Accuracy and clinical utility of comprehensive dysphagia screening assessments in acute stroke: a systematic review and meta-analysis

Authors:

Jacqueline K. Benfield (BSc (Hons), MSc)

Email: jacqueline.benfield@nottingham.ac.uk

Academic affiliation: Division of Medical Science and Graduate Entry Medicine, School of Medicine, University of Nottingham, United Kingdom

Phone: +44 1332 785891

Lisa F. Everton (BSc, MRes)

Email: lisa.everton@nottingham.ac.uk

Academic affiliation: Division of Clinical Neuroscience, School of Medicine, University of Nottingham, United Kingdom

Phone: +44 115 823 1808

Philip M. Bath (DSc FMedSci)

Email: philip.bath@nottingham.ac.uk

Academic affiliation: Stroke Trials Unit, Division of Clinical Neuroscience, School of Medicine, University of Nottingham, United Kingdom

Second address: Stroke, Nottingham University Hospitals NHS Trust, Nottingham UK

Phone: +44 115 823 1765

Timothy J. England MBChB PhD FRCP

Email: timothy.england@nottingham.ac.uk

Academic affiliation: Division of Medical Science and Graduate Entry Medicine, School of Medicine, University of Nottingham, United Kingdom

Phone: +44 1332 724668

Work performed: School of Medicine, University of Nottingham, United Kingdom

Funding/Grants: PB is a NIHR Senior Investigator.

Jacqueline K. Benfield's PhD Studentship is funded by the National Institute for Health Research (NIHR) and Collaboration for Leadership in Applied Health Research and Care (CLAHRC). The views expressed are those of the author(s) and not necessarily those of the National Health Service, the NIHR, or the Department of Health.

There are no conflicts of interest.

Corresponding Author:

Jacqueline K. Benfield

Vascular Medicine, School of Medicine, University of Nottingham

Royal Derby Hospital Centre

Uttoxeter Road,

Derby

DE22 3DT

Short title: Review of dysphagia screening assessments

Accuracy and clinical utility of comprehensive dysphagia screening assessments in acute stroke: a systematic review and meta-analysis

Abstract

Introduction: Nurses and other non-specialists in dysphagia are often trained to screen swallowing post-stroke. There are many basic tools that test water only, they are usually conservative and patients that fail the test remain nil by mouth until a speech and language therapy assessment. More comprehensive tests also allow non-specialists to recommend modified oral intake. Little is known about the accuracy, clinical utility and cost effectiveness of these tests.

Methods: Following PRISMA guidelines, a systematic review was conducted to describe comprehensive swallowing tests that are available for use in acute stroke by nurses or other non-specialists in dysphagia. A meta-analysis was performed to evaluate accuracy and considered their clinical utility. Searches and analyses, conducted by two reviewers, included MEDLINE, EMBASE, trial registries, and grey literature up to December 2018. Validated studies were assessed for quality and risk of bias using QUADAS-2.

Results: Twenty studies were included, describing five different tests, three of which had undergone validation. The tests varied in content, recommendations and use. There was no test superior in accuracy and clinical utility. Three studies validating the Gugging Swallow Screen provided sufficient data for meta-analysis, demonstrating high sensitivity; 96% (95%CI 0.90-0.99) but low specificity, 65% (95%CI 0.47-0.79) in line with many water swallow tests. Results should be interpreted with caution as study quality and applicability to the acute stroke population was poor.

Conclusions: There is no comprehensive nurse dysphagia assessment tool that has robustly demonstrated good accuracy, clinical utility and cost effectiveness in acute stroke.

Relevance to Clinical Practice: Nurses and other clinicians can develop competencies in screening swallowing and assessing for safe oral intake in those with post stroke dysphagia. It is important to use a validated assessment tool that demonstrates good accuracy, clinical utility and cost effectiveness.

What does this paper contribute to the wider global clinical community?

- A description of how nurses and other clinicians are involved in screening and assessment of swallowing after acute stroke

- A summary and critique of the available tools for nurses and other clinicians to screen and assess swallowing within the acute stroke pathway
- An idea of how nurse-based screening and assessment of swallow post acute stroke might impact on patient outcomes

Key words: swallowing, dysphagia, assessment, screening, stroke, multidisciplinary, nurse

Introduction

Post stroke dysphagia is common, affecting around 50% of acute stroke patients (Martino et al., 2005). Early identification is key to reduce rates of stroke associated pneumonia and mortality (Bray et al., 2016; Yeh et al., 2011). Speech and Language Therapists (SLT) are, in many countries, considered to be the specialists in assessment and management of dysphagia. However, swallow screening tools such as water swallow tests are often used by non-specialists in dysphagia, including nurses, to identify patients at risk of aspiration and refer patients for further assessment by SLT. There are a multitude of screening tools described in the literature and systematic reviews have demonstrated that some of the best tools have good sensitivity but often lower specificity (Schepp, Tirschwell, Miller, & Longstreth, 2012). This translates to many patients unnecessarily remaining nil by mouth (NBM) for prolonged periods, with or without nasogastric tube feeding, until they are assessed by a SLT, which can have negative consequences (Langdon, Lee, & Binns, 2007; Langmore, Krisciunas, Miloro, Evans, & Cheng, 2012). Water swallow tests have been criticised because swallowing water is not the same as swallowing food (Marques, De Rosso, & Andre, 2008) and the tools have often been validated for screening aspiration, one of the possible consequences of dysphagia, rather than for the presence of dysphagia itself (Sasaki & Leder, 2014). Reduced efficiency or uncontrolled oral and pharyngeal transit and clearance, impaired mastication and reduced sensation may result in other symptoms such as choking and sub-optimal nutrition (Serra-Prat et al., 2012; Smithard et al., 1996). Aside from water swallow tests, there are several more comprehensive swallowing tests that mean non-specialists can screen for dysphagia and also assess various diet and fluid consistencies, so safe oral intake may be commenced earlier. To date there has been no review identifying, describing or comparing these more comprehensive tests.

It is essential that dysphagia screening tools have adequate accuracy and are safe to use clinically. The UK National Institute of Health and Care Excellence (NICE) 2016 Stroke Guidelines say that swallowing should be assessed using a validated tool (Royal College of Physicians Intercollegiate Stroke Working

Party, 2016). There is also a move to demonstrate the clinical utility of screening and diagnostic tests, not only the technical performance in accurately screening for and diagnosing a condition (Bossuyt, Reitsma, Linnet, & Moons, 2012; Thompson, Plüddemann, Price, & Heneghan, 2013). In the case of patients with dysphagia, clinical utility refers to how the tests improve the clinical outcomes of the patients such as pneumonia rates and be more cost effective than other tools or pathways.

