Water Swallow Screening Test for Patients After Surgery for Head and Neck Cancer: Early Identification of Dysphagia, Aspiration and Limitations of Oral Intake

CHRISTIANE HEY¹, BENJAMIN P. LANGE², SILVIA EBERLE¹, YEVGEN ZARETSKY¹, ROBERT SADER³, TIMO STÖVER⁴ and JENS WAGENBLAST⁴

Departments of ¹Phoniatrics and Pediatric Audiology, ³Oral, Cranio-, Maxillofacial and Facial Plastic Surgery, and ⁴Oto-Rhino-Laryngology, Head and Neck Surgery, University of Frankfurt/Main, Frankfurt/Main, Germany;

²Institute of Medical Psychology and Medical Sociology, Center of Psychosocial Medicine, Faculty of Medicine, University of Göttingen, Göttingen, Germany

Abstract. Patients with head and neck cancer (HNC) are at high risk for oropharyngeal dysphagia (OD) following surgical therapy. Early identification of OD can improve outcomes and reduce economic burden. This study aimed to evaluate the validity of a water screening test using increasing volumes postsurgically for patients with HNC (N=80) regarding the early identification of OD in general, and whether there is a need for further instrumental diagnostics to investigate the presence of aspiration as well as to determine the limitations of oral intake as defined by fiberoptic endoscopic evaluation of swallowing. OD in general was identified in 65%, with aspiration in 49%, silent aspiration in 21% and limitations of oral intake in 56%. Despite a good sensitivity, for aspiration of 100% and for limitations of oral intake of 97.8%, the presented water screening test did not satisfactorily predict either of these reference criteria due to its low positive likelihood ratio (aspiration=2.6; limitations of oral intake=3.1). However, it is an accurate tool for the early identification of OD in general, with a sensitivity of 96.2% and a positive likelihood ratio of 5.4 in patients after surgery for HNC.

Oropharyngeal dysphagia (OD) is a common sequela in approximately 75% of patients after treatment for head and neck cancer (HNC) (1, 2). Its prevalence increases with tumor size and extent of resection (3-5). As a crucial contributing factor to

Correspondence to: Christiane Hey, MD, Department of Phoniatrics and Pediatric Audiology, University of Frankfurt/Main, Theodor-Stern-Kai 7, House 7A, D-60590 Frankfurt/Main, Germany. Tel: +49 06963015775, Fax: +49 06963015002, e-mail: Christiane.Hey@kgu.de

Key Words: Dysphagia, swallowing disorder, head and neck cancer, postsurgical, screening.

malnutrition and aspiration, inadequately managed OD increases mortality and overall healthcare costs due to severe comorbidities, tube dependency and prolonged length of hospital stay (6, 7). The early identification of patients who need intervention with a valid OD screening tool is the first critical step to improve outcomes (8) and reduce the economic and social burden (9). For patients with acute and chronic stroke, water swallow tests (WST) are mainly used as a screening tool for OD, with aspiration as the main reference criterion.

There is one OD screening tool published for HNC based on the 100 ml WST of Wu *et al.* (10) who reported swallowing speed as being a sensitive indicator for identifying patients at risk for swallowing dysfunction. However, this OD screening tool is validated for patients with HNC following (chemo)radiotherapy (11-13) and focuses only on the reference criterion of aspiration, whereas severe OD, especially in patients with oral cancer, can exist without aspiration (8, 14).

To our knowledge, no postsurgical screening tool exists for patients with HNC, especially none which focuses not only on aspiration but on limitations of oral intake and OD in general with the need for further instrumental diagnostics.

Hence, the aim of our study was to examine the screening validity of a WST using progressively greater volumes in postsurgical HNC by three reference criteria, namely OD in general with the need for further instrumental diagnostics; aspiration; and limitations of oral intake defined by fiberoptic endoscopic evaluation of swallowing (FEES).

Patients and Methods

Patients. In this prospective study, 98 patients were recruited between November 2010 and February 2013 after surgery for HNC. The inclusion criteria were defined as Union internationale contre le cancer (UICC) stage II-IV, age 18 to 99 years and written informed consent. Patients with neurological diseases (n=5) or pre-existing OD (n=8) were excluded. Eighty-five patients fulfilled the clinical inclusion

0250-7005/2013 \$2.00+.40 4017

Table	I.	Patient	characteristics	and	prevalence	of	or opharynge al
dyspho	agi	a in gene	eral (ODG), aspi	ratio	ı (ASP), and	tub	e dependency.

