REVIEW ARTICLE

Systemic review on highly qualified screening tests for swallowing disorders following stroke: Validity and reliability issues

Marziyeh Poorjavad, Shohreh Jalaie[1]
Departments of Speech Therapy, [1]Physiotherapy, School of Rehabilitation, Tehran University of Medical Sciences, Tehran, Iran

Background: Oropharyngeal dysphagia following stroke enhances the risk of dehydration, malnutrition, aspiration pneumonia, persistent disablement, and even death. Screening of dysphagia has been shown to positively change health outcomes. The aim of the present study was to systematically introduce the published swallowing screening methods in patients with stroke and their appropriateness for detecting swallowing disorders following stroke with an emphasis on the methodological quality of their research studies. **Materials and Methods:** A computerized search through the Medline (PubMed), Embase, Scopus, and Google Scholar; databases from 1990 through 20 July 2013 was performed. In addition, the related citations and reference lists of the selected articles were considered. **Results:** A total of 264 papers were retrieved and 19 articles finally met inclusion criteria. Sixty-eight percent of included papers did not have a sufficient quality and only six articles were scored as having evidence level 'I' and were reported descriptively. The most prevalent bias in the included studies was probably a kind of spectrum bias that could lead to select just a subgroup of admitted stroke patients. The screening tests' sensitivities ranged from 47 to 100%, while their specificities ranged from about 63 to 100%. Strengths and limitations of each test have been discussed. **Conclusion:** We ultimately found four simple, valid, reliable, sensitive, and specific tests for screening swallowing disorders in the almost all acute alert stroke patients. Further validation and reliability assessing of screening tests need to follow a very accurate and well-established method in a large sample of the almost all acute alert stroke patients admitted to the hospitals.

Keywords: Swallowing, stroke, screening tests, systemic review, valid, reliable

How to cite this article: Poorjavad M, Jalaie Sh. Systemic review on highly qualified screening tests for swallowing disorders following stroke: Validity and reliability issues. J Res Med Sci 2014;19:776-85.

INTRODUCTION

Oropharyngeal swallowing dysfunction is one of the most significant problems after stroke.[1] The reported prevalence of oropharyngeal dysphagia following stroke varies between 22 and 65%,[2,3] depending on different sampling methods,[4] methods and timing of assessment,[5,6] and definition of dysphagia. Persistent oropharyngeal dysphagia is a marker of poor prognosis of stroke patients.[6] It can enhance the risk of dehydration,[3] malnutrition,[3] aspiration pneumonia,[3,7] and persistent disablement.[3,8,9] Aspiration pneumonia is one of the most life-threatening consequences of dysphagia in stroke.[7] Patients with dysphagia are 3-11 times more likely to develop pneumonia than stroke patients with reserved swallowing ability, depending on severity of dysphagia and presence or absence of aspiration.[7] Also, mortality risk is higher in stroke patients with dysphagia.[3] It is shown that about 20% of stroke victims will die from aspiration pneumonia in the 1st year post onset.[10]

Dysphagia screening methods and dysphagia assessment procedures (clinical and/or instrumented)[11] are usually used with different purposes. Clinical and instrumental assessment methods are administrated to find the underlying anatomic and/or physiologic abnormalities leading to swallowing problems and finally to design the appropriate treatment plan.[12] But swallowing screening methods, according to the American Speech-Language-Hearing Association (ASHA),[13] are the pass/fail procedures to identify individuals who may need a comprehensive assessment of swallowing function. Screening of swallowing abnormalities, the first step in an appropriate management plan,[14] has been shown to reduce risk of developing pneumonia,[7,15] percutaneous endoscopic gastrostomy (PEG) insertion rates, and mortality in patients with stroke.[7] Hinchey *et al*.,[15] showed that systematic use of a formal dysphagia screening protocol can decrease pneumonia rates from 5.4 to 2.4%. So management dysphagia guidelines, developed by the Heart and Stroke Foundation of Ontario (HSFO), emphasize that all patients with acute stroke have to be kept 'nil by mouth' (NPO) including medications until their swallowing safety has been

Address for correspondence: Asst. Prof. Shohreh Jalaie, Department of Physiotherapy, School of Rehabilitation, Tehran University of Medical Sciences, Tehran, Iran. E-mail: Jalaiesh@sina.tums.ac.ir
Received: 27-08-2013; **Revised:** 17-12-2013; **Accepted:** 26-02-2014

established. According to these guidelines, swallowing ability of all stroke patients should be screened as soon as they are awake and alert.[16] Nowadays, most acute care settings use a kind of dysphagia screening protocol, especially in the developed countries. But differences in the accuracy of the used screening protocols and specialists' training levels can impact on the results. Some patients may be consequently underdiagnosed and the risk of developing aspiration pneumonia may be enhanced. On the other hand, some patients may be kept NPO for a period of time without any swallowing dysfunction. So, the implementation of a simple, valid, and reliable screening test that is sensitive and specific to the swallowing problems[11] in the acute care settings is necessary to reduce stroke-related costs and some of the resulting preventable consequences.[7] However, there are very delicate biases impacting on the authenticity of a test, even though its psychometric values seem very reasonable. Considering the quality of the research study[11] is therefore suggested as an important factor when selecting a screening test. So the aim of the present study was to systematically introduce the published swallowing screening methods in patients with stroke and their appropriateness for the detection of swallowing disorders following stroke with an emphasis on the methodological quality of their research studies. The following question was formulated: What are the psychometric and feasibility properties of the available highly qualified screening tests to detect swallowing disorders following stroke? Although there are some other published systematic reviews in this regard,[6,14,17] the present study was focused on the methodological quality of the research studies and some common biases like spectrum and verification biases. In fact this review was interested in the well-qualified screening tests that can be administrated in the almost all acute alert stroke patients by the frontline professionals who have the earliest contact with the patients.

