The semisolid bolus swallow test for clinical diagnosis of oropharyngeal dysphagia: a prospective randomised study

Corinna Schultheiss · Ricki Nusser-Müller-Busch · Rainer O. Seidl

Received: 28 January 2011 / Accepted: 2 May 2011 / Published online: 24 May 2011 © The Author(s) 2011. This article is published with open access at Springerlink.com

Abstract  Contrary to clinical experience, clinical swallow tests are predominantly performed using water (water swallow tests, WST). In this study, we examine whether swallow tests performed using a bolus of semisolid food (bolus swallow test, BST) offer benefits. In a prospective, randomised, blind study, the results of a standardised saliva swallow test (SST), WST, BST, combinations of these tests and an endoscopic swallow test (FEES) in patients with oropharyngeal swallowing disorders of neurological (NEU) and non-neurological (NNEU) origin were compared. Sensitivity, specificity, test accuracy and inter-rater reliability were analysed. 62 patients (mean age = 64.68; range = 22–84) were included in the study (NEU = 40; NNEU = 22). A sensitivity of 70.7% (NEU = 70.3%, NNEU = 71.4%) and specificity of 82.5% (NEU = 92.3%; NNEU = 100%) were determined for the WST. The BST + SST was found to have a sensitivity of 89.6% (NEU = 66.7%; NNEU = 90.9%) and a specificity of 72.7% (NEU = 87.5%; NNEU = 90.9%). Analysis of test accuracy showed a statistically significant correlation between FEES and BST + SST. Only BST + SST exhibited statistically significant inter-rater reliability. BST in combination with SST was the sensitive clinical instrument for detecting aspiration both over the patient population as a whole and over the two sub-populations.

Introduction

Diagnosing and treating swallowing disorders represents a major challenge in everyday clinical practice. The gold standard diagnostic procedures for oropharyngeal dysphagia are videofluoroscopy and fibre-optic endoscopic swallowing examination (FEES). Because these procedures are technically demanding, the tendency in everyday practice is to try to obtain meaningful information on a patient’s swallowing ability by using standardised screening procedures. This is primarily achieved using swallow tests with water, modified in a variety of ways [1, 4, 6–9, 17].

Under current recommendations, clinical swallow tests for oropharyngeal dysphagia are carried out using water [1, 4, 6–9, 17]. This, however, runs contrary to clinical experience which shows that semisolid food causes fewer problems in patients with swallowing disorders. Thus, semisolid food is generally used in the early stages of swallowing therapies. Investigations using water, which is more difficult to swallow, have the advantage that test results have a higher sensitivity where the result is negative. Smaller changes which could nevertheless endanger the patient are more likely to be identified. This method also, however, classifies many patients who are able to swallow simpler foods as having a swallowing disorder. This is reflected in the high sensitivity and low specificity of the water swallow test [1, 4, 6–9, 17].
The objective of this study was to examine whether, by changing the consistency of the food used, it is possible to achieve a better balance between sensitivity and specificity and thus achieve a better degree of predictability with regard to swallowing disorders.

Materials and methods

The study was carried out as a prospective, randomised, blind study after being scrutinised by Berlin ethics committee (EA1/087/06). Comparison was made between standardised swallow tests carried out using saliva (SST), water (WST) and a bolus (BST), combinations of the above and the results of FEES [13]. The clinical swallow tests were performed by three therapists trained in the technique. FEES was carried out within 24 h of the clinical swallow test by an ENT specialist who was not present during the clinical tests and was not aware of the results of these tests.

Test subjects

The study was carried out on a mixed population of patients (62 patients; mean age = 64.68; range = 22–84) with swallowing disorders of varied origin who were undergoing in-patient treatment in the ENT department, stroke unit or early rehabilitation clinic of an acute care hospital. Patients were aged over 18 and had sufficient vigilance to ensure that they were able to take food and adequate situational understanding to be able to follow instructions. Patients who were pregnant, were fitted with a non-deblockable tracheostomy tube, had limited vigilance, inadequate situational understanding or clinical indications of an acute infection were excluded from the study. The tests were carried out sitting at a table or sitting up in bed depending on the patient.

