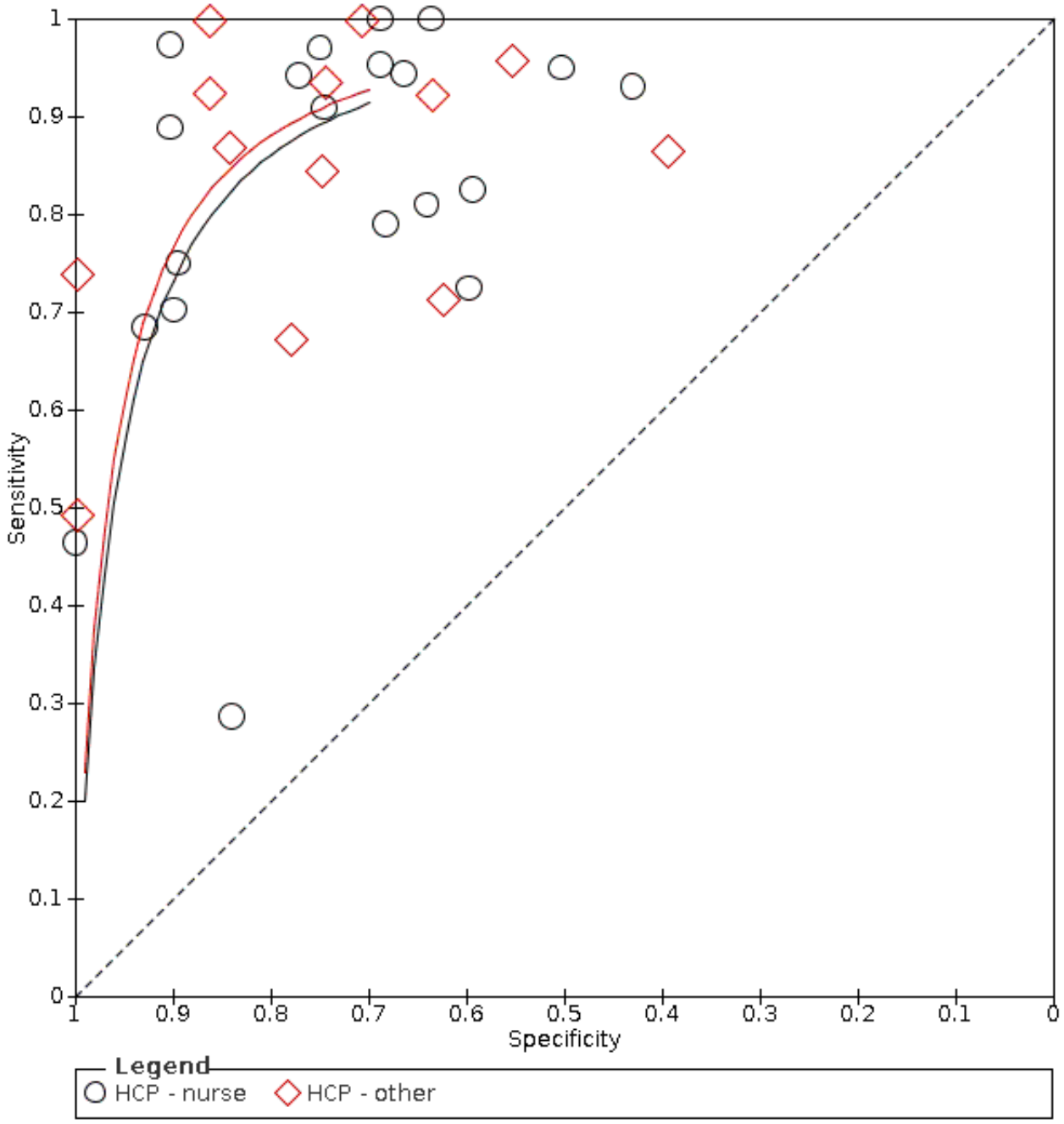


Figure 5. Summary ROC plot of tests: grouping index tests by outcome - aspiration and dysphagia.

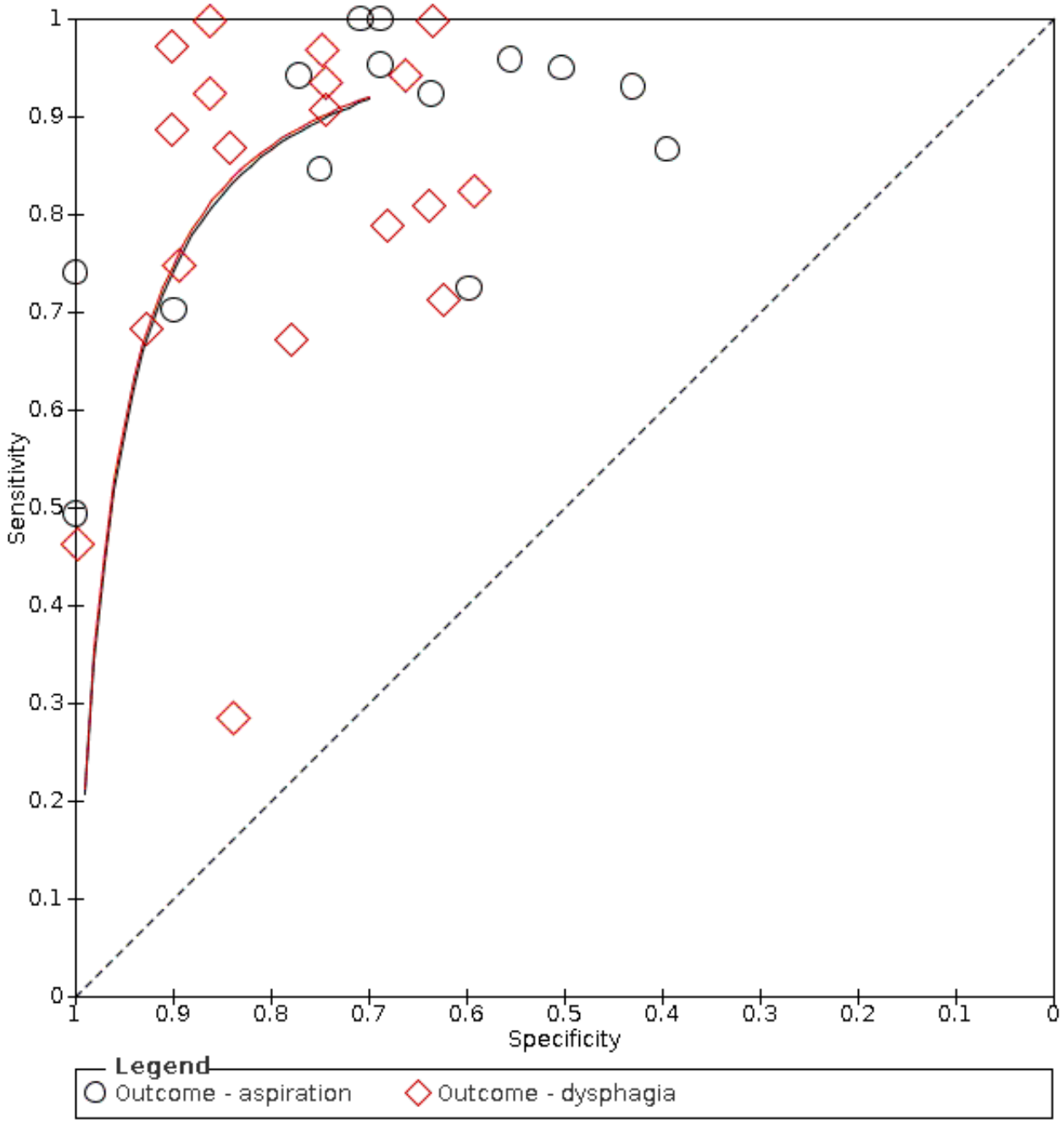
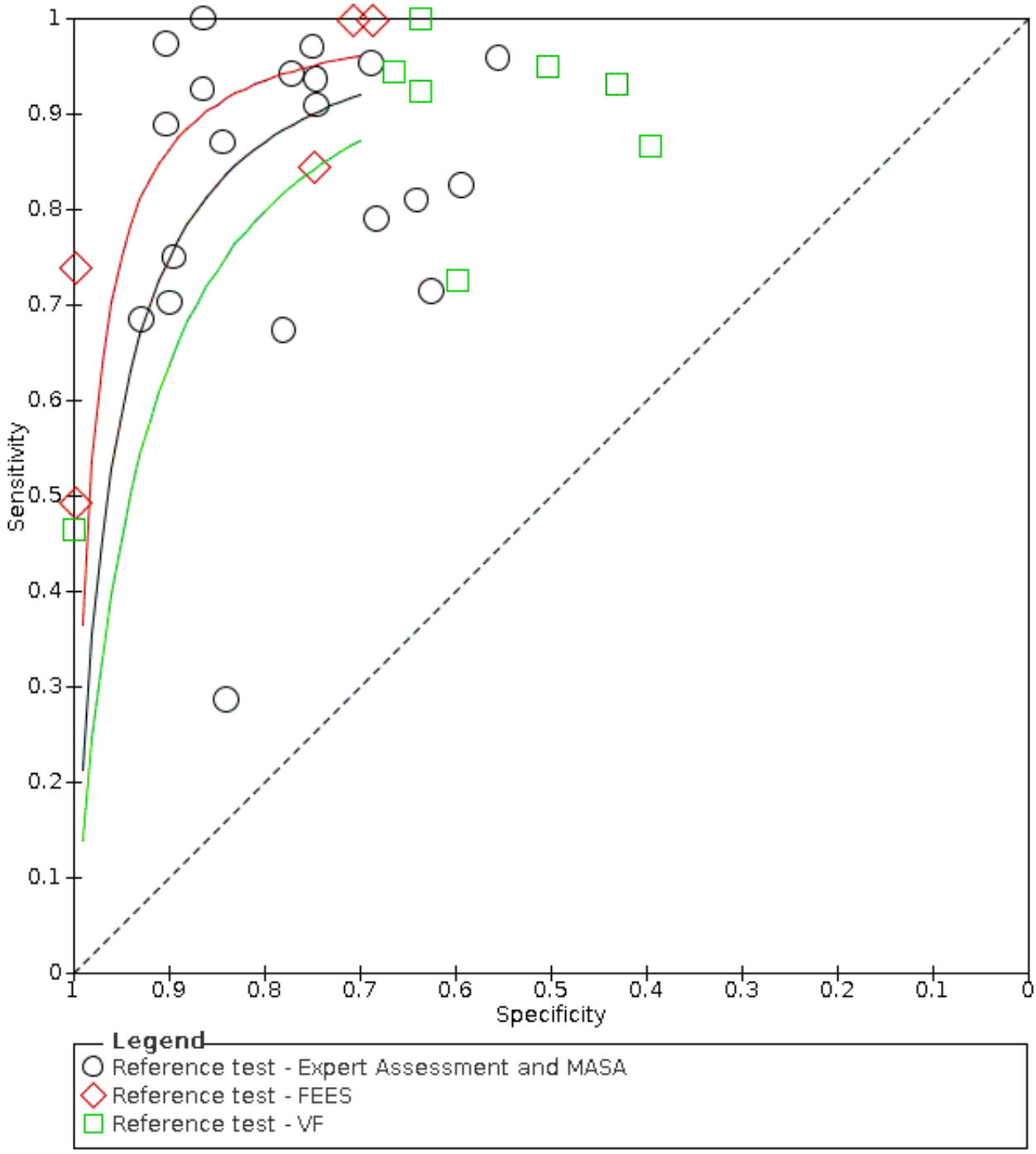


Figure 6. Summary ROC plot of tests: index tests grouped by reference test used - Expert Assessment, FEES and VS.

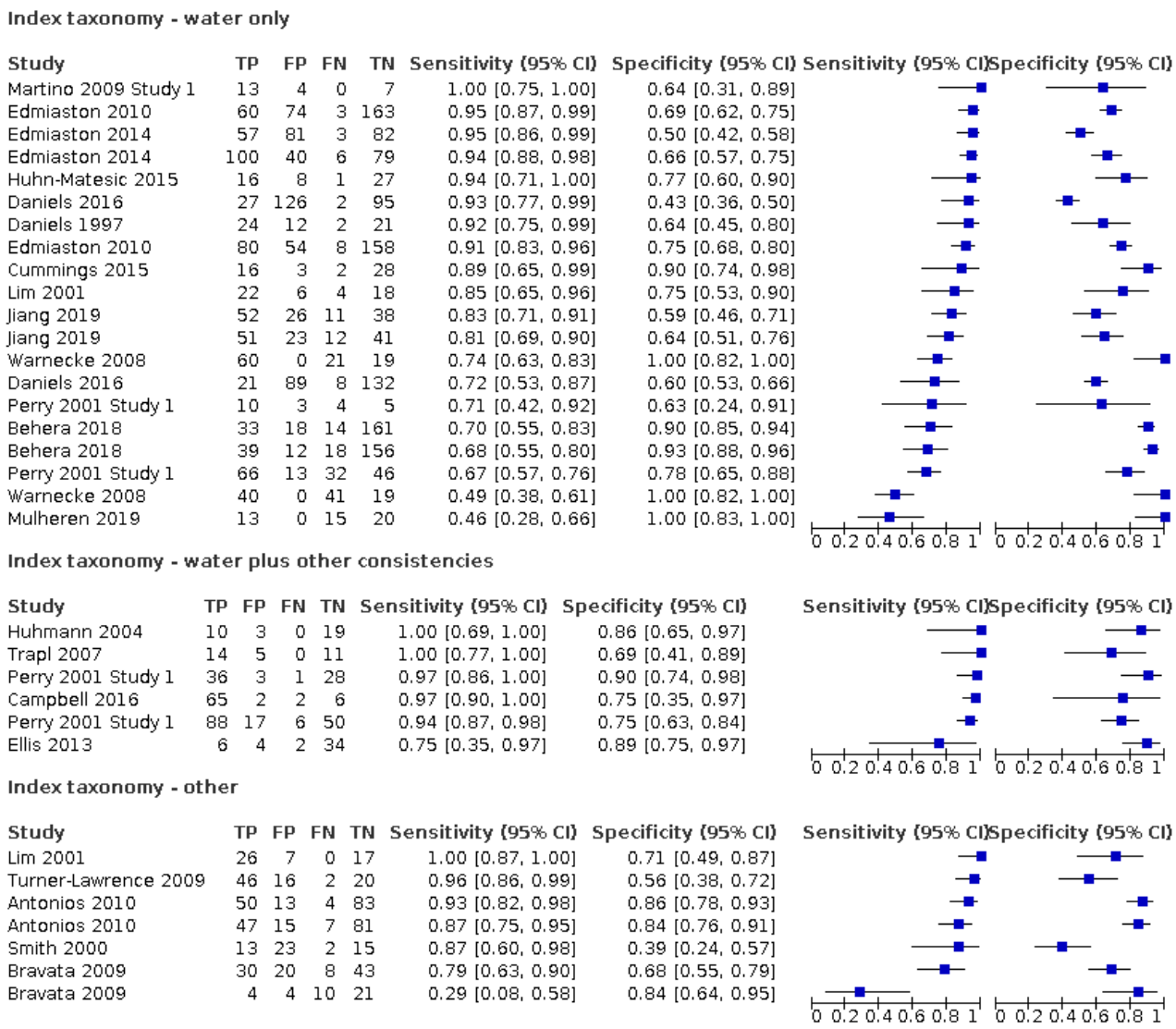


Type of test (water only tests versus water plus other consistencies tests versus other tests)

Screening tools that used water and other consistencies (accuracy ranged from sensitivity of 0.75 (95% CI 0.35 to 0.97) with specificity of 0.89 (95% CI 0.75 to 0.97), to sensitivity of 1.00 (95% CI 0.69 to 1.00) with specificity of 0.86 (95% CI 0.65 to 0.97)) generally were more accurate than screening tests that used only water (accuracy ranged from sensitivity of 0.46 (95% CI 0.28 to 0.66) with specificity of 1.00 (95% CI 0.83 to 1.00) to sensitivity of 1.00 (95% CI 0.75 to 1.00) with specificity of 0.64 (95% CI 0.31 to 0.89)). Those that used methods other than water only and water and other consistencies had mixed results; some performed as well as the water-only tests, with the most accurate scoring sensitivity of 1.00 (95% CI 0.87 to 1.00) with specificity of 0.71 (95% CI 0.49 to 0.87)) (Figure 7). Others performed worse, with the least accurate scoring having sensitivity of 0.29 (95% CI 0.08 to 0.58) with specificity of

0.84 (95% CI 0.64 to 0.95)). Pooling of data by the screening test taxonomy confirmed that screening tests that used water plus other consistencies (participants n = 412) may be more sensitive with similar or better specificity than screening tests that used water only (participants n = 2914) or other methods (participants n = 627) (Figure 3). However, only six tests used this method. Screening tests that used only water had similar accuracy (in terms of sensitivity and specificity) compared with screening tests using other methods.

Figure 7. Forest plot of tests grouped by Index Test Taxonomy - Water only, Water plus other consistencies, and Other tests.

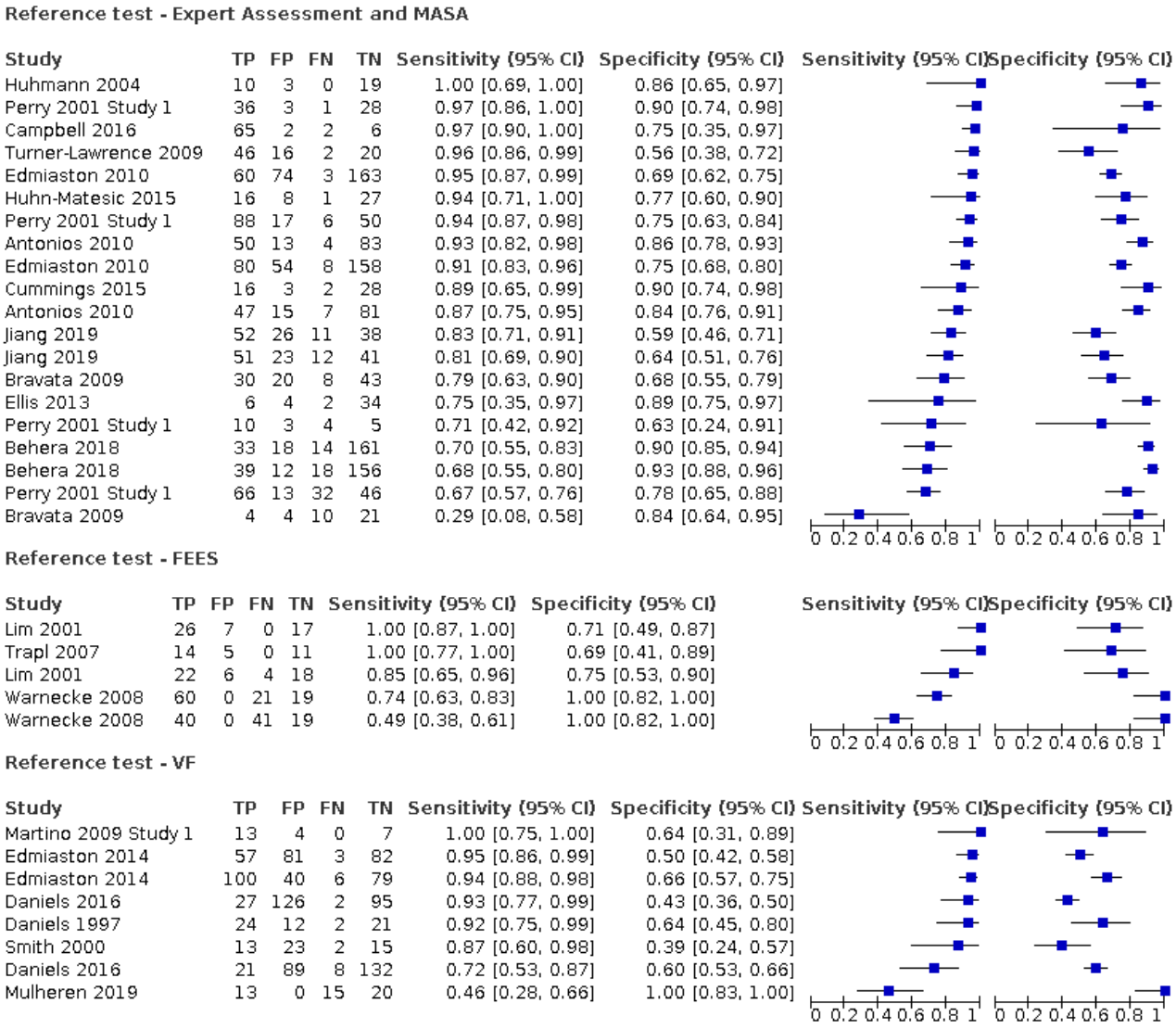


Type of reference standard (expert assessment or MASA versus VF versus FEES)

Screening tools showed better agreement with expert assessment or the MASA (accuracy ranged from sensitivity of 0.29 (95% CI 0.08 to 0.58) with specificity of 0.84 (95% CI 0.64 to 0.95) to sensitivity of 1.00 (95% CI 0.69 to 1.00) with specificity of 0.86 (95% CI 0.65 to 0.97)) than they did with the instrument-based reference test of FEES (accuracy ranged from sensitivity of 0.49 (95% CI 0.38 to 0.61) with specificity of 1.00 (95% CI 0.82 to 1.00) to sensitivity of 1.00 (95% CI 0.87 to 1.00) with specificity of 0.71 (95% CI 0.49 to 0.87)) or VF (accuracy ranged from sensitivity of 0.46 (95% CI 0.28 to 0.66) with specificity of 1.00 (95% CI 0.83 to 1.00) to sensitivity of 1.00 (95% CI 0.75 to 1.00) with specificity of 0.64 (95% CI 0.31 to 0.89) (Figure 8). Pooling of screening tests by reference test (Figure 6) confirmed that screening tests that used expert assessment or the MASA (participants n = 2491) had the highest accuracy. This may be expected, as the screening tests will be more closely aligned to expert assessment methods than to instrumental assessment. Screening tests that used VF as a reference test had high sensitivity but lower specificity than those using expert assessment or the MASA. However, the number of

index tests compared to instrument-based reference tests was low: five were compared to FEES (participants n = 330) and eight were compared to VF (participants n = 1132).

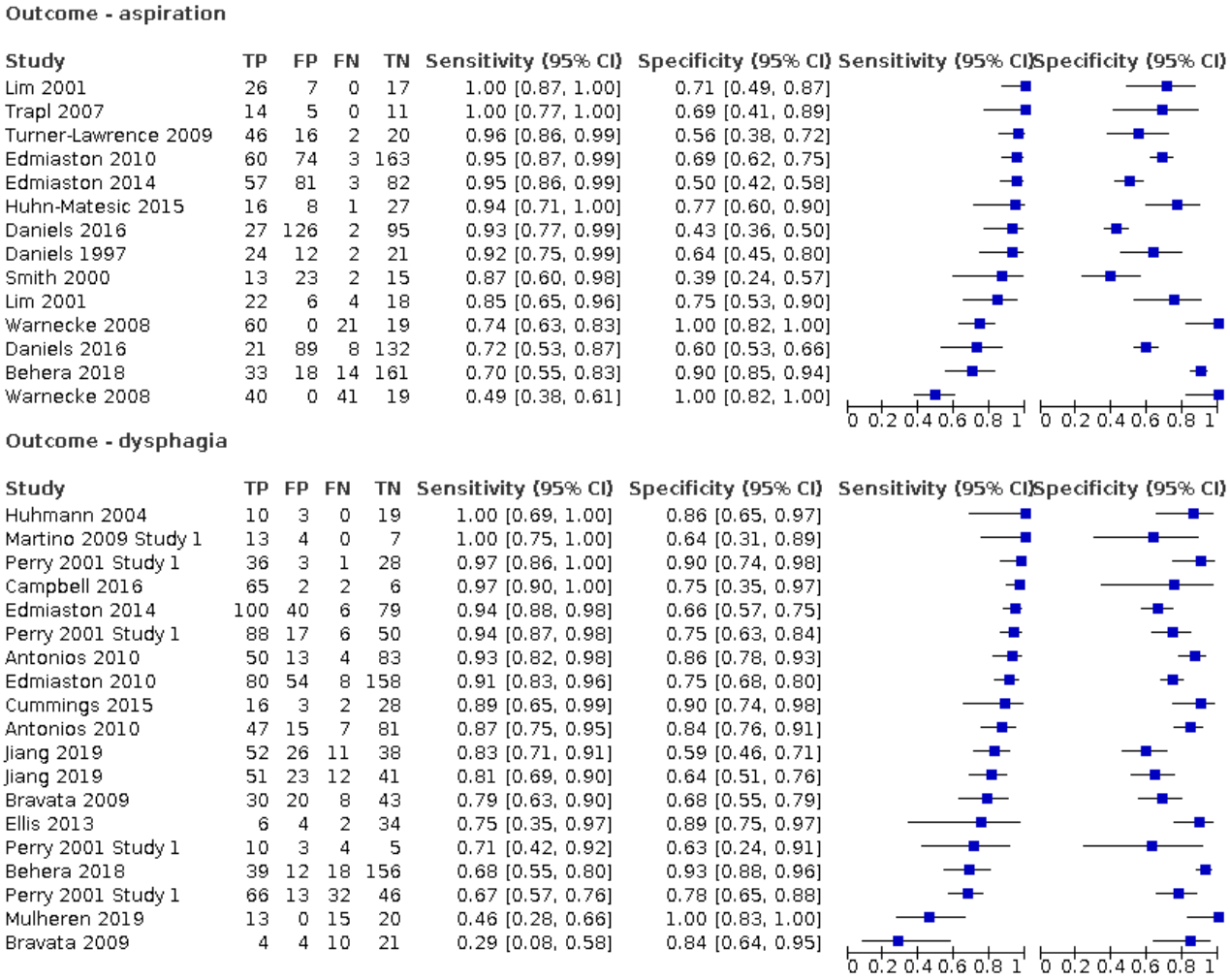
Figure 8. Forest plot of tests grouped by Reference Test - Expert Assessment and MASA, FEES, and VF.