MATERIALS AND METHODS

Literature search

A computerized search through the Medline (PubMed), Embase, Scopus, and Google Scholar databases from 1990 through 20 July 2013 was performed. It was limited to published articles on humans. In addition, the related citations and reference lists of the selected articles were considered. Table 1 shows used terms and a flowchart for identified abstracts.

Study selection and eligibility criteria

After elimination of duplicate ones, the outcome of search strategy was 264 papers. From all retrieved sources, just original studies focused on the development and validation

Table 1: Search terms and a flowchart for identified abstracts

Database	Search terms	Identified articles	Excluded
PubMed	Dysphagia [Title] OR		
	Swallowing [Title] OR		
	Aspiration [Title] OR		
	Deglutition [Title]	26,700	
	Stroke [Title]	48,621	
	#a AND #b	307	
	screening [Title/Abstract] OR test [Title/Abstract] OR clinical assessment [Title/Abstract] OR bedside assessment [Title/Abstract] OR tool [Title/Abstract]	1,285,687	
	#c AND #d	54	
Google Scholar	Stroke swallowing OR dysphagia OR aspiration "screening"	18	
Embase	TITLE (swallowing OR dysphagia OR aspiration) and TITLE (stroke) and TITLE-ABSTR-KEY (screening OR assessment OR test)	36	
Scopus	TITLE (swallowing OR dysphagia OR aspiration) AND TITLE (stroke) AND TITLE-ABS-KEY (screening OR assessment OR test OR tool)	201	
Reference check		32	
Total identified unique abstract		264	
Original articles, focused on the development and validation of a screening tool for swallowing disorders following stroke		118	
Exclusion criteria:			99
No comparison with VF or FEES as a reference test			
Different assessments for different patients as a reference test			
Patients with other kinds of etiologies apart from stroke			
Full text in the other languages a' side from English or Persian			
Final included papers		19	

of a screening tool for swallowing disorders following stroke were included. Therefore, reviews, editorials, or letters and those articles that were unrelated to our mentioned purpose were not reviewed. Then the abstracts (or full text in doubtful cases) of included articles (118 papers) were reviewed based on the exclusion criteria. Those articles that had studied patients with other kinds of etiologies apart from diagnosed stroke were excluded. Besides, this review was interested in clinical screening tests that were compared with a videofluoroscopic (VF) assessment (or modified barium swallow) or fiberoptic endoscopic evaluation of swallowing (FEES). So, articles that had used a clinical assessment, speech, and language pathologists' judgments about swallowing function, or patients' clinical features and outcomes as a gold standard were excluded. Also there were some studies that had used varied criteria and tests as a reference test. In other words, authors had not used a unique test for all their patients. These articles were excluded, because different tests will lead to different results. Also only publications with full text in English and Persian were reviewed. These criteria are presented in Table 1 in detail.

Study quality of every included article (19 papers) was assessed using the 12-step criteria adapted from Jaeschke *et al*., 1994.[18,19] This form considered following three broad issues for appraising a diagnostic test:

a. Are the results of the study valid?
b. What are the results?
c. Will the results help me and my patient/population?

Table 2 shows a description of these criteria in brief with a little modification in some questions' grammar. Most questions were answered with a 'Yes', 'No', or 'Can't tell" except questions 7, 8, and 12 that should be described. The first two questions were "screening questions" and could be answered fast. Even if the answer to one of them was "No" or "can't tell", it was not worth continuing to the remaining questions. It seemed we could not be sure about an article's results (Question 8) if the reference test and the index test were not carried out blindly (Question 4), and/or all patients did not get the index and the reference test regardless of the results of the index test (Question 3), and/or there was a kind of spectrum bias in selection of stroke patients leading to choose only a subgroup of stroke patients (Question 5), and/or there were other confidence limits in the methodology. In addition, a diagnostic test could not be useful for patients and could not help to identify swallowing disorders following stroke (Questions 11 and 12), unless we could be sure about its results at least approximately (Question 8), and its psychometric features (e.g., sensitivity and specificity) were acceptable. Based on these criteria, the evidence level of every article was categorized as level I or II:

I. Blinded comparison (Question 4) with no verification and spectrum biases (Questions 3 and 5 answered Yes or at least Can't tell), and with reported or at least calculable results (Question 7).
II. Studies which did not have at least one of the four above conditions. Table 3 shows the results of articles' quality assessment.

Data extraction and abstraction

Tables 4 and 5 show some data extracted from the studies with evidence level I. This information can be divided into two categories: Information about characteristics of the studied population and information about the used index and reference tests.

RESULTS

As mentioned above, 19 articles met our inclusion criteria[2-4,20,24,25,30-32] and level of evidence of 32% of

Table 2: The 12-steps criteria adapted from Jaeschke *et al*., (1994) in brief

Items	
Issue (a)	
1	Was there a clear question for the study to address?
2	Was there a comparison with an appropriate reference standard?
3	Did all patients get the diagnostic test and the reference standard? (verification bias)
4	Could the results of the test of interest have been influenced by the results of the reference standard? (review bias)
5	Is the disease status of the tested population clearly described? (spectrum bias)
6	Were the methods for performing the test described in sufficient detail?
Issue (b)	
7	What are the results?
8	Are we sure about these results?
Issue (c)	
9	Can the results be applied to your patients/the population of interest?
10	Can the test be applied to your patient or population of interest? (availability of resources, expertise, and opportunity costs)
11	Were all outcomes important to the individual or population considered?
12	What would be the impact of using this test on your patients/population?