Swallow tests

A detailed explanation was given before carrying out all tests, and the written consent of the patient or their carer was obtained.

Clinical variables

To evaluate a patient’s swallowing ability, the clinical variables breathing (airway obstruction, breath noises, raised breathing rate, etc.), voice and coughing (with and without follow-up swallowing) were evaluated after each swallow during the clinical swallow tests [3, 10]. During the saliva swallow test, vigilance (restricted or unrestricted) and swallowing following oral stimulation according to the facio-oral tract therapy (F.O.T.T.®) [14] were also assessed (see Fig. 1).

Saliva

The clinical examination commenced with an assessment of saliva swallowing. Spontaneous swallowing of saliva and swallowing frequency were assessed while preparing for the test and positioning the patient. If spontaneous swallowing did not occur, oral stimulation was performed using F.O.T.T.®. If it proved impossible to facilitate a swallowing attempt, the examination was terminated.

Liquid and semisolid food

For the remainder of the examination, the sequence of food consistencies was selected at random. The WST involved testing two volumes each of 5, 10 and 20 ml water in ascending order. The patient was instructed to drink each volume of liquid in one go. The BST involved testing 1 g (1/3 teaspoon), 2.5 g (1/2 teaspoon) and 5 g (1 teaspoon) of jelly in ascending order. In order to monitor voice tone, patients were asked to phonate an ‘ah’ sound after each individual swallow attempt. A break of 1 min was allowed between each swallow, during which the investigator observed whether the patient was able to remove any residues (check for coughing with or without follow-up swallow). The test was interrupted if the patient was not capable of attempting to swallow, coughed during three swallow attempts or exhibited confused coordination of breathing and swallowing. Swallowing of liquids and semisolid food was assessed using the penetration–aspiration scale [12] (see Fig. 1; liquids were assessed as in the BST).

Fibre-optic endoscopic swallowing examination

The endoscopic control examinations were carried out independently of the clinical tests by an ENT specialist and a therapist. The results of swallowing saliva, liquid and semisolid food were evaluated using the penetration–aspiration scale [12].

Inter-rater reliability

To check inter-rater reliability, the tests were evaluated by two independent investigators observing the clinical tests simultaneously as they were carried out by a third investigator.
Statistical analysis

The individual clinical tests (SST, WST, BST) and combinations of these tests with the saliva swallowing test (WST\textsubscript{SST}, BST\textsubscript{SST}) were subjected to statistical analysis. To examine the accuracy of the test, the sensitivity, specificity, confidence interval (CI) and positive and negative predictive values (PPV, NPV) were determined. Sensitivity and specificity were calculated using a $2 \times 2$ contingency table. The calculation was based on a comparison between the results of the clinical tests and FEES. A 95% confidence interval was used for testing.

The accuracy of the clinical tests was checked using McNemar’s $\chi^2$ test. The difference between the individual clinical tests and FEES was also analysed. There was a difference between the results with $p < 0.05$.

The reliability of the clinical tests was checked using Cohen’s correlation coefficient ($\kappa$). Values greater than 0.60 are evaluated as acceptable, values greater than 0.75 as very good correlation.
Results

During February–August 2008, investigations were carried out on 70 patients, 62 of whom were able to be included in the study. The clinical tests required an average of 20 min, FEES 15 min.

Test subjects

62 patients, 38 men and 24 women (mean age = 64.68, range = 22–84), were included in the study. 20 (32.3%) ENT patients with neoplastic diseases following treatment (surgery or radiotherapy), 16 (25.8%) patients with a CVA, 8 (12.9%) patients with a cerebral haemorrhage, 4 (6.5%) with cerebral contusion following trauma and 2 (3.2%) patients with tetraplegia were studied. The test subjects were divided into two sub-groups according to the aetiology of their disease: patients with a neurological disorder (NEU, n = 40) and patients with a disorder of non-neurological origin (NNEU, n = 22) (see Table 1).

Inter-rater reliability

The clinical tests on 20 randomly selected patients were simultaneously evaluated by two independent investigators. In order to compare the two investigators, the kappa value (Cohen’s $\kappa$) was calculated. The test results for the SST, BST and BST SST showed statistically significant correlations between the investigators ($\kappa > 0.75$; see Table 2).