Type of primary outcome (aspiration versus dysphagia)

Screening tests with dysphagia as the primary outcome (accuracy ranged from sensitivity of 0.29 (95% CI 0.08 to 0.58) with specificity of 0.84 (95% CI 0.64 to 0.95) to sensitivity of 1.00 (95% CI 0.69 to 1.00) with specificity of 0.86 (95% CI 0.65 to 0.97)) generally performed better than screening tests for which the primary outcome was aspiration (accuracy ranged from sensitivity of 0.49 (95% CI 0.38 to 0.61) with specificity of 1.00 (95% CI 0.82 to 1.00) to sensitivity of 1.00 (95% CI 0.87 to 1.00) with specificity of 0.71 (95% CI 0.49 to 0.87)) (Figure 9). Pooling of screening tests by outcome (Figure 5) showed that screening tests with the greatest accuracy, with both high sensitivity and high specificity, had an outcome of dysphagia (participants n = 2126). This was to be expected, as dysphagia is more easily observed than aspiration. Several tests that detected aspiration risk (participants n = 1827) had high sensitivity but slightly lower specificity.

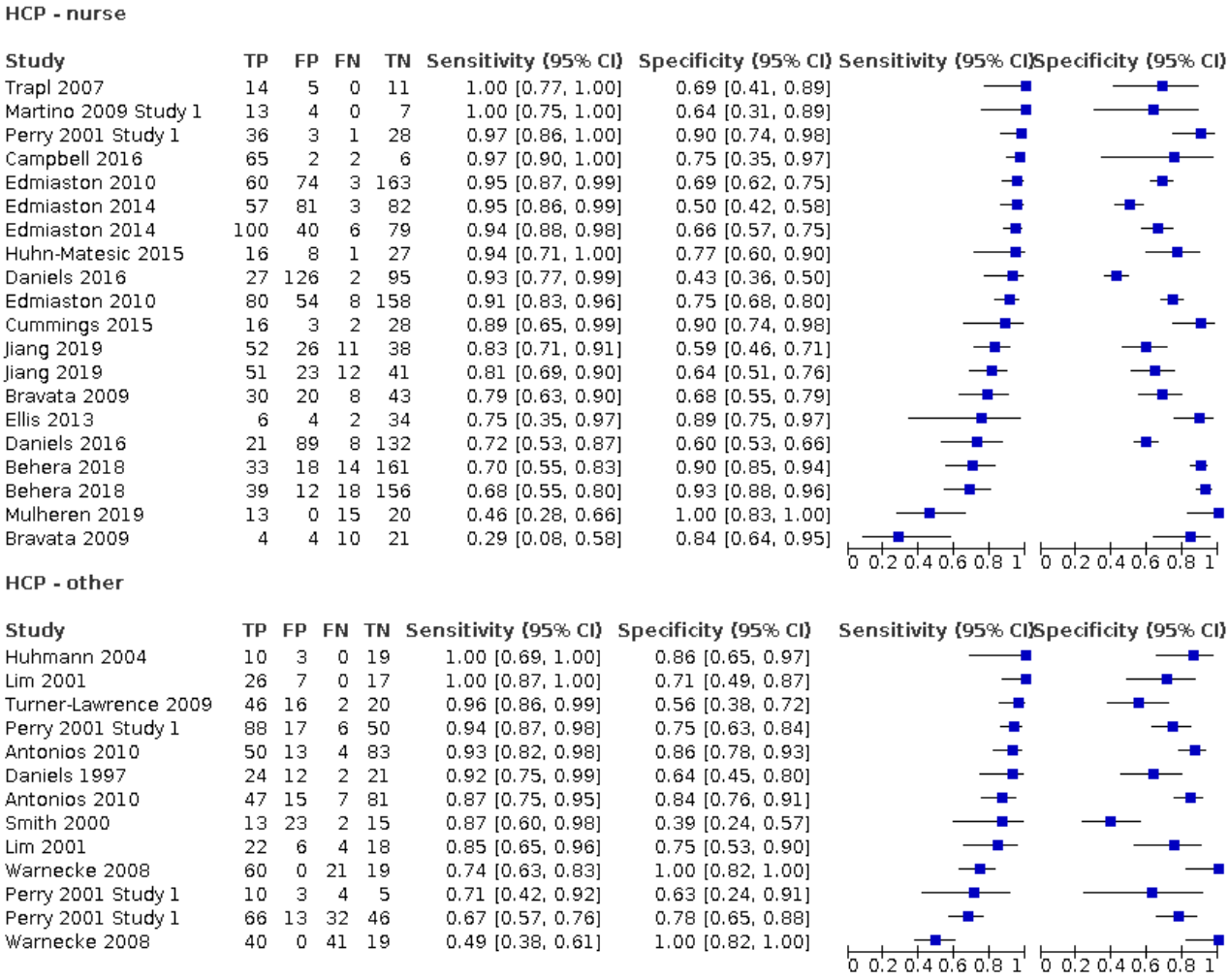
Figure 9. Forest plot of tests grouped by Index Test Outcome - aspiration and dysphagia.



Tests designed for nurses compared to tests designed for other healthcare professionals

Screening tools carried out by nurses (accuracy ranged from sensitivity of 0.29 (95% CI 0.08 to 0.58) with specificity of 0.84 (95% CI 0.64 to 0.95) to sensitivity of 1.00 (95% CI 0.77 to 1.00) with specificity of 0.69 (95% CI 0.41 to 0.89)) performed consistently better than those carried out by other HCPs (excluding SLTs) (accuracy ranged from sensitivity of 0.49 (95% CI 0.38 to 0.61) with specificity of 1.00 (95% CI 0.82 to 1.00) to sensitivity of 1.00 (95% CI 0.69 to 1.00) with specificity of 0.86 (95% CI 0.65 to 0.97)) (Figure 10). Generally, screening tests designed for use by nurses (participants n = 2785) performed equally well as those designed for use by other HCPs (excluding SLTs) (participants n = 1168) in terms of sensitivity and specificity (Figure 4).

Figure 10. Forest plot of tests grouped by HCP - nurse and other.



Other considerations

Screening tests with larger sample sizes (n > 200) were Behera 2018a, Behera 2018b, Daniels 2016a, Daniels 2016b, Edmiaston 2010a, Edmiaston 2010b, Edmiaston 2014a, and Edmiaston 2014b. Of these, screening tests with the greatest accuracy (both high sensitivity and high specificity) were Edmiaston 2010a, Edmiaston 2010b, and Edmiaston 2014b. All of these had narrow 95% CIs for estimates of sensitivity and specificity. Their lower 95% CIs for sensitivity had a higher value than for some of the overall best performing tests - Lim 2001a, Trapl 2007b, and Martino 2009 Study 1. None of the tests with a larger sample size were assessed as low risk for all four QUADAS-2 risk of bias (ROB) domains. Edmiaston 2010a and Edmiaston 2010b were unclear for three domains in the ROB and were low for one domain. Edmiaston 2014b was unclear for one domain in the ROB and was low for the other three domains. In summary, even though 95% CIs for estimates of sensitivity and specificity are narrower for the best performing screening tests, which used a larger sample size, their QUADAS-2 risk of bias is not as low as that for the overall best performing tests.

DISCUSSION

Numerous variables need to be systematically reported regarding training required by the healthcare professional undertaking the swallow screening tool, consistencies offered as part of the tool, and food and drink consistency management options available to the person undertaking the test post screen. From a clinical perspective, some bedside swallow screening tests are criticised for making recommendations for consistencies that have not been assessed. For example, water-only tests allow normal dietary intake without assessment. Similarly water and other consistency screening tests show a discrepancy between diet and fluid consistencies tested and those recommended. Future studies should take these concerns into consideration.

Summary of main results

This review aimed to summarise the evidence regarding accuracy of swallow screening tests to detect aspiration risk associated with dysphagia in acute stroke. Studies did not always report sufficient information regarding study characteristics; therefore, it is not possible to explore the influence of sources of heterogeneity.

Results show that a number of screening tools offered high sensitivity. However, we were unable to identify a swallow screening tool with concomitant high specificity. Furthermore, we were unable to identify a diagnostically accurate swallow screening tool that tests a variety of consistencies and offers immediate management advice for healthcare settings that do not have the facility to offer expert comprehensive assessments.

Population and setting

Although this review focused on acute stroke patients, some studies did not define the time period (from stroke onset or from admission) in which use of swallow screening tool was undertaken. Owing to fluctuation in swallow function within a 24-hour period, it is incumbent on researchers to specify the time period to direct clinical practice (i.e. use of different screening tools at different time periods post stroke) and to allow for comparisons of sensitivity versus other studies. Screening tools that were used in the outpatient setting or in the rehabilitation setting were excluded from the review.

Index tests

The ideal balance of sensitivity and specificity may vary according to the care setting. For example, a very sensitive screening test may be appropriate for untrained staff who have professional backup to rapidly identify false-positives. No screening test identified or excluded swallowing difficulties with perfect accuracy. We were able to identify tests with the best ability to detect people with and without risk of aspiration from studies providing good quality evidence. The best performing swallow screening tools were the Bedside Aspiration test (Lim 2001a), the Gugging Swallowing Screen (GUSS) (Trapl 2007b), and the Toronto Bedside Swallowing Screening Test (TOR-BSST) (Martino 2009 Study 1). Although screening tests with larger sample sizes (n > 200) had smaller confidence intervals (CIs) in estimates of sensitivity and specificity, they had higher risk of bias than the best performing tests so are not among the recommended screening tests. For clinicians with a variety of consistencies available, the best performing water plus other consistencies test was GUSS (Trapl 2007b). Although GUSS offers different consistencies as part of the assessment and offers a consistency management plan for clinicians to implement, the consistencies used are limited to puree, water, and solids, and the test does not include all the consistencies identified in the International Dysphagia Diet Standardised Initiative (International Dysphagia Diet Std Initiative 2017). Therefore, this test may not be suitable for implementation across all healthcare settings. The best water-only swallow screening test was TOR-BSST (Martino 2009 Study 1). Although all tests demonstrate a combined highest sensitivity and specificity and low risk of bias across all domains, clinicians should be cautious in their interpretation of these findings, as these tests are based on single studies with small sample sizes, which limits the interpretation of estimates of reliability for these screening tests.

Sensitivity is similar for tools with a primary outcome of dysphagia or aspiration, but tools with dysphagia as their outcome perform better in terms of specificity. This may be due to the greater number of clinical signs of difficulty (e.g. chewing function) that may be observed in dysphagia compared to aspiration, making it more easily detected.

There is general agreement between index tests that some degree of behavioural assessment (e.g. alertness), together with oromotor assessment, was fundamental before progression to a test of oral intake (either various sizes of water bolus or water plus various consistencies). Generally, water plus various consistencies (n = 6) performed better than water-only tools in terms of sensitivity and specificity. The small number of index tests that did not use direct patient assessment (e.g. used national stroke scales to identify risk of swallowing difficulty) did not consistently perform well.

Only a few studies gave direction on what food and drink consistencies should be given to an individual following the screen (e.g. GUSS; Trapl 2007b). These studies will be useful in healthcare settings, nationally and internationally, when an expert assessment or an instrumental assessment is not available.

The best performing index tests are included in the group carried out by nurses. Index tests carried out by other HCPs are less consistent. This may be due to confounding factors regarding the amount of training received by nursing staff within studies rather than a reflection on the type of index test or professional group undertaking the test. Of the 21 index tests undertaken by nurses, 17 (81.0%) recorded some form of training; for other HCPs, only four (25.0%) papers reported a training element. Furthermore, papers that reported training was required often failed to present the elements included in the training. In future studies, the type of training delivered should be reported.

Reference standard

Reference tests were defined as either expert clinical assessments (usually undertaken by speech and language therapists (SLTs)) or instrumental assessments. A variety of reference tests were included to capture patients who could access different instrumental assessments. However, the reliability of instrumental assessments used to diagnostically identify aspiration is of clinical concern.

Following the SLT opinion, we considered the Mann Assessment of Swallowing Ability (MASA) as equivalent to an expert clinical assessment and included the seven studies that used MASA as a reference test. Tools that used the MASA or expert assessment for reference testing had a better combined sensitivity and specificity than tools that used instrumental tests of fibreoptic endoscopic evaluation of swallowing (FEES) or videofluoroscopy (VF). This may be due to the similarity between index tests and reference tests of MASA/expert assessment in that both rely on clinical observations rather than on instrumental assessment.

From summary receiver operating characteristic (SROC) plots, which plot sensitivity versus (1-specificity), index tests that used FEES as the reference test generally had higher sensitivity but similar specificity than index tests that used VF or MASA as the reference test combined with expert assessment. Most of the studies that used FEES or VF report sensitivity greater than 85% but poorer specificity than those that used MASA or expert assessments as the reference test. No studies were compared to scintigraphy, which remains mainly a research tool.

Quality of the evidence

The quality of evidence varied across included studies. The most common reasons for reduced quality was considerable heterogeneity between studies that affected our ability to combine and directly compare results.

Studies often failed to distinguish between dysphagia and aspiration as the primary outcome. Some did not routinely identify time between stroke onset or admission to hospital and time the screen was undertaken. Similarly, some studies failed to report time between index test and reference test, and the swallow may have improved or deteriorated in the interval between tests. Many studies failed to report the training required by different healthcare professionals to implement the screening tool.

Strengths and weaknesses of the review

Strengths

This DTA undertook the literature search from an International perspective and included all relevant studies irrespective of language of publication. Therefore we can be assured that no relevant publications were excluded from our searches. We were able to identify the best performing tools within each category of swallow screening tools that used water alone, water plus other bolus consistencies, and other methods. We included as narrative studies trials that did not report statistical results.

Weaknesses

Apart from one screening tool that was assessed by two studies, all remaining screening tools were assessed by single studies. This hampered the possibility to perform any meaningful meta-analysis. Only descriptive analyses grouped by general categories were presented. Studies often failed to distinguish between dysphagia and aspiration as the primary outcome. Most studies had small sample sizes, which limits interpretation of the estimates of reliability of the screening tests. Some did not routinely identify time between stroke onset or admission to hospital and time the screen was undertaken. Similarly, some studies failed to report time between the index test and the reference test, and the swallow may have improved or deteriorated in the interval between tests. Many studies failed to report the training required by different healthcare professionals to implement the screening tool. This has had an impact on the quality of this review. We performed no formal assessment of the overall quality of the evidence, for example, by using GRADE. We have not included any ongoing studies.

Applicability of findings to the review question

This review demonstrates that currently, high-quality evidence from research is insufficient to provide conclusive results regarding the accuracy of bedside swallow screening tools in identifying aspiration associated with dysphagia in the acute stage of stroke.

Applicability of evidence

The review authors are aware of ongoing international studies that are currently unpublished and are not included as part of this review. However, this review does identify currently available swallow screening tools that screen with water and those that screen with a variety of consistencies to offer an ongoing food and drink consistency management plan that may be implemented in the clinical setting.

Implications for research

High-quality, appropriately statistically powered studies are needed to evaluate the accuracy of a swallow screening tool in identifying aspiration associated with dysphagia in acute stages of stroke.

Future research studies should clearly define their primary outcomes. They should provide more detail regarding participant inclusion and exclusion criteria, and when a study includes individuals with differing underlying aetiologies that precipitate swallowing difficulties, these data should be reported separately.

Future studies should provide greater detail on participant location and timing of the swallow screening tool used post stroke or post admission. This, together with a description of participant characteristics and swallow severity, would allow researchers to identify the accuracy of a screening tool within different time frames, revealing alterations in swallow function post stroke onset.

Studies should clarify the reference tests used, and when several tests are undertaken, these should be reported separately.

The level of training provided to healthcare professionals undertaking the index test in dysphagia and/or using the index test has been underreported, and studies that detailed the training show variation in both length and content of training. Future research should detail the amount and content of training offered and should examine the impact of training on study fidelity and outcomes.

In addition to technical performance in accurately identifying patients with dysphagia with its associated risk of aspiration, bedside swallow screening tools should demonstrate clinical utility. Evaluation of both short-term and long-term participant outcomes at defined periods would facilitate assessment of whether incidence of aspiration pneumonia, quality of life, and participant mortality were influenced by timing of the screening test post stroke or post admission.

Bedside swallow screening test cost-effectiveness data should be reported to support healthcare services with clinical implementation.

Future studies should apply the better performing screening tests that we have identified, making minor changes in their design to ensure that they have low Quality Assessment of Studies of Diagnostic Accuracy (QUADAS-2) risk of bias and using larger sample sizes, rather than continuing to develop local screening tests. This would build up a body of evidence, would reduce heterogeneity between screening tests, and ultimately would allow a meta-analysis of different tests to be conducted.

Researchers should be encouraged to consider inclusion of these issues to when performing future comparative studies.

Agreements and disagreements with other studies or reviews

This is the largest, most inclusive international and current review on identification of risk of aspiration associated with dysphagia as the main symptom of concern in the acute phase of stroke. Several recent systematic reviews have explored the accuracy of screening tools. However, these studies have limited the scope of their review owing to the restricted range of languages accepted by the search (Brodsky 2016; Daniels 2012; Jiang 2016; Oliveira 2001; Schepp 2012), the heterogeneity of included disorders (Brodsky 2016; Jiang 2016; O'Horo 2015), the paucity of information available regarding the healthcare professional undertaking the study (Brodsky 2016; Chen 2016; Daniels 2012; O'Horo 2015), comparison of only water swallow tools (Brodsky 2016; Chen 2016), exclusion of instrumental assessments from reference tests (Jiang 2016), consideration of

only dysphagia as the primary outcome (Jiang 2016; Oliveira 2001; Schepp 2012), and specification that no training in dysphagia should be required (Schepp 2012). However, this current review had no language restrictions, focuses on acute stages of stroke, identifies the range of healthcare professionals undertaking the test, explores the diagnostic accuracy of tools including water and water with other consistencies, and includes instrumental and expert assessments as reference tests with aspiration associated with dysphagia as the focus of the assessment. Therefore, direct comparisons are difficult to make.

AUTHORS' CONCLUSIONS

Implications for practice

This review demonstrates that current high-quality evidence from research is insufficient to provide conclusive results regarding the accuracy of bedside swallow screening tools to identify aspiration associated with dysphagia in the acute stage of stroke.

Implications for research

We have identified the need for high-quality research. Study authors should be encouraged to include information regarding patient demographics (including comorbidities that may impact outcome and stroke location or classification) and definitions regarding dysphagia and aspiration and time intervals (admission to index test and time from index test to reference test). We recommend that future studies should apply better performing screening tests, enhanced to ensure they have low QUADAS-2 risk of bias domains, and should use larger sample sizes.

ACKNOWLEDGEMENTS

The review authors acknowledge the contributions of Brigit Chesworth, Kerry Hannah, Dawn Doran, Janet Reed, Hazel Fraser, and Josh Cheyne.

REFERENCES

References to studies included in this review

Antonios 2010 *{published data only}*10.1016/j.jstrokecerebrovasdis.2009.03.007

Antonios N, Carnaby-Mann G, Crary M, Miller L, Hubbard H, Hood K, et al. Analysis of a physician tool for evaluating dysphagia on an inpatient stroke unit: the Modified Mann Assessment of Swallowing Ability. *Journal of Stroke and Cerebrovascular Diseases* 2010;**19**(1):49-57.

Behera 2018 *{published data only}*10.1016/j.jstrokecerebrovasdis.2018.02.038

Behera A, Read D, Jackson N, Saour B, Alshekhlee D, Mosier AK. A validated swallow screener for dysphagia and aspiration in patients with stroke. *Journal of Stroke and Cerebrovascular Diseases* 2018;**27**(7):1897-904. [DOI: 10.1016/j.jstrokecerebrovasdis.2018.02.038]

Bravata 2009 *{published data only}*

Bravata DM, Daggett VS, Woodward-Hagg H, Damush T, Plue L, Russell S, et al. Comparison of two approaches to screen for dysphagia among acute ischemic stroke patients: nursing admission screening tool versus National Institutes of Health Stroke Scale. *Journal of Rehabilitation Research and Development* 2009;**46**(9):1127-33.

Campbell 2016 *{published data only}*

Campbell GB, Carter T, Kring D, Martinez C. Nursing bedside dysphagia screen: is it valid? *Journal of Neuroscience Nursing* 2016;**48**(2):75-9.

Cummings 2015 *{published data only}*

Cummings J, Soomans D, O'Laughlin J, Snapp V, Jodoin A, Proco H, et al. Sensitivity and specificity of a nurse dysphagia screen in stroke patients. *MEDSURG Nursing* 2015;**24**(4):219-63.

Daniels 1997 *{published data only}*10.1044/1058-0360.0604.17

Daniels SK, McAdam CP, Brailey K, Foundas AL. Clinical assessment of swallowing and prediction of dysphagia severity. *American Journal of Speech-Language Pathology* 1997;**6**(4):17-24.

Daniels 2016 *{published data only}*

Daniels SK, Pathak S, Rosenbek JC, Morgan RO, Anderson JA. Rapid aspiration screening for suspected stroke: part 1: development and validation. *Archives of Physical Medicine and Rehabilitation* 2016;**97**(9):1440-8.

Edmiaston 2010 *{published data only}*

Edmiaston J, Connor LT, Loehr L, Nassief A. Validation of a dysphagia screening tool in acute stroke patients. *American Journal of Critical Care* 2010;**19**(4):357-64.

Edmiaston 2014 *{published data only}*

Edmiaston J, Connor LT, Steger-May K, Ford AL. A simple bedside stroke dysphagia screen, validated against videofluoroscopy, detects dysphagia and aspiration with high sensitivity. *Journal of Stroke and Cerebrovascular Diseases* 2014;**23**(4):712-6.

Ellis 2013 *{published data only}*

Ellis AL, Hannibal RR. Nursing swallow screens: why is testing water only not enough? *Journal of Neuroscience Nursing* 2013;**45**(5):244-53.

Eren 2019 *{published data only}*10.5152/NSN.2019.10815

Eren Y, Karaca EU, Saylam G, Yaylaci A, Alicura S, Çomoğlu SS. Validity and reliability of the Turkish version of the Barnes-Jewish Hospital stroke dysphagia screen test in patients with acute stroke. *Neurological Sciences and Neurophysiology* 2019;**36**(2):78-83. [DOI: 10.5152/NSN.2019.10815]

Huhmann 2004 *{published data only}*

Huhmann M, Decker RT, Byham-Gray L, Maillet JO, VonHagen S. Comparison of dysphagia screening by a registered dietitian in acute stroke patients to speech language pathologist's evaluation. *Topics in Clinical Nutrition* 2004;**19**(3):239-49.