Table 3: The results of articles' quality assessment

Reference	Items												Evidence level
	Issue (a)						Issue (b)		Issue (c)				
	1	2	3	4	5	6	7	8	9	10	11	12	
DePippo *et al*., 1992[20]	Yes	Yes	Yes	Not blind	No	Yes	R	No	No	Yes	No	CID	II
Horner *et al*., 1993[21]	Yes	Yes	No	Not blind	No	Yes	NR	No	No	Yes	No	CID	II
Kidd *et al*., 1993[22]	Yes	Yes	Yes	CNT	Yes	Yes	R	CNT	Yes	Yes	CNT	CNT	II
Collins and Bakheit, 1997[23]	Yes	Yes	Yes	Blind	No	Yes	R	No	No	Yes	No	CID	II
Daniels *et al*., 1997[24]	Yes	Yes	Yes	Blind	Yes	Yes	R	Yes	Yes	Yes	Yes	ID	I
Daniels *et al*., 1998[2]	Yes	Yes	Yes	CNT	Yes	Yes	R	CNT	Yes	Yes	CNT	CNT	II
Smithard *et al*., 1998[3]	Yes	Yes	Yes	Blind	Yes	Yes	R	No	Yes	Yes	No	CID	I
a) Medical bedside assessment	Yes	Yes	Yes	Blind	Yes	No	R	No	Yes	Yes	No	CID	I
b) speech therapy assessment													
Smith *et al*., 2000[4]	Yes	Yes	Yes	Blind	No	No	R	No	No	Yes	No	CID	II
Lim *et al*., 2001[25]	Yes	Yes	Yes	Blind	Yes	Yes	R	Yes	Yes	Yes	Yes	ID	I
Leder and Espinosa 2002[26]	Yes	Yes	Yes	Blind	No	Yes	R	No	No	Yes	No	CID	II
Chong *et al*., 2003[27]	Yes	Yes	Yes	Blind	No	Yes	R	No	No	Yes	No	CID	II
Nishiwaki *et al*., 2005[28]	Yes	Yes	Yes	CNT	No	Yes	R	No	No	Yes	No	CID	II
Ramsey *et al*., 2006[29]	Yes	Yes	CNT	Blind	No	Yes	R	No	No	Yes	No	CID	II
Trapl *et al*., 2007[30]	Yes	Yes	Yes	Blind	CNT	Yes	R	Yes	Yes	Yes	Yes	ID	I
Warnecke *et al*., 2008[31]	Yes	Yes	Yes	Blind	Yes	Yes	R	Yes	Yes	Yes	Yes	CNT	I
Martino *et al*., 2009[32]	Yes	Yes	Yes	Blind	Yes	No	R	Yes	Yes	Yes	Yes	ID	I
Zhou *et al*., 2011[33]	Yes	Yes	Yes	Blind	No	Yes	R	No	No	Yes	CNT	CNT	II
Umay *et al*., 2013[34]	Yes	Yes	Yes	CNT	No	Yes	R	No	No	Yes	No	CID	II
Osawa *et al*., 2013[35]	Yes	Yes	Yes	Not blind	No	Yes	R	No	No	Yes	No	CID	II

CNT = Can't tell; R = Reported; NR = Not reported; ID = Identification of disorder accurately; CID = Can't identify disorder accurately

them (six articles)[3,24,25,30-32] was I according to the performed quality assessment.

Included studies

VF evaluation had been carried out as a "gold standard" in most studies (68.5%).[2-4,20-24,28,29,32,33,35] All other included studies had used FEES as a reference test.[25-27,30,31,34]

A variety of tests were used to screen swallowing disorders in bedside. In eight of included studies (42%), screening protocols consisted of a combination of a sensorimotor examination and clinical swallowing assessment.[2,3,22,24,26,30,32,33] There was a large variety in tasks assessed in sensorimotor examinations. In these studies, clinical swallowing assessments usually included water swallowing in different volumes.[3,22,24,26,32,33] Different consistencies had been used as swallow materials just in two articles.[7,32] Four studies (21%)[20,28,31,35] used just a kind of water swallow test as a screening tool and one[23] involved measurement of oxygen desaturation alone. Five papers described a combination of pulse oximetry and trial swallows.[4,25,27,29,34] Finally, Horner *et al*.,[21] examined some clinical features to assess risk of aspiration, such as age, lesion site, abnormal gag, volitional cough, and voice.

About one-third of included papers did not have or did not report a blind design,[2,20-22,28,34,35] and so their evidence levels were scored II. But this kind of bias was not so popular in the more recent years.

The most prevalent bias in the included studies probably was a kind of spectrum bias that could lead to selection of a subgroup of admitted stroke patients and so could influence on the test's generalizability. Eleven papers[4,20,21,23,26-29,33-35] had such bias to some extent. Some studies included those stroke patients who referred for swallowing evaluations[23,26-28,33] and no all consecutive stroke patients admitted to the hospital. Also, some other researchers[4,29] excluded patients with probably more severe disabilities because of some problems with sitting balance and poor medical condition.

Studies with evidence level I

Tables 4 and 5 show some properties and psychometric features of tests that met quality criteria. A half of these studies used VF[3,24,32] and the other half used FEES[25,30,31] as a reference test. Daniels *et al*.,[24] and Lim *et al*.,[25] described their used volumes and consistencies of swallow materials in the reference standard in detail. Smithard *et al*.,[3] reported the using of an adaptation of Logemann standard protocol for videofluoroscopy.[12] But other researchers[30-32] did not present a description of their used protocol for the reference test in sufficient detail.