Aims

A systematic review was conducted to describe the comprehensive tools that are available for nurses or other members of the multidisciplinary team (MDT) to screen swallowing and assess for safe oral intake post stroke. The clinical utility of the tests is described, the results of a meta-analysis are presented and the quality of the tools that had undergone validation is discussed.

Methods

A systematic review of the English and Spanish literature was completed by searching databases; MEDLINE, EMBASE, CINAHL, Web of Science, Trial databases; Clinicaltrials.gov, ICTRP and grey literature from start to October 2018. See Appendix 1 for an example of the search criteria used in EMBASE. The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (Moher, Liberati, Tetzlaff, & Altman, 2009) guidance was followed (see completed checklist in Supplementary File 1). Identified studies at different stages of the process were managed in folders on EndNote.

Inclusion criteria

Inclusion criteria for the narrative review was broad as the number of published tools was estimated to be small. Studies were included in the narrative review if they had sufficient information in English or Spanish to establish that they described a comprehensive nursing or MDT assessment of swallowing to screen for dysphagia in stroke patients. Comprehensive assessment was defined as a screening test for dysphagia that included assessing more than one diet or fluid texture allowing for recommendations of modified diet and fluids where appropriate. For the quantitative analysis, studies were included if they gave data regarding the accuracy of the assessment tool such as sensitivity and specificity. Studies were also included that reported the cost effectiveness or clinical utility of a test.

Study selection

One reviewer (JB) searched the titles and abstracts and excluded non-relevant studies. Full text was requested for relevant studies that could be included in a narrative review and, in the case of validation studies, a quantitative review. Data extraction and assessment of quality were carried out by the same reviewer (JB). Decisions for inclusion and exclusion, based on eligibility criteria were discussed and agreed with a second reviewer (LE). The second reviewer (LE) also reviewed and

1 agreed the data extraction and quality assessments. Any disagreements were discussed with a third
2 reviewer (TE).

3 Data extraction

4 Data were extracted using a predesigned form (Appendix 2) including information on the content of
5 the tests, possible outcomes, who administers the test and what training they require. For validation
6 papers data were collected using the Revised Quality Assessment of Diagnostic Accuracy Studies
7 (QUADAS-2) (Whiting, Rutjes, Westwood, & et al., 2011) on the gold standard reference test used,
8 the time between assessments, whether blinding occurred for example (Appendix 3). Authors were
9 contacted where details were unclear or data were not present.

10 Risk of Bias

11 Risk of bias and applicability of primary diagnostic accuracy studies were assessed using the four
12 domains of the QUADAS-2 (Whiting et al., 2011). 1. Patient selection; were patients recruited
13 consecutively? Were they representative of an acute stroke setting? 2. Index test; who carried out
14 the index test (the test being validated)? Were they blinded to other tests? 3. Reference standard;
15 was the gold standard an acceptable assessment to compare to? Were the assessors blinded to the
16 results of the index text? 4. Flow and timing; what was the time between the index and reference
17 test? Was all data (including missing data) reported? Prior to the quality assessment it was decided
18 that to be classed as low concern for applicability to an acute stroke population, over 50% of
19 participants in the sample needed to be representative of acute stroke patients; defined as newly
20 admitted (less than one week post stroke), including all types and severities of stroke and who may
21 or may not have dysphagia. Overall quality was summarised using the GRADE guidelines
22 (Schunemann et al., 2008).

23 Statistical Analysis

24 Diagnostic accuracy data for the studies validating an assessment tool were summarised. Sensitivity,
25 specificity, positive and negative predictive values, receiver operating characteristic curve, inter and
26 intra rater reliability were included where available as were the respective confidence intervals
27 which gives an indication of consistency. Studies that reported a 2x2 table detailing numbers of true
28 and false positives and negatives were included in a meta-analysis(Macaskill P, 2010). The data were
29 analysed in STATA using a hierarchical model (HSROC) accounting for both within- and between
30 study heterogeneity (Lee, Kim, Choi, Huh, & Park, 2015; Macaskill P, 2010; Takwoingi, April 2016).

Results

Database searches identified 868 studies and grey literature searches found a further 48 studies. See PRISMA flow diagram (Fig. 1). After duplicates were removed and titles and abstracts were screened, 60 full texts were requested and reviewed. After exclusions, 20 met the criteria for the narrative review.

Identified tests

Five tests were identified and are summarised in Table 1. They are described as tests, screening tools and assessments. They all met the criteria as a screening tool for dysphagia and included testing different consistencies so that those who fail with water but can safely manage some oral intake can be recommended modified diet and fluids whilst they wait for further assessment by SLT. The Gugging Swallow Screening (GUSS, n=11 publications) and Volume Viscosity Swallowing Test (VVST, n=4) advise the use of instrumental assessments if dysphagia is identified on the test (Clave et al., 2008; Trapl et al., 2007). The VVST and the Dysphagia Trained Nurse Assessment (DTNAX, n=3 publications) suggest they can also be used to review patient's swallowing (Guillen-Sola et al., 2013; Mary Heritage, 2003).

Non-swallow section

The GUSS, Bedside Swallow Screening Test (BESST, n=1) and DTNAX include a non-swallow section at the beginning before offering any oral trials (Benfield, 2018; E. Boaden, 2011; Trapl et al., 2007). This varies from observation of respiration, swallowing and alertness levels to direct testing of oromotor function. If this section is failed in the GUSS and the BESST then the rest of the assessment is not administered and the patient remains NBM. In the DTNAX, whether the non-swallow section was passed or failed the assessment proceeds to swallow trials. The VVST and Two Volume Three Texture Test (2v/3t-P, n=1) do not include a non-swallow section and the papers are not clear about whether there are any patients who are not suitable to be tested (Cocho et al., 2017; Rofes, Arreola, Mukherjee, & Clave, 2014).

Oral trials

Many countries have not adopted the International Dysphagia Descriptors Standardisation Initiative (IDDSI) framework (The International Dysphagia Diet Standardisation Initiative, 2016) and many of these tests were devised before IDDSI was launched in 2015. The DTNAX and the GUSS have been converted to the IDDSI framework (M. Heritage, 2001; Trapl et al., 2007). The oral trials will be described within in the IDDSI framework, levels (L) 1 to 7, where possible.

The tests vary in what is given orally. BESST evaluates two consistencies only, thin fluids (L0) and puree diet (L4). Whereas the GUSS trials thin fluids (L0), regular diet (L7) and a 'semi-solid' texture (L3). The 2v/3t-P tests different volumes (5 & 10mls) of thin fluids (L0), 'semisolid' (estimated L3 or L4) textures and regular diet (L7). The VVST tests different volumes (5, 10, 20mls) of thin fluids (L0), puree diet (L4) and nectar fluids (could be approximated to L2 fluids). The DTNAx is more comprehensive and tests a range of fluid volumes (5, 10, 100mls) and viscosities (L0, L2, and L3) and food textures (L4, L5, L6, L7). Several of the tests (Clave et al., 2008; Cocho et al., 2017; Trapl et al., 2007) comment on the order of the oral trials and argue that starting with thin fluids may result in aspiration and therefore they begin with puree diet (L4) or thickened fluids.