Huhn-Matesic 2015 *{published data only}*10.1891/1939-2095.8.1.117

Huhn-Matesic EK. Developing and testing a bedside aspiration screen to protect stroke patients from aspiration and mortality. *Clinical Scholars Review* 2015;**8**(1):117-24. [DOI: 10.1891/1939-2095.8.1.117]

Jiang 2019 *{published data only}*10.4103/tcmj.tcmj_158_18

Jiang J, Yu J, Wang J, Wang Y, Wang W. Evaluation of the Chinese version of the swallowing screen in stroke patients with dysphagia. *Tzu Chi Medical Journal* 2019;**31**(4):270-5. [DOI: 10.4103/tcmj.tcmj_158_18]

Lim 2001 *{published data only}*

Lim SHB, Lieu PK, Phua SY, Seshadri R, Venketasubramanian N, Lee SH, et al. Accuracy of bedside clinical methods compared with fiberoptic endoscopic examination of swallowing (FEES) in determining the risk of aspiration in acute stroke patients. *Dysphagia* 2001;**16**(1):1-6.

Martino 2009 Study 1 *{published data only}*10.1161/STROKEAHA.107.510370

Martino R, Silver F, Teasell R, Bayley M, Nicholson G, Streiner DL, et al. The Toronto Bedside Swallowing Screening Test (TOR-BSST) development and validation of a dysphagia screening tool for patients with stroke. *Stroke* 2009;**40**(2):555-61.

Martino 2014 *{published data only}*10.3109/17549507.2014.902995.

Martino R, Maki E, Diamant N. Identification of dysphagia using the Toronto Bedside Swallowing Screening Test (TOR-BSST®): are 10 teaspoons of water necessary? *International Journal of Speech-Language Pathology* 2014;**16**(3):193-8.

Mulheren 2019 *{published data only}*10.1016/j.apmr.2018.12.032

Mulheren RW, González-Fernández M. Swallow screen associated with airway protection and dysphagia after acute stroke. *Archives of Physical Medicine and Rehabilitation* 2019;**100**(7):1289-93. [DOI: 10.1016/j.apmr.2018.12.032]

Nishiwaki 2005 *{published data only}*

Nishiwaki K, Tsuji T, Liu M, Hase K, Tanaka N, Fujiwara T. Identification of a simple screening tool for dysphagia in patients with stroke using factor analysis of multiple dysphagia. *Journal of Rehabilitation Medicine* 2005;**37**:247-51.

Perry 2001 Study 1 *{published data only}*

Perry L. Screening swallowing function of patients with acute stroke. Part one: identification, implementation and initial evaluation of a screening tool for use by nurses. *Journal of Clinical Nursing* 2001;**10**(4):463-73.

Smith 2000 *{published data only}*

Smith HA, Lee SH, O'Neill PA, Connolly MJ. The combination of bedside swallowing assessment and oxygen saturation monitoring of swallowing in acute stroke: a safe and humane screening tool. *Age and Ageing* 2000;**29**(6):495-9.

Trapl 2007 *{published data only}*10.1161/STROKEAHA.107.483933

Trapl M, Enderle P, Nowotny M, Teuschl Y, Matz K, Dachenhausen A, et al. Dysphagia bedside screening for acute-stroke patients: the Gugging Swallowing Screen. *Stroke* 2007;**38**(11):2948-52.

Turner-Lawrence 2009 *{published data only}*10.1016/j.annemergmed.2009.03.007

Turner-Lawrence DE, Peebles M, Price MF, Singh SJ, Asimos AW. A feasibility study of the Sensitivity of Emergency Physician Dysphagia Screening in acute stroke patients. *Annals of Emergency Medicine* 2009;**54**(3):344-8.e1.

Warnecke 2008 *{published data only}*10.1136/jnnp.2007.134551

Warnecke T, Teismann I, Meimann W, Olenberg S, Zimmermann J, Kramer C, et al. Assessment of aspiration risk in acute ischaemic stroke - evaluation of the simple swallowing provocation test. *Journal of Neurology, Neurosurgery and Psychiatry* 2008;**79**(3):312-4.

Zhou 2011 *{published data only}*

Zhou Z, Salle JY, Daviet JC, Stuit A, Nguyen CL. Combined approach in bedside assessment of aspiration risk post stroke: PASS. *European Journal of Physical and Rehabilitation Medicine* 2011;**47**(3):441-6.

References to studies excluded from this review

Addington 1999 Study 1 *{published data only}*

Addington WR, Stephens RE, Gilliland K, Rodriguez M. Assessing the laryngeal cough reflex and the risk of developing pneumonia after stroke. *Archives of Physical Medicine and Rehabilitation* 1999;**80**(2):150-4.

Addington 1999 Study 2 *{published data only}*

Addington WR, Stephens RE, Gilliland KA. Assessing the laryngeal cough reflex and the risk of developing pneumonia after stroke: an interhospital comparison. *Stroke* 1999;**30**(6):1203-7.

Alsibai 2014 *{published data only}*

Alsibai A, Voytas J. Use of electrical stimulation in treatment of dysphagia, a nursing home experience. *Journal of the American Geriatrics Society* 2014;**62**:S37.

Anderson 2016 *{published data only}*

Anderson JA, Pathak S, Rosenbek JC, Morgan RO, Daniels SK. Rapid aspiration screening for suspected stroke: Part 2: initial and sustained nurse accuracy and reliability. *Archives of Physical Medicine and Rehabilitation* 2016;**97**(9):1449-55.

Archer 2015 Study 1 *{published data only}*

Archer S, Smith C, Newham D. Surface electromyography biofeedback in dysphagia therapy: do patients think it helps? *International Journal of Stroke* 2015;**10**:63.

Archer 2015 Study 2 *{published data only}*

Archer S, Smith C, Newham D. Surface electromyographic biofeedback and the effortful swallow exercise for stroke-related dysphagia. *International Journal of Stroke* 2015;**10**:63.

Armesto 2015 *{published data only}*

Armesto A, Howald R, Roseman E, Kitchen-Clark T, Convery C, Boyle K, et al. Swallowing screening in the emergency department: a collaborative effort to improve patient care. *International Journal of Stroke* 2015;**10**:59-60.

Aviv 1997 *{published data only}*

Aviv JE, Sacco RL, Mohr JP, Thompson JLP, Levin B, Sunshine S, et al. Laryngopharyngeal sensory testing with modified barium swallow as predictors of aspiration pneumonia after stroke. *Laryngoscope* 1997;**107**(9):1254-60.

Aydogdu 2015 *{published data only}*

Aydogdu I, Kiylioglu N, Tarlaci S, Tanriverdi Z, Alpaydin S, Acarer A, et al. Diagnostic value of "dysphagia limit" for neurogenic dysphagia: 17 years of experience in 1278 adults. *Clinical Neurophysiology* 2015;**126**(3):634-43.

Babi 2014 *{published data only}*

Babi MA, Gorman M, Commichau C, Kenney S, Boyle L, Comeau M, et al. Failed bedside dysphagia screen predicts pneumonia treatment in acute stroke: a resident quality initiative project in a JC-certified primary stroke center. Neurology 2014;**82**(10 Suppl 1).

Bahia 2014 *{published data only}*

Bahia MM, Mourao LF, Lima FO, Li LM. Clinical predictors of oropharyngeal dysphagia following acute stroke. *Dysphagia* 2014;**29**(6):774.

Bahia 2016 *{published data only}*

Bahia MM, Mourao LF, Chun RYS. Dysarthria as a predictor of dysphagia following stroke. *NeuroRehabilitation* 2016;**38**(2):155-62.

Bax 2013 *{published data only}*

Bax L, McFarlane M, Green E, Miles A. Introducing a speech and language therapy (SLT) led fibre optic endoscopic evaluation of swallowing (FEES): functional outcomes for acute stroke patients. *International Journal of Stroke* 2013;**8**:40.

Bax 2014 *{published data only}*

Bax L, McFarlane M, Green E, Miles A. Speech-language pathologist-led fiberoptic endoscopic evaluation of swallowing: functional outcomes for patients after stroke. *Journal of Stroke and Cerebrovascular Diseases* 2014;**23**(3):e195-e200.

Benjamin 2015 *{published data only}*

Bax L, McFarlane M, Green E, Miles A. Introducing a speech and language therapy (SLT) led fibre optic endoscopic evaluation of swallowing (FEES): functional outcomes for acute stroke patients. *International Journal of Stroke* 2013;**8**:40.

Burks 2011 *{published data only}*

Burks C, Meyer DM. Validation of the TOR-BSST swallow screening tool. *Stroke* 2011;**42**(3):e129-e130.

Carnaby 2014 *{published data only}*

Carnaby GD. Relationships among videofluoroscopic scales and stroke outcome. *Dysphagia* 2014;**29**(6):751.

Carrington 2013 *{published data only}*

Carrington S, Archer S, Smith C. Evaluating the impact of fibreoptic endoscopic (FEES) evaluation of swallowing with dysphagic patients on a stroke unit. *International Journal of Stroke* 2013;**8**:27.

Chavarria 2015 *{published data only}*

Chavarria M, Arauz A, Barboza M, Patino M, Becerril M. Scale for severe dysphagia in lateral medullary stroke. *Journal of the Neurological Sciences* 2015;**357**:e108.

Choi 2010 *{published data only}*

Choi MG, Sung HY, Kim JS, Cho YK, Park JM, Lee IS, et al. Usefulness of dysphagia screening protocol including high resolution manometry in acute cerebrovascular accident. *Gastroenterology* 2010;**138**(5 Suppl 1):S346.

Chong 2003 *{published data only}*

Chong MS, Lieu PK, Sitoh YY, Meng YY, Leow LP. Bedside clinical methods useful as screening test for aspiration in elderly patients with recent and previous strokes. *Annals of the Academy of Medicine Singapore* 2003;**32**(6):790-4.

Cichero 2009 *{published data only}*

Cichero JA, Heaton S, Bassett L. Triaging dysphagia: nurse screening for dysphagia in an acute hospital. *Journal of Clinical Nursing* 2009;**18**(11):1649-59.

Clayton 2006 *{published data only}*

Clayton J, Jack CIA, Ryall C, Tran J, Hilal E, Gosney M. Tracheal pH monitoring and aspiration in acute stroke. *Age and Ageing* 2006;**35**(1):47-53.

Collins 1997 *{published data only}*

Collins MJ, Bakheit AMO. Does pulse oximetry reliably detect aspiration in dysphagic stroke patients? *Stroke* 1997;**28**(9):1773-5.

Crary 2013 Study 1 *{published data only}*

Crary M, Carnaby G, Sura L, Sia I, Rolon S, Levien E, et al. Spontaneous swallow frequency identifies dysphagia in acute stroke. *Dysphagia* 2013;**28**(4):612.

Crary 2013 Study 2 *{published data only}*

Crary M, Carnaby-Mann G, Sura L, Sia I. Spontaneous swallow frequency as a screening protocol for dysphagia in acute stroke. Neurology 2013;**80**(1 Meeting Abstracts).

Crary 2014 Study 1 *{published data only}*

Crary M, Carnaby G. Multidisciplinary clinical rehabilitation spontaneous swallow frequency and dysphagia/stroke related outcomes in acute stroke. *International Journal of Stroke* 2014;**9**:220.

Crary 2014 Study 2 *{published data only}*

Crary MA, Carnaby GD, Sia I. Spontaneous swallow frequency compared with clinical screening in the identification of dysphagia in acute stroke. *Journal of Stroke and Cerebrovascular Diseases* 2014;**23**(8):2047-53.

Daniels 2009 *{published data only}*

Daniels SK, Schroeder MF, DeGeorge PC, Corey DM, Foundas AL, Rosenbek JC. Defining and measuring dysphagia following stroke. *American Journal of Speech-Language Pathology* 2009;**18**(1):74-81.

Davis-De Geus 2009 *{published data only}*

Davis-De Geus M. Using an evidence-based bedside swallow screen for all stroke patients: how health care providers can improve stroke patients' outcomes. *Dysphagia* 2009;**24**(4):463.

DePippo 1992 *{published data only}*

DePippo KL, Holas MA, Reding MJ. Validation of the 3-oz water swallow test for aspiration following stroke. *Archives of Neurology* 1992;**49**(12):1259-61.

Farneti 2015 *{published data only}*

Farneti D, Prencipe RS, Montano L, Nicolini L, Genovese E. Reliability of screening and bedside procedures in the management of acute stroke patients in a community hospital. *Dysphagia* 2015;**30**(5):626.

Giraldo-Cadavid 2015 *{published data only}*

Giraldo-Cadavid LF, Burguete J, Rueda F, Galvis AM, Castaneda N, Arbulu M, et al. Accuracy of the laryngeal adductor reflex explored by a new laryngo-pharyngeal esthesiometer during the fiberoptic endoscopic evaluation of swallowing. *American Journal of Respiratory and Critical Care Medicine* 2015;**191**:A2123.

Guerra 2011 *{published data only}*

Guerra L, Ottaviani C, Grazia Scucchi M, Amendola L, Macchiaolo S, Capecchi D, et al. Assessment of dysphagia in acute stroke: analysis of degree of agreement between nurses. *L'Infermiere* 2011;**48**(2):e9-13.

Guillén-Solà 2011 *{published data only}*

Guillén-Solà A, Martínez-Orfila J, Gómez RB, Castelló SM, Marco E. Screening of swallowing disorders in stroke: utility

of clinical signs and clinical examination method of volume-viscosity in comparison with videofluoroscopy. *Rehabilitacion* 2011;**45**(4):293-300.

Guillén-Solà 2013 *{published data only}*

Guillen-Sola A, Marco E, Martinez-Orfila J, Donaire Mejias MF, Depolo Passalacqua M, Duarte E, et al. Usefulness of the volume-viscosity swallow test for screening dysphagia in subacute stroke patients in rehabilitation income. *NeuroRehabilitation* 2013;**33**(4):631-8.

Gurcay 2018 *{published data only}*

Gurcay E, Bahceci K, Ozturk E, Yilmaz V, Gundogdu I, Ceylan T, et al. Validity and reliability of Turkish version of the gugging swallowing screen test in the early period of hemispheric stroke. *Neurological Sciences and Neurophysiology* 2018;**35**(1):6-13.

Higo 2003 *{published data only}*

Higo R, Tayama N, Watanabe T, Nito T. Pulse oximetry monitoring for the evaluation of swallowing function. *European Archives of Oto-Rhino-Laryngology* 2003;**260**(3):124-7.

Joundi 2017 *{published data only}*

Joundi RA, Martino R, Saposnik G, Giannakeas V, Jiming F, Kapral MK, et al. Predictors and outcomes of dysphagia screening after acute ischemic stroke. *Stroke* 2017;**48**(4):900-6.

Karaca 2018 *{published data only}*

Karaca Umay E, Gundogdu I, Gurcay E, Ozturk E, Yilmaz V, Karaahmet O, et al. The psychometric evaluation of the Turkish version of the Mann assessment of swallowing ability in patients in the early period after stroke. *Turkish Journal of Medical Sciences* 2018;**48**(6):1153-61.

Logemann 1999 *{published data only}*

Logemann JA, Veis S, Colangelo L. A screening procedure for oropharyngeal dysphagia. *Dysphagia* 1999;**14**(1):44-51.

Mandysova 2015 *{published data only}*

Mandysova P, Ehler E, Skvrnakova J, Cerny M, Bartova I, Pellant A. Development of the Brief Bedside Dysphagia Screening Test - Revised: a cross-sectional Czech study. *Acta Medica (Hradec Kralove)* 2015;**58**:49-55.

Marques 2008 *{published data only}*

Marques CHD, De Rosso ALZ, Andre C. Bedside assessment of swallowing in stroke: water tests are not enough. *Topics in Stroke Rehabilitation* 2008;**15**(4):378-83.

Martino 2009 Study 2 *{published data only}*

Martino R, Streiner D, Maki E, Diamant N. A sensitivity analysis to determine whether ten teaspoons of water are really necessary. *Dysphagia* 2009;**24**(4):473.

Massey 2002 *{published data only}*

Massey R, Jedlicka D. The Massey Bedside Swallowing Screen. *Journal of Neuroscience Nursing* 2002;**34**(5):252-60.

McCullough 2001 *{published data only}*

McCullough GH, Wertz RT, Rosenbek JC. Sensitivity and specificity of clinical/bedside examination signs for detecting aspiration in adults subsequent to stroke. *Journal of Communication Disorders* 2001;**34**(1-2):55-72.

McCullough 2005 *{published data only}*

McCullough GH, Rosenbek JC, Wertz RT, McCoy S, Mann G, McCullough K. Utility of clinical swallowing examination measures for detecting aspiration post-stroke. *Journal of Speech, Language, and Hearing Research* 2005;**48**(6):1280-93.

Moalli 2016 *{published data only}*

Moalli D, Dell'Angelo D, Bedard L. CVAS: a rapid dysphagia evaluation tool. *Connecticut Medicine* 2016;**80**(4):205-7.

Moon 2010 *{published data only}*

Moon KH, Sohn HS, Lee ES, Paek EK, Kang EJ, Lee SH, et al. Comparison for risk estimate of aspiration between the revised dysphagia assessment tool and videofluoroscopy in post-stroke patients. *Journal of Korean Academy of Nursing* 2010;**40**(3):359-66.

Mozzanica 2017 *{published data only}*

Mozzanica F, Scarponi L, Pedrali S, Pizzorni N, Pinotti C, Foieni F, et al. Dysphagia screening in subacute care settings using the Italian version of the Royal Brisbane and Women's Hospital (I-RBWH) dysphagia screening tool [Screening della disfagia nelle unita di cure sub-acute utilizzando la versione italiana del Royal Brisbane and Women's Hospital (I-RBWH) dysphagia screening tool]. *Acta Otorhinolaryngologica Italica* 2017;**37**(1):25-31.

Murray 2011 *{published data only}*

Murray J, Milich A, Ormerod D. Screening for dysphagia. *Australian Nursing Journal* 2011;**18**(11):44-6.

NCT00306501 *{published data only}*

NCT00306501. Volitional swallowing in stroke patients with chronic dysphagia. https://clinicaltrials.gov/ct2/show/record/NCT00306501 (first received 23 March 2006).

NCT00580138 *{published data only}*

NCT00580138. Real-time Internet evaluation of swallowing. https://clinicaltrials.gov/ct2/show/NCT00580138 (first received 24 December 2007).

NCT01765673 *{published data only}*

NCT01765673. The Passy Muir Swallowing Self Training Device (PMSST). https://clinicaltrials.gov/ct2/show/NCT01765673 (first received 10 January 2013).

NCT02080988 *{published data only}*

NCT02080988. Reflexive coughing force in severe aspirators. https://clinicaltrials.gov/ct2/show/NCT02080988 (first received 7 March 2014).

NCT02848664 *{published data only}*

NCT02848664. Use of the Passy Muir Swallowing Self Trainer. https://clinicaltrials.gov/ct2/show/NCT02848664 (first received 28 July 2016).

NCT03167892 *{published data only}*

NCT03167892. Oral Screens in Post Stroke Training. https://clinicaltrials.gov/ct2/show/NCT03167892 (first received 30 May 2017).

Nomura 2014 *{published data only}*

Nomura JT, Poole CL, Boyer BW, Papas MA, Thomas J, Gregore DR, et al. Customized swallow screening tool to exclude aspiration pneumonia risk in acute stroke patients. *Academic Emergency Medicine* 2014;**21**(5 Suppl 1):S257.

Oh 2016 *{published data only}*

Oh JC, Park JH, Jung MY, Yoo EY, Chang KY, Lee TY. Relationship between quantified instrumental swallowing examination and comprehensive clinical swallowing examination. *Occupational Therapy International* 2016;**23**(1):3-10.

Ohira 2013 *{published data only}*

Ohira M, Ishida R, Maki Y, Okubo M, Sugiyama T, Sakayori T, et al. Evaluation of swallowing ability using Mann assessment of swallowing ability (MASA) of dependent elderly in Japan. *Dysphagia* 2013;**28**(4):635.

Ohira 2017 *{published data only}*

Ohira M, Ishida R, Maki Y, Ohkubo M, Sugiyama T, Sakayori T, Sato T. Evaluation of a dysphagia screening system based on the Mann Assessment of Swallowing Ability for use in dependent older adults. *Geriatrics & Gerontology International* 2017;**17**(4):561-7.

Osawa 2013 *{published data only}*

Osawa A, Maeshima S, Tanahashi N. Water-swallowing test: screening for aspiration in stroke patients. *Cerebrovascular Diseases* 2013;**35**(3):276-81.

Ostrofsky 2016 *{published data only}*

Ostrofsky C, Seedat J. The South African dysphagia screening tool (SADS): a screening tool for a developing context. *South African Journal of Communication Disorders* 2016;**63**(1):1-9.

Palli 2017 *{published data only}*

Palli C, Fandler S, Doppelhofer K, Niederkorn K, Enzinger C, Vetta C, et al. Early dysphagia screening by trained nurses reduces pneumonia rate in stroke patients: a clinical intervention study. *Stroke* 2017;**48**(9):2583-5.

Pennsylvania Patient Safety Authority 2009 *{published data only}*

Pennsylvania Patient Safety Authority. Does your admission screening adequately predict aspiration risk? *Pennsylvania Patient Safety Authority* 2009;**6**(4):115-21. [http://patientsafety.pa.gov/ADVISORIES/Pages/200912_115.aspx]

Perry 2000 *{published data only}*

Perry L, McLaren SM. Dysphagia screening, assessment and outcomes in stroke patients: Phase 1 of an evidence-based practice project. *Journal of Clinical Excellence* 2000;**1**(4):201-8.

Perry 2001 Study 2 *{published data only}*

Perry L. Screening swallowing function of patients with acute stroke. Part two: detailed evaluation of the tool used by nurses. *Journal of Clinical Nursing* 2001;**10**(4):474-81.

Radhakrishnan 2013 *{published data only}*

Radhakrishnan S, Menon UK, Anandakuttan A. A combined approach of bedside clinical examination and flexible endoscopic evaluation of swallowing in poststroke dysphagia: a pilot study. *Annals of Indian Academy of Neurology* 2013;**16**(3):388-93.

Rosenbek 2004 *{published data only}*

Rosenbek JC, McCullough GH, Wertz RT. Is the information about a test important? Applying the methods of evidence-based medicine to the clinical examination of swallowing. *Journal of Communication Disorders* 2004;**37**(5):437-50.

Sato 2012 *{published data only}*

Sato M, Tohara H, Iida T, Wada S, Inoue M, Ueda K. Simplified cough test for screening silent aspiration. *Archives of Physical Medicine and Rehabilitation* 2012;**93**(11):1982-6.

Schrock 2011 *{published data only}*

Schrock JW, Bernstein J, Glasenapp M, Drogell K, Hanna J. A novel emergency department dysphagia screen for patients presenting with acute stroke. *Academic Emergency Medicine* 2011;**18**(6):584-9.

Sellars 1998 *{published data only}*

Sellars C, Dunnet C, Carter R. A preliminary comparison of videofluoroscopy of swallow and pulse oximetry in the identification of aspiration in dysphagic patients. *Dysphagia* 1998;**13**(2):82-6.

Smith Hammond 2009 *{published data only}*

Smith Hammond CA, Goldstein LB, Horner RD, Ying J, Gray L, Gonzalez-Rothi L, et al. Predicting aspiration in patients with ischemic stroke: comparison of clinical signs and aerodynamic measures of voluntary cough. *Chest* 2009;**135**(3):769-77.

Sohn 2018 *{published data only}*

Sohn D, Park G-Y, Koo HJ, Jang YJ, Han Y, Im S. Determining peak cough flow cutoff values to predict aspiration pneumonia among patients with dysphagia using the Citric Acid Reflexive Cough Test. *Archives of Physical Medicine & Rehabilitation* 2018;**99**(12):2532-9.

Soria 2013 *{published data only}*

Soria F, Furkim AM, Nunes MC, Pinto G, Lange M, Zetola V. Identification risk of aspiration pneumonia in two groups of patients with stroke. *Dysphagia* 2013;**28**(2):325-6.

Stroke and Chewing 2012 *{published data only}*

Stroke and chewing. Dental Abstracts 2012;**57**(3):165.

Sung 2018 *{published data only}*

Sung HY, Lee KS, Choi MG, Kim JS. Dysphagia screening using high-resolution impedance manometry in acute stroke. *International Journal of Gerontology* 2018;**12**(2):110-5.

Sørensen 2013 *{published data only}*

Sørensen RT, Rasmussen RS, Overgaard K, Lerche A, Johansen AM, Lindhardt T. Dysphagia screening and intensified oral hygiene reduce pneumonia after stroke. *Journal of Neuroscience Nursing* 2013;**45**(3):139-46.