All final selected tests[3,24,25,30-32] consisted of a clinical swallowing assessment part. Trapl *et al*.,[30] used a variety of consistencies (semisolid, liquid, and solid) in their clinical swallowing trials; and according to the points in different consistencies, they could suggest a special diet for each

Table 4: Properties of articles with evidence level I

References	Population	Exclusion criteria	Sample size and time of assessments	RT and endpoint	IT and endpoint	Average time between RT and IT	Time needed for administration	Administrator
Daniels *et al*., 1997[24] (USA)	Stroke patients consecutively admitted with a new neurological deficit; mean age ± SD: 66 ± 11 years	Obtunded and agitated patients, a prior history of oropharyngeal dysphagia, oropharyngeal structural damage, or neurological disease other than stroke that may produce dysphagia	59; time: Within 5 days of admission	VF Swallow material: 3, 5, 10, and 20 mL liquid barium, 1/2 teaspoon barium paste, half of a cookie, unregulated amounts of thinned liquid barium Endpoint: Moderate/severe dysphagia (risk of aspiration): Scores 2-4 versus mild dysphagia/ normal swallowing (no risk of aspiration): Scores 0-1	Oropharyngeal assessment: Examination of gag reflex, volitional cough, speech, and voice + Clinical swallowing assessment: Twice ingestion of 5, 10, and 20 ml of water (for a total of 70 ml) Endpoint: Two or more variables from six predictive variables (dysphonia, dysarthria, abnormal volitional cough, abnormal gag reflex, cough after swallow, and voice change after swallow) versus fewer than two variables from six predictive variable	48 h	NR	SLPs
Smithard *et al*., 1998[3] (UK)	Patients admitted within 24 h of the onset of acute stroke; median age 79 years (range 40-93)	Admission after 24 h, failure to obtain consent or the presence of serious intercurrent illness (e.g., advanced malignancy)	94; time: Within 3 days of the stroke (median time: 2 days)	VF Swallow material: Different consistencies and volumes of barium. a standard protocol adapted from that of Logemann[12] Endpoint: Aspiration versus no aspiration	a) Medical bedside assessment (*n*=94): Clinical swallowing assessment (2 stages): 1. Swallow a 5 ml spoonful of water three times 2. Swallowing 60 ml of water within 2 min Endpoint: Unsafe swallow: If coughing and/or choking on more than one occasion out of three attempts of stage 1 and or a wet voice; or coughing and/or choking during stage 2; or the presence of a wet voice versus safe swallow b) Speech therapy assessment (*n*=83): A protocol consists of oropharyngeal examination and swallow trials Endpoint: Unsafe swallow: The speech and language therapist's overall clinical judgment versus safe swallow	24 h	NR	a) Physicians b) SLPs
Lim *et al*., 2001[25] (Singapore)	All acute stroke patients admitted to the stroke unit; mean age±SD: 67.5±11.73 years	Severe peripheral vascular disease, a consciousness level not sufficient to give informed consent, no CT evidence of a stroke or no significant neurological deficit (e.g., new hemiplegia) lasting more than 24 h, and insufficient lip seal to retain 10 ml of water in the mouth	50; Time: Mean time from stroke onset: 5.14 days (SD: 3.62)	FEES Swallow material: Three or more boluses of 5 mL of water thickened to a pudding consistency with "thick and easy" (puree) and three or more boluses of 5 mL of water thickened to a honey consistency and with water alone Endpoint: Aspiration/ or penetration versus no aspiration/or penetration	Clinical swallowing assessment: Swallow 50 ml of water in 10-ml aliquots + Pulse oximetry: The highest and lowest oxygen saturation readings during and for up to 2 min after swallowing 10 ml of water for three times Endpoint: Bedside aspiration: Coughing or choking during the water drinking test or had a change in voice quality following the swallow (up to 5 min after swallowing) and/or a desaturation of ≥2% versus no bedside aspiration	90% within 24 h 10% within the next 24 h	NR	Physicians (but not specific to physicians)

(Continued)