Swallow tests

Analysis was carried out on the numerical results of the individual tests (see Table 3). To determine sensitivity, specificity and predictive values, the results were compared to the results obtained using FEES.

Table 1  Patient characteristics

| Cause of disease       | Number | Mean age | Sex (F:M) |
|------------------------|--------|----------|-----------|
| NNEU                   |        |          |           |
| ENT                    | 20     | 62.47    | 5:15      |
| Tetraplegia            | 2      | 81.5     | 1:1       |
| NEU                    |        |          |           |
| CVA                    | 16     | 70       | 11:5      |
| Cerebral haemorrhage   | 8      | 66.38    | 2:6       |
| Cerebral contusion after trauma | 4 | 46.75 | 2:2 |
| Other disease          | 12     | 63.5     | 3:8       |

$NEU$ underlying neurological disorder, $NNEU$ non-neurological underlying disorder, $M$ male, $F$ female

Table 2  Inter-rater reliability of the examination procedure

|                      | Investigators 1 and 2 |         | Investors 2 and 3 |         |
|----------------------|-----------------------|---------|-------------------|---------|
|                      | Cohen’s $\kappa$      | Exact sig. | Cohen’s $\kappa$   | Exact sig. |
| SST                  | 0.750                 | 0.000*** | 0.739             | 0.05*   |
| WST                  | 1.000                 | 0.071    | 0.667             | 0.400   |
| WST SST              | 0.632                 | 0.143    | 0.667             | 0.400   |
| BST                  | 0.770                 | 0.011*   | 1.000             | 0.017*  |
| BST SST              | 1.000                 | 0.05*    | 0.739             | 0.033*  |
| WST SST + BST SST    | 0.632                 | 0.143    | 0.250             | 1.000   |

The results for the SST, BST and BST SST showed statistically significant correlations ($\kappa > 0.75$)

$BST$ bolus swallow test, $BST SST$ saliva and bolus swallow test, $WST$ water swallow test, $WST SST$ saliva and water swallow test

Difference of the two investigators is significant $p = 0.000–0.05$ (bold values)

Saliva swallow test

The saliva swallow test achieved a value of 44.4% of the total number of test subjects in the study and a positive predictive value of 40%. The specificity was 72.7% with a negative predictive value of 76.2%.

Water swallow test

The WST for all test subjects showed a sensitivity of 70.7%. Where the SST was included in the analysis (WST SST), sensitivity was reduced to 60.5%. The positive predictive value (PPV) for the WST was 93.5%, for the WST SST 92.5%. The sensitivity of the WST in the non-neurology group was 71.4%, reducing to 57.1% for the WST SST. The PPV was 100% for both tests (WST, WST SST). The calculated PPV achieved values of 95% (WST) and 90% (WST SST) (see Tables 3, 4).

The WST achieved a specificity of 95.2% for all test subjects. When the SST (WST SST) was included, specificity fell to 89.5%. The negative predictive values (NPV) were 60% (WST) and 50% (WST SST). The specificity for the NNEU group was 100% (WST), for the NEU group 92.3% (WST). Where the SST was included, the specificity for the non-neurology group remained unchanged at 100% (WST SST), whilst the value for the neurology group fell to 81.8% (WST SST). The negative predictive values were 66.7% (WST) and 57.1% (WST SST) in the non-neurology group and 63.2% (WST) and 45% (WST SST) in the neurology group.

Bolus swallow test

The sensitivity of the BST in the patient population as a whole was 62.5%, increasing to 89.6% when taken in conjunction with the SST. The positive predictive value for...
the BST was 71.4% and for the BSTSST 74.3% (see Table 5).

The sensitivity was similar in both patient groups (NNEU = 60%, NEU = 64.3%). The sensitivity increased where the BST was evaluated in conjunction with the SST (NNEU = 90.9%, NEU = 88.9%). The positive predictive values for the non-neurology group were 85.7% (BST) and 90.9% (BSTSST). The PPVs for the neurology group were 64.3% (BST) and 66.7% (BSTSST).