Justification for the inclusion of different textures was a theme that emerged from the literature. Umay et al 2018 points out that water is not the only thing that patients' swallow thereby only testing water may result in false positives (Umay et al., 2018). Boaden 2011 argues that a sufficient quantity of thin fluids needs to be included in the test because small amounts of water are not representative of normal swallowing (E. Boaden, 2011). Ferreira et al 2018 suggests that assessing different consistencies is more representative of normal eating habits (Ferreira et al., 2018). St John et al 2015 et al describe how the GUSS has replaced a water swallow test in one stroke centre because the team were concerned about the safety of starting patients on diet after just being tested with water (John & Berger, 2015).

Criteria for detecting aspiration or dysphagia

Most of the tests use clinical judgements to determine presence of aspiration or dysphagia. In particular, all tests use presence of cough and voice changes and most use lack of laryngeal elevation (E. Boaden, 2011; Cocho et al., 2017; M. Heritage, 2001; Trapl et al., 2007). In addition, the VVST and 2v/3t-P use a drop in oxygen saturations of >2% to detect silent aspiration. Other criteria varied between tests, see Table 2 for details of the full criteria each test uses to determine aspiration or dysphagia.

Outcomes

A common theme was highlighted in the literature; non-expert professionals can use the tests to commence patients on safe oral intake who would otherwise remain NBM from failing a water swallow test (E. Boaden, 2011; Cocho et al., 2017; Guillen-Sola et al., 2013; Mary Heritage, 2003; John & Berger, 2015; Trapl et al., 2007).

The outcome of the tests can be: 1. Pass - where normal diet and fluids are recommended, 2. Fail – where the patient is recommended to remain NBM or 3. Fail - with recommendations of a modified diet and fluids. The more comprehensive the test the wider the range of modified diet and fluids

recommendations. The BESST only recommends puree/pudding (L4) consistency as the modified option. The VVST and 2v/3t-P can recommend different volumes of thin fluids (L0), thickened fluid and pureed diet (L4). The DTNAX can recommend several different thickened fluids and a range of modified diets. Several tests recommend textures that have not been directly tested; the BESST and VVST allow recommendations of normal diet (L7) when only puree texture is assessed and the GUSS recommends 'soft food' and different levels of thickened fluids when only thin (L0), puree/pudding (L4) and normal diet (L7) are tested.

None of the studies validating the tools collected outcomes of the patients following the initial index and reference tests.

Administration time

The GUSS, VVST, BESST are reported to take between 5-10 minutes to administer, the DTNAX takes around 20 minutes and there is no information on the 2v/3t-P test.

Pathway

Three of the tests (E. Boaden, 2011; Cocho et al., 2017; M. Heritage, 2001) have been designed and, in some cases (E. Boaden, 2011), validated to be the initial swallow test an acute stroke patient receives before commencing oral intake. The other two (Rofes et al., 2014; Trapl et al., 2007) are intended to be used after an initial screening to identify those at risk of dysphagia who need a more detailed test.

Profession, training and competence

The tests are designed to be carried out by non-specialists in dysphagia, in most cases nurses (E. Boaden, 2011; Cocho et al., 2017; Mary Heritage, 2003; Trapl et al., 2007) but also physicians and dietitians (Rofes et al., 2014). The GUSS and the VVST papers suggest it can also be used by SLTs as a standardised bedside assessment (Ferreira et al., 2018; Rofes et al., 2014).

Little is known about the training required in order to be able to administer the tests, from what has been documented the training received is very variable. The BESST requires no training. The GUSS required a short theory session and an observation of the test being administered. The VVST and DTNAX require theory and practical sessions using the test. The DTNAX includes an assessment of competency in administering the test by an SLT.

Accuracy

Three of the identified tools (E. Boaden, 2011; Clave et al., 2008; Trapl et al., 2007) have undergone validity and reliability testing (Table 3) and the DTNAX is currently being validated – (Clinical trials.gov NCT03700853). The GUSS and the VVST used an instrumental assessment (FEES or VFS) as

the gold standard to validate the tests, the BESST was validated against an experienced SLT performing a clinical bedside assessment. All tests demonstrated good sensitivity (87.5% - 100%) and variable specificity (28% - 96.1%). The lower levels of specificity came from the VVST for identifying aspiration (Clave et al., 2008; Guillen-Sola et al., 2013; Rofes et al., 2014) but sensitivity (94%, 95% confidence interval, CI, 0.87–0.98) and specificity (88%, 95% CI, 0.50–0.99) for identifying dysphagia was higher (Rofes et al., 2014).

Only three of the studies (AbdelHamid & Abo-Hasseba, 2017; Trapl et al., 2007; Warnecke et al., 2017), all validating GUSS, reported detailed data that could be included in a meta-analysis. Figure 2 compares validation data across these studies; overall, GUSS sensitivity and specificity was 0.96 (CI 95% 0.90 - 0.99) and 0.65 (CI 95% 0.47 - 0.79) respectively. The summary receiving operator curve (SROC) could not be estimated due to there being less than 4 studies.

Quality

Most studies demonstrated very low (AbdelHamid & Abo-Hasseba, 2017; Clave et al., 2008; Ferreira et al., 2018; Guillen-Sola et al., 2013; Rofes et al., 2014; Samia E S B, 2017; Umay et al., 2018; Warnecke et al., 2017) or low quality (Trapl et al., 2007) according to the QUADAS-2 (Whiting et al., 2011) and GRADE criteria (Schunemann et al., 2008). Table 4 shows the risk of bias and concern for applicability of each test along with the level of quality. Reduced quality was due to concern or uncertainty regarding risk of bias or applicability of index test, reference test, patient selection methods or flow and timing. The study validating BESST (E. Boaden, 2011) demonstrated good study design, accuracy and reliability but was scored as moderate quality due to lack of a gold standard reference test and imprecise results with wide confidence intervals..

Clinical utility and cost effectiveness

We did not find any studies evaluating the cost effectiveness of these tools over other tools or pathways. However, several studies evaluated the effect of using these more comprehensive tests on the clinical outcomes of patients.