Tanuma 2002 *{published data only}*

Tanuma S, Sudo E, Takahashi Y, Yoshida A, Kobayashi C, Ohama Y. The usefulness of the water swallowing test and videofluorography in swallowing rehabilitation in patients with cerebrovascular disease. *Nippon Ronen Igakkai Zasshi Japanese Journal of Geriatrics* 2002;**39**(4):427-32.

Teramoto 2000 *{published data only}*

Teramoto S, Fukuchi Y. Detection of aspiration and swallowing disorder in older stroke patients: simple swallowing provocation test versus water swallowing test. *Archives of Physical Medicine and Rehabilitation* 2000;**81**(11):1517-9.

Toscano 2019 *{published data only}*

Toscano M, Vigano A, Rea A, Verzina A, Sasso D'Elia T, Puledda F, et al. Sapienza Global Bedside Evaluation of Swallowing after Stroke: the GLOBE-3S study. *European Journal of Neurology* 2019;**26**(4):596-602.

Tuncay 2011 *{published data only}*

Tuncay F, Tasbas Ö, Borman P, Geçene M, Coskun Ö. The bedside clinical evaluation of swallowing function of stroke patients in acute stage. *Journal of Physical Medicine & Rehabilitation Sciences* 2011;**14**(2):33-8.

Umay 2013 *{published data only}*

Umay E, Unlu E, Kilic G, Cakci A, Korkmaz H. Evaluation of dysphagia in early period stroke patients by bedside screening test, endoscopic and electrophysiological methods. *Dysphagia* 2013;**28**(3):395-403.

Virvidaki 2019 *{published data only}*

Virvidaki IE, Giannopoulos S, Nasios G, Dimakopoulos G, Michou E, Milionis H. Predictive value of a novel pragmatic tool for post-stroke aspiration risk: the Functional Bedside Aspiration Screen. *Neurogastroenterology and Motility* 2019;**31**(10):e13683.

Warnecke 2017 *{published data only}*

Warnecke T, Im S, Kaiser C, Hamacher C, Oelenberg S, Dziewas R. Aspiration and dysphagia screening in acute stroke - the Gugging Swallowing Screen revisited. *European Journal of Neurology* 2017;**24**(4):594-601.

Weinhardt 2008 *{published data only}*

Weinhardt J, Hazelett S, Barrett D, Lada R, Enos T, Keleman R. Accuracy of a bedside dysphagia screening: a comparison of registered nurses and speech therapists. *Rehabilitation Nursing* 2008;**33**(6):247-52.

Ye 2018 *{published data only}*

Ye T, Huang SY, Dong Y, Dong Q. Bedside modified volume-viscosity swallow test might be reliably to predict risk of aspiration pneumonia after stroke. *Stroke* 2019;**50 Suppl 1**:AWP479.

Yeh 2011 *{published data only}*

Yeh SJ, Huang KY, Wang TG, Chen YC, Chen CH, Tang SC, et al. Dysphagia screening decreases pneumonia in acute stroke patients admitted to the stroke intensive care unit. *Journal of the Neurological Sciences* 2011;**306**(1-2):38-41.

Zhang 2004 *{published data only}*

Zhang J, Wang YJ, Cui T. Reliability and validity of nine rating scales for dysphagia following stroke. *Chinese Journal of Clinical Rehabilitation* 2004;**8**(7):1201-3.

Additional references

Almeida 2015

Almeida TM, Cola PC, Pernambuco LA, Magalhaes HV Jr, Silva RG. Screening tools for oropharyngeal dysphagia in stroke. *Audiology - Communication Research* 2015;**20**(4):361-70.

Avan 2019

Avan A, Digaleh H, Di Napoli M, Stranges S, Behrouz R, Shojaeianbabaei G, et al. Socioeconomic status and stroke incidence, prevalence, mortality, and worldwide burden: an ecological analysis from the Global Burden of Disease Study. BMC Medicine 2019;**17**(191). [DOI: 10.1186/s12916-019-1397-3]

Benfield 2020

Benfield JK, Everton LF, Bath PM, England TJ. Accuracy and clinical utility of comprehensive dysphagia screening assessments in acute stroke: a systematic review and meta-analysis. *Journal of Clinical Nursing* 2020;**29**(9-10):1527-38. [DOI: 10.1111/jocn.15192]

Blitzer 1988

Blitzer A, Krespi YP, Oppenheimer RW, Levine TM. Surgical management of aspiration. *Otolaryngologic Clinics of North America* 1988;**4**:743–50.

Boaden 2011

Boaden EE. Improving the identification and management of aspiration after stroke [Doctoral thesis]. University of Central Lancashire 2011.

Bray 2016

Bray BD, Smith CJ, Cloud GC, Enderby P, James M, Paley L, et al. The association between delays in screening for and assessing dysphagia after acute stroke, and the risk of stroke-associated pneumonia. *Journal of Neurology, Neurosurgery and Psychiatry* 2017;**88**(1):25-30.

Brodsky 2016

Brodsky MB, Suiter DM, Gonzalez-Fernandez M, Michtalik HJ, Frymark TB, Venediktov R, et al. Screening accuracy for aspiration using bedside water swallow tests: a systematic review and meta-analysis. *Chest* 2016;**150**(1):148-63.

CADTH 2018

CADTH. Grey matters: a practical tool for searching health-related grey literature [Internet]. Available from https://www.cadth.ca/resources/finding-evidence 2018.

Chen 2016

Chen PC, Chuang CH, Leong CP, Guo SE, Hsin YJ. Systematic review and meta-analysis of the diagnostic accuracy of the water swallow test for screening aspiration in stroke patients. *Journal of Advanced Nursing* 2016;**72**(11):2575-86. [DOI: 10.1111/jan.13013]

Daniels 2012

Daniels SK, Anderson JA, Willson PC. Valid items for screening dysphagia risk in patients with stroke. *Stroke* 2012;**43**(3):892-7.

Donovan 2013

Donovan NJ, Daniels SK, Edmiaston J, Weinhardt J, Summers D, Mitchell PH. Dysphagia screening: state of the art: invitational conference proceeding from the State-of-the-Art Nursing Symposium, International Stroke Conference 2012. *Stroke* 2013;**44**(4):e24-31.

European Stroke Organisation 2015

European Stroke Organisation. The European Stroke Organisation Guidelines: a standard operating procedure. *International Journal of Stroke* 2015;**10**(SA100):128-35. [DOI: 10.1111/ijs.12583]

Gupta 2016

Gupta D, Hull ML, Fraser I, Miller L, Bossuyt PMM, Johnson N, et al. Endometrial biomarkers for the non-invasive diagnosis ofendometriosis. *Cochrane Database of Systematic Reviews* 2016, Issue 4. Art. No: CD012165. [DOI: 10.1002/14651858.CD012165]

Intercollegiate Stroke Working Party 2012

Intercollegiate Stroke Working Party. National Sentinel Stroke Clinical Audit 2010. Royal College of Physicians, 2012.

Intercollegiate Stroke Working Party 2016

Intercollegiate Stroke Working Party. Intercollegiate Stroke Working Party National Clinical Guideline for Stroke. 5th edition. London: Royal College of Physicians, 2016.

International Dysphagia Diet Std Initiative 2017

International Dysphagia Diet Standardisation Initiative. Complete IDDSI Framework Detailed definitions. 2017. https://iddsi.org/framework/. [http://iddsi.org/framework/]

Jiang 2016

Jiang JL, Fu SY, Wang WH, Ma YC. Validity and reliability of swallowing screening tools used by nurses for dysphagia: a systematic review. *Ci Ji Yi Xue Za Zhi* 2016;**28**(2):41-8. [DOI: 10.1016/j.tcmj.2016.04.006]

Johnson 2019

Johnson CO, Nguyen M, Roth GA, Nichols E, Alam T, Abate D, et al. Global, regional, and national burden of stroke, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet Neurology* 2019;**18**(5):439-58.

Katan 2018

Katan M, Luft A. Global burden of stroke. *Seminars in Neurology* 2018;**38**(2):208-11.

Kumar 2010

Kumar S, Selim MH, Caplan LR. Medical complications after stroke. *Lancet Neurology* 2010;**9**:105-18.

Martino 2005

Martino R, Foley N, Bhogal S, Diamant N, Speechley M, Teasell R. Dysphagia after stroke: incidence, diagnosis, and pulmonary complications. *Stroke* 2005;**36**(12):2756-63.

National Stroke Foundation 2010

National Stroke Foundation. Clinical guidelines for stroke management, 2010. https://extranet.who.int/ncdccs/Data/AUS_D1_Clinical%20Guidelines%20for%20Stroke%20Management.pdf.

O'Horo 2015

O'Horo JC, Rogus-Pulia N, Garcia-Arguello L, Robbins J, Safdar N. Bedside diagnosis of dysphagia: a systematic review. *Journal of Hospital Medicine* 2015;**10**(4):256-65.

Oliveira 2001

Oliveira I, Mota L, Freitas S, Ferreira P. Dysphagia screening tools for acute stroke patients available for nurses: a systematic review. *Nursing Practice Today* 2019;**6**(3):103-15. [DOI: 10.18502/npt.v6i3.1253]

Powers 2019

Powers WJ, Rabinstein AA, Ackerson T, Adeoye OM, Bambakidis NC, Becker K, et al, on behalf of the American Heart Association Stroke Council. Guidelines for the early management of patients with acute ischemic stroke: 2019 update to the 2018 guidelines for the early management of acute ischemic stroke: a guideline for healthcare professionals from the American Heart Association/American Stroke Association. *Stroke* 2019;**50**:e344–e418. [DOI: 10.1161/STR.0000000000000211]

Review Manager 2014 [Computer program]

The Nordic Cochrane Centre, the Cochrane Collaboration Review Manager 5 (RevMan 5). Version 5.3. Copenhagen: The Nordic Cochrane Centre, the Cochrane Collaboration, 2014.

Rofes 2011

Rofes L, Arreola V, Almirall J. Diagnosis and management of oropharyngeal dysphagia and its nutritional and respiratory complications in the elderly. *Gastroenterology Research and Practice* 2011;**2011**:13.

Schepp 2012

Schepp K, Tirschwell DL, Miller RM, Longstreth WT Jr. Swallowing screens after acute stroke: a systematic review. *Stroke* 2012;**43**(3):869-71.

Teasell 2019

Teasell R, Salbach NM, Acerra N, Bastasi D, Carter SL, Fung J, et al, on behalf of the Rehabilitation and Recovery following Stroke Writing Group. Assessment and Management of Dysphagia and Malnutrition following Stroke. In Lindsay MP, Mountain A, Gubitz G, Dowlatshahi D, Casaubon L, de Jong A, Smith EE (Editors); on behalf of the Canadian Stroke Best

Practices Advisory Committee. Canadian Stroke Best Practice Recommendations 2019:71-5.

WHF 2016

World Heart Federation. The global burden of stroke, 2016. www.world-heart-federation.org/cardiovascular-health/stroke/.

References to other published versions of this review

Boaden 2017

Boaden E, Doran D, Burnell J, Clegg A, Dey P, Hurley M, et al. Screening for aspiration risk associated with dysphagia in acute stroke. *Cochrane Database of Systematic Reviews* 2017, Issue 6. Art. No: CD012679. [DOI: 10.1002/14651858.CD012679]

CHARACTERISTICS OF STUDIES

Characteristics of included studies *[ordered by study ID]*

Antonios 2010

Study characteristics	
Patient Sampling	Consecutive
Patient characteristics and setting	Inclusion criteria: AIS, diagnosed clinically by a neurologist; radiographic (CT or MRI) confirmation of AIS Exclusion criteria: unresolved dysphagia before stroke; history of head/neck surgery or trauma that may impact swallowing ability; any other concomitant neurological disorder that could impact oropharyngeal swallowing ability Setting: 1 × tertiary care academic medical centre with a certified comprehensive stroke centre
Index tests	Modified MASA (MMASA)
Target condition and reference standard(s)	Dysphagia MASA
Flow and timing	All patients received the index test(s) and reference standard. All patients were included in the 2 × 2 table Interventions between index/reference tests: not reported Interval between tests ≤ 4 hours
Comparative	
Notes	

Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Was a two-gate design avoided?	Yes		

Antonios 2010 *(Continued)*

Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (All tests)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	No		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Behera 2018

Study characteristics	
Patient Sampling	All patients were admitted under a standard stroke protocol
Patient characteristics and setting	Inclusion criteria: patients admitted to a comprehensive stroke centre Exclusion criteria: unclear Setting: 1 × hospital designated comprehensive stroke centre

Behera 2018 *(Continued)*

Index tests	The DePaul Hospital Swallow Screening tool includes a combination of non-swallow and swallow items. The 8 non-swallow items (including tests of alertness, presence of a feeding tube, presence of a tracheostomy tube, presence of drooling, facial asymmetry, abnormal tongue movement, abnormal vocal quality, and abnormal voluntary cough) are scored, and if patients pass (score ≤ 5), they progresses onto the 3 oz WST. If patients fail the non-swallow items (score ≥ 6), or if they fail the WST (clinical signs of aspiration), they are kept nil by mouth
Target condition and reference standard(s)	Two analyses: dysphagia and aspiration risk MASA
Flow and timing	374 were admitted to the stroke unit; 225 received the DHSS Index test. The reason for the drop in numbers is not reported. 224 received at least 1 MASA assessment 2 × 2 table is reported for 224 patients Interventions between index/reference tests: not reported Interval between index and reference tests: not reported
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Unclear		
Was a two-gate design avoided?	Yes		
Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			Unclear
DOMAIN 2: Index Test (All tests)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			

Behera 2018 *(Continued)*

Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Unclear risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Unclear risk	

Bravata 2009

Study characteristics	
Patient Sampling	Patients who had received at least 1 Index test plus the reference test
Patient characteristics and setting	Inclusion criteria: discharge diagnosis with ischaemic stroke (ICD9 434.X and 436) Exclusion criteria: none recorded Setting: 1 × tertiary care - Department of Veterans Affairs Medical Centre
Index tests	Nurse Screening Tool applied during admission process. NIHSS applied retrospectively
Target condition and reference standard(s)	Dysphagia SLP: expert assessment
Flow and timing	Flow is not possible to report. Study reported only on those who had at least 1 Index test + the reference test Timing not reported. In fact, study authors list possible time differences as a limitation
Comparative	
Notes	QUADAS-2 is different for each index test. Stroke Severity NIHSS (test2) reported here, as it is the 'worst', compared to the Nurse Screening Tool (test1)

Bravata 2009 *(Continued)*

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Unclear		
Was a two-gate design avoided?	Yes		
Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (All tests)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
If a threshold was used, was it pre-specified?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Unclear risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		High risk	

Campbell 2016

Study characteristics	
Patient Sampling	Convenience sample
Patient characteristics and setting	Inclusion criteria: diagnosis of stroke or TIA; admitted between 8:30am and 4:30pm; no change in initial NIHSS score between screening by the nurse and evaluation by the speech pathologist Exclusion criteria: non-stroke and non-TIA admissions; stroke and TIA admissions after 4:30pm; NIHSS score change from initial screening by the nurse, and subsequent evaluation performed by speech pathologist Setting: 1 × neuroscience unit - part of a Joint Commission Disease Specific certified stroke centre
Index tests	NBDS
Target condition and reference standard(s)	Dysphagia SLP: expert assessment
Flow and timing	3 patients not screened by Nurse 2 Interventions between index/reference tests: not reported Interval between index and reference tests: < 1 hour
Comparative	
Notes	

Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Yes		
Was a two-gate design avoided?	Yes		
Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		

Campbell 2016 *(Continued)*

Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Unclear		
Could the patient flow have introduced bias?		Unclear risk	

Cummings 2015

Study characteristics	
Patient Sampling	Convenience sample
Patient characteristics and setting	Inclusion criteria: medical diagnosis of stroke; age 18 or older; ability to follow commands Exclusion criteria: individuals who were NPO for any reason other than swallowing problems; history of previous swallowing problems; mechanical ventilation and/or intubation longer than 24 hours during current admission; inability to follow commands Setting: 2 × hospital - 20-bed medical neurology unit and 9-bed neurology intermediate care unit
Index tests	Nurse dysphagia screen
Target condition and reference standard(s)	Dysphagia Dysphagia evaluation by SLP
Flow and timing	All patients received index test(s) and reference standard. All patients were included in 2 × 2 table Interventions between index/reference tests: not reported

Cummings 2015 *(Continued)*

	Index to reference tests interval: ≤ 2 hours		
Comparative			
Notes			
Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Unclear		
Was a two-gate design avoided?	Yes		
Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	No		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Unclear risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Unclear
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		

Cummings 2015 *(Continued)*

Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Daniels 1997

Study characteristics			
Patient Sampling	Consecutive		
Patient characteristics and setting	Inclusion criteria: new neurological deficit; stroke confirmed by documentation of acute infarct by CT or MRI scan Exclusion criteria: obtunded and agitated patients; those with prior history of oropharyngeal dysphagia, oropharyngeal structural damage, or neurological disease other than stroke that may produce dysphagia Setting: 1 × Veterans Affairs Medical Center in New Orleans		
Index tests	Oropharyngeal assessment to include clinical swallowing assessment		
Target condition and reference standard(s)	Dysphagia severity or risk of aspiration VSS		
Flow and timing	All patients received index test(s) and reference standard. All patients included in 2 × 2 table apparently, but numbers do not really add up Interventions between index/reference tests: not reported Average time between test: 48 hours		
Comparative			
Notes	Flow and timing high risk due to 48-hour average time gap		
Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Was a two-gate design avoided?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (All tests)			

Daniels 1997 *(Continued)*

Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		Unclear risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Unclear
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	No		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		High risk	

Daniels 2016

Study characteristics	
Patient Sampling	Consecutive
Patient characteristics and setting	Inclusion criteria: suspected stroke Exclusion criteria: non-stroke neurological disease; head/neck structural changes; dysphagia unrelated to stroke; TIA/stroke ruled out; not competent/no legally authorised representative; medically unstable; repeat stroke; previously participated in the study; > 3 weeks from symptom onset; > 5 days from admission; childbearing potential female; deceased Setting: 1 × Michael E DeBakey Veterans Affairs Medical Center – comprehensive stroke centre
Index tests	RAS3 and WST from RAS3
Target condition and reference standard(s)	Aspiration in patients with suspected stroke

Daniels 2016 *(Continued)*

	VFSS
Flow and timing	8 patients who received the index test were not reported on. Reasons for no VFSS included C-spine surgery evident on X-ray (n = 1) and no oral intake before surgery (n = 1) Reasons for exclusion from VFSS analyses included dysphagia history prior to stroke (n = 2), equipment failure (n = 1), and C-spine damage (n = 3) Interventions between index/reference tests: not reported Index to reference test time interval within 2 hours of screening; mean (SD) 0.49 (0.24) hours
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Was a two-gate design avoided?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (All tests)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		

Daniels 2016 *(Continued)*

Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	No		
Could the patient flow have introduced bias?		Low risk	

Edmiaston 2010

Study characteristics			
Patient Sampling	Not recorded		
Patient characteristics and setting	Inclusion criteria: admitted to stroke service Exclusion criteria: none recorded Setting: 1 × Stroke Service at Barnes-Jewish Hospital, Washington University Medical Centre		
Index tests	ASDS		
Target condition and reference standard(s)	Aspiration risk and dysphagia MASA		
Flow and timing	All patients received index test(s) and reference standard. All patients included in 2 × 2 table Interventions between index/reference tests: not reported Interval between tests: approximately 24 hours		
Comparative			
Notes			
Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Unclear		
Was a two-gate design avoided?	Yes		

Edmiaston 2010 *(Continued)*

Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (All tests)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
If a threshold was used, was it pre-specified?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Unclear risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Unclear risk	

Edmiaston 2014

Study characteristics	
Patient Sampling	Acute stroke patients prospectively enrolled from Barnes-Jewish Hospital inpatient stroke service (p2)
Patient characteristics and setting	Inclusion criteria: clinical diagnosis of stroke (ischaemic or haemorrhagic); age ≥ 18 years Exclusion criteria: decreased level of alertness preventing participation in VFSS (defined as a score of 2 on the 'Alertness' component of the MASA); physical limitations preventing ability to sit up-

Edmiaston 2014 *(Continued)*

	right (excluding intubation, or if treating physician had ordered the patient to have the head of the bed flat); confirmed or suspected pregnancy Setting: 1 × Barnes-Jewish Hospital inpatient stroke service (urban, tertiary care referral centre admitting 1300 stroke patients annually)
Index tests	BJH-SDS
Target condition and reference standard(s)	Dysphagia and aspiration VFSS
Flow and timing	Study included 2 tests: 1 for dysphagia and 1 for aspiration. For the dysphagia test, all patients received the reference test; for the aspiration test, 2 patients were excluded and did not receive the reference test Interventions between index/reference tests: not reported Interval between index and reference test: 0 to 8 hours (mean interval 2 hours)
Comparative	
Notes	

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Yes		
Was a two-gate design avoided?	Yes		
Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (All tests)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern

Edmiaston 2014 *(Continued)*

DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Unclear		
Could the patient flow have introduced bias?		Low risk	

Ellis 2013

Study characteristics	
Patient Sampling	Not recorded
Patient characteristics and setting	Inclusion criteria: patient suffered stroke or TIA Exclusion criteria: not recorded Setting: 1 × small community hospital - certified primary stroke centre
Index tests	Ellis Nursing Bedside Swallowing Screen
Target condition and reference standard(s)	Dysphagia Complete dysphagia evaluation by SLT
Flow and timing	53 participants were included, baseline data for 49 participants are given, and 46 participants received both screening tests. Outline data are provided for patients who were excluded or withdrawn; however numbers do not necessarily specify the difference between original sample and those receiving the tests Interventions between index/reference tests: not reported Interval between tests: not reported
Comparative	
Notes	

Ellis 2013 *(Continued)*

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Unclear		
Was a two-gate design avoided?	Yes		
Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			Unclear
DOMAIN 2: Index Test (All tests)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		
Were the reference standard results interpreted without knowledge of the results of the index tests?	No		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Unclear
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Unclear		
Were all patients included in the analysis?	No		
Could the patient flow have introduced bias?		High risk	

Eren 2019

Study characteristics	
Patient Sampling	Unclear
Patient characteristics and setting	Inclusion criteria: diagnosis of acute stroke based on clinical and MRI results; aged > 18 years; normal cognitive function (Mini-Mental Test Score > 24 points) Exclusion criteria: presence of previous stroke; neurodegenerative or muscular disease history potentially associated with swallowing disorder; malignancy; history of surgery in the head and neck region; bilateral cranial infarction; psychiatric disorder; presence of infectious disease such as HIV, hepatitis B or C, decompensated heart failure, and nasal obstruction (exclusion criteria for FEES) Setting: 1 × neurology department
Index tests	Turkish version of BJH-SDS consists of 5 items, each with 2 choices (i.e. present = yes, absent = no). The first 4 items assess consciousness and asymmetry or weakness in facial, tongue, and palatal muscles. Level of consciousness is assessed by the Glasgow Coma Scale, and presence of dysarthria is identified together with the use of other items. The fifth item consists of the 3-oz water test, and the abnormality was defined as coughing, choking, or breathlessness while swallowing, or wet/gurgled voice after swallowing. The screening tool is a pass/fail
Target condition and reference standard(s)	Dysphagia FEES
Flow and timing	It is not clear whether the same number of patients received index and reference tests 2 × 2 table: not reported Interventions between index/reference test: not reported Interval between tests: ≤ 24 hours
Comparative	
Notes	

Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Yes		
Was a two-gate design avoided?	Yes		
Could the selection of patients have introduced bias?		Unclear risk	

Eren 2019 *(Continued)*

Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (All tests)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Unclear		
Could the patient flow have introduced bias?		Unclear risk	

Huhmann 2004

Study characteristics	
Patient Sampling	Convenience sample
Patient characteristics and setting	Inclusion criteria: none recorded Exclusion criteria: none recorded Setting: 1 × admitted to the stroke team
Index tests	Registerd Dietician Dysphagia Screening Tool
Target condition and reference standard(s)	Dysphagia risk

Huhmann 2004 *(Continued)*

	SLT bedside swallow assessment
Flow and timing	All patients received index test(s) and reference standard. All patients were included in the 2 × 2 table Interventions between index/reference tests: not reported Interval between tests: not reported
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Unclear		
Was a two-gate design avoided?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		Unclear risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			High
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			