Table 4: Properties of articles with evidence level I *(continued)*

References	Population	Exclusion criteria	Sample size and time of assessments	RT and endpoint	IT and endpoint	Average time between RT and IT	Time needed for administration	Administrator
Trapl *et al*., 2007[30] (Austria)	Consecutive patients with first ever acute stroke and suspected dysphagia who were admitted to the acute stroke unit; mean ages±SD: 76.8±1.85 years	Multiple infarcts, dysphagia of other known cause, and somnolence or coma within 24 h	30; time: Within 24 h of stroke onset	FEES Graded according to the Penetration Aspiration Scale (PAS) of Rosenbek *et al*.[36] Endpoint: Aspiration risk (PAS: 5-8) versus minimal or no aspiration risk (PAS: 1-4)	Combination of two parts: Part 1: Vigilance, voluntary, cough, throat clearing, and saliva swallowing Part 2: Clinical swallowing assessment: 1. Semisolid Swallowing Trial: 1/3-1/2 teaspoon of pudding, followed by five more half teaspoons. Abort the investigation if one of the four aspiration signs (deglutition, cough, drooling, and voice change) is positive 2. Liquid Swallowing Trial: 3 mL aqua bi. If it is successful; 5, 10, and 20 mL of aqua bi. Finally, drinking the 50 mL as fast as possible 3. Solid Swallowing Trial: A small piece of dry bread that was repeated five times. Endpoint: Aspiration risk (score 0-14) versus no aspiration risk (score 15-20)	Within 24 h	NR	Stroke nurses and therapists
Warnecke *et al*., 2008[31] (Germany)	Consecutive patients with first ever stroke; mean age±SD: 71.43±11.81 years	Non-ischemic stroke, admittance later than 24 h after symptom onset, a history of a preexisting dysphagia or disease probably causing dysphagia, a severely reduced state of consciousness (i.e., stupor or coma). Patients had to have either an NIH-SS > 3 points and/or had to present with a facial palsy and/or dysarthria to be eligible for advanced dysphagia assessment	100; time: Within 72 h of stroke onset	FEES Swallow material: teaspoon-wise of pureed food, liquid, and soft solid food Endpoint: Aspiration risk (aspiration and/or penetration of the patient's own saliva or any feeding consistency) versus no aspiration risk	Clinical swallowing assessment: Insertion of a thin 4-charriere catheter through the nostril into the oropharynx and injection of 0.4 ml (first step) and 2.0 ml (second step) of distilled water. Then measurement of the latent time from the beginning of water bolus injection to onset of swallowing (i.e., laryngeal elevation) by a stopwatch Endpoint: Abnormal swallowing (each of the two steps was classified as abnormal when the latent time was more than 3 s) versus normal swallowing (latent time <3 s in both steps)	Immediately one after another	NR	NR

(Continued)

Table 4: Properties of articles with evidence level I *(continued)*

References	Population	Exclusion criteria	Sample size and time of assessments	RT and endpoint	IT and endpoint	Average time between RT and IT	Time needed for administration	Administrator
Martino *et al*., 2009[32] (Canada)	All consecutive patients newly admitted to hospital with the confirmed diagnosis of a brain stem stroke or cerebellar stroke and all other stroke patients with a NIH-SS ≥ 4	Non-brainstem and non-cerebellar stroke patients with low NIH-SS scores, current respiratory compromise, a nonoral feeding regime, or a history of one or more of the following: Nonstroke neurological disorder, surgery to the head or neck, a history of previous oropharyngeal dysphagia, dementia, or decreased level of consciousness	59; time: The mean time from hospital admission to screening was 4.6 days in acute sites and 4.9 days in rehabilitation sites	VF; judgments of VF findings using three measures: The (PAS)[36], the MASA dysphagia subscore,[37] and the MASA aspiration subscore[37] Endpoint: Dysphagia (aspiration or any physiological abnormality on VF) versus no dysphagia	Oropharyngeal examination: Tongue movement+clinical swallowing assessment: Voice before swallow trial, Kidd 50-mL water swallow test,[22] and voice after swallow trial Endpoint: Dysphagia risk (failure on any item) versus no dysphagia risk	24 h	10 min	Any healthcare professional trained in the clinical assessment of post-stroke patients

RT = Reference test; IT = Index test; SD = Standard deviation; CT = Computed tomography, NR = Not reported; FEES = Fiberoptic endoscopic evaluation of swallowing; NIH-SS = National institutes of health stroke scale; MASA = The mann assessment of swallowing ability, VF = Videofluoroscopy; SLP = Speech-language pathologists

Table 5: Psychometric features of final selected screening tests

References	Prevalence according to endpoint of RT (%)	Prevalence according to endpoint of IT (%)	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	LR+	LR-	Reliability of IT (%)
Daniels *et al*., 1997[24]	44.1	61	92.3	66.7	NR	NR	2.1	0.3	NR
Smithard *et al*., 1998[3]	21	a) 41 b) 22	a) 70 b) 47	a) 66 b) 86	a) 36 b) 50	a) 89 b) 85	a) 2.0 b) 3.35	a) 0.45 b) 0.61	a) 75% agreement between DOCs (*n*=65) (K=0.5, 95% CI=0.26-0.73) b) 93% agreement between SLPs (*n*=74) (*K*=0.79, 95% CI=0.55-1.0)
Lim *et al*., 2001[25]	52	66	100	70.8	78.8	100	3.33	0	NR; the highest degree of agreement with FEES (*K*=0.716)
Trapl *et al*., 2007[30]	46.6	63.3	100	69	74	100	3.22	0	Interrater reliability of IT: Excellent agreement (*K*=0.835)
Warnecke *et al*., 2008[31]	81	1st step of IT: 60 2nd step of IT: 40	1st step: 74.1 2nd step: 49.4	1st step: 100 2nd step: 100	1st step: 100 2nd step: 100	1st step: 47.5 2nd step: 31.7	1st step: ∞ 2nd step: ∞	1st step: 0.26 2nd step: 0.51	NR
Martino *et al*., 2009[32]	39 (in subset of acute patients (*n*=24): 54.2; in subset of rehabilitation patients (*n*=35): 28.6)	55.9 (in subset of acute patients (*n*=24): 70.8; in subset of rehabilitation patients (*n*=35) :45.7)	91.3 (in subset of acute patients (*n*=24): 96.3; in subset of rehabilitation patients (*n*=35): 80)	66.7 (in subset of acute patients (*n*=24): 63.6; in subset of rehabilitation patients (*n*=35): 68)	(in subset of acute patients (*n*=24): 76.5; in subset of rehabilitation patients (*n*=35): 50)	(in subset of acute patients (*n*=24): 93.3; in subset of rehabilitation patients (*n*=35): 89.5)	2.7 (in subset of acute patients (*n*=24): 2.6; in subset of rehabilitation patients (*n*=35): 2.5)	0.1 (in subset of acute patients (*n*=24): 0.1; in subset of rehabilitation patients (*n*=35): 0.3)	Interrater reliability for administration of IT by trained nurse screeners (*n*=50): excellent; ICC=0.92 (95% CI: 0.85-0.96)