In a retrospective study (N=384) (Palli et al., 2017), the GUSS test was introduced into a stroke service during out of hours periods where no SLTs were available to assess and manage swallowing. This resulted in significantly reduced pneumonia rates from 11.6% before the introduction to 3.8% after (p=0.004). Median length of hospital stay also decreased from 9 days to 8 days (p=0.033). However, in another retrospective database study (N=1394) (Teuschl et al., 2018) there were no differences in pneumonia rates between patients admitted with a stroke and assessed with GUSS (5.0%) and those not assessed (5.5%). Due its methodological design, groups were not matched therefore limited conclusions can be drawn. The 2v/3t-P test also resulted in a significant reduction

in pneumonia rates (6.2% before vs. 2.1% after, $p = 0.05$) in a prospective analysis of consecutively admitted patients ($N=418$) to the stroke unit when it replaced a water swallow test (Cocho et al., 2017). A published clinical audit ($N=61$) described how acute patients were seen quicker and the number of days they spent NBM dropped by over 30% following a fivefold increase in the number of nurses trained to perform the DTNAx (M. Heritage, 2001).

Discussion

Nurses and other non-specialists in dysphagia assess swallowing and recommend diet and fluid intake in post stroke patients. Little is known about the content, accuracy or the way these assessments are carried out. It is important that the tools used during these assessments have undergone validation to ensure they are accurate in identifying dysphagia and that patients are being recommended safe oral intake to prevent complications such as aspiration pneumonia, choking or undernourishment.

We conducted a systematic review to identify and describe the available tools and compare their accuracy and clinical utility where this had been tested. Five different tests were identified from the literature. The tests differed in content, the recommendations generated, the professionals administering the test and the training and competency requirements. Only three of the tests have been validated against a gold standard swallowing assessment. There was no single test that was highly accurate, backed up with a high-quality study design and that demonstrated clinical utility.

The GUSS has undergone the most validation testing of all the tests and was the only test in the studies identified that was eligible for the meta-analysis. Over all it demonstrated good sensitivity (96%) and lower specificity (65%). These pooled results represent the overall ability of the GUSS to identify risk of aspiration rather than dysphagia as not all of the studies validated the test for identification of dysphagia (AbdelHamid & Abo-Hasseba, 2017; Trapl et al., 2007). It is possible therefore that some of the patients who pass the test in fact have dysphagia and are at risk of choking or undernutrition. The VVST had the highest accuracy for identification of dysphagia (Rofes et al., 2014).

The accuracy results for the meta-analysis must be interpreted cautiously due to the limited number of studies and the mostly poor or very poor quality or applicability. Two of the studies selected patients who were already suspected as having dysphagia (AbdelHamid & Abo-Hasseba, 2017; Trapl et al., 2007) and one excluded mild strokes (Warnecke et al., 2017) therefore they did not represent the broad range of the acute stroke population in whom the test may be used. In two of the studies (AbdelHamid & Abo-Hasseba, 2017; Warnecke et al., 2017) the GUSS was performed by experts rather than non-specialists which is not applicable to the clinical use of the test. In one of the

1 studies (AbdelHamid & Abo-Hasseba, 2017) the timing was unclear between the GUSS and the
2 reference test (FEES) and there was no reporting of any missing data. Individually and to some
3 extent in the pooled data the studies demonstrated imprecise results with wide confidence intervals
4 especially with specificity. The issues with quality could have skewed the results; for example, the
5 high sensitivity may in part be due to the test only being carried out on participants already
6 identified as being at risk of dysphagia (AbdelHamid & Abo-Hasseba, 2017; Trapl et al., 2007) or with
7 more severe strokes (Warnecke et al., 2017). The strict non-swallow section which results in a test
8 failure for those with reduced oromotor function and places the patient NBM until further
9 assessment might explain the low specificity (Warnecke et al., 2017; Wirth et al., 2013). This
10 specificity is comparable to some of the best water swallow tests (Schepp, Tirschwell, Miller, &
11 Longstreth Jr, 2012). From a clinical utility perspective the GUSS may be better than no test (Palli et
12 al., 2017) but not better than a water swallow test (Teuschl et al., 2018) at reducing pneumonia
13 rates. There is also a jump between the diet and fluid consistencies tested to those recommended;
14 for example, a patient can be recommended IDDSI L1 or L2 fluids and L5 diet without having been
15 tested with any of these. In the same way, water swallow tests are also criticised for allowing normal
16 diet intake without assessment (Marques et al., 2008). Given it may not be any more accurate, safe
17 or clinically effective than water swallow tests, and training and administration time is greater, the
18 GUSS may be less cost effective.

19 The BESST was of moderate quality and had acceptable sensitivity and negative predictive value with
20 lower specificity to identifying dysphagia. However, the reference test used was a clinical bedside
21 assessment (CBA) which could be argued is not a gold standard assessment of swallowing, especially
22 because a validated CBA was not used. CBA have been shown to be less effective at describing
23 dysphagia and identifying aspiration (Splaingard, Hutchins, Sulton, & Chaudhuri, 1988) than gold
24 standard instrumental assessments and the author acknowledges this as a limitation with the BESST
25 validation.

26 The construct validity of the tests has not been reported. This pertains to how well a test is
27 constructed to identify dysphagia based on what is known about dysphagia. There are some
28 common characteristics across the tests that suggests good construct validity: all of the tests
29 evaluate liquids and solids; and they all have criteria for judging both the oral stage and pharyngeal
30 stages of swallowing. This includes specifics on identifying signs of aspiration such as cough and
31 voice change which have been shown to be the most reliable signs in water swallow tests (Brodsky
32 et al., 2016). Progressive volumes of thin fluids also increases accuracy of identifying aspiration
33 (Brodsky et al., 2016), most of the tests do this to some degree. However, there are limitations in
34 some of the tests that reduce their construct validity. Two of the tests do not include food

1 textures that are part of regular diet (E. Boaden, 2011; Clave et al., 2008). Also, it has been
2 established that bedside assessments are limited in detecting silent aspiration (McCullough et al.,
3 2005). Two of the tests have tried to address this by including pulse oximetry to measure a drop in
4 oxygen saturation, however more recently this measure has been found not to be reliable in
5 detecting aspiration (Wang, Chang, Chen, & Hsiao, 2005). These tests are designed to identify
6 dysphagia with aspiration being one aspect of that and silent aspirators may present with other signs
7 of dysphagia (Ramsey, Smithard, & Kalra, 2005). This may limit the potential of any bedside test to
8 attain high accuracy scores for identification of aspiration as to date there is no non-instrumental
9 test that has been found to identify aspiration reliably.

10 Both the VVST and the GUSS follow on from a preliminary screening component to identify those
11 who may be at risk of aspiration or dysphagia. The whole pathway (preliminary screen and test) has
12 not been validated with consecutively admitted acute stroke patients for either of these tests.
13 Perhaps this could be a more cost-effective pathway if both preliminary screening and then
14 dysphagia testing are shown to be acceptable in diagnostic accuracy in methodologically robust
15 studies.