The specificity of the BST for all test subjects was 84.2%, reducing to 72.7% for the BSTSST. The NPVs were 78% (BST) and 89.5% (BSTSST). In the non-neurological group, the specificities were 91.7% (BST) and 90.9% (BSTSST). The NPVs were 73.3% (BST) and 90.9% (BSTSST). In the neurology group, the specificities fell to 80.8% (BST) and 63.6% (BSTSST). The NPVs were 80.8% (BST) and 87.5% (BSTSST).

**Sum of the clinical tests**

For the test population as a whole, combining the WST and BST by addition achieved a sensitivity of 76.2% and PPV of 86.5%. A combination of all three clinical tests by addition showed a sensitivity of 84.4% and a PPV of 90.5%. Specificity for the test population as a whole was 75%. The NPV was 60%. A combination of all three clinical tests showed a specificity of 70.9% and an NPV of 63.2%.

**Accuracy of the test**

In order to examine the accuracy of the tests, the results of the clinical swallow tests were tested against the results of the endoscopic examination using McNemar’s $\chi^2$ test ($p < 0.05$). There were statistically significant differences for the results of the WST ($p = 0.000$), WSTSST ($p = 0.000$), BST ($p = 0.027$) and the overall clinical test ($p = 0.001$). The difference between the clinical and endoscopic examinations was not statistically significant for the SST ($p = 0.480$) and BSTSST ($p = 0.481$).

**Discussion**

Patients with swallowing disorders represent a particular challenge in everyday clinical practice. The issue of an appropriate clinical screening instrument for obtaining a comprehensive picture of a patient’s swallowing ability has been a hot topic for several years.

This study has been able to demonstrate that the BST and BSTSST are suitable clinical diagnostic instruments for patients with conditions of both neurological and non-neurological origin (see Fig. 2).

The primary tests used in everyday clinical practice are various modified versions of the WST. The literature describes various versions of the WST with varying results. DePippo et al. [5, 6] described the Burke Dysphagia Screening Test. Their 1992 study examined 44 stroke patients. It found a sensitivity of 76% and a specificity of 94%. Hinds et al. [8] adopted DePippo’s 3 oz WST [5] for

---

**Table 3** Comparison of the results of clinical and endoscopic examinations

| Number | WST | WST (FEES) | BST | BST (FEES) |
|--------|-----|-----------|-----|------------|
| All    | 62  | 41/62 (66.1%) | 30/62 (48.4%) | 25/62 (40.3%) | 20/62 (32.3%) |
| NNEU   | 22  | 14/22 (63.6%) | 10/22 (45.5%) | 11/22 (50%)   | 6/22 (27.3%)  |
| NEU    | 40  | 27/40 (67.5%) | 20/20 (50%)   | 17/40 (42.5%) | 11/40 (27.5%) |

WST water swallow test, BST bolus swallow test, NNEU underlying neurological disorder, NNEU non-neurological underlying disorder

**Table 4** Comparison of the water swallow test

| Test   | Sn  | Sp  | PPV | NPV |
|--------|-----|-----|-----|-----|
| WST    | 70.7| 95.2| 93.5| 62.5|
| NNEU   | 71.4| 100 | 100 | 66.7|
| NEU    | 70.3| 92.3| 95  | 60  |

**Table 5** Comparison of the bolus swallow test

| Test   | Sn  | Sp  | PPV | NPV |
|--------|-----|-----|-----|-----|
| BST    | 62.5| 84.2| 71.4| 78  |
| NNEU   | 60  | 91.7| 85.7| 73.3|
| NEU    | 64.3| 80.8| 80.8| 73.3|

BST bolus swallow test, BSTSST saliva and bolus swallow test, NNEU underlying neurological disorder, NNEU non-neurological disorder, NPV negative predictive value, PPV positive predictive value, Sn sensitivity, Sp specificity
their study of 115 stroke patients and added the clinical variables swallow capacity and volume per swallow. With these two clinical variables taken into consideration, they found a sensitivity of 97% and a specificity of 69%. Excluding these two clinical variables, sensitivity fell to 73%, specificity to 67%. Daniels et al. [4] studied 59 stroke patients using a 70-ml WST. Patients drank two 5 ml, 10 ml and 20 ml volumes of water. Sensitivity was 92.3%, specificity 66.7% (see Fig. 3).