RT = Reference test; IT = Index test; PPV = Positive predictive value; NPV = Negative predictive value; LR+ = Positive likelihood ratio; LR- = Negative likelihood ratio; NR = Not reported; ICC = Interclass correlation coefficient; CI = Confidence interval

patient.[30] All other researchers[3,24,25,31,32] just assessed patients' ability to swallow liquids (different volumes of water). Used liquid volumes ranging from 2.4[31] to 88 ml[30] in different selected tests. Most researchers divided liquid volumes into smaller aliquots that gradually progress to larger volumes, and discontinued their test if a patient developed some signs of swallowing disorders or discomfort during each step.[3,24,30,32] The test proposed by Trapl *et al*.,[30] however, consisted of a timed swallow test of a relatively large amount of water (50 ml) that should be administered cautiously.

Most tests had been assessed for their accuracies to identify only aspiration and/or penetration.[3,25,26,31] But Daniels *et al*.,[24] and Martino *et al*.,[32] paid attention to dysphagia as a global term that may include any abnormal physiology of oropharyngeal swallowing, regardless of the presence or absence of aspiration.[12] The reported endpoints by Daniels *et al*.,[24] Martino *et al*.,[32] and Trapl *et al*.,[30] for the index tests included at least one variable that was exclusively associated with oral phase of swallowing. But other tests[3,25,31] did not include any indicator of the oral phase and so could not detect disorders in patients with a predominantly impaired oral phase of swallowing and a relatively intact pharyngeal phase.[31]

Except Daniels *et al*.,[24] all other researchers[3,25,30-32] administered the index and the reference test within 24 h.

There has been a trend towards developing the screening tests that could be administrated by various healthcare specialists and not just speech-language pathologists (SLPs) or physicians.[30,32]

The screening tests' sensitivities ranged from 47 to 100%, while their specificities ranged from about 63 to 100%. The test proposed by Lim *et al*.,[25] achieved the highest sensitivity and specificity (100 and 70.8%, respectively). Fifty percent of studies did not report the tests' reliability.[24,25,31] Interrater reliability varied from moderate to excellent agreement in the remaining screening tests.[3,30,32] [Table 5].

DISCUSSION

This review showed that there are a large variety of screening tests for swallowing disorders following stroke that are different in types, methods, endpoints, and their psychometric values. There were many differences in selected population, time of the test administration, and other aspects of methodology. In this systematic review, the methodological quality of every included article was assessed using criteria adapted from Jaeschke *et al*., 1994[18,19] and was scored according to our predefined values as either having evidence level I or II. Bases on our relatively strict judgment, 68% of included papers did not have a sufficient quality. It emphasizes the importance of considering methodological limitations of studies, and of improving study design standards in such studies. A blind design for validation of diagnostic tests, as an instance, is vital. Because if the reference test and the index test are not interpreted independently, the results of tests may be influenced by each other. This kind of bias was not so popular in the more recently validated screening tests. But the results of many reviewed studies,[4,20,21,23,26-29,33-35] according to our quality assessment, had been influenced by a sort of spectrum bias. We were hoping to find the tests that can be administered in almost all patients with acute alert stroke admitted to hospital newly. When patients are selected based on lots of inclusion and exclusion criteria, the selected patients may represent only a subgroup of stroke patients (spectrum bias) and no all alert patients admitted to hospital with acute stroke. Although some criteria such as consciousness and being able to follow some simple instruction are necessary for swallowing assessment, but some other features like receptive dysphasia or inability to sit upright without support must not deprive patients of assessing of swallowing mechanism. This manner of selecting patients can have biased the assessed population toward patients with mild and moderate strokes. Also selecting patients from those referred to SLPs to assess swallowing[23,26-28,33] or from those having some features indicating possible dysphagia,[20,28,34,35] may lead to select patients who more likely suffer from dysphagia or have more obvious swallowing disorders. A very accurate screening test may be not necessary for identification of such disorder.[18,19] In addition, silent aspiration is a serious concern in acute stroke[2] and patients with this kind of aspiration may not refer to speech-language therapist for evaluation of swallowing function due to absence of clinical symptoms. The whole spectrum of patients with acute stroke, therefore, was not included in these studies. So in this review, those articles with the least selection on admitted patients met quality criteria and are reported in detail.

A half of six tests with evidence level I had used VF as a reference test. Although VF evaluation is almost accepted as a gold standard for assessing swallowing disorders,[38] some limitations are reported for it. Interrater reliability of VF is often poor[39,40] and it assesses the patients' ability in swallowing of small amounts of foods and in an optimal situation that does not usually reflect the natural situation of patient's feeding.[41] These limitations may impact on the calculated validities of the index tests. So Smithard *et al*.,[3] recommended that the use of VF as a gold standard in the validation studies should be critically explored in the further studies.

As mentioned above, just a half of the high qualified studies[3,24,25] provided a description of their used protocol for the reference test in detail. Different protocols will examine patients' swallowing mechanism in different

levels and with different accuracies. It may be one of the reasons of various reported prevalence of swallowing disorders following stroke. Making a description of the parts of performed 'gold standard', therefore, can help readers to make a more accurate judgment about the study.

Only Trapl *et al*.,[30] used different consistencies in the swallowing trails. Although it may increase needed time and equipment for the test administration, but can lead to a more accurate picture of patients' swallowing abilities. This screening test[30] also consisted of a swallowing trial of a relatively large amount of water. Large volumes of liquids may introduce a high risk of aspiration and airway obstruction to the patients.[42] Although the authors[30] warned about cautious administration of this part, but the administration of a "screening test" must not be dangerous for patients.