16 Heritage 2003 argues that to manage dysphagia effectively SLTs need to share their skills,
17 responsibility and workload with nurses (Mary Heritage, 2003). Several publications suggested
18 screening tests were not designed to replace the role of the SLT (23). Instead they were meant as
19 easy-to-follow tools for those best placed (30) with the best skills (21) to identify patients with
20 dysphagia so that SLT resources could be better directed to assessment and management of those
21 most in need (20). The Interprofessional Dysphagia Framework (IDF) sets out how non-SLTs can
22 develop skills in dysphagia assessment and management at different levels (E Boaden & Davies,
23 2008). The Foundation Level of training allows those competent, to carry out a protocol-guided
24 swallowing assessment for which training and competency verification is required. The level of
25 training required is set at a high standard because these tests involve making clinical judgements on
26 signs of dysphagia and aspiration that may be subtle. In the UK, SLTs develop these skills by
27 completing at least an undergraduate module and post graduate training in the theory of dysphagia
28 and must accumulate 40 hours of clinical experience to be competent to practice (RCSLT, 2014).
29 Training must therefore be essential if non-SLTs are assessing dysphagia. Whether training was
30 required to use the tests identified in this review appeared variable and the DTNAX is the only tool
31 that has described a training and competency assessment that meets the IDF's criteria.

Limitations

This review only included studies published in English or Spanish, therefore published and non-published studies in other languages describing assessment tools may have been missed. There are likely to be many other nurse dysphagia assessments that have been developed by individual services that have not been published or described in the literature and therefore have not been included in this review. It is unlikely, however, that these in-house assessments have undergone rigorous validation without publication.

Future directions

To make decisions around which test is superior in diagnostic accuracy, further validation using robust study design is required. Information regarding clinical utility and cost effectiveness is also desirable to use with accuracy data to determine which tools should be used as standard in routine clinical practice. All the tests and gold standard comparators evaluate only small volumes of oral intake in order to make appropriate recommendations; however, little is known about how the recommendations are tolerated over time and whether there are any negative consequences such as pneumonia, choking incidents and malnutrition. Further studies should consider comparing tools using clinical outcomes at later time-points to ensure the tools are safe and effective. Future hyper-acute clinical trials may benefit from a robustly validated outcome tool that can be used by non-specialists to identify dysphagia (Cohen et al., 2016).

Conclusions

There are several tools used by nurses and other non-specialists to screen for dysphagia and recommend oral intake for acute stroke patients with mild to moderate dysphagia. Three have been validated and show that they are good at identifying patients at risk of aspiration and dysphagia, but often over diagnose, resulting in patients unnecessarily being kept NBM or on modified oral intake. Overall, however, the quality of studies in this review was graded as poor or showing low applicability for use by non-specialists to assess for dysphagia within the acute stroke setting. There is limited variable quality evidence that these tests may reduce pneumonia, reduce length of time patients are NBM and awaiting a swallowing assessment compared to no test. Further validation is required with robust study design to discover the accuracy, clinical utility and cost effectiveness of these tests so that they can be evaluated and compared.

References

- AbdelHamid, A., & Abo-Hasseba, A. (2017). Application of the GUSS test on adult Egyptian dysphagic patients. *The Egyptian Journal of Otolaryngology*, 33(1), 103-110. doi:10.4103/1012-5574.199419
- Benfield, J. K. (2018). Validation and Reliability testing of a Dysphagia Trained Nurse Assessment.
- Boaden, E. (2011). *Improving the identification and management of aspiration after stroke* (Doctorate in Philosophy). Improving the identification and management of aspiration after stroke,
- Boaden, E., & Davies, S. (2008). Interprofessional Dysphagia Framework. *National Dysphagia Competence Steering Group*.
- Bossuyt, P. M. M., Reitsma, J. B., Linnet, K., & Moons, K. G. M. (2012). Beyond diagnostic accuracy: the clinical utility of diagnostic tests. *Clinical chemistry*, 58(12), 1636. doi:10.1373/clinchem.2012.182576
- Bray, B. D., Smith, C. J., Cloud, G. C., Enderby, P., James, M., Paley, L., . . . Rudd, A. G. (2016). The association between delays in screening for and assessing dysphagia after acute stroke, and the risk of stroke-associated pneumonia. *Journal of Neurology, Neurosurgery & Psychiatry*, 88(1). doi:10.1136/jnnp-2016-313356
- Brodsky, M. B., Suiter, D. M., Gonzalez-Fernandez, M., Michtalik, H. J., Frymark, T. B., Venediktov, R., & Schooling, T. (2016). Screening Accuracy for Aspiration Using Bedside Water Swallow Tests: A Systematic Review and Meta-Analysis. *Chest*, 150(1), 148-163. doi:10.1016/j.chest.2016.03.059
- Clave, P., Arreola, V., Romea, M., Medina, L., Palomera, E., & Serra-Prat, M. (2008). Accuracy of the volume-viscosity swallow test for clinical screening of oropharyngeal dysphagia and aspiration. *Clin Nutr*, 27(6), 806-815. doi:10.1016/j.clnu.2008.06.011
- Cocho, D., Sagales, M., Cobo, M., Homs, I., Serra, J., Pou, M., . . . Otermin, P. (2017). Lowering bronchoaspiration rate in an acute stroke unit by means of a 2 volume/3 Acute ischaemic texture dysphagia screening test with pulse oximetry. *Neurologia*, 32(1), 22-28. doi:10.1016/j.nrl.2014.12.005
- Cohen, D. L., Roffe, C., Beavan, J., Blackett, B., Fairfield, C. A., Hamdy, S., . . . Bath, P. M. (2016). Post-stroke dysphagia: A review and design considerations for future trials. In (Vol. 11, pp. 399-411). London, England.
- Ferreira, A. M. d. S., Pierdevara, L., Ventura, I. M., Gracias, A. M. B., Marques, J. M. F., & Reis, S. M. P. M. (2018). The Gugging Swallowing Screen : A contribution to the cultural and linguistic validation for the Portuguese context. *Revista de Enfermagem Referência*, IV - n.º 16.
- Guillen-Sola, A., Marco, E., Martinez-Orfila, J., Mejias, M. F. D., Passalacqua, M. D., Duarte, E., & Escalada, F. (2013). Usefulness of the volume-viscosity swallow test for screening dysphagia in subacute stroke patients in rehabilitation income. *NeuroRehabilitation*, 33(4), 631-638. doi:10.3233/nre-130997
- Heritage, M. (2001). A collaborative approach to the assessment and management of dysphagia. *International Journal of Language & Communication Disorders*, 36 Suppl, 369.
- Heritage, M. (2003). Swallowing Caseloads. *RCSLT Bulletin*, April 2003.
- John, J. S., & Berger, L. (2015). Using the gugging swallowing screen (GUSS) for dysphagia screening in acute stroke patients. *J Contin Educ Nurs*, 46(3), 103-104. doi:10.3928/00220124-20150220-12
- Langdon, C., Lee, A., & Binns, C. (2007). High incidence of respiratory infections in nil by mouth acute stroke patients. *Dysphagia*, 22(4), 358-358.
- Langmore, S., Krisciunas, G. P., Miloro, K. V., Evans, S. R., & Cheng, D. M. (2012). Does PEG use cause dysphagia in head and neck cancer patients? *Dysphagia*, 27(2), 251-259. doi:10.1007/s00455-011-9360-2
- Lee, J., Kim, K. W., Choi, S. H., Huh, J., & Park, S. H. (2015). Systematic Review and Meta-Analysis of Studies Evaluating Diagnostic Test Accuracy: A Practical Review for Clinical Researchers-Part