Huhmann 2004 *(Continued)*

Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		High risk	

Huhn-Matesic 2015

Study characteristics			
Patient Sampling	Unclear		
Patient characteristics and setting	Regional non-critical care stroke unit. Patient inclusion and exclusion criteria not reported		
Index tests	EHMBAS followed by simple WST; interpretation not reported		
Target condition and reference standard(s)	MASA		
Flow and timing	All patients received the index test(s) and reference standard. All patients were included in the 2 × 2 table Interventions between index/reference tests: not reported Interval between tests: not reported		
Comparative			
Notes			
Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Unclear		
Was a two-gate design avoided?	Yes		
Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			Unclear
DOMAIN 2: Index Test (All tests)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		

Huhn-Matesic 2015 *(Continued)*

If a threshold was used, was it pre-specified?	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		Unclear risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Unclear
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Unclear risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Unclear		
Could the patient flow have introduced bias?		Unclear risk	

Jiang 2019

Study characteristics	
Patient Sampling	Consecutive
Patient characteristics and setting	Inclusion criteria: diagnosis of stroke, clinical diagnosis from a neurologist, radiographic (CT or MRI) confirmation of stroke Exclusion criteria: history of head/neck surgery, trauma; brain tumour that could influence swallowing abilities; any other concomitant neurological disorder that could influence oropharyngeal swallowing ability Setting: 1 × neurology unit
Index tests	The 8-item Chinese version of the modified Standardized Swallowing Assessment instrument is divided into 2 parts. In part 1, patients must (1) voluntarily cough, (2) control saliva, (3) lick top and bottom lips, (4) breathe freely, and (5) cannot have a wet or hoarse voice. In part 2 (items 6 to 8), patients are first given 1 mL of water. Problems are identified if no attempts are made to swallow or water leaks straight out of the mouth, if coughing, choking, or breathlessness is observed, or if a wet/gurgly voice develops. In the absence of any problems, the process is repeated with second and third tests with 1 mL of water. Gradually, increasing volumes ranging from 1 to 10 mL sequential swallowing are used based on patient tolerance. If no prob-

Jiang 2019 *(Continued)*

	lems are evident, half a glass of water (75 mL) is administered. The screening instrument is a pass vs fail procedure The 6-item Chinese version of the modified Standardized Swallowing Assessment includes the following items from the 8-item instrument: items 1, 3, and 5 of the oromotor examinations, and items 6, 7, and 8 of the WST
Target condition and reference standard(s)	Dysphagia SLP expert assessment
Flow and timing	All patients received both index and reference tests. All patients were included in the analysis Interventions between index/reference tests: none reported Interval between tests: ≤ 48 hours
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Was a two-gate design avoided?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (All tests)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	No		
Could the conduct or interpretation of the index test have introduced bias?		Unclear risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		

Jiang 2019 *(Continued)*

Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Unclear risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Lim 2001

Study characteristics	
Patient Sampling	Consecutive
Patient characteristics and setting	Inclusion criteria: acute stroke patients admitted to stroke unit of Tan Tock Seng Hospital Exclusion criteria: severe peripheral vascular disease; consciousness level not sufficient to give informed consent; no CT evidence of stroke or no significant neurological deficit (e.g. new hemiplegia) lasting > 24 hours; insufficient lip seal to retain 10 mL of water in the mouth Setting: 1 × stroke unit
Index tests	WST and bedside aspiration (combined WST and oxygen saturation test)
Target condition and reference standard(s)	Aspiration FEES
Flow and timing	Patient flow is unclear - 58 patients were eligible, 50 patients received reference and index tests, 8 patients subsequently declined FEES Interventions between index/reference tests: not reported Time interval between tests: < 24 hours for 45 patients; < 48 hours for 5 people due to 'technical issues'
Comparative	
Notes	
Methodological quality	

Lim 2001 *(Continued)*

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Was a two-gate design avoided?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (All tests)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Martino 2009 Study 1

Study characteristics	
Patient Sampling	Consecutive. However 1 in 5 patients receiving a negative screen and all receiving a positive screen on the index test received the reference test
Patient characteristics and setting	Inclusion criteria: newly admitted to hospital; confirmed diagnosis of brain stem stroke or cerebellar stroke and all other stroke types with NIHSS score ≥ 4; stroke confirmed from physician's clinical note, CT scan or MRI Exclusion criteria: non-brain stem and non-cerebellar stroke patients with low NIHSS scores; patients with current respiratory compromise; a non-oral feeding regime; history of non-stroke neurological disorder; surgery to the head or neck; history of previous oropharyngeal dysphagia, dementia, or decreased level of consciousness; not alert; cannot be supported to sit upright; cannot follow simple instructions Setting: 2 acute settings will be used for analysis, originally 4. Inpatient stroke units at 2 acute (used in data extract) and 2 rehabilitation tertiary care hospitals (not used in data extract) in southern Ontario
Index tests	TOR-BSST
Target condition and reference standard(s)	Dysphagia risk Videofluoroscopic assessment of swallowing
Flow and timing	Of the 311 patients recruited, only 103 were from an acute setting, and of these only 24 received both tests. Remaining patients were therefore excluded from the analysis Interventions between index/reference tests: not reported Interval between tests ≤ 24 hours
Comparative	
Notes	

Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Was a two-gate design avoided?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (All tests)			

Martino 2009 Study 1 *(Continued)*

Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	No		
Could the patient flow have introduced bias?		Low risk	

Martino 2014

Study characteristics	
Patient Sampling	Consecutive
Patient characteristics and setting	Inclusion criteria: eligible and consented stroke patients in the original study Exclusion criteria: not reported Setting: 2 × acute and 2 × rehabilitation facilities
Index tests	TOR-BSST
Target condition and reference standard(s)	Dysphagia VFSS

Martino 2014 *(Continued)*

Flow and timing	All patients recruited received both index and reference tests. All patients were included in the analysis, but no 2 × 2 table was reported Interventions between index/reference tests: none reported Interval between tests < 24 hours
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
Was a two-gate design avoided?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (All tests)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			

Martino 2014 *(Continued)*

Was there an appropriate interval between index test and reference standard?	Yes	
Did all patients receive the same reference standard?	Yes	
Were all patients included in the analysis?	Yes	
Could the patient flow have introduced bias?		Low risk

Mulheren 2019

Study characteristics	
Patient Sampling	Acute stroke patients admitted to the unit
Patient characteristics and setting	Inclusion criteria: acute stroke patients on the units Exclusion criteria: history of swallowing dysfunction or neurological disease that might lead to swallowing dysfunction; medical compromise barring transport to fluoroscopy; insufficient level of alertness to follow directions during VFSS as determined by the care team Setting: 1 × neuroscience critical and acute care units of a tertiary care facility
Index tests	The Johns Hopkins Hospital Brain Rescue Unit Modified 3 oz Swallow Screen is based on the 3-oz water swallow test by DePippo 1992. The screen consists of 3 steps: (1) medical history and status; (2) 1 tsp water swallow with palpation for hyolaryngeal excursion and observation for signs of aspiration; (3) 3 oz water swallow with palpation for hyolaryngeal excursion and observation for signs of aspiration. Observation for signs of aspiration is continuous during and up to 1 minute after each water trial. Wet vocal quality is assessed objectively during the vocalisation trial by trained nurses. Each step is administered sequentially, and failure of 1 step results in discontinuation and failure of the entire screen (positive aspiration risk) with nil-by-mouth status and referral to speech and language pathology for bedside evaluation
Target condition and reference standard(s)	Dysphagia VFSS
Flow and timing	All patients recruited received both index and reference tests. All patients were included in the analysis Interventions between index/reference tests: none reported Interval between tests < 72 hours
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			

Mulheren 2019 *(Continued)*

Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Yes		
Was a two-gate design avoided?	Yes		
Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (All tests)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Nishiwaki 2005

Study characteristics

Nishiwaki 2005 *(Continued)*

Patient Sampling	Consecutive patients
Patient characteristics and setting	Inclusion criteria: stroke was diagnosed with CT or MRI. Patients presented with ≥ 1 of the following: (1) bilateral or brain stem stroke; (2) history of aspiration pneumonia or increased sputum secretion; (3) cough associated with feeding and/or drinking; (4) weight loss, decreased oral intake, or prolonged feeding times; (5) complaint of difficulty in swallowing; (6) need for a therapeutic diet or non-oral feeding Exclusion criteria: patients who could not follow commands, who had a tracheostomy, who had a prior history of oropharyngeal impairment, or who had active respiratory infection Setting: 5 × university hospital and affiliated hospitals
Index tests	Clinical swallowing tests: 6 oromotor examinations
Target condition and reference standard(s)	Aspiration VFSS
Flow and timing	All patients recruited received both index and reference tests. All patients were included in the analysis Interventions between index/reference tests: none reported Interval between tests < 7 days, but < 24 hours for acute patients
Comparative	
Notes	

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Was a two-gate design avoided?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
If a threshold was used, was it pre-specified?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Unclear risk	

Nishiwaki 2005 *(Continued)*

Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Unclear
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Unclear risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	No		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		High risk	

Perry 2001 Study 1

Study characteristics	
Patient Sampling	Consecutive
Patient characteristics and setting	Inclusion criteria: admissions with a clinical diagnosis of acute stroke (ICD 10 codes 160-164, with or without CT confirmation) Exclusion criteria: none recorded Setting: 1 × study wards of Mayday University Hospital (all wards that normally care for stroke patients with the exception of 2 female care of the elderly wards where another project was in progress)
Index tests	Standardized Swallowing Assessment tool (test 1) Standardized Swallowing Assessment tool (test 2) (subset of test 1) Gag reflex (test 3) (subset of test 4) Gag reflex data from the 20-month audit period used (test 4)
Target condition and reference standard(s)	Dysphagia Summative clinical judgement - used VFS when requested by SLT to clarify decision

Perry 2001 Study 1 *(Continued)*

Flow and timing	Flow of patients is unclear from the paper. Although participants are discussed in the paper, results are presented as episodes. The results also include repeat screenings for participants; participants are excluded after recruitment and are not included in the analysis Interventions between index/reference tests: not reported Interval between tests: not reported
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Was a two-gate design avoided?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (All tests)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
If a threshold was used, was it pre-specified?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Unclear risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Unclear risk	

Perry 2001 Study 1 *(Continued)*

Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	No		
Were all patients included in the analysis?	No		
Could the patient flow have introduced bias?		High risk	

Smith 2000

Study characteristics			
Patient Sampling	Consecutive		
Patient characteristics and setting	Inclusion criteria: acute haemorrhagic or infarctive stroke; stroke confirmed by CT scan; age 18 to 90 years Exclusion criteria: impaired consciousness level; cognitive impairment or receptive dysphasia sufficient to prevent informed consent; inability to sit upright without minimal support; current lower respiratory infection; additional neurological condition; terminal illness; any medical condition that precluded VF Setting: 2 × university teaching hospitals in Manchester		
Index tests	Smith Bedside Swallowing, pulse oximetry (Minolta Pulsox 7), or both		
Target condition and reference standard(s)	Aspiration and penetration VF		
Flow and timing	All patients received the index test(s) and the reference standard. All patients were included in the 2 × 2 table, but for tests 1 and 5, the 2 × 2 was recovered in the RevMan calculator. This gave test 5 with N = 52, 1 patient missing Interventions between index/reference tests: not reported Interval between tests ≤ 24 hours (all tests carried out on the same day)		
Comparative			
Notes			
Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			

Smith 2000 *(Continued)*

Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Was a two-gate design avoided?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (All tests)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		Unclear risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Trapl 2007

Study characteristics			
Patient Sampling	Consecutive		

Trapl 2007 *(Continued)*

Patient characteristics and setting	Inclusion criteria: patients with first-ever acute stroke and suspected dysphagia; admitted to the acute stroke unit on weekdays between Monday and Thursday Exclusion criteria: multiple infarcts visible on CT or MRI scan; dysphagia of other known cause; somnolence or coma within 24 hours Setting: 1 × acute stroke unit
Index tests	GUSS
Target condition and reference standard(s)	Aspiration FEES
Flow and timing	In group 1, 1 patient refused FEES (reference test) and so was not included in the 2 × 2 table In group 2, all patients received the index test(s) and the reference standard. All patients were included in the 2 × 2 table Interventions between index/reference tests: not reported Interval between tests < 24 hours
Comparative	
Notes	

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Was a two-gate design avoided?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (All tests)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern

Trapl 2007 *(Continued)*

DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	No		
Could the patient flow have introduced bias?		Low risk	

Turner-Lawrence 2009

Study characteristics	
Patient Sampling	Convenience sample
Patient characteristics and setting	Inclusion criteria: presumptive stroke diagnosis; age ≥ 18 years; presenting to the emergency department within 24 hours of symptom onset; Glasgow Coma Scale score > 12; hospital admission Exclusion criteria: non-stroke diagnosis (based on admitting or consultant physician evaluation); primary brain lesion as cause of stroke; intubation before completion of ED screening or SLP evaluation; history of dysphagia or modified feeding route; pre-existing neuromuscular disorder; history of head and neck cancer or radiation; preexisting pneumonia on ED chest radiography; pregnancy Setting: 1 × 850-bed, tertiary-care, TJC-certified, primary stroke centre located in the southeast
Index tests	ED dysphagia screen (2-tiered approach)
Target condition and reference standard(s)	Aspiration SLP dysphagia assessment
Flow and timing	All patients received the index test(s) and the reference standard. All patients were included in the 2 × 2 table Interventions between index/reference tests: not reported Interval between tests ≤ 24 hours

Turner-Lawrence 2009 *(Continued)*

Comparative
Notes

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Did the study avoid inappropriate exclusions?	Unclear		
Was a two-gate design avoided?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Unclear risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		

Turner-Lawrence 2009 *(Continued)*

Could the patient flow have introduced bias?	Low risk

Warnecke 2008

Study characteristics			
Patient Sampling	Consecutive		
Patient characteristics and setting	Inclusion criteria: first-ever stroke; admitted within 24 hours of onset of symptoms; patients eligible for advanced dysphagia assessment as defined by NIHSS score ≥ 3 points and/or who had to present with facial palsy and/or dysarthria Exclusion criteria: patients with history of a pre-existing dysphagia or disease probably causing dysphagia; severely reduced state of consciousness (i.e. stupor or coma) Setting: not recorded		
Index tests	2-step Swallowing Provocation Test		
Target condition and reference standard(s)	Aspiration FEES		
Flow and timing	All patients received the index test(s) and the reference standard. All patients were included in the 2 × 2 table Interventions between index/reference tests: not reported Interval between tests ≤ 24 hours, reported as immediately following		
Comparative			
Notes			
Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Was a two-gate design avoided?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (All tests)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		

Warnecke 2008 *(Continued)*

If a threshold was used, was it pre-specified?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Zhou 2011

Study characteristics	
Patient Sampling	Unclear
Patient characteristics and setting	Inclusion criteria: patients suffering a first-ever hemispheric stroke; symptoms lasting at least 24 hours; objective lesions on cerebral imaging (CT and/or MRI) confirmed Exclusion criteria: patients with subarachnoid haemorrhage, subtentorial stroke or TIA, patients with sufficient loss of consciousness to prevent oral feeding; patients with history of otolaryngological cancer; < 18 years of age; those who did not expressly agree to participate in the study 1 × Neurology Department at university hospital. 1 × Department of Physical Medicine and Readaptation
Index tests	Clinical Predicative Scale of Aspiration; Practical Aspiration Screening Scheme

Zhou 2011 *(Continued)*

Target condition and reference standard(s)	Reference standard: videofluoroscopic examination
Flow and timing	
Comparative	
Notes	This paper is classified as narrative

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Unclear		
Was a two-gate design avoided?	Yes		
Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			Unclear
DOMAIN 2: Index Test (All tests)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
If a threshold was used, was it pre-specified?	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		Unclear risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Unclear
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Unclear risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Unclear
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		

Zhou 2011 *(Continued)*

Did all patients receive the same reference standard?	Unclear	
Were all patients included in the analysis?	Unclear	
Could the patient flow have introduced bias?		Unclear risk

AIS: acute ischaemic stroke.
ASDS: Acute Stroke Dysphagia Screening.
BJH-SDS: Barnes-Jewish Hospital-Stroke Dysphagia Screen.
CT: computed tomography.
ED: emergency department.
EHMBAS: Edith-Huhn-Matesic Bedside Aspiration Screen.
FEES: fibreoptic endoscopic evaluation of swallowing.
GUSS: Gugging Swallowing Screen.
MASA: Mann Assessment of Swallowing Ability.
MRI: magnetic resonance imaging.
NBDS: Nursing Bedside Dysphagia Screen.
NIHSS: National Institutes of Health Stroke Scale.
NPO: nil by mouth.
RAS3: Rapid Aspiration Screening for Suspected Stroke.
SLP: speech and language pathologist.
SLT: speech and language therapist.
TIA: transient ischaemic attack.
TOR-BSST: Toronto Bedside Swallowing Screening Test.
VF: videofluoroscopy.
VFSS: videofluoroscopic swallowing study.
WST: water-swallowing test.

Characteristics of excluded studies *[ordered by study ID]*

Study	Reason for exclusion
Addington 1999 Study 1	No acute stroke
Addington 1999 Study 2	No acute stroke
Alsibai 2014	No outcome of interest
Anderson 2016	Does not compare tests
Archer 2015 Study 1	No index or reference test of interest
Archer 2015 Study 2	No index or reference test of interest
Armesto 2015	No outcome of interest
Aviv 1997	No outcome of interest
Aydogdu 2015	Not limited to stroke patients
Babi 2014	Does not compare tests
Bahia 2014	Does not compare tests
Bahia 2016	Not an acute setting

Study	Reason for exclusion
Bax 2013	Does not compare tests
Bax 2014	No index or reference test of interest
Benjamin 2015	Not a single-gate or 2-gate study
Burks 2011	No screening for dysphagia
Carnaby 2014	No index or reference test of interest
Carrington 2013	Does not compare tests
Chavarria 2015	No screening for dysphagia
Choi 2010	No index or reference test of interest
Chong 2003	Not an acute setting
Cichero 2009	Not limited to stroke patients
Clayton 2006	No index or reference test of interest
Collins 1997	Not an acute setting
Crary 2013 Study 1	No index or reference test of interest
Crary 2013 Study 2	No index or reference test of interest
Crary 2014 Study 1	No index or reference test of interest
Crary 2014 Study 2	No index or reference test of interest
Daniels 2009	Does not compare tests
Davis-De Geus 2009	No outcome of interest
DePippo 1992	Not an acute setting
Farneti 2015	Index test undertaken by experts
Giraldo-Cadavid 2015	No outcome of interest
Guerra 2011	Does not compare tests
Guillén-Solà 2011	No screening for dysphagia
Guillén-Solà 2013	Not an acute setting
Gurcay 2018	Sample not restricted to acute stages of stroke
Higo 2003	No index or reference test of interest
Joundi 2017	Not a single-gate or 2-gate study
Karaca 2018	Reference test MASA considered an SLT dysphagia assessment

Study	Reason for exclusion
Logemann 1999	Not limited to stroke patients
Mandysova 2015	Not limited to stroke patients
Marques 2008	Does not compare tests
Martino 2009 Study 2	No index or reference test of interest
Massey 2002	Not all patients received the same reference test
McCullough 2001	SLT dysphagia assessment, not screening
McCullough 2005	No acute stroke
Moalli 2016	Does not compare tests
Moon 2010	No consecutive stroke patients
Mozzanica 2017	Heterogeneous population. Stroke not analysed separately
Murray 2011	No outcome of interest
NCT00306501	No accuracy of screening
NCT00580138	No index or reference test of interest
NCT01765673	No accuracy of screening
NCT02080988	No outcome of interest
NCT02848664	Not limited to stroke patients
NCT03167892	No accuracy of screening
Nomura 2014	Does not compare tests
Oh 2016	All patients had dysphagia or aspiration; therefore sensitivity/specificity cannot be calculated
Ohira 2013	Not limited to stroke patients
Ohira 2017	Not an acute setting
Osawa 2013	No consecutive stroke patients
Ostrofsky 2016	Index tests carried out by SLT/SLP
Palli 2017	No reference test performed on same patients who had nurse swallow screen – the comparison was with case note data from historic SLT assessment
Pennsylvania Patient Safety Authority 2009	Not limited to stroke patients
Perry 2000	Does not compare tests
Perry 2001 Study 2	Does not compare tests

Study	Reason for exclusion
Radhakrishnan 2013	No screening for dysphagia
Rosenbek 2004	SLT dysphagia assessment, not screening
Sato 2012	Not an acute setting
Schrock 2011	No index or reference test of interest
Sellars 1998	All patients had dysphagia or aspiration; therefore sensitivity/specificity cannot be calculated
Smith Hammond 2009	Index tests carried out by SLT/SLP
Sohn 2018	Reference test is a retrospective evaluation of case records
Soria 2013	Not limited to stroke patients
Stroke and Chewing 2012	No outcome of interest
Sung 2018	Does not compare tests
Sørensen 2013	Does not compare tests
Tanuma 2002	Not screening for dysphagia
Teramoto 2000	Not an acute setting
Toscano 2019	Index test undertaken by experts
Tuncay 2011	Not index or reference test of interest
Umay 2013	No acute stroke
Virvidaki 2019	Index test undertaken by experts
Warnecke 2017	Index tests carried out by SLT/SLP
Weinhardt 2008	Does not compare tests
Ye 2018	No reference test was included
Yeh 2011	Does not compare tests
Zhang 2004	All patients had dysphagia or aspiration; therefore sensitivity/specificity cannot be calculated

MASA: Mann Assessment of Swallowing Ability.
SLP: speech and language pathologist.
SLT: speech and language therapist.

DATA

Presented below are all the data for all of the tests entered into the review.