In our study, the WST did not achieve the high levels of sensitivity described in the literature, but did show a higher specificity. This result applied equally to the neurological group. No values for the remaining test subjects are to be found in the literature. The differences in results between test procedures are likely to be due to modifications to the examination procedure and the variables selected.

The clinical variables used to assess swallowing ability in the individual test procedures vary between studies. Following a study by Daniels et al. [2, 4], most test procedures take account of evaluation of voice and the occurrence of coughing after swallowing. Logemann et al. [11] introduced the Northern Dysphagia Patient Check Sheet and cited an aspiration sensitivity of 78% and specificity of 58% for the variable “cough during trail swallows”. On the basis of this study, we added the terms “coughing with follow-up swallow” and “coughing with no follow-up swallow” to the variable ‘cough’ on our examination sheet. This subdivision was intended to help assess whether patients were able to perceive and deal independently with any residue. In addition to assessing voice, assessment of breathing was also adopted.

Extending the procedure previously described in the literature, an SST has been added to the clinical examination. The additional evaluation of saliva should allow a better assessment of the patient’s everyday abilities, making the results more meaningful. Combining the WST and SST reduced both sensitivity and specificity. Sensitivity was also reduced in each of the sub-groups although specificity remained unchanged in the non-neurological group (see Table 3). Contrary to our hypothesis, the comparability of the consistencies of saliva and water meant that combining the SST and WST results did not allow more meaningful conclusions to be drawn.

In contrast to the clinical experience that, at least for patients with disorders of neurological origin, semisolid food is more easily swallowed, BSTs have been studied in
clinical diagnostics only rarely. Tohara et al. [15] studied 63 stroke patients using 3 ml water and 4 g pudding. Their BST had a sensitivity of 72% and a specificity of 66%. A study by Trapl et al. [16] described the Gugging Swallowing Screen (GUSS). 50 stroke patients were examined using the GUSS, which required them to swallow varying volumes of water (3, 5, 10, 20, 50 ml) and 1/2 teaspoon of pudding. The GUSS achieved a sensitivity of 100% and specificity of 50–69% (see Fig. 4).

The BST used in our study achieved a lower sensitivity but significantly higher specificity than comparable studies. Patients who did not have a swallowing disorder were identified with a greater degree of certainty. Where the BST was combined with an SST, sensitivity was increased, whilst specificity was reduced, a finding which applied across both sub-groups (non-neurological and neurological). In this case, the combination of two consistencies appears to improve the quality of the conclusion reached.

In order to allow comparison with the GUSS, the results of our tests were summarised in a comparable fashion. Combining all of the tests did not achieve the sensitivity of the GUSS. The result of the BST\textsubscript{SST} did, however, exceed that of the GUSS.

This study confirms that a bolus swallow test offers advantages over a water swallow test. For patients with neurological disorders, who generally have altered sensitivity in addition to motor disorders, the benefit offered by the BST lies in improved perception of the bolus in the oral and pharyngeal cavities and thus better control during the oral phase of swallowing. Water is the more difficult consistency for patients with a swallowing disorder of neurological origin. Despite this fact having been known for many years, previous clinical dysphagia screening tests have primarily tested liquids. One reason for this is concern about the potential risk posed to patients from aspiration of a bolus. This ignores the fact that a bolus such as jelly consists largely of water and is thus more or less equivalent to a liquid. Various studies have shown that the primary factor influencing the occurrence of pneumonia is not the consistency of the substance aspirated, but the bacterial flora within the oral cavity through which the food passes [8]. Another reason for using a liquid is the desire to increase the usefulness of the conclusions drawn from the test by increasing its difficulty. However, our investigations show that a combination of an SST and a BST allows the most reliable assessment of aspiration risk to be made.

The study is limited by its small sample size. Further studies with larger populations are required. The use of endoscopic examinations to assess aspiration is known to have limitations. Other studies also, however, use endoscopy as a control examination. In many cases, in view of the severity of the patient’s underlying condition and the point at which the examination is carried out on the intensive care unit, this is the only option available.