1 II. Statistical Methods of Meta-Analysis. *Korean journal of radiology*, 16(6), 1188.
2 doi:10.3348/kjr.2015.16.6.1188

3 Macaskill P, G. C., Deeks JJ, Harbord RM, Takwoingi Y. . (2010). *Analysing and Presenting Results*.
4 Retrieved from Available from: <http://srdta.cochrane.org/>.

5 Marques, C. H. D., De Rosso, A. L. Z., & Andre, C. (2008). Bedside assessment of swallowing in stroke:
6 Water tests are not enough. *Topics in Stroke Rehabilitation*, 15(4), 378-383.

7 Martino, R., Foley, N., Bhogal, S., Diamant, N., Speechley, M., & Teasell, R. (2005). Dysphagia after
8 stroke: incidence, diagnosis, and pulmonary complications. *Stroke*, 36(12), 2756-2763.
9 doi:10.1161/01.STR.0000190056.76543.eb

10 McCullough, G. H., Rosenbek, J. C., Wertz, R. T., McCoy, S., Mann, G., & McCullough, K. (2005). Utility
11 of Clinical Swallowing Examination Measures for Detecting Aspiration Post-Stroke. *Journal of*
12 *Speech, Language, and Hearing Research*, 48(6), 1280-1293. doi:10.1044/1092-
13 4388(2005/089)

14 Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic
15 reviews and meta- analyses: the PRISMA statement. *BMJ*, 339. doi:10.1136/bmj.b2535

16 Palli, C., Fandler, S., Doppelhofer, K., Niederkorn, K., Enzinger, C., Vetta, C., . . . Gattringer, T. (2017).
17 Early Dysphagia Screening by Trained Nurses Reduces Pneumonia Rate in Stroke Patients A
18 Clinical Intervention Study. *Stroke*, 48(9), 2583-2585. doi:10.1161/strokeaha.117.018157

19 Ramsey, D., Smithard, D., & Kalra, L. (2005). Silent aspiration: what do we know? *Dysphagia*, 20(3),
20 218-225. doi:10.1007/s00455-005-0018-9

21 RCSLT. (2014). *Dysphagia Training & Competency Framework*. Retrieved from
22 Rofes, L., Arreola, V., Mukherjee, R., & Clave, P. (2014). Sensitivity and specificity of the Eating
23 Assessment Tool and the Volume-Viscosity Swallow Test for clinical evaluation of
24 oropharyngeal dysphagia. *Neurogastroenterol Motil*, 26(9), 1256-1265.
25 doi:10.1111/nmo.12382

26 Rofes, L., Muriana, D., Palomeras, E., Vilardell, N., Palomera, E., Alvarez-Berdugo, D., . . . Clavé, P.
27 (2018). Prevalence, risk factors and complications of oropharyngeal dysphagia in stroke
28 patients: A cohort study. *Neurogastroenterology & Motility*, 30(8), n/a-n/a.
29 doi:10.1111/nmo.13338

30 Royal College of Physicians Intercollegiate Stroke Working Party. (2016). *National Clinical Guideline*
31 *for Stroke*.

32 Ruiz Merino, L., Hochsprung, A., & Garcia Quesada, M. A. (2014). Acute stroke management results
33 of the adoption of an interdisciplinary dysphagia protocol in the kick-off of a new stroke unit
34 in a dedicated ward. *International Journal of Stroke*, 3), 81.

35 Samia E S B, S. N. A., Rasha H S, Shaimaa A S. (2017). Assessment of Dysphagia in Acute Stroke
36 Patients by the Gugging Swallowing screen. *Global Journal of Otolaryngology*, 9(4), 555766.

37 Sasaki, C. T., & Leder, S. B. (2014). Comments on selected recent dysphagia literature. *Dysphagia*,
38 29(5), 632-636.

39 Schepp, S. K., Tirschwell, D. L., Miller, R. M., & Longstreth Jr, W. T. (2012). Swallowing screens after
40 acute stroke: A systematic review. *Dysphagia*, 27(4), 562-563.

41 Schepp, S. K., Tirschwell, D. L., Miller, R. M., & Longstreth, W. T., Jr. (2012). Swallowing screens after
42 acute stroke: a systematic review. *Stroke*, 43(3), 869-871.
43 doi:10.1161/STROKEAHA.111.638254

44 Schunemann, H. J., Oxman, A. D., Brozek, J., Glasziou, P., Jaeschke, R., Vist, G. E., . . . Guyatt, G. H.
45 (2008). Grading quality of evidence and strength of recommendations for diagnostic tests
46 and strategies. *Bmj*, 336(7653), 1106-1110. doi:10.1136/bmj.39500.677199.AE

47 Serra-Prat, M., Palomera, M., Gomez, C., Sar-Shalom, D., Saiz, A., Montoya, J., . . . Clave, P. (2012).
48 Oropharyngeal dysphagia as a risk factor for malnutrition and lower respiratory tract
49 infection in independently living older persons: a population- based prospective study. *Age*
50 *Ageing*, 41(3), 376-381. doi:10.1093/ageing/afs006