Table Tests. Data tables by test

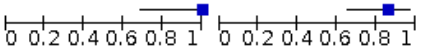
Test	No. of studies	No. of participants
1 Registered Dietitian (RD) Dysphagia Screening tool - Huhmann (2004)	1	32
2 Bedside aspiration - Combined WST & Oxygen Saturation - Lim (2001)	1	50
3 Gugging Swallowing Screen (GUSS) - Group2 - Trappl (2007)	1	30
4 Toronto Bedside Swallowing Screening Test (TOR-BSST) - Martino (2009)	1	24
5 Standardized Swallowing Assessment tool (SSA) - Test2 - Perry (2001) Test2	1	68
6 Nursing Bedside Dysphagia Screen (NBDS) - Campbell (2016)	1	75
7 Emergency Department (ED) dysphagia screen - Turner-Lawrence (2009)	1	84
8 Acute Stroke Dysphagia Screening (ASDS) - Aspiration - Edmiaston (2010)	1	300
9 Barnes-Jewish Hospital-Stroke Dysphagia Screen (BJH-SDS) Aspiration - Edmiaston (2014)	1	223
10 Edith-Huhn-Matesic Bedside Aspiration Screen (EHMBAS) plus water swallow test - Huhn-Matesic (2015)	1	52
11 Standardized Swallowing Assessment tool (SSA) - Test1 - Perry (2001)	1	161
12 Barnes-Jewish Hospital-Stroke Dysphagia Screen (BJH-SDS) Dysphagia - Edmiaston (2014)	1	225
13 Modified MASA (MMASA) Neurologist 1 - Antonios (2010)	1	150
14 Rapid Aspiration Screening for Suspected Stroke (RAS3) - Daniels (2016)	1	250
15 Clinical examination - Daniels (1997)	1	59
16 Acute Stroke Dysphagia Screening (ASDS) Dysphagia - Edmiaston (2010)	1	300
17 Nurse Dysphagia Screen - Cummings (2015)	1	49
18 Modified MASA (MMASA) Neurologist 2 - Antonios (2010)	1	150
19 Oxygen saturation ≥ 2% - Test2 for Aspiration - Smith (2000)	1	53
20 Bedside swallow test - WST only - Lim (2001)	1	50
21 Chinese version of the modified SSA original 8 items - Jiang (2019)	1	127
22 Chinese version of the modified SSA reduced 6 items - Jiang (2019)	1	127
23 Stroke Severity using National Institutes of Health Stroke Scale (NIHSS) - Bravata (2009)	1	101
24 Nursing Bedside Swallowing Screen (NBSS) - Ellis (2013)	1	46
25 2-step Swallowing Provocation Test (SPT) - step 1 - 0.4 mL - Warneke (2008)	1	100

Test	No. of studies	No. of participants
26 Rapid Aspiration Screening for Suspected Stroke (RAS3) - WST only - Daniels (2016)	1	250
27 Gag function - Test3 - Perry (2001)	1	22
28 DePaul Hospital Swallow Screener (DHSS) for Aspiration Risk - Behera (2018)	1	226
29 DePaul Hospital Swallow Screener (DHSS) for Dysphagia - Behera (2018)	1	225
30 Gag function - Test4 - Perry (2001)	1	157
31 2-step swallowing provocation test (SPT) step 2 - 2.0 mL - Warneke (2008)	1	100
32 Johns Hopkins Hospital Brain Rescue Unit Modified 3 oz Swallow Screen - Mulheren (2019)	1	48
33 Nursing Screening Tool - Bravata (2009)	1	39
34 Barnes-Jewish Hospital Stroke Dysphagia Screen (BJH-SDS) – Turkish version (T-BJH) - Eren (2019)	1	0
35 Clinical Predicative Scale of Aspiration (CPSA) - Zhou (2011)	1	0
36 TOR-BSST water swallow item - Martino (2014)	1	0
37 Clinical swallowing tests - 6 oromotor examinations - Nishiwaki (2005)	1	0
38 Index taxonomy - water only	13	2914
39 Index taxonomy - water plus other consistencies	5	412
40 Index taxonomy - other	5	627
41 Outcome - aspiration	11	1827
42 Outcome - dysphagia	13	2126
43 Reference test - Expert Assessment and MASA	12	2491
44 Reference test - FEES	3	330
45 Reference test - VF	6	1132
46 HCP - nurse	14	2785
47 HCP - other	8	1168

Test 1. Registered Dietitian (RD) Dysphagia Screening tool - Huhmann (2004)

Registered Dietitian (RD) Dysphagia Screening tool - Huhmann (2004)

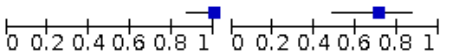
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Huhmann 2004	10	3	0	19	1.00 [0.69, 1.00]	0.86 [0.65, 0.97]		



Test 2. Bedside aspiration - Combined WST & Oxygen Saturation - Lim (2001)

Bedside aspiration - Combined WST & Oxygen Saturation - Lim (2001)

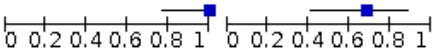
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Lim 2001	26	7	0	17	1.00 [0.87, 1.00]	0.71 [0.49, 0.87]		



Test 3. Gugging Swallowing Screen (GUSS) - Group2 - Trappl (2007)

Gugging Swallowing Screen (GUSS) - Group2 - Trappl (2007)

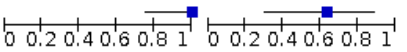
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Trapl 2007	14	5	0	11	1.00 [0.77, 1.00]	0.69 [0.41, 0.89]		



Test 4. Toronto Bedside Swallowing Screening Test (TOR-BSST) - Martino (2009)

Toronto Bedside Swallowing Screening Test (TOR-BSST) - Martino (2009)

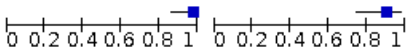
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Martino 2009 Study 1	13	4	0	7	1.00 [0.75, 1.00]	0.64 [0.31, 0.89]		



Test 5. Standardized Swallowing Assessment tool (SSA) - Test2 - Perry (2001) Test2

Standardized Swallowing Assessment tool (SSA) - Test2 - Perry (2001) Test2

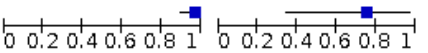
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Perry 2001 Study 1	36	3	1	28	0.97 [0.86, 1.00]	0.90 [0.74, 0.98]		



Test 6. Nursing Bedside Dysphagia Screen (NBDS) - Campbell (2016)

Nursing Bedside Dysphagia Screen (NBDS) - Campbell (2016)

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Campbell 2016	65	2	2	6	0.97 [0.90, 1.00]	0.75 [0.35, 0.97]		



Test 7. Emergency Department (ED) dysphagia screen - Turner-Lawrence (2009)

Emergency Department (ED) dysphagia screen - Turner-Lawrence (2009)

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Turner-Lawrence 2009	46	16	2	20	0.96 [0.86, 0.99]	0.56 [0.38, 0.72]	0 0.2 0.4 0.6 0.8 1	0 0.2 0.4 0.6 0.8 1

Test 8. Acute Stroke Dysphagia Screening (ASDS) - Aspiration - Edmiaston (2010)

Acute Stroke Dysphagia Screening (ASDS) - Aspiration - Edmiaston (2010)

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Edmiaston 2010	60	74	3	163	0.95 [0.87, 0.99]	0.69 [0.62, 0.75]	0 0.2 0.4 0.6 0.8 1	0 0.2 0.4 0.6 0.8 1

Test 9. Barnes-Jewish Hospital-Stroke Dysphagia Screen (BJH-SDS) Aspiration - Edmiaston (2014)

Barnes-Jewish Hospital-Stroke Dysphagia Screen (BJH-SDS) Aspiration - Edmiaston (2014)

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Edmiaston 2014	57	81	3	82	0.95 [0.86, 0.99]	0.50 [0.42, 0.58]	0 0.2 0.4 0.6 0.8 1	0 0.2 0.4 0.6 0.8 1

Test 10. Edith-Huhn-Matesic Bedside Aspiration Screen (EHMBAS) plus water swallow test - Huhn-Matesic (2015)

Edith-Huhn-Matesic Bedside Aspiration Screen (EHMBAS) plus water swallow test - Huhn-Matesic (2015)

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Huhn-Matesic 2015	16	8	1	27	0.94 [0.71, 1.00]	0.77 [0.60, 0.90]	0 0.2 0.4 0.6 0.8 1	0 0.2 0.4 0.6 0.8 1

Test 11. Standardized Swallowing Assessment tool (SSA) - Test1 - Perry (2001)

Standardized Swallowing Assessment tool (SSA) - Test1 - Perry (2001)

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Perry 2001 Study 1	88	17	6	50	0.94 [0.87, 0.98]	0.75 [0.63, 0.84]	0 0.2 0.4 0.6 0.8 1	0 0.2 0.4 0.6 0.8 1

Test 12. Barnes-Jewish Hospital-Stroke Dysphagia Screen (BJH-SDS) Dysphagia - Edmiaston (2014)

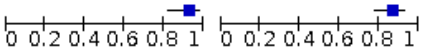
Barnes-Jewish Hospital-Stroke Dysphagia Screen (BJH-SDS) Dysphagia - Edmiaston (2014)

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Edmiaston 2014	100	40	6	79	0.94 [0.88, 0.98]	0.66 [0.57, 0.75]	0 0.2 0.4 0.6 0.8 1	0 0.2 0.4 0.6 0.8 1

Test 13. Modified MASA (MMASA) Neurologist 1 - Antonios (2010)

Modified MASA (MMASA) Neurologist 1 - Antonios (2010)

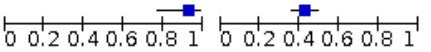
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Antonios 2010	50	13	4	83	0.93 [0.82, 0.98]	0.86 [0.78, 0.93]		



Test 14. Rapid Aspiration Screening for Suspected Stroke (RAS3) - Daniels (2016)

Rapid Aspiration Screening for Suspected Stroke (RAS3) - Daniels (2016)

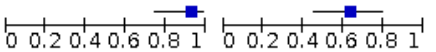
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Daniels 2016	27	126	2	95	0.93 [0.77, 0.99]	0.43 [0.36, 0.50]		



Test 15. Clinical examination - Daniels (1997)

Clinical examination - Daniels (1997)

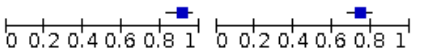
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Daniels 1997	24	12	2	21	0.92 [0.75, 0.99]	0.64 [0.45, 0.80]		



Test 16. Acute Stroke Dysphagia Screening (ASDS) Dysphagia - Edmiaston (2010)

Acute Stroke Dysphagia Screening (ASDS) Dysphagia - Edmiaston (2010)

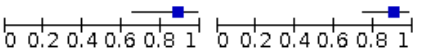
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Edmiaston 2010	80	54	8	158	0.91 [0.83, 0.96]	0.75 [0.68, 0.80]		



Test 17. Nurse Dysphagia Screen - Cummings (2015)

Nurse Dysphagia Screen - Cummings (2015)

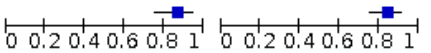
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Cummings 2015	16	3	2	28	0.89 [0.65, 0.99]	0.90 [0.74, 0.98]		



Test 18. Modified MASA (MMASA) Neurologist 2 - Antonios (2010)

Modified MASA (MMASA) Neurologist 2 - Antonios (2010)

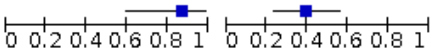
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Antonios 2010	47	15	7	81	0.87 [0.75, 0.95]	0.84 [0.76, 0.91]		



Test 19. Oxygen saturation ≥ 2% - Test2 for Aspiration - Smith (2000)

Oxygen saturation ≥ 2% - Test2 for Aspiration - Smith (2000)

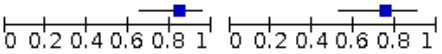
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Smith 2000	13	23	2	15	0.87 [0.60, 0.98]	0.39 [0.24, 0.57]		



Test 20. Bedside swallow test - WST only - Lim (2001)

Bedside swallow test - WST only - Lim (2001)

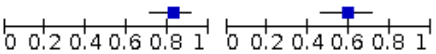
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Lim 2001	22	6	4	18	0.85 [0.65, 0.96]	0.75 [0.53, 0.90]		



Test 21. Chinese version of the modified SSA original 8 items - Jiang (2019)

Chinese version of the modified SSA original 8 items - Jiang (2019)

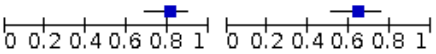
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Jiang 2019	52	26	11	38	0.83 [0.71, 0.91]	0.59 [0.46, 0.71]		



Test 22. Chinese version of the modified SSA reduced 6 items - Jiang (2019)

Chinese version of the modified SSA reduced 6 items - Jiang (2019)

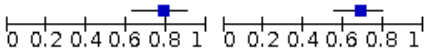
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Jiang 2019	51	23	12	41	0.81 [0.69, 0.90]	0.64 [0.51, 0.76]		



Test 23. Stroke Severity using National Institutes of Health Stroke Scale (NIHSS) - Bravata (2009)

Stroke Severity using National Institutes of Health Stroke Scale (NIHSS) - Bravata (2009)

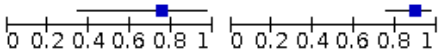
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Bravata 2009	30	20	8	43	0.79 [0.63, 0.90]	0.68 [0.55, 0.79]		



Test 24. Nursing Bedside Swallowing Screen (NBSS) - Ellis (2013)

Nursing Bedside Swallowing Screen (NBSS) - Ellis (2013)

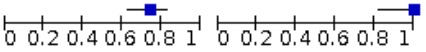
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Ellis 2013	6	4	2	34	0.75 [0.35, 0.97]	0.89 [0.75, 0.97]		



Test 25. 2-step Swallowing Provocation Test (SPT) - step 1 - 0.4 mL - Warneke (2008)

2-step Swallowing Provocation Test (SPT) - step 1 - 0.4 mL - Warneke (2008)

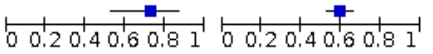
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Warnecke 2008	60	0	21	19	0.74 [0.63, 0.83]	1.00 [0.82, 1.00]		



Test 26. Rapid Aspiration Screening for Suspected Stroke (RAS3) - WST only - Daniels (2016)

Rapid Aspiration Screening for Suspected Stroke (RAS3) - WST only - Daniels (2016)

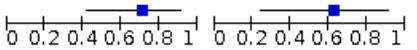
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Daniels 2016	21	89	8	132	0.72 [0.53, 0.87]	0.60 [0.53, 0.66]		



Test 27. Gag function - Test3 - Perry (2001)

Gag function - Test3 - Perry (2001)

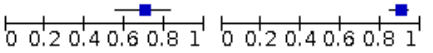
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Perry 2001 Study 1	10	3	4	5	0.71 [0.42, 0.92]	0.63 [0.24, 0.91]		



Test 28. DePaul Hospital Swallow Screener (DHSS) for Aspiration Risk - Behera (2018)

DePaul Hospital Swallow Screener (DHSS) for Aspiration Risk - Behera (2018)

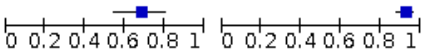
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Behera 2018	33	18	14	161	0.70 [0.55, 0.83]	0.90 [0.85, 0.94]		



Test 29. DePaul Hospital Swallow Screener (DHSS) for Dysphagia - Behera (2018)

DePaul Hospital Swallow Screener (DHSS) for Dysphagia - Behera (2018)

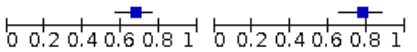
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Behera 2018	39	12	18	156	0.68 [0.55, 0.80]	0.93 [0.88, 0.96]		



Test 30. Gag function - Test4 - Perry (2001)

Gag function - Test4 - Perry (2001)

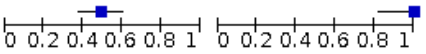
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Perry 2001 Study 1	66	13	32	46	0.67 [0.57, 0.76]	0.78 [0.65, 0.88]		



Test 31. 2-step swallowing provocation test (SPT) step 2 - 2.0 mL - Warneke (2008)

2-step swallowing provocation test (SPT) step 2 - 2.0 mL - Warneke (2008)

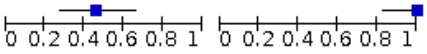
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Warnecke 2008	40	0	41	19	0.49 [0.38, 0.61]	1.00 [0.82, 1.00]		



Test 32. Johns Hopkins Hospital Brain Rescue Unit Modified 3 oz Swallow Screen - Mulheren (2019)

Johns Hopkins Hospital Brain Rescue Unit Modified 3 oz Swallow Screen - Mulheren (2019)

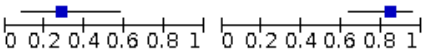
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Mulheren 2019	13	0	15	20	0.46 [0.28, 0.66]	1.00 [0.83, 1.00]		



Test 33. Nursing Screening Tool - Bravata (2009)

Nursing Screening Tool - Bravata (2009)

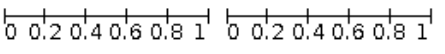
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Bravata 2009	4	4	10	21	0.29 [0.08, 0.58]	0.84 [0.64, 0.95]		



Test 34. Barnes-Jewish Hospital Stroke Dysphagia Screen (BJH-SDS) – Turkish version (T-BJH) - Eren (2019)

Barnes-Jewish Hospital Stroke Dysphagia Screen (BJH-SDS) – Turkish version (T-BJH) - Eren (2019)

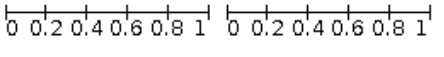
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Eren 2019	0	0	0	0	Not estimable	Not estimable		



Test 35. Clinical Predicative Scale of Aspiration (CPSA) - Zhou (2011)

Clinical Predicative Scale of Aspiration (CPSA) - Zhou (2011)

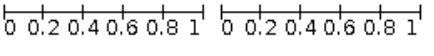
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Zhou 2011	0	0	0	0	Not estimable	Not estimable		



Test 36. TOR-BSST water swallow item - Martino (2014)

TOR-BSST water swallow item - Martino (2014)

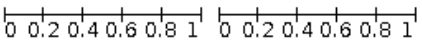
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Martino 2014	0	0	0	0	Not estimable	Not estimable		



Test 37. Clinical swallowing tests - 6 oromotor examinations - Nishiwaki (2005)

Clinical swallowing tests - 6 oromotor examinations - Nishiwaki (2005)

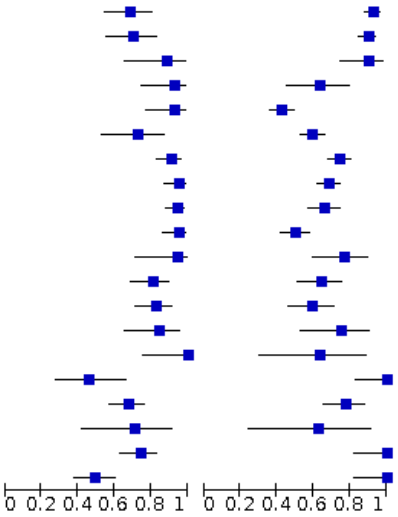
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Nishiwaki 2005	0	0	0	0	Not estimable	Not estimable		



Test 38. Index taxonomy - water only

Index taxonomy - water only

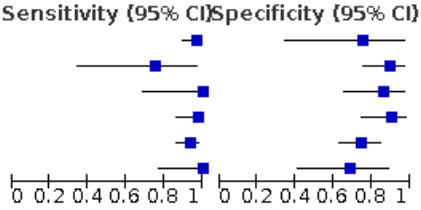
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Behera 2018	39	12	18	156	0.68 [0.55, 0.80]	0.93 [0.88, 0.96]		
Behera 2018	33	18	14	161	0.70 [0.55, 0.83]	0.90 [0.85, 0.94]		
Cummings 2015	16	3	2	28	0.89 [0.65, 0.99]	0.90 [0.74, 0.98]		
Daniels 1997	24	12	2	21	0.92 [0.75, 0.99]	0.64 [0.45, 0.80]		
Daniels 2016	27	126	2	95	0.93 [0.77, 0.99]	0.43 [0.36, 0.50]		
Daniels 2016	21	89	8	132	0.72 [0.53, 0.87]	0.60 [0.53, 0.66]		
Edmiaston 2010	80	54	8	158	0.91 [0.83, 0.96]	0.75 [0.68, 0.80]		
Edmiaston 2010	60	74	3	163	0.95 [0.87, 0.99]	0.69 [0.62, 0.75]		
Edmiaston 2014	100	40	6	79	0.94 [0.88, 0.98]	0.66 [0.57, 0.75]		
Edmiaston 2014	57	81	3	82	0.95 [0.86, 0.99]	0.50 [0.42, 0.58]		
Huhn-Matesic 2015	16	8	1	27	0.94 [0.71, 1.00]	0.77 [0.60, 0.90]		
Jiang 2019	51	23	12	41	0.81 [0.69, 0.90]	0.64 [0.51, 0.76]		
Jiang 2019	52	26	11	38	0.83 [0.71, 0.91]	0.59 [0.46, 0.71]		
Lim 2001	22	6	4	18	0.85 [0.65, 0.96]	0.75 [0.53, 0.90]		
Martino 2009 Study 1	13	4	0	7	1.00 [0.75, 1.00]	0.64 [0.31, 0.89]		
Mulheren 2019	13	0	15	20	0.46 [0.28, 0.66]	1.00 [0.83, 1.00]		
Perry 2001 Study 1	66	13	32	46	0.67 [0.57, 0.76]	0.78 [0.65, 0.88]		
Perry 2001 Study 1	10	3	4	5	0.71 [0.42, 0.92]	0.63 [0.24, 0.91]		
Warnecke 2008	60	0	21	19	0.74 [0.63, 0.83]	1.00 [0.82, 1.00]		
Warnecke 2008	40	0	41	19	0.49 [0.38, 0.61]	1.00 [0.82, 1.00]		



Test 39. Index taxonomy - water plus other consistencies

Index taxonomy - water plus other consistencies

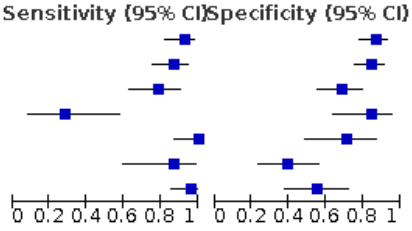
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Campbell 2016	65	2	2	6	0.97 [0.90, 1.00]	0.75 [0.35, 0.97]		
Ellis 2013	6	4	2	34	0.75 [0.35, 0.97]	0.89 [0.75, 0.97]		
Huhmann 2004	10	3	0	19	1.00 [0.69, 1.00]	0.86 [0.65, 0.97]		
Perry 2001 Study 1	36	3	1	28	0.97 [0.86, 1.00]	0.90 [0.74, 0.98]		
Perry 2001 Study 1	88	17	6	50	0.94 [0.87, 0.98]	0.75 [0.63, 0.84]		
Trapl 2007	14	5	0	11	1.00 [0.77, 1.00]	0.69 [0.41, 0.89]		



Test 40. Index taxonomy - other

Index taxonomy - other

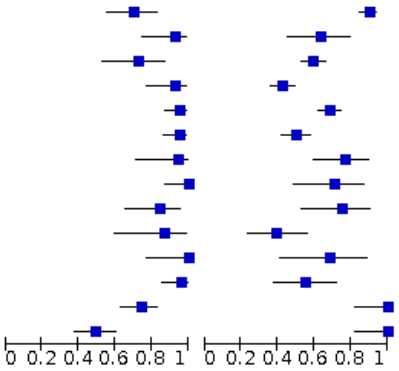
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Antonios 2010	50	13	4	83	0.93 [0.82, 0.98]	0.86 [0.78, 0.93]		
Antonios 2010	47	15	7	81	0.87 [0.75, 0.95]	0.84 [0.76, 0.91]		
Bravata 2009	30	20	8	43	0.79 [0.63, 0.90]	0.68 [0.55, 0.79]		
Bravata 2009	4	4	10	21	0.29 [0.08, 0.58]	0.84 [0.64, 0.95]		
Lim 2001	26	7	0	17	1.00 [0.87, 1.00]	0.71 [0.49, 0.87]		
Smith 2000	13	23	2	15	0.87 [0.60, 0.98]	0.39 [0.24, 0.57]		
Turner-Lawrence 2009	46	16	2	20	0.96 [0.86, 0.99]	0.56 [0.38, 0.72]		



Test 41. Outcome - aspiration

Outcome - aspiration

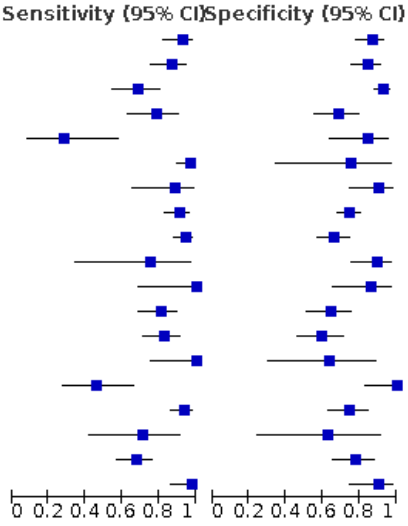
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Behera 2018	33	18	14	161	0.70 [0.55, 0.83]	0.90 [0.85, 0.94]
Daniels 1997	24	12	2	21	0.92 [0.75, 0.99]	0.64 [0.45, 0.80]
Daniels 2016	21	89	8	132	0.72 [0.53, 0.87]	0.60 [0.53, 0.66]
Daniels 2016	27	126	2	95	0.93 [0.77, 0.99]	0.43 [0.36, 0.50]
Edmiaston 2010	60	74	3	163	0.95 [0.87, 0.99]	0.69 [0.62, 0.75]
Edmiaston 2014	57	81	3	82	0.95 [0.86, 0.99]	0.50 [0.42, 0.58]
Huhn-Matesic 2015	16	8	1	27	0.94 [0.71, 1.00]	0.77 [0.60, 0.90]
Lim 2001	26	7	0	17	1.00 [0.87, 1.00]	0.71 [0.49, 0.87]
Lim 2001	22	6	4	18	0.85 [0.65, 0.96]	0.75 [0.53, 0.90]
Smith 2000	13	23	2	15	0.87 [0.60, 0.98]	0.39 [0.24, 0.57]
Trapl 2007	14	5	0	11	1.00 [0.77, 1.00]	0.69 [0.41, 0.89]
Turner-Lawrence 2009	46	16	2	20	0.96 [0.86, 0.99]	0.56 [0.38, 0.72]
Warnecke 2008	60	0	21	19	0.74 [0.63, 0.83]	1.00 [0.82, 1.00]
Warnecke 2008	40	0	41	19	0.49 [0.38, 0.61]	1.00 [0.82, 1.00]