Patient characteristic		ODG	ASP	Tube dependency
	n	n (%)	n (%)	n (%)
Age				
<60	38	24/38 (63.2)	18/38 (47.4)	22/38 (57.9)
>60	42	28/42 (66.7)	21/42 (50.0)	23/42 (54.8)
Tumor stage				
II	23	9/23 (39.1)	5/23 (12.8)	6/23 (26.1)
III	9	5/9 (55.6)	3/9 (33.3)	4/9 (44.4)
IV	48	38/48 (79.2)	31/48 (79.4)	35/48 (72.9)
Tumor site				
Cavitas oris	30	19/30 (63.3)	11/30 (36.7)	16/30 (53.3)
Oropharynx	34	25/34 (73.5)	21/34 (61.8)	23/34 (67.7)
Hypopharynx/larynx	16	8/16 (50.0)	7/16 (43.8)	6/16 (37.5)
Total	80	52/80 (65.0)	39/80 (48.8)	45/80 (56.3)

criteria, however, five patients refused to give written consent. Eighty patients were included in our study, 58 males and 22 females, ranging from 18.92 to 87.75 years of age, with a mean age of 60.96 years (SD=12.93). Refer to Table I for patients' characteristics. The Ethics Committee of the University Hospital Frankfurt/Main, Germany, approved the study protocol (approval # 240/10).

Procedure. Before the first postsurgical oral intake, all 80 patients underwent the WST administered by one of two speech and language pathologists, both with over five years' experience of swallowing disorders. FEES (Langmore standard) (15) was followed as reference directly after the WST for which the maximum time frame was one hour. The evaluation was performed by the first Author, a phoniatrician, with more than 10 years' experience in FEES, blinded and independently of the speech and language pathologists.

Water swallow test. For safety reasons, the WST was administered with increasing calibrated volumes of water, starting with 2 ml given with a spoon, followed by 5 ml via a cup, then 10 ml and finally 20 ml. The first two volumes were offered twice [2 ml=swallow (S) 1a and S1b, 5 ml=S2a and S2b]. In cases of a failure of one of these two attempts, a third was completed (2 ml=S1c; 5 ml=S2c), whereas 10 ml (S3) and 20 ml (S4) were offered only once (Figure 1). At maximum, a total of 51 ml was given.

Three criteria for failure were defined: i) wet voice before swallowing, ii) wet voice/voice change after swallowing and iii) cough or throat clearing after swallowing. The WST were recorded as failed and discontinued if one of the three criteria were fulfilled and passed if none were fulfilled during the two attempts at 2 and 5 ml and one attempt at 10 and 20 ml. The WST was not started and therewith was recorded as failed when a patient presented with wet voice before the test and was not able to clear his/her throat sufficiently.

The inter- and intra-rater reliability was determined on 20 additional patients with HNC. The entire WST was administered, scored and documented in about five minutes.

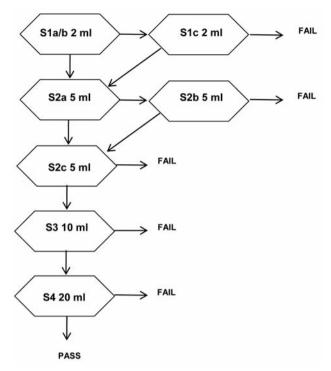


Figure 1. Screening flow chart of the water swallow test.

Reliability. The overall intra-rater reliability (Cohen's K=1, p<0.001) and the overall inter-rater reliability (Cohen's K=1, p<0.001) were both excellent.

FEES. FEES was performed following the Langmore standard (16) with a transnasal flexible endoscope 11101 RP2 (Karl Storz GmbH, Germany) and recorded with an ENT video endoscopy system EndoStrob-DX (Xion medical GmbH, Germany).

Swallowing was evaluated with calibrated volumes starting with 2 ml of water and proceeding with 5, 10 and 20 ml. In addition, puree and solid consistencies with progressive volumes were tested. In cases of aspiration which were unchangeable by therapeutic intervention, FEES was terminated.