Although swallowing disorders in the pharyngeal phase are common in patients with stroke,[7,12] but dysphagia is described as any kind of difficulty moving food from mouth to stomach.[12] Oropharyngeal dysphagia screening tests therefore should consider both oral and pharyngeal phases of swallowing process. The screening test reported by Daniels *et al*.,[24] Martino *et al*.,[32] and Trapl *et al*.,[30] included at least one indicator of the oral phase.

Since the severity of dysphagia changes during acute phase after stroke rapidly, a 24-h interval between administration of the reference and index tests seems short enough to be sure that the patient's condition will not change between the two tests significantly.[17] The average time between the two tests was more than 24 h only in the study of Daniels *et al*.[24]

SLPs are in short supply in many hospitals.[11] So screening tests that can be conducted by various healthcare professionals may accelerate the screening process of newly admitted acute stroke patients.[11] Screening tests developed by Lim *et al*.,[25] Trapl *et al*.,[30] and Martino *et al*.,[32] could be administrated by a variety of healthcare specialists.

Regarding serious consequences of swallowing disorders, it seems a valid clinical examination for detecting such disorders after stroke must have a high sensitivity. Such a screening test will miss just a few patients with swallowing disorders.[32] Since the main purpose of administration of a swallowing screening, according to ASHA,[13] is identification of patients who need to refer for a more comprehensive swallowing assessment and not designing treatment plan, a moderate-high specificity may be enough. In such circumstances, some patients without dysphagia may be referred to speech-language therapists for assessment and before starting of any kind of treatment, will be probably identified as patients with safe and intact swallowing abilities.[32] In this review, we could find four well-qualified screening tests with high sensitivity.[24,25,30,32] Specificities of these tests[24,25,30,32] were almost near to each other and ranged from 66.7 to 70.8%. The test proposed by Lim *et al*.,[25] that was a combination of water swallow and pulse oximetry, achieved the highest sensitivity and specificity (100 and 70.8%, respectively).

Systematic reviews are prone to the selection bias, especially if they were limited to studies in English. It means a systematic review is not probably included of all available studies about a specific subject.[17] This kind of bias was likely the most significant limitation of the present systematic review. In addition, our search strategy was restricted to a few databases and did not include a manual search of available books in swallowing disorders or stroke. We cannot assert that we searched all available articles on swallowing screening following stroke. Also we focused on tests that compared with VF or FEES. It resulted in the exclusion of some popular tests like the Burk dysphagia screening test[43] because of its used reference test. In addition, we were strict about spectrum bias in the quality appraisal assessment in order to be able to generalize the results to the almost all admissions with acute stroke to the hospitals. Some other articles, therefore, were assessed as having an evidence level II and so were not reported in detail.

CONCLUSION

We were hoping to find simple, valid, reliable, sensitive, and specific tests for screening swallowing disorders in almost all acute alert stroke patients. It seems the four reported high qualified screening tests including Oral Pharyngeal and Clinical Swallowing Examination,[24] Bedside Aspiration Test,[25] The Gugging Swallowing Screen,[30] and The Toronto Bedside Swallowing Screening Test (TOR-BSST),[32] have almost all of these characteristics. Further researches are needed to investigate the effects of the administration of these tests upon stroke patients' outcomes. Also, further validation and reliability assessing of screening tools need to follow a very accurate and well-established method in a large sample of almost all stroke patients admitted to the hospitals. Only such screening tools could ultimately lead to the reduction of the consequences of swallowing disorders in the patients with stroke.

AUTHORS' CONTRIBUTIONS

SJ contributed in the conception of the work, design of the work, analysis and interpretation of data, conducting the study, revising the draft, approval of the final version of the manuscript, and agreed for all aspects of the work. MP contributed in the conception of the work, design of the work, interpretation of data, revising the draft, approval of the final version of the manuscript, and agreed for all aspects of the work.