Test 42. Outcome - dysphagia

Outcome - dysphagia

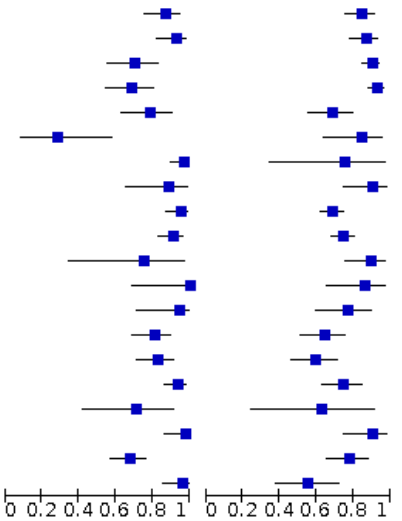
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Antonios 2010	50	13	4	83	0.93 [0.82, 0.98]	0.86 [0.78, 0.93]
Antonios 2010	47	15	7	81	0.87 [0.75, 0.95]	0.84 [0.76, 0.91]
Behera 2018	39	12	18	156	0.68 [0.55, 0.80]	0.93 [0.88, 0.96]
Bravata 2009	30	20	8	43	0.79 [0.63, 0.90]	0.68 [0.55, 0.79]
Bravata 2009	4	4	10	21	0.29 [0.08, 0.58]	0.84 [0.64, 0.95]
Campbell 2016	65	2	2	6	0.97 [0.90, 1.00]	0.75 [0.35, 0.97]
Cummings 2015	16	3	2	28	0.89 [0.65, 0.99]	0.90 [0.74, 0.98]
Edmiaston 2010	80	54	8	158	0.91 [0.83, 0.96]	0.75 [0.68, 0.80]
Edmiaston 2014	100	40	6	79	0.94 [0.88, 0.98]	0.66 [0.57, 0.75]
Ellis 2013	6	4	2	34	0.75 [0.35, 0.97]	0.89 [0.75, 0.97]
Huhmann 2004	10	3	0	19	1.00 [0.69, 1.00]	0.86 [0.65, 0.97]
Jiang 2019	51	23	12	41	0.81 [0.69, 0.90]	0.64 [0.51, 0.76]
Jiang 2019	52	26	11	38	0.83 [0.71, 0.91]	0.59 [0.46, 0.71]
Martino 2009 Study 1	13	4	0	7	1.00 [0.75, 1.00]	0.64 [0.31, 0.89]
Mulheren 2019	13	0	15	20	0.46 [0.28, 0.66]	1.00 [0.83, 1.00]
Perry 2001 Study 1	88	17	6	50	0.94 [0.87, 0.98]	0.75 [0.63, 0.84]
Perry 2001 Study 1	10	3	4	5	0.71 [0.42, 0.92]	0.63 [0.24, 0.91]
Perry 2001 Study 1	66	13	32	46	0.67 [0.57, 0.76]	0.78 [0.65, 0.88]
Perry 2001 Study 1	36	3	1	28	0.97 [0.86, 1.00]	0.90 [0.74, 0.98]



Test 43. Reference test - Expert Assessment and MASA

Reference test - Expert Assessment and MASA

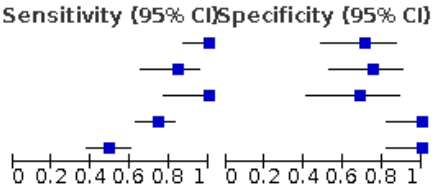
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Antonios 2010	47	15	7	81	0.87 [0.75, 0.95]	0.84 [0.76, 0.91]
Antonios 2010	50	13	4	83	0.93 [0.82, 0.98]	0.86 [0.78, 0.93]
Behera 2018	33	18	14	161	0.70 [0.55, 0.83]	0.90 [0.85, 0.94]
Behera 2018	39	12	18	156	0.68 [0.55, 0.80]	0.93 [0.88, 0.96]
Bravata 2009	30	20	8	43	0.79 [0.63, 0.90]	0.68 [0.55, 0.79]
Bravata 2009	4	4	10	21	0.29 [0.08, 0.58]	0.84 [0.64, 0.95]
Campbell 2016	65	2	2	6	0.97 [0.90, 1.00]	0.75 [0.35, 0.97]
Cummings 2015	16	3	2	28	0.89 [0.65, 0.99]	0.90 [0.74, 0.98]
Edmiaston 2010	60	74	3	163	0.95 [0.87, 0.99]	0.69 [0.62, 0.75]
Edmiaston 2010	80	54	8	158	0.91 [0.83, 0.96]	0.75 [0.68, 0.80]
Ellis 2013	6	4	2	34	0.75 [0.35, 0.97]	0.89 [0.75, 0.97]
Huhmann 2004	10	3	0	19	1.00 [0.69, 1.00]	0.86 [0.65, 0.97]
Huhn-Matesic 2015	16	8	1	27	0.94 [0.71, 1.00]	0.77 [0.60, 0.90]
Jiang 2019	51	23	12	41	0.81 [0.69, 0.90]	0.64 [0.51, 0.76]
Jiang 2019	52	26	11	38	0.83 [0.71, 0.91]	0.59 [0.46, 0.71]
Perry 2001 Study 1	88	17	6	50	0.94 [0.87, 0.98]	0.75 [0.63, 0.84]
Perry 2001 Study 1	10	3	4	5	0.71 [0.42, 0.92]	0.63 [0.24, 0.91]
Perry 2001 Study 1	36	3	1	28	0.97 [0.86, 1.00]	0.90 [0.74, 0.98]
Perry 2001 Study 1	66	13	32	46	0.67 [0.57, 0.76]	0.78 [0.65, 0.88]
Turner-Lawrence 2009	46	16	2	20	0.96 [0.86, 0.99]	0.56 [0.38, 0.72]



Test 44. Reference test - FEES

Reference test - FEES

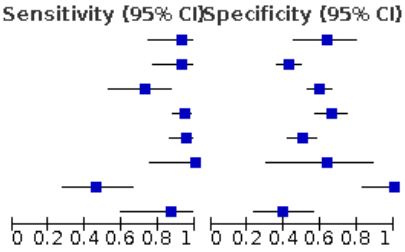
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Lim 2001	26	7	0	17	1.00 [0.87, 1.00]	0.71 [0.49, 0.87]
Lim 2001	22	6	4	18	0.85 [0.65, 0.96]	0.75 [0.53, 0.90]
Trapl 2007	14	5	0	11	1.00 [0.77, 1.00]	0.69 [0.41, 0.89]
Warnecke 2008	60	0	21	19	0.74 [0.63, 0.83]	1.00 [0.82, 1.00]
Warnecke 2008	40	0	41	19	0.49 [0.38, 0.61]	1.00 [0.82, 1.00]



Test 45. Reference test - VF

Reference test - VF

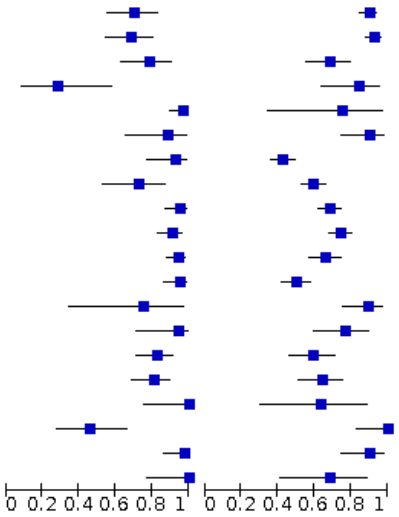
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Daniels 1997	24	12	2	21	0.92 [0.75, 0.99]	0.64 [0.45, 0.80]
Daniels 2016	27	126	2	95	0.93 [0.77, 0.99]	0.43 [0.36, 0.50]
Daniels 2016	21	89	8	132	0.72 [0.53, 0.87]	0.60 [0.53, 0.66]
Edmiaston 2014	100	40	6	79	0.94 [0.88, 0.98]	0.66 [0.57, 0.75]
Edmiaston 2014	57	81	3	82	0.95 [0.86, 0.99]	0.50 [0.42, 0.58]
Martino 2009 Study 1	13	4	0	7	1.00 [0.75, 1.00]	0.64 [0.31, 0.89]
Mulheren 2019	13	0	15	20	0.46 [0.28, 0.66]	1.00 [0.83, 1.00]
Smith 2000	13	23	2	15	0.87 [0.60, 0.98]	0.39 [0.24, 0.57]



Test 46. HCP - nurse

HCP - nurse

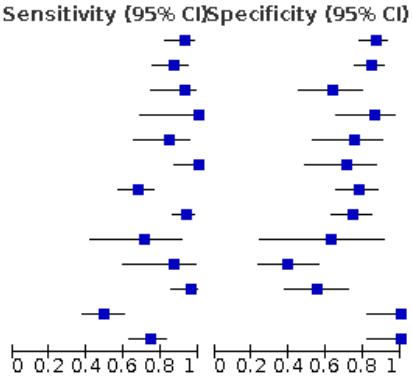
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Behera 2018	33	18	14	161	0.70 [0.55, 0.83]	0.90 [0.85, 0.94]
Behera 2018	39	12	18	156	0.68 [0.55, 0.80]	0.93 [0.88, 0.96]
Bravata 2009	30	20	8	43	0.79 [0.63, 0.90]	0.68 [0.55, 0.79]
Bravata 2009	4	4	10	21	0.29 [0.08, 0.58]	0.84 [0.64, 0.95]
Campbell 2016	65	2	2	6	0.97 [0.90, 1.00]	0.75 [0.35, 0.97]
Cummings 2015	16	3	2	28	0.89 [0.65, 0.99]	0.90 [0.74, 0.98]
Daniels 2016	27	126	2	95	0.93 [0.77, 0.99]	0.43 [0.36, 0.50]
Daniels 2016	21	89	8	132	0.72 [0.53, 0.87]	0.60 [0.53, 0.66]
Edmiaston 2010	60	74	3	163	0.95 [0.87, 0.99]	0.69 [0.62, 0.75]
Edmiaston 2010	80	54	8	158	0.91 [0.83, 0.96]	0.75 [0.68, 0.80]
Edmiaston 2014	100	40	6	79	0.94 [0.88, 0.98]	0.66 [0.57, 0.75]
Edmiaston 2014	57	81	3	82	0.95 [0.86, 0.99]	0.50 [0.42, 0.58]
Ellis 2013	6	4	2	34	0.75 [0.35, 0.97]	0.89 [0.75, 0.97]
Huhn-Matesic 2015	16	8	1	27	0.94 [0.71, 1.00]	0.77 [0.60, 0.90]
Jiang 2019	52	26	11	38	0.83 [0.71, 0.91]	0.59 [0.46, 0.71]
Jiang 2019	51	23	12	41	0.81 [0.69, 0.90]	0.64 [0.51, 0.76]
Martino 2009 Study 1	13	4	0	7	1.00 [0.75, 1.00]	0.64 [0.31, 0.89]
Mulheren 2019	13	0	15	20	0.46 [0.28, 0.66]	1.00 [0.83, 1.00]
Perry 2001 Study 1	36	3	1	28	0.97 [0.86, 1.00]	0.90 [0.74, 0.98]
Trapl 2007	14	5	0	11	1.00 [0.77, 1.00]	0.69 [0.41, 0.89]



Test 47. HCP - other

HCP - other

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Antonios 2010	50	13	4	83	0.93 [0.82, 0.98]	0.86 [0.78, 0.93]
Antonios 2010	47	15	7	81	0.87 [0.75, 0.95]	0.84 [0.76, 0.91]
Daniels 1997	24	12	2	21	0.92 [0.75, 0.99]	0.64 [0.45, 0.80]
Huhmann 2004	10	3	0	19	1.00 [0.69, 1.00]	0.86 [0.65, 0.97]
Lim 2001	22	6	4	18	0.85 [0.65, 0.96]	0.75 [0.53, 0.90]
Lim 2001	26	7	0	17	1.00 [0.87, 1.00]	0.71 [0.49, 0.87]
Perry 2001 Study 1	66	13	32	46	0.67 [0.57, 0.76]	0.78 [0.65, 0.88]
Perry 2001 Study 1	88	17	6	50	0.94 [0.87, 0.98]	0.75 [0.63, 0.84]
Perry 2001 Study 1	10	3	4	5	0.71 [0.42, 0.92]	0.63 [0.24, 0.91]
Smith 2000	13	23	2	15	0.87 [0.60, 0.98]	0.39 [0.24, 0.57]
Turner-Lawrence 2009	46	16	2	20	0.96 [0.86, 0.99]	0.56 [0.38, 0.72]
Warnecke 2008	40	0	41	19	0.49 [0.38, 0.61]	1.00 [0.82, 1.00]
Warnecke 2008	60	0	21	19	0.74 [0.63, 0.83]	1.00 [0.82, 1.00]



ADDITIONAL TABLES

Table 1. Description of screening tests

		Tests (percentage) (n = 37 tests)
Complete or narrative	Complete - 2 × 2 table data extracted	33 (89%)
	Narrative - 2 × 2 table data not available	4 (11%)
Participants		
N[a]	Median (IQR) 100 (50 to 161)	Min = 22, max = 300

Table 1. Description of screening tests *(Continued)*

N with dysphagia/aspiration[a]	Median (IQR) 38 (17 to 63)	Min = 8, max = 106
Mean age	Median (IQR) 67.6 (64.75 to 71.4)	Min = 58.6, max = 76.8
Sex % male	Median (IQR) 54.7 (49 to 65.6)	Min = 35, max = 100
Year recruitment started		Min = 1995, max = 2016
Country	UK and Europe	8 (22%)
	USA and Canada	22 (59%)
	Other	6 (16%)
	Not recorded	1 (3%)
Study design		
Admission to index test	≤ 24 hours	16 (43%)
	> 24 hours and < 72 hours	6 (16%)
	≥ 72 hours	2 (5%)
	Not recorded	13 (35%)
Order tests applied	Index then reference	28 (76%)
	Reference then index	1 (3%)
	Mixed or not specified	5 (14%)
	Not recorded	3 (8%)
Index/reference time interval	≤ 24 hours	18 (49%)
	> 24 hours	10 (27%)
	Not recorded	9 (24%)
Index test training given	Yes	21 (57%)
	Not recorded	16 (43%)
Index test type	Water only	24 (65%)
	Water plus other consistencies	6 (16%)
	Other	7 (19%)
Index test HCP	Nurse	21 (57%)
	Other	16 (43%)
Outcome	Aspiration	15 (41%)

Table 1. Description of screening tests *(Continued)*

	Dysphagia	20 (54%)
	NR	2 (5%)
Reference test	Expert assessment and MASA	20 (54%)
	FEES	6 (16%)
	VF	11 (30%)

[a]Using complete studies only.
FEES: fibreoptic endoscopic evaluation of swallowing.
HCP: healthcare professional.
IQR: interquartile range.
MASA: Mann Assessment of Swallowing Ability.
VF: videofluoroscopy.

Table 2. Index test reference IDs

Test ID	Test name
Antonios 2010a	Modified MASA (MMASA) Neurologist 1: Antonios 2010
Antonios 2010b	Modified MASA (MMASA) Neurologist 2: Antonios 2010
Behera 2018a	DePaul Hospital Swallow Screener (DHSS) Aspiration: Behera 2018
Behera 2018b	DePaul Hospital Swallow Screener (DHSS) Dysphagia: Behera 2018
Bravata 2009a	Nursing Screening Tool: Bravata 2009
Bravata 2009b	Stroke Severity using National Institutes of Health Stroke Scale (NIHSS): Bravata 2009
Campbell 2016	Nursing Bedside Dysphagia Screen (NBDS): Campbell 2016
Cummings 2015	Nurse Dysphagia Screen: Cummings 2015
Daniels 1997	Clinical examination: Daniels 1997
Daniels 2016a	Rapid Aspiration Screening for Suspected Stroke (RAS3): Daniels 2016
Daniels 2016b	Rapid Aspiration Screening for Suspected Stroke (RAS3) - WST only: Daniels 2016
Edmiaston 2010a	Acute Stroke Dysphagia Screening (ASDS) Aspiration: Edmiaston 2010
Edmiaston 2010b	Acute Stroke Dysphagia Screening (ASDS) Dysphagia: Edmiaston 2010
Edmiaston 2014a	Barnes-Jewish Hospital-Stroke Dysphagia Screen (BJH-SDS) Aspiration: Edmiaston 2014
Edmiaston 2014b	Barnes-Jewish Hospital-Stroke Dysphagia Screen (BJH-SDS) Dysphagia: Edmiaston 2014
Ellis 2013	Nursing Bedside Swallowing Screen (NBSS): Ellis 2013
Eren 2019	Barnes-Jewish Hospital Stroke Dysphagia Screen – Turkish version (BJH-SDS) or (T-BJH): Eren 2019

Table 2. Index test reference IDs *(Continued)*

Huhmann 2004	Registered Dietitian (RD) Dysphagia Screening tool: Huhmann 2004
Huhn-Matesic 2015	Edith-Huhn-Matesic Bedside Aspiration Screen (EHMBAS) followed by simple water swallow test: Huhn-Matesic 2015
Jiang 2019a	Chinese version of the modified SSA – original 8 items: Jiang 2019
Jiang 2019b	Chinese version of the modified SSA – reduced 6 items: Jiang 2019
Lim 2001a	Bedside aspiration - Combined WST and Oxygen Saturation: Lim 2001
Lim 2001b	Bedside swallow test - WST only: Lim 2001
Martino 2009	Toronto Bedside Swallowing Screening Test (TOR-BSST): Martino 2009 Study 1
Martino 2014	TOR-BSST water swallow item: Martino 2014
Mulheren 2019	Johns Hopkins Hospital Brain Rescue Unit Modified 3 oz Swallow Screen: Mulheren 2019
Nishiwaki 2005	Clinical swallowing tests: Nishiwaki 2005
Perry 2001a	Gag function - Test3: Perry 2001 Study 1
Perry 2001b	Gag function - Test4: Perry 2001 Study 1
Perry 2001c	Standardized Swallowing Assessment tool (SSA) - Test1: Perry 2001 Study 1
Perry 2001d	Standardized Swallowing Assessment tool (SSA) – Test2: Perry 2001 Study 1
Smith 2000	Oxygen saturation ≥ 2% - Test2 for Aspiration: Smith 2000
Trapl 2007b	Gugging Swallowing Screen (GUSS) - Group2: Trapl 2007
Turner-Lawrence 2009	Emergency Department (ED) dysphagia screen: Turner-Lawrence 2009
Warnecke 2008a	2-step swallowing provocation test (SPT) - step 1 - 0.4 mL: Warnecke 2008
Warnecke 2008b	2-step swallowing provocation test (SPT) - step 2 - 2.0 mL: Warnecke 2008
Zhou 2011	Clinical predicative scale of aspiration (CPSA): Zhou 2011

APPENDICES

Appendix 1. MEDLINE search strategy

1. cerebrovascular disorders/ or basal ganglia cerebrovascular disease/ or exp brain ischemia/ or exp carotid artery diseases/ or exp cerebral small vessel diseases/ or exp intracranial arterial diseases/ or exp "intracranial embolism and thrombosis"/ or exp intracranial hemorrhages/ or stroke/ or exp brain infarction/ or stroke, lacunar/ or vasospasm, intracranial/ or vertebral artery dissection/

2. (stroke$ or poststroke or apoplex$ or cerebral vasc$ or brain vasc$ or cerebrovasc$ or cva$ or SAH).tw.

3. ((brain or cerebr$ or cerebell$ or vertebrobasil$ or hemispher$ or intracran$ or intracerebral or infratentorial or supratentorial or middle cerebral artery or MCA$ or anterior circulation or posterior circulation or basilar artery or vertebral artery or space-occupying) adj5 (isch? emi$ or infarct$ or thrombo$ or emboli$ or occlus$ or hypoxi$)).tw.

4. ((brain$ or cerebr$ or cerebell$ or intracerebral or intracran$ or parenchymal or intraparenchymal or intraventricular or infratentorial or supratentorial or basal gangli$ or putaminal or putamen or posterior fossa or hemispher$ or subarachnoid) adj5 (h?emorrhag$ or h?ematoma$ or bleed$)).tw.

5. hemiplegia/ or exp paresis/ or exp Gait Disorders, Neurologic/

6. (hemipleg$ or hemipar$ or paresis or paraparesis or paretic).tw.

7. or/1-6

8. Deglutition/

9. exp Deglutition Disorders/

10. ((swallow$ or deglutit$ or dysphag$) adj3 (disturbance$ or disorder$ or difficult$ or dysfunction$ or impair$ or condition$ or abnormal$ or damage$ or injur$)).tw.

11. ((swallow$ or deglutit$ or dysphag$) adj3 (scale$ or screen$ or checklist$ or assess$ or exam$ or identif$ or recogni$ or evaluat$ or diagnos$ or detect$ or hazard or risk or test)).tw.

12. exp Respiratory Aspiration/ or exp Pneumonia, Aspiration/

13. ((inhal$ or aspirat$ or ingest$) adj3 (scale$ or screen$ or checklist$ or assess$ or exam$ or identif$ or recogni$ or evaluat$ or diagnos$ or detect$ or hazard or risk or test)).tw.

14. Pharynx/ or pharyngeal muscles/ or esophageal sphincter, upper/ or exp Esophagus/

15. ((throat or oesophag$ or esophag$ or pharyn$ or oropharyn$) adj3 (disturbance$ or disorder$ or difficult$ or dysfunction$ or impair$ or condition$ or abnormal$ or damage$ or injur$)).tw.

16. exp Fluoroscopy/ or bronchoscopy/ or laryngoscopy/ or endoscopy/ or manometry/

17. Fiber Optic Technology/

18. (fluoroscop$ or videofluroscop$ or fluorescence radiation or fluorescent scan$ or fluorophotograph$ or photofluoroscop$ or radiofluoroscop$ or roentgenfluoroscop$ or bronchoscop$ or bronchial endoscop$ or laryngotracheobronchoscop$ or tracheobronchoscop$).tw.

19. or/8-18

20. 7 and 19

Appendix 2. Cochrane Central Register of Controlled Trials (CENTRAL)

June 2017 search – was applied to CENTRAL and HTA

December 2019 search –applied separately to CENTRAL and HTA. HTA was only available until March 2018.

ID Search Hits

#1 MeSH descriptor: [Cerebrovascular Disorders] this term only

#2 MeSH descriptor: [Basal Ganglia Cerebrovascular Disease] this term only

#3 MeSH descriptor: [Brain Ischemia] explode all trees

#4 MeSH descriptor: [Carotid Artery Diseases] explode all trees

#5 MeSH descriptor: [Cerebral Small Vessel Diseases] explode all trees

#6 MeSH descriptor: [Intracranial Arterial Diseases] explode all trees

#7 MeSH descriptor: [Intracranial Embolism and Thrombosis] explode all trees

#8 MeSH descriptor: [Intracranial Hemorrhages] explode all trees

#9 MeSH descriptor: [Stroke] explode all trees

#10 MeSH descriptor: [Vasospasm, Intracranial] this term only

#11 MeSH descriptor: [Vertebral Artery Dissection] this term only

#12 (stroke* or poststroke or apoplex* or cerebral vasc* or brain vasc* or cerebrovasc* or cva* or SAH):ti,ab,kw (Word variations have been searched)

#13 ((brain or cerebr* or cerebell* or vertebrobasil* or hemispher* or intracran* or intracerebral or infratentorial or supratentorial or middle cerebral artery or MCA* or anterior circulation or posterior circulation or basilar artery or vertebral artery or space-occupying) near/5 (isch? emi* or infarct* or thrombo* or emboli* or occlus* or hypoxi*)):ti,ab,kw (Word variations have been searched)

#14 ((brain* or cerebr* or cerebell* or intracerebral or intracran* or parenchymal or intraparenchymal or intraventricular or infratentorial or supratentorial or basal gangli* or putaminal or putamen or posterior fossa or hemispher* or subarachnoid) near/5 (h?emorrhag* or h? ematoma* or bleed*)):ti,ab,kw (Word variations have been searched)

#15 MeSH descriptor: [Hemiplegia] this term only

#16 MeSH descriptor: [Paresis] explode all trees

#17 MeSH descriptor: [Gait Disorders, Neurologic] explode all trees

#18 (hemipleg* or hemipar* or paresis or paraparesis or paretic):ti,ab,kw (Word variations have been searched)

#19 {or #1-#18}

#20 MeSH descriptor: [Deglutition] this term only

#21 MeSH descriptor: [Deglutition Disorders] explode all trees

#22 ((swallow* or deglutit* or dysphag*) near/3 (disturbance* or disorder* or difficult* or dysfunction* or impair* or condition* or abnormal* or damage* or injur*)):ti,ab,kw (Word variations have been searched)

#23 ((swallow* or deglutit* or dysphag*) near/3 (scale* or screen* or checklist* or assess* or exam* or identif* or recogni* or evaluat* or diagnos* or detect* or hazard or risk or test)):ti,ab,kw (Word variations have been searched)

#24 MeSH descriptor: [Respiratory Aspiration] explode all trees

#25 MeSH descriptor: [Pneumonia, Aspiration] explode all trees

#26 ((inhal* or aspirat* or ingest*) near/3 (scale* or screen* or checklist* or assess* or exam* or identif* or recogni* or evaluat* or diagnos* or detect* or hazard or risk or test)):ti,ab,kw (Word variations have been searched)

#27 MeSH descriptor: [Pharynx] this term only

#28 MeSH descriptor: [Pharyngeal Muscles] this term only

#29 MeSH descriptor: [Esophageal Sphincter, Upper] this term only

#30 MeSH descriptor: [Esophagus] explode all trees

#31 ((throat or oesophag* or esophag* or pharyn* or oropharyn*) near/3 (disturbance* or disorder* or difficult* or dysfunction* or impair* or condition* or abnormal* or damage* or injur*)):ti,ab,kw (Word variations have been searched)

#32 MeSH descriptor: [Fluoroscopy] explode all trees

#33 MeSH descriptor: [Bronchoscopy] this term only

#34 MeSH descriptor: [Laryngoscopy] this term only

#35 MeSH descriptor: [Endoscopy] this term only

#36 MeSH descriptor: [Manometry] this term only

#37 MeSH descriptor: [Fiber Optic Technology] this term only

#38 (fluoroscop* or videofluroscop* or fluorescence radiation or fluorescent scan* or fluorophotograph* or photofluoroscop* or radiofluoroscop* or roentgenfluoroscop* or bronchoscop* or bronchial endoscop* or laryngotracheobronchoscop* or tracheobronchoscop*):ti,ab,kw (Word variations have been searched)

#39 {or #20-#38}

#40 #19 and #39

Appendix 3. Embase search strategy

1. cerebrovascular disease/ or brain disease/ or exp basal ganglion hemorrhage/ or exp brain hemangioma/ or exp brain hematoma/ or exp brain hemorrhage/ or exp brain infarction/ or exp brain ischemia/ or exp carotid artery disease/ or exp cerebral artery disease/ or exp cerebrovascular accident/ or exp cerebrovascular malformation/ or exp intracranial aneurysm/ or exp occlusive cerebrovascular disease/ or exp vertebrobasilar insufficiency/

2. (stroke$ or poststroke or apoplex$ or cerebral vasc$ or brain vasc$ or cerebrovasc$ or cva$ or SAH).tw.