- 1 Smithard, D. G., Oneill, P. A., Park, C., Morris, J., Wyatt, R., England, R., & Martin, D. (1996).
2 Complications and outcome after acute stroke - Does dysphagia matter? *Stroke*, 27(7), 1200-
3 1204.
- 4 Splaingard, M. L., Hutchins, B., Sulton, L. D., & Chaudhuri, G. (1988). Aspiration in rehabilitation
5 patients: Videofluoroscopy vs bedside clinical assessment. *Archives of Physical Medicine and*
6 *Rehabilitation*, 69(8), 637-640.
- 7 Takwoingi, Y. (April 2016). *Meta-analysis of test accuracy studies in Stata: a bivariate model*
8 *approach*. Retrieved from <http://methods.cochrane.org/sdt/>.
- 9 Teuschl, Y., Trapl, M., Ratajczak, P., Matz, K., Dachenhausen, A., & Brainin, M. (2018). Systematic
10 dysphagia screening and dietary modifications to reduce stroke-associated pneumonia rates
11 in a stroke-unit. *PLoS One*, 13(2). doi:10.1371/journal.pone.0192142
- 12 The International Dysphagia Diet Standardisation Initiative. (2016). IDDSI.
- 13 Thompson, M., Plüddemann, A., Price, C. P., & Heneghan, C. (2013). Emerging diagnostic
14 technologies in primary care: what's on the horizon? *The British journal of general practice : the journal of the Royal College of General Practitioners*, 63(609), 177.
15 doi:10.3399/bjgp13X665107
- 16
- 17 Trapl, M., Enderle, P., Nowotny, M., Teuschl, Y., Matz, K., Dachenhausen, A., & Brainin, M. (2007).
18 Dysphagia bedside screening for acute-stroke patients: the Gugging Swallowing Screen.
19 *Stroke*, 38(11), 2948-2952. doi:10.1161/STROKEAHA.107.483933
- 20 Trapl, M., Firlinger, B., Teuschl, Y., Dachenhausen, A., & Brainin, M. (2017). International
21 dissemination and usage of the gugging swallowing screen (GUSS). A prospective cross
22 sectional study. *Dysphagia*, 32 (1), 191.
- 23 Umay, E. K., Gurcay, E., Bahceci, K., Ozturk, E., Yilmaz, V., Gundogdu, I., . . . Cakci, A. (2018). Validity
24 and reliability of Turkish version of the gugging swallowing screen test in the early period of
25 hemispheric stroke. *Journal of Neurological Sciences-Turkish*, 35(1), 6-13.
26 doi:10.5152/nsn.2018.10436
- 27 Wang, T. G., Chang, Y. C., Chen, S. Y., & Hsiao, T. Y. (2005). Pulse oximetry does not reliably detect
28 aspiration on videofluoroscopic swallowing study. *Arch Phys Med Rehabil*, 86(4), 730-734.
29 doi:10.1016/j.apmr.2004.10.021
- 30 Warnecke, T., Im, S., Kaiser, C., Hamacher, C., Oelenberg, S., & Dziewas, R. (2017). Aspiration and
31 dysphagia screening in acute stroke – the Gugging Swallowing Screen revisited. *European*
32 *Journal of Neurology*, 24(4), 594-601. doi:10.1111/ene.13251
- 33 Whiting, P. F., Rutjes, A. S., Westwood, M. E., & et al. (2011). Quadas-2: A revised tool for the quality
34 assessment of diagnostic accuracy studies. *Annals of Internal Medicine*, 155(8), 529-536.
35 doi:10.7326/0003-4819-155-8-201110180-00009
- 36 Wirth, R., Smoliner, C., Jäger, M., Warnecke, T., Leischker, A. H., & Dziewas, R. (2013). Guideline
37 clinical nutrition in patients with stroke. *Experimental & translational stroke medicine*, 5(1),
38 14. doi:10.1186/2040-7378-5-14
- 39 Yeh, S. J., Huang, K. Y., Wang, T. G., Chen, Y. C., Chen, C. H., Tang, S. C., . . . Jeng, J. S. (2011).
40 Dysphagia screening decreases pneumonia in acute stroke patients admitted to the stroke
41 intensive care unit. *Journal of the Neurological Sciences*, 306(1-2), 38-41.

Tables

Table 1. Summary of the multidisciplinary comprehensive swallowing assessment tools used in acute stroke

Test	Studies identified	Test description	Test recommendations	Who can administer	Time of administration	Training required
Bedside Swallow Screening Tool (BESST)	Boaden 2011	Pre-screening & test with L0 and L4 consistencies.	3 options: 1. L0 fluids & L7 diet 2. L4 diet and fluids 3. NBM	Nurses	10 minutes	None
Dysphagia Trained Nurse Assessment DTNAx (Previously named 'Screening for dysphagia')	Heritage 2001, Heritage 2003 & Benfield 2018	Pre-screening checklist, Oromotor test, test of thin fluids progressing to level 2 & 3 fluids if unsafe and test L4, L5, L6, L7 diet as safe.	13 options: 1. L0 fluids & L7 diet 2. Any combination of L0, L2 or L3 fluids and L4, L5, L6, L7 diet 3. NBM	Nurses	20 minutes	1 day theory and practical. 4 x assessments completed independently then competency Ax with SLT
Gugging Swallow Screen (GUSS)	Trapl 2007, Ruiz Merino 2014 (Ruiz Merino, Hochsprung, & Garcia Quesada, 2014), John 2015, AbdelHamid 2017, Samia 2017, Palli 2017, Trapl 2017 (Trapl, Firlinger, Teuschl, Dachenhausen, & Brainin, 2017), Warneke 2017, Teuschl 2018, Ferreira 2018, Umay 2018	Preliminary indirect assessment – cough & swallow function. Direct assessment with 3-5 tsps (L3), 3, 5, 10, 20, 50 mls L0, 1.5cm piece of L7 diet x 5	4 options 1. L0 fluids & L7 diet 2. Level 1-2 fluids and L5 or L6 diet 3. L2-L3 fluids and L4 diet 4. NBM	Nurses & SLTs	5-10 minutes	10-15 minute theory, demonstration of GUSS by experienced nurse.
Volume Viscosity Swallowing Test (V-VST)	Rofes 2014, Clave 2008, Guillen-Sola 2013, Rofes 2018 (Rofes et al., 2018)	Assess 5, 10, 20 mls ~L2 fluids, then 5, 10, 20mls thin fluids as safe, then 5, 10 & 20mls L4 diet. Observation of signs & pulse oximetry.	26 options: 1. L0 fluids and L4 diet 2. Any combination of 5, 10 or 20mls of L0 or ~L2 fluids and/or 5, 10 or 20mls of L4 diet 3. NBM	Nurses, Physicians, Dietitians and SLTs.	5-8 minutes	Specific V-VST courses including theory (description, validation, algorithm, clinical cases) and practice with real patients
2 Volume, 3 texture test (2v/3t-P)	Cocho 2015	5mls then 10mls of ~L3/L4. 5mls then 10mls of L0 Then 1.5 cm piece L7 diet Observation of signs & pulse oximetry.	Unclear but likely 6 options: 1. 5 or 10mls L0 fluids & L7 diet 2. Any combination of 5 or 10mls L0 and 5 or 10mls of L3/4 or L7 diet 3. NBM	Nurses	No details	No details

International Dysphagia Diet Standardisation Initiative (IDDSI) Levels (L) are used. ~ is used to denote when the level is an estimation from another descriptor classification.

Table 2. Criteria for detecting aspiration and/or dysphagia on each of the comprehensive swallowing tests.