REFERENCES

1. McLaren S. Eating disabilities following stroke. Br J Community Health Nurs 1997;2:9-19.
2. Daniels SK, Brailey K, Priestly DH, Herrington LR, Weisberg LA, Foundas AL. Aspiration in patients with acute stroke. Arch Phys Med Rehabil 1998;79:14-9.
3. Smithard DG, O'neill PA, Park C, England R, Renwick DS, Wyatt R, *et al*. North West Dysphagia Group. Can bedside assessment reliably exclude aspiration following acute stroke? Age Ageing 1998;27:99-106.
4. 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 Ageing 2000;29:495-9.
5. 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. J Clin Nurs 2001;10:463-73.
6. Ramsey DJ, Smithard DG, Kalra L. Early assessments of dysphagia and aspiration risk in acute stroke patients. Stroke 2003;34:1252-7.
7. Martino R, Foley N, Bhogal S, Diamant N, Speechley M, Teasell R. Dysphagia after stroke incidence, diagnosis, and pulmonary complications. Stroke 2005;36:2756-63.
8. Smithard DG, O'Neill PA, England RE, Park CL, Wyatt R, Martin DF, *et al*. The natural history of dysphagia following a stroke. Dysphagia 1997;12:188-93.
9. Ramsey DJ, Smithard DG. Assessment and management of dysphagia. Hosp Med 2004;65:274-9.
10. Scmidt EV, Smirnov VE, Ryabova VS. Results of the seven-year prospective study of stroke patients. Stroke 1988;19:942-9.
11. Donovan NJ, Daniels SK, Edmiaston J, Weinhardt J, Summers D, Mitchell PH. American Heart Association Council on Cardiovascular Nursing and Stroke Council. Dysphagia Screening: State of the art: Invitational conference proceeding from the State-of-the-Art Nursing Symposium, International Stroke Conference 2012. Stroke 2013;44:e24-31.
12. Logemann JA. Evaluation and treatment of swallowing disorders. San Diego: College-Hill Press; 1983, 1st ed.
13. Association AS-L-H. Preferred practice patterns for the profession of speech-language pathology 2004.
14. Perry L, Love CP. Screening for dysphagia and aspiration in acute stroke: A systematic review. Dysphagia 2001;16:7-18.
15. Hinchey JA, Shephard T, Furie K, Smith D, Wang D, Tonn S. Stroke Practice Improvement Network Investigators. Formal dysphagia screening protocols prevent pneumonia. Stroke 2005;36:1972-6.
16. Fedorak A, Finestone HM, Little J, MacGarvie D, Martino R. Improving recognition and management of dysphagia in acute stroke; 2002.
17. Bours GJ, Speyer R, Lemmens J, Limburg M, de Wit R. Bedside screening tests vs. videofluoroscopy or fibreoptic endoscopic evaluation of swallowing to detect dysphagia in patients with neurological disorders: Systematic review. J Adv Nurs 2009;65:477-93.
18. Jaeschke R, Guyatt G, Sackett D. Users Guides to the medical literature: How to use an article about a diagnostic test. B. What are the results and will they help me in caring for my patients? The Evidence-Based Medicine Working Group. JAMA 1994;271:703-7.
19. Jaeschke R, Guyatt G, Sackett DL. Group E-BMW. Users' guides to the medical literature: III. How to use an article about a diagnostic test: A. Are the results of the study valid? JAMA 1994;271:389-91.
20. DePippo KL, Holas MA, Reding MJ. Validation of the 3-oz water swallow test for aspiration following stroke. Arch Neurol 1992;49:1259-61.
21. Horner J, Brazer SR, Massey EW. Aspiration in bilateral stroke patients: A validation study. Neurology 1993;43:430.
22. Kidd D, Lawson J, Nesbitt R, MacMahon J. Aspiration in acute stroke: A clinical study with videofluoroscopy. QJM 1993;86:825-9.
23. Collins MJ, Bakheit AM. Does pulse oximetry reliably detect aspiration in dysphagic stroke patients? Stroke 1997;28:1773-5.
24. Daniels SK, McAdam CP, Brailey K, Foundas AL. Clinical assessment of swallowing and prediction of dysphagia severity. Am J Speech-Lang Pathol 1997;6:17.
25. Lim SH, 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;16:1-6.
26. Leder SB, Espinosa JF. Aspiration risk after acute stroke: Comparison of clinical examination and fiberoptic endoscopic evaluation of swallowing. Dysphagia 2002;17:214-8.
27. 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. Ann Acad Med Singapore 2003;32:790-4.
28. 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 variables. J Rehabil Med 2005;37:247-51.
29. Ramsey DJ, Smithard DG, Kalra L. Can pulse oximetry or a bedside swallowing assessment be used to detect aspiration after stroke? Stroke 2006;37:2984-8.
30. 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;38:2948-52.
31. Warnecke T, Teismann I, Meimann W, Oelenberg S, Zimmermann J, Krämer C, *et al*. Assessment of aspiration risk in acute ischaemic stroke—evaluation of the simple swallowing provocation test. J Neurol Neurosurgery Psychiatry 2008;79:312-4.
32. 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;40:555-61.
33. Zhou Z, Salle J, Daviet J, Stuit A, Nguyen C. Combined approach in bedside assessment of aspiration risk post stroke: PASS. Eur J Phys Rehabil Med 2011;47:441-6.
34. Umay EK, Unlu E, Saylam GK, Cakci A, Korkmaz H. Evaluation of dysphagia in early stroke patients by bedside, endoscopic, and electrophysiological methods. Dysphagia 2013;28:395-403.
35. Osawa A, Maeshima S, Tanahashi N. Water-swallowing test: Screening for aspiration in stroke patients. Cerebrovasc Dis 2013;35:276-81.
36. Rosenbek JC, Robbins JA, Roecker EB, Coyle JL, Wood JL. A penetration-aspiration scale. Dysphagia 1996;11:93-8.
37. Mann G. MASA: The Mann assessment of swallowing ability. CengageBrain. com; 2002.
38. Rao N, Brady S, Chaudhuri G, Donzelli J, Wesling M. Gold-standard? Analysis of the videofluoroscopic and fiberoptic endoscopic swallow examinations. J Appl Res 2003;3:89-96.
39. Scott A, Perry A, Bench J. A study of interrater reliability when using videofluoroscopy as an assessment of swallowing. Dysphagia 1998;13:223-7.
40. Stoeckli SJ, Huisman TA, Seifert B, Martin–Harris BJ. Interrater reliability of videofluoroscopic swallow evaluation. Dysphagia 2003;18:53-7.
41. Wright F. A swallow assessment clinic-why and how. Radiology 1991;8:55-7.
42. Logemann JA, Veis S, Colangelo L. A screening procedure for oropharyngeal dysphagia. Dysphagia 1999;14:44-51.
43. DePippo KL, Holas MA, Reding MJ. The Burke dysphagia screening test: Validation of its use in patients with stroke. Arch Phys Med Rehabil 1994;75:1284-6.

Source of Support: Nil, **Conflict of Interest:** None declared.