3. ((brain or cerebr$ or cerebell$ or vertebrobasil$ or hemispher$ or intracran$ or intracerebral or infratentorial or supratentorial or middle cerebral artery or MCA$ or anterior circulation or posterior circulation or basilar artery or vertebral artery or space-occupying) adj5 (isch? emi$ or infarct$ or thrombo$ or emboli$ or occlus$ or hypoxi$)).tw.

4. ((brain$ or cerebr$ or cerebell$ or intracerebral or intracran$ or parenchymal or intraparenchymal or intraventricular or infratentorial or supratentorial or basal gangli$ or putaminal or putamen or posterior fossa or hemispher$ or subarachnoid) adj5 (h?emorrhag$ or h? ematoma$ or bleed$)).tw.

5. exp hemiplegia/ or exp paresis/ or neurologic gait disorder/

6. (hemipleg$ or hemipar$ or paresis or paraparesis or paretic).tw.

7. or/1-6

8. dysphagia/

9. swallowing/

10. ((swallow$ or deglutit$ or dysphag$) adj3 (disturbance$ or disorder$ or difficult$ or dysfunction$ or impair$ or condition$ or abnormal $ or damage$ or injur$)).tw.

11. ((swallow$ or deglutit$ or dysphag$) adj3 (scale$ or screen$ or checklist$ or assess$ or exam$ or identif$ or recogni$ or evaluat$ or diagnos$ or detect$ or hazard or risk or test)).tw.

12. aspiration pneumonia/ or aspiration/

13. acid aspiration/

14. ((inhal$ or aspirat$ or ingest$) adj3 (scale$ or screen$ or checklist$ or assess$ or exam$ or identif$ or recogni$ or evaluat$ or diagnos $ or detect$ or hazard or risk or test)).tw.

15. exp pharynx/

16. exp esophagus/ or esophagus muscle/ or upper esophagus sphincter/

17. ((throat or oesophag$ or esophag$ or pharyn$ or oropharyn$) adj3 (disturbance$ or disorder$ or difficult$ or dysfunction$ or impair $ or condition$ or abnormal$ or damage$ or injur$)).tw.

18. fluoroscopy/ or fluoroscopy system/

19. exp bronchoscopy/ or bronchus examination/

20. exp laryngoscope/ or exp laryngoscopy/

21. endoscopy/ or fiberscope endoscopy/ or videoendoscopy/

22. manometry/

23. (fluoroscop$ or videofluroscop$ or fluorescence radiation or fluorescent scan$ or fluorophotograph$ or photofluoroscop $ or radiofluoroscop$ or roentgenfluoroscop$ or bronchoscop$ or bronchial endoscop$ or laryngotracheobronchoscop$ or tracheobronchoscop$).tw.

24. or/8-23

25. 7 and 24

Appendix 4. CINAHL search strategy

S1 (MH "Cerebrovascular Disorders") OR (MH "Basal Ganglia Cerebrovascular Disease+") OR (MH "Carotid Artery Diseases+") OR (MH "Cerebral Ischemia+") OR (MH "Cerebral Vasospasm") OR (MH "Intracranial Arterial Diseases+") OR ((MH "Intracranial Embolism and Thrombosis")) OR (MH "Intracranial Hemorrhage+") OR (MH "Stroke") OR (MH "Vertebral Artery Dissections") OR (MH "Stroke Patients") OR (MH "Stroke Units")

S2 TI (stroke or poststroke or post-stroke or cerebrovasc* or brain vasc* or cerebral vasc or cva or apoplex or SAH) or AB (stroke or poststroke or post-stroke or cerebrovasc* or brain vasc* or cerebral vasc or cva or apoplex or SAH)

S3 TI (brain or cerebr* or cerebell* or vertebrobasil* or hemispher* or intracran* or intracerebral or infratentorial or supratentorial or middle cerebral artery or MCA* or anterior circulation or posterior circulation or basilar artery or vertebral artery or space-occupying) N5 TI (ischemi* or ischaemi* or infarct* or thrombo* or emboli* or occlus*)

S4 AB (brain or cerebr* or cerebell* or vertebrobasil* or hemispher* or intracran* or intracerebral or infratentorial or supratentorial or middle cerebral artery or MCA* or anterior circulation or posterior circulation or basilar artery or vertebral artery or space-occupying) N5 AB (ischemi* or ischaemi* or infarct* or thrombo* or emboli* or occlus*)

S5 TI (brain* or cerebr* or cerebell* or intracerebral or intracran* or parenchymal or intraparenchymal or intraventricular or infratentorial or supratentorial or basal gangli* or putaminal or putamen or posterior fossa or hemispher* or subarachnoid) N5 TI (haemorrhage* or hemorrhage* or haematoma* or hematoma* or bleed*)

S6 AB (brain* or cerebr* or cerebell* or intracerebral or intracran* or parenchymal or intraparenchymal or intraventricular or infratentorial or supratentorial or basal gangli* or putaminal or putamen or posterior fossa or hemispher* or subarachnoid) N5 AB (haemorrhage* or hemorrhage* or haematoma* or hematoma* or bleed*)

S7 (MH "Hemiplegia") or (MH "Gait Disorders, Neurologic+")

S8 TI ((hemipleg* or hemipar* or paresis or paretic)) OR AB ((hemipleg* or hemipar* or paresis or paretic))

S9 S1 OR S2 OR S3 OR S4 OR S5 OR S6 OR S7 OR S8

S10 (MH "Deglutition") OR (MH "Gagging")

S11 (MH "Deglutition Disorders")

S12 TI ((swallow* or deglutit* or dysphag*) N5 (disturbance* or disorder* or difficult* or dysfunction* or impair* or condition* or abnormal* or damage* or injur*)) OR AB ((swallow* or deglutit* or dysphag*) N5 (disturbance* or disorder* or difficult* or dysfunction* or impair* or condition* or abnormal* or damage* or injur*))

S13 TI ((swallow* or deglutit* or dysphag*) N5 (scale* or screen* or checklist* or assess* or exam* or identif* or recogni* or evaluat* or diagnos* or detect* or hazard or risk or test)) OR AB ((swallow* or deglutit* or dysphag*) N5 (scale* or screen* or checklist* or assess* or exam* or identif* or recogni* or evaluat* or diagnos* or detect* or hazard or risk or test))

S14 (MH "Aspiration")

S15 (MH "Pneumonia, Aspiration")

S16 (MH "Risk for Aspiration (NANDA)")

S17 TI ((inhal* or aspirat* or ingest*) N5 (scale* or screen* or checklist* or assess* or exam* or identif* or recogni* or evaluat* or diagnos* or detect* or hazard or risk or test)) OR AB ((inhal* or aspirat* or ingest*) N5 (scale* or screen* or checklist* or assess* or exam* or identif* or recogni* or evaluat* or diagnos* or detect* or hazard or risk or test))

S18 (MH "Pharynx+")

S19 (MH "Esophagus")

S20 TI ((throat or oesophag* or esophag* or pharyn* or oropharyn*) N5 (disturbance* or disorder* or difficult* or dysfunction* or impair* or condition* or abnormal* or damage* or injur*)) OR AB ((throat or oesophag* or esophag* or pharyn* or oropharyn*) N5 (disturbance* or disorder* or difficult* or dysfunction* or impair* or condition* or abnormal* or damage* or injur*))

S21 (MH "Fiber Optics")

S22 TI ((fluoroscop* or videofluroscop* or fluorescence radiation or fluorescent scan* or fluorophotograph* or photofluoroscop* or radiofluoroscop* or roentgenfluoroscop* or bronchoscop* or bronchial endoscop* or laryngotracheobronchoscop* or tracheobronchoscop*)) OR AB ((fluoroscop* or videofluroscop* or fluorescence radiation or fluorescent scan* or fluorophotograph* or

photofluoroscop* or radiofluoroscop* or roentgenfluoroscop* or bronchoscop* or bronchial endoscop* or laryngotracheobronchoscop* or tracheobronchoscop*))

S23 S10 OR S11 OR S12 OR S13 OR S14 OR S15 OR S16 OR S17 OR S18 OR S19 OR S20 OR S21 OR S22

S24 S9 AND S23

Appendix 5. HTA search strategy

1. cerebrovascular disorders/ or exp brain ischemia/ or exp carotid artery diseases/ or exp cerebral small vessel diseases/ or exp intracranial arterial diseases/ or exp "intracranial embolism and thrombosis"/ or exp intracranial hemorrhages/ or stroke/ or exp brain infarction/ or vasospasm, intracranial/ or vertebral artery dissection/

2. (stroke or strokes or poststroke or cerebral vasc* or brain vasc$ or cerebrovasc* or cva or SAH) any field

3. ((brain or cerebr* or cerebell* or vertebrobasil* or hemisphere* or intracran* or intracerebral or infratentorial or supratentorial or middle cerebral artery or MCA or anterior circulation or posterior circulation or basilar artery or vertebral artery or space-occupying) and (ischemi* or ischaemi* or infarct* or thrombo* or emboli* or occlus* or hypoxi*)) any field

4. ((brain or cerebr* or cerebell* or intracerebral or intracran* or parenchymal or intraparenchymal or intraventricular or infratentorial or supratentorial or posterior fossa or hemisphere* or subarachnoid) and (haemorrhag* or hemorrhag* or hematoma* or haematoma* or bleed*)) any field

5. exp paresis/ or exp Gait Disorders, Neurologic/

6. (hemipleg* or paresis or paraparesis or paretic) any field

7. or/1-6

8. Deglutition/

9. exp Deglutition Disorders/

10. ((swallow* or deglutit* or dysphag*) and (disturbance* or disorder* or difficult* or dysfunction* or impair* or condition* or abnormal* or damage* or injur*)) any field

11. ((swallow* or deglutit* or dysphag*) and (scale* or screen* or checklist* or assess* or exam* or identif* or recogni* or evaluat* or diagnos* or detect* or hazard or risk or test)) any field

12. exp Pneumonia, Aspiration/

13. ((inhal* or aspirat* or ingest*) and (scale* or screen* or checklist* or assess* or exam* or identif* or recogni* or evaluat* or diagnos* or detect* or hazard or risk or test)) any field

14. Pharynx/ or pharyngeal muscles/ or exp Esophagus/

15. ((throat or oesophag* or esophag* or pharyn* or oropharynx*) and (disturbance* or disorder* or difficult* or dysfunction* or impair* or condition* or abnormal* or damage* or injur*)) any field

16. exp Fluoroscopy/ or bronchoscopy/ or laryngoscopy/ or endoscopy/ or manometry/

17. (fluoroscop* or videofluoroscop* or fluorescence radiation or fluorescent scan* or fluorophotograph* or photofluoroscop* or radiofluoroscop* or roentgenfluoroscop* or bronchoscop* or bronchial endoscop* or laryngotracheobronchoscop* or tracheobronchoscop*) any field

18. or/8-17

29. 7 and 18

Appendix 6. List of targeted grey literature sources (CADTH, 2018)

- Health Quality Council of Alberta (HQCA) http://hqca.ca/studies-and-reviews/completed-reviews/
- Institut national d'excellence en santé et en services sociaux (INESSS) [formerly AETMIS] http://www.inesss.qc.ca/en/publications/publications.html
- Manitoba Centre for Health Policy (MCHP) http://mchp-appserv.cpe.umanitoba.ca/deliverablesList.html
- NLCAHR : Newfoundland and Labrador Centre for Applied Health Research. Contextualized Health Research Synthesis Program (CHRSP) http://www.nlcahr.mun.ca/CHRSP/CompletedCHRSP.php

- Ottawa Hospital Research Institute (OHRI) http://www.ohri.ca/ksgroup/publications.asp
- Therapeutics Initiative. Therapeutics Letter http://www.ti.ubc.ca/TherapeuticsLetter
- University of British Columbia. Centre for Health Services and Policy Research http://chspr.ubc.ca/publications/
- World Health Organization Regional Office for Europe. Health Evidence Network (WHO HEN) http://www.euro.who.int/en/data-and-evidence/evidence-informed-policy-making/publications/by-keyword
- Australian Government. Department of Health and Ageing. Australia and New Zealand Horizon Scanning Network (ANZHSN). http://www.horizonscanning.gov.au/internet/horizon/publishing.nsf/Content/technologies-assessed-lp-2
- Australian Government Department of Health and Ageing. Medical Services Advisory Committee (MSAC). http://www.msac.gov.au/internet/msac/publishing.nsf/Content/completed-assessments
- Joanna Briggs Institute (JBI). The Joanna Briggs Institute EBP Database http://connect.jbiconnectplus.org/Search.aspx
- Monash Health. Centre for Clinical Effectiveness (CCE). Centre for Clinical Effectiveness – Publications http://monashhealth.org/health-professionals/cce/cce-publications/
- Queensland Government (Australia). Health Technology Reference Group. Health Technologies Evaluated-Reports and Briefs (COAG Health Council). https://www.coaghealthcouncil.gov.au/AHMAC/Health-Technology-Reference-Group/Reports-and-Briefs
- Kenniscentrum voor de Gezondheidszorg / Le Centre d'expertise des soins de santé. Belgian Health Care Knowledge Centre (KCE) https://kce.fgov.be/en/all-reports
- Sundhedsstyrelsen. Danish Health and Medicines Authority (DHMA). http://sundhedsstyrelsen.dk/en/publications
- Haute Autorité de santé/ French National Authority for Health (HAS). Haute Autorité de santé http://www.has-sante.fr/portail/jcms/c_946986/en/english-toutes-nos-publications-ligne-principale?portal=r_1457306
- Health Service Executive. Irish Health Repository (Lenus) http://www.lenus.ie/hse/
- De Gezondheidsraad (GR). Health Council of the Netherlands http://www.gezondheidsraad.nl/en/publications
- Zorginstituut Nederland. National Health Care Institute Netherlands https://english.zorginstituutnederland.nl/publications
- Folkehelseinstituttet. Norwegian Institute of Public Health. Publications https://www.fhi.no/en/publ/
- Agència de Qualitat i Avaluació Sanitàries de Catalunya (AQuAS). Agency for Health Quality and Assessment of Catalonia http://aquas.gencat.cat/ca/publicacions/
- Sahlgrenska Universitetssjukhuset. Sahlgrenska University Hospital. Regional activity-based HTA. https://www.sahlgrenska.se/forskning/htacentrum/
- Healthcare Improvement Scotland. http://www.healthcareimprovementscotland.org/
- NHS National Institute for Health and Care Excellence (NICE). http://www.nice.org.uk/
- UK Department of Health (NHS). International Resource for Infection Control (INRIC) https://www.nric.org.uk/resources
- Agency for Healthcare Research and Quality (AHRQ). Technology Assessments https://www.ahrq.gov/research/findings/ta/index.html
- Institute for Clinical and Economic Review (ICER). https://icer-review.org/materials/

Appendix 7. QUADAS-2 tool: risk of bias and applicability judgements

Domain 1. Patient selection	
1. **Risk of bias**	
Describe methods of patient selection:	
· **Was a consecutive or random sample of patients enrolled?**	Yes/No/Unclear
• **Was a case-control design avoided?**	Yes/No/Unclear
• **Did the study avoid inappropriate exclusions?**	Yes/No/Unclear
· **Could the selection of patients have introduced bias?**	RISK: LOW/HIGH/UNCLEAR
1. **Concerns regarding applicability**	
Describe included patients (prior testing, presentation, intended use of index test and setting):	

(Continued)

Is there concern that the included patients do not match the review question?	CONCERN: LOW/HIGH/UNCLEAR
Domain 2. Index test(s) ***(if more than 1 index test was used, please complete for each test)***	
1. **Risk of bias**	
Describe the index test and how it was conducted and interpreted:	
• **Were the index test results interpreted without knowledge of results of the reference standard?**	Yes/No/Unclear
• **If a threshold was used, was it prespecified?**	Yes/No/Unclear
· **Could conduct or interpretation of the index test have introduced bias?**	RISK: LOW/HIGH/UNCLEAR
1. **Concerns regarding applicability**	
Is there concern that the index test, its conduct, or its interpretation differ from the review question?	CONCERN: LOW/HIGH/UNCLEAR
Domain 3. Reference standard	
1. **Risk of bias**	
Describe the reference standard and how it was conducted and interpreted:	
• **Is the reference standard likely to correctly classify the target condition?**	Yes/No/Unclear
• **Were the reference standard results interpreted without knowledge of results of the index test?**	Yes/No/Unclear
· **Could the reference standard, its conduct, or its interpretation have introduced bias?**	RISK: LOW/HIGH/UNCLEAR
1. **Concerns regarding applicability**	
Is there concern that the target condition as defined by the reference standard does not match the review question?	CONCERN: LOW/HIGH/UNCLEAR
Domain 4. Flow and timing	
1. **Risk of bias**	
Describe any patients who did not receive the index test(s) and/or reference standard or who were excluded from the 2 ×2 table (refer to flow diagram): **Describe the time interval and any interventions between index test(s) and reference standard:**	
• **Was there an appropriate interval between index test(s) and reference standard?**	Yes/No/Unclear
• **Did all patients receive a reference standard?**	Yes/No/Unclear
• **Did patients receive the same reference standard?**	Yes/No/Unclear
• **Were all patients included in the analysis?**	Yes/No/Unclear
· **Could the patient flow have introduced bias?**	RISK: LOW/HIGH/UNCLEAR

Appendix 8. Description of screening tests

Bedside Aspiration Test **(** Lim 2001 **a)**

Stage 1: the Water Swallow Test assessed patient's ability to swallow 50 mL of water in 10-mL aliquots. The patient is seated upright and is asked to drink all of the 50 mL. Patients were deemed to be clinically aspirating if they coughed or choked during the water drinking test or had a change in their voice quality following the swallow. Patients were monitored for these signs of possible aspiration for up to 5 minutes after swallowing.

Stage 2: after a rest period of 10 minutes, the Oxygen Saturation Test was conducted. The finger probe of a Novametrix Oxypleth pulse oximeter (model 520A) was placed on the index finger of the unaffected hand of the patient, with the patient in an upright position. After equilibrating for 5 minutes, a baseline oxygen saturation reading was taken. A drop > 2% in arterial oxygen saturation was considered clinically significant (Collins and Bakheit 1997). Maximum and minimum oxygen saturation readings during and for up to 2 minutes after swallowing 10 mL of water were noted 3 times, and highest and lowest readings were taken. If oxygen saturation fell by 5% or more, the test was stopped immediately to stop an additional decline in arterial blood oxygenation levels, especially in patients who may have compromised pulmonary function. Aspiration was deemed present if overt signs of aspiration and/or oxygen desaturation of 2% or more was noted.

Gugging Swallowing Screen **(** Trapl 2007 **b)**

The Bedside Gugging Swallow Screen (GUSS) consists of two parts. Part 1, the preliminary assessment, involved a simple saliva swallow test. Patients who were unable to produce enough saliva because of dry mouth were given saliva spray as a substitute. Vigilance, voluntary cough, throat clearing, and saliva swallowing were assessed. Part 2 consisted of 3 sequentially performed subtests, starting with semisolid (thickened distilled water), then liquid (increasing amounts of 3, 5, 10, 20, and 50 mL), and finally solid textures (5 tests with dry bread). Ten seconds is the time limit for a small solid bolus, including the oral preparatory phase. The test was stopped if 1 of the 4 aspiration signs (deglutition, cough, drooling, and voice change) was positive. Patients must successfully complete all repetitions in the subtest to achieve the full score of 5 points. If a subtest results in < 5 points, the examination is stopped and a special oral diet and/or further investigation by videofluoroscopy or fibreoptic endoscopy is recommended. Diet recommendations are then offered, aligned to the functional oral intake scale, and are given according to points reached in the GUSS.

The Toronto Bedside Swallow Screening Test (Martino 2009 Study 1 ***)***

The Toronto Bedside Swallowing Screening Test (TOR-BSST) included the following tests: the Kidd 50 mL water swallow test, pharyngeal sensation, tongue movement, and general dysphonia split into 2 items - 'voice before' and 'voice after'. A pass/fail response was assigned to each item, so that failure on any item constitutes a positive screen result and therefore increased risk for dysphagia. Administration continues only until the first TOR-BSST item is failed.

WHAT'S NEW

Date	Event	Description
13 July 2021	Amended	Paper submitted with amendments in response to reviewers' comments

HISTORY

Protocol first published: Issue 6, 2017

Date	Event	Description
26 February 2021	Amended	Paper submitted with amendments in response to reviewers' comments
30 September 2020	Amended	Paper submitted
30 January 2017	Amended	Protocol amended in response to reviewers' comments

CONTRIBUTIONS OF AUTHORS

EB, JB, and LH drafted the review and developed the background section with support from CW. PD, AC, MH, and CW, who also provided advice regarding the methodological perspective and content. EB, JB, LH, PD, LL, ML, HR, EM, AA, and CW provided support for data collection and analysis, and MH provided statistical expertise. All review authors commented on all sections of the review and reviewed the final version prior to submission.

DECLARATIONS OF INTEREST

Elizabeth Boaden: none known.
Jane Burnell: none known.
Lucy Hives: none known.
Paola Dey: "my institution has received funding for consultancies and grants in aid from NHS, government and research charities. I have also been in receipt of a travel bursary from the organisers of the United European Gastroenterology Week in Berlin in 2006"
Andrew Clegg: none known.
Mary W Lyons: none known.
C Elizabeth Lightbody: none known.
Margaret A Hurley: none known.
Hazel Roddam: none known.
Elizabeth McInnes: none known.
Anne Alexandrov: none known.
Caroline L Watkins: none known.

SOURCES OF SUPPORT

Internal sources

- No sources of support provided

External sources

- No sources of support provided

DIFFERENCES BETWEEN PROTOCOL AND REVIEW

We did not perform meta-analyses, as the number of studies for each individual index test was insufficient, that is, did not include at least four per group. We have presented only descriptive analyses grouped by general categories. Hence we performed no analysis of the influence of sources of heterogeneity, no subgroup analysis, and no sensitivity analysis. None of the included studies used scintigraphy as the reference test. We included studies in which the index test was carried out by any healthcare professionals with the exception of speech and language therapists (SLTs).