Test	Oral residue	Drooling	Ability to chew	No laryngeal elevation	Reduced laryngeal elevation	Delayed swallow	Voice change	Breath change	Cough post swallow	Throat clear	Multiple Swallows	Reported pharyngeal residue	Drop O2 saturation >2%
GUSS		✓		✓		✓	✓		✓				
DTNax	✓		✓	✓		✓	✓	✓	✓				
BESST				✓	✓		✓	✓	✓	✓	✓		
2v/3t-P	✓	✓		✓			✓	✓	✓				✓
VVST	✓	✓					✓		✓		✓	✓	✓

GUSS – Gugging Swallow Screen, DTNax – Dysphagia Trained Nurse Assessment, BESST – Bedside Swallow Screening Test, 2v/3t-P – 2 Volume, 3 Texture Pulse oximetry Test, VVST – Volume Viscosity Swallow Test.

Table 3. Diagnostic accuracy of multidisciplinary dysphagia assessments that have undergone validation.

Test	Study	N	Reference test?	Validated for?	Sensitivity % (and CI if reported)	Specificity % (and CI if reported)	PPV % (and CI if reported)	NPV % (and CI if reported)	ROC (and CI if reported)	Inter-rater reliability (and CI if reported)
Bedside Swallow Screening Test (BESST)	Boden 2011	136	SLT bedside assessment	Dysphagia	From 87.5 (76.0-99.0) to 92.9 (85.1-100)	From 70.1 (59.9 -80.4) to 81.6 (72.9 -90.3)	From 62.9 (50.9-74.9) to 71.4 (58.8-84.1)	From 92.3 (58.8-84.1) to 94.7 (88.9-100)	Not reported	81% agreement Kw=0.61 (0.45-0.77)
Volume -Viscosity Swallowing Test (V-VST)	Clave 2008	85	VFS	Aspiration	100	28.8	28.8	100	Not reported	Not assessed
	Guillen-Sola 2013	52	VFS	Aspiration	88.2	71.4	60	92.6	Not reported	Not assessed
	Rofes 2014	134	VFS	Aspiration	91 (0.78–0.99)	28 (0.17–0.34)	21	94	Not reported	k=0.628 (0.45–0.78)
				Dysphagia	94 (0.87–0.98)	88 (0.50–0.99)	98	70		
Gugging Swallow Screen (GUSS)	Trapl 2007	50	FEES	Aspiration	100	50-69	74-81	100	Group 1: 0.77 (0.53 to 1.02) Group 2: 0.933 (0.833 to 1.033)	k =0.835, P<0.001
	Abdelhamed 2017	42	FEES	Aspiration	93.3	83.3	93.3	83.3	0.94 (0.85–1)	k=0.84, P>0.05, PO=91%
	Warneke 2017	100	FEES	Aspiration	96.5 (87.8–99.5)	55.8 (39.8–70.9)	74.3 (62.8–83.7)	92.3 (74.6–98.9)	0.76 (0.67–0.84)	Not assessed
				Dysphagia	98.5 (92.3–99.6)	53.3 (34.3–71.6)	83.1 (73.3–90.4)	94.1 (71.3–99.8)	0.76 (0.66–0.84)	
	Samia 2017	40	FEES	Aspiration	93.8	96.1	96.2	93.7	Not reported	Not assessed
	Ferriera 2018	174	GUSS	GUSS score	100	43	not reported – no data to calculate	not reported – no data to calculate	Nurse 1 = 0.987 Nurse 2 = 0.991	k= 0.818 – 0.905 with p<0.001
	Umay 2018	113	FEES	Aspiration	95.3-97.5	75.2-76.2	84.3	95.1-95.3	0.885-0.913	ICC = 0.955 (0.935-0.969) p< 0.001
				Dysphagia	95.3-97.5	69.6-72.2	73.6-78.4	80.0-81.3	0.791– 0.822	
Dysphagia Trained Nurse Assessment (DTNAx)				Undergoing Validation – ClinicalTrials.gov Identifier: NCT03700853						
2 Volume,3 texture test (2v/3t-P)				No validation studies found						

CI =95% Confidence Interval, k= Kappa, kw = Weighted Kappa, ICC = Intra-class Correlation Coefficient, PPV = Positive predictive value, NPV = negative predictive value. ROC = Region under the Curve. The shaded areas indicated the studies that met the criteria for inclusion in meta-analysis.

Table 4. QUADAS-2 scores for risk of bias and concern for applicability of the diagnostic accuracy of the studies included in the systematic review multidisciplinary swallowing assessments.

Assessment	Study	Patient selection		Index test		Reference standard		Flow and timing	Overall quality of evidence based on GRADE criteria
		Risk of bias	Concern for applicability	Risk of bias	Concern for applicability	Risk of bias	Concern for applicability	Risk of bias	
Bedside Swallow Screening Test (BESST)	Boden 2011	Low	Low	Low	Low	Low	Low	Low	Moderate †%
Volume -Viscosity Swallowing Test (V-VST)	Clave 2008	High	High	Low	Unclear	Low	Low	Unclear	Very Low ***†‡¥
	Guillen-Sola 2013	High	High	Unclear	Unclear	Unclear	Low	High	Very Low ***%†
	Rofes 2014	High	High	Low	Unclear	Low	Low	Low	Very low ***†‡
Gugging Swallow Screen (GUSS)	Trapl 2007	High	High	Low	Low	Low	Low	Low	Low ***†
	Abdelhamed 2015	High	High	Low	High	Low	Low	Unclear	Very low ***‡¥†
	Warneke 2017	High	High	Low	High	Low	Low	Low	Very low ***††
	Samia 2017	Unclear	Unclear	Unclear	Unclear	Unclear	Low	Unclear	Very low ***‡¥†%
	Ferriera 2018	Unclear	High	Unclear	Low	High	Low	Unclear	Very low ***‡%¥†
	Umay 2018	High	High	Unclear	Unclear	Unclear	Low	Low	Very low***‡% †
Dysphagia Trained Nurse Assessment (DTNAx)		Undergoing Validation – ClinicalTrials.gov Identifier: NCT03700853							
2 Volume,3 texture test (2v/3t-P)		No validation studies found							

GRADE rating downgraded due to: %concern or uncertainty regarding risk of bias or applicability of reference test †imprecise results *concern or uncertainty regarding risk of bias or applicability of patient selection methods ‡concern or uncertainty regarding risk of bias and/or applicability of the index test ¥risk of bias in flow and or timing

Figures

Fig.1 PRISMA flow diagram showing number of records identified, screened, eligible and included

Fig.2 Forest plot comparing and pooling the sensitivity and specificity of the three studies validating the Gugging Swallow Screening test (GUSS)

CI = Confidence Intervals.

Appendices

Appendix 1: The search strategy used in EMBASE database

Appendix 2: Data extraction form

Appendix 3: QUADAS 2 tool