Previous studies have considered only patients with a neurological disorder and have excluded patients with swallowing disorders of non-neurological origin. Patients who have undergone surgery to remove tumours have in most cases substantial anatomical changes due to resection and, in general, a lesser degree of altered sensitivity. In the non-neurological group, the WST also achieved good results in addition to the good results achieved with the BST. The results presented here show for the first time that a single test procedure can be used to investigate all patients. Other patients groups, such as geriatrics and children, must in future be included in scientific studies of clinical procedures.

**Conclusion**

The results of this study show that the BST offers significant benefits for identifying aspiration for both the test population as a whole and for neurological and non-neurological sub-groups. As well as being highly sensitive and specific, a combination of BST and SST was found to be the only test with a statistically significant correlation with the endoscopic examination. The tests have adequate inter-rater reliability for everyday clinical use. A bolus swallow test should in future form an additional component during clinical diagnosis of dysphagia and dysphagia screening.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.
References

1. Daniels S (2000) Patterns of care for dysphagic stroke patients. Semin Speech Lang 21:323–331
2. Daniels SK, Brailey K, Priestly DH, Herrington LR, Weisberg LA, Foundas AL (1998) Aspiration in patients with acute stroke. Arch Phys Med Rehabil 79:14–19
3. Daniels SK, Ballo L, Mahoney MC, Foundas AL (2000) Clinical predictors of dysphagia and aspiration risk: outcome measures in acute stroke patients. Arch Phys Med Rehabil 81:1030–1033
4. Daniels SK, McAdam CP, Brailey K, Foundas AL (1997) Clinical assessment of swallowing and prediction of dysphagia severity. Am J Speech Lang Pathol 6:17–24
5. DePippo KL, Holas MA, Reding MJ (1994) The Burke dysphagia screening test: validation of its use in patient with stroke. Arch Phys Med Rehabil 75:1284–1286
6. DePippo KL, Holas MA, Reding MJ (1992) Validation of the 3-oz water swallow test for aspiration following stroke. Arch Neurol 49:1259–1261
7. Gottlieb D, Kipnis M, Sister E, Vardi Y, Brill S (1996) Validation of the 50 ml drinking test for evaluation of post-stroke dysphagia. Disabil Rehabil 18:529–532
8. Hinds NP, Wiles CM (1998) Assessment of swallowing and referral to speech and language therapists in acute stroke. Q J Med 91:829–835
9. Kidd D, Lawson J, Neshitt R, McMahon J (1993) Aspiration in acute stroke: a clinical study with videofluoroscopy. Q J Med 86:825–829
10. Langmore SE, Terpenning M, Schork A, Chen Y, Murray J, Lopatin D, Loesche W (1998) Predictors of aspiration pneumonia: how important is dysphagia? Dysphagia 13:69–81
11. Logemann JA, Veis S, Colangelo L (1999) Screening procedure for oropharyngeal dysphagia. Dysphagia 14:44–51
12. Rosenbeck JC, Robbins JA, Roecker EB, Coyle JL, Wood JL (1996) A penetration-aspiration scale. Dysphagia 11:93–98
13. Seidl RO, Nusser-Müller-Busch R (2001) Der Berliner Dysphagie Index: Eine Einführung, pp 1–43
14. Seidl RO, Nusser-Müller-Busch R, Hollweg W, Westhofen M, Ernst A (2007) Pilot study of a neurophysiological dysphagia therapy for neurological patients. Clin Rehabil 21:686–697
15. Tohara H, Saitho E, Mays KA, Kuhlmeier K, Palmer JB (2003) Three tests for predicting aspiration without videofluorography. Dysphagia 18:126–134
16. Trapl M, Enderle P, Nowotny M, Teuschl Y (2007) Dysphagia bedside screening for acute-stroke patients: the Gugging swallowing screen. Stroke 38:1–7
17. Wu MC, Chang YC, Wang TG, Lin LC (2004) Evaluating swallowing dysfunction using a 100-ml water swallowing test. Dysphagia 19:43–47