Aspiration was defined by a level ≥ 6 on the penetration aspiration scale (PAS) of Rosenbek *et al.* (17) (6=material enters the airway, passes below the vocal folds, and is ejected into the larynx or out of the airway), limitations of oral intake by a level ≤ 4 on the functional oral intake scale (FOIS) of Crary *et al.* (18) (4=total oral diet of a single consistency). OD in general was defined by PAS ≥ 4 (4=material enters the airway, contacts the vocal folds, and is ejected from the airway), or FOIS ≤ 4 .

FEES reliability. The digitally recorded examinations were rated, with 25 percent of them were re-rated two months later to analyze the intra-rater reliability. In order to determine the inter-rater reliability, the same recordings were rated by a second examiner with experience of over 10 years in FEES who was blinded to the WST screening results and the ratings of the first author.

The intra-rater reliability was excellent (Cohen's κ =0.898, p<0.001) for OD in general, for aspiration (Cohen's κ =0.935,

Table II. Diagnostic accuracy measurements of the water swallow test for oral dysphagia in general (ODG), aspiration (ASP), and limitations of oral intake (L IO)

Reference criteria	Results (tp/fp/fn/tn)	Sensitivity (%) 95% CI	Specificity (%) 95% CI	LR+ 95% CI	Efficiency 95% CI
ODG	50/5/2/23	96.2 (86.8-99.5)	82.1 (63.1-93.9)	5.4 (2.4-11.9)	91.3 (82.8-96.4)
ASP	39/16/0/25	100 (91.0-100)	61.0 (44.5-75.8)	2.6 (1.8-3.8)	80.0 (69.6-88.1)
L IO	44/11/1/24	97.8 (88.2-99.94)	68.6 (50.7-83.2)	3.1 (1.9-5.1)	85.0 (75.3-92.0)

tp: True-positive, fp: false-positive, fn: false-negative, tn: true-negative. LR+: Positive likelihood ratio, CI: confidence interval, efficiency: number of correct decisions.

p<0.001) and for limitations of oral intake (Cohen's K=0.925, p<0.001). The inter-rater reliability was also excellent for OD in general (Cohen's K=0.898, p<0.001), for aspiration (Cohen's K=0.936, p<0.001) and for limitations of oral intake (Cohen's K=1, p<0.001).

Statistical analyses. The prevalence of OD in general, aspiration and limitations of oral intake were analyzed using descriptive statistics. An ANOVA was carried out to test the effect of tumor site, stage and patient age.

The screening accuracy was determined by means of sensitivity, specificity, positive likelihood ratio and efficiency.

To analyze the possibility of test item reduction, the summarized pass/fail decisions for all the individual attempts at 2 ml, 5 ml, 10 ml and 20 ml were compared by crosstabs with the final outcome of the WST and the chi-square (χ^2) with significance values were calculated. Similarly, the necessity of S1c and S2c, that is the third swallowing attempts of 2 ml and 5 ml, was examined. To identify the most powerful of the three failure criteria, namely wet voice before, wet voice/voice change after swallowing and cough/throat clearing, the results for these criteria regarding the four volume levels were also compared with the final outcome of the WST by crosstabs and chisquare.

All statistical analyses were performed using SPSS (Version 20, International Business Machines Corp., New York, USA).

Results

Prevalence of OD in general, aspiration and limitations of oral intake. An OD in general was found in 65.0% (52/80) of our study population. Out of all patients, 48.8% (39/80) aspirated, silently in 21.3% (17/80), and 56.3% (45/80) were tube-dependent. None of our patients were scored at a level 4 on the FOIS. The rate of aspiration and tube dependency was significantly influenced by tumor stage, aspiration more so $(p=0.005; \, \eta_p^2=0.138)$ than limitations of oral intake $(p=0.010; \, \eta_p^2=0.122)$. In contrast, age and tumor site had no influence on aspiration and limitations of oral intake $(p>0.4; \, \eta_p^2<0.03)$, even though the patients with oropharyngeal carcinoma had the highest percentage of OD in general at 73.5% (25/34), aspiration at 61.8% (21/34) and tube dependency at 67.7% (23/34).

Refer to Table I for prevalence as well as patient age, tumor stage and tumor site-related details.

Test characteristics: sensitivity, specificity, likelihood ratio and efficiency. Efficiency, i.e. the number of correct decisions (the percentage of test results including true-positive and true-negative ones correctly identified by the test) for the WST was satisfying, with values from 80.0% to 91.3%. The sensitivity was higher (OD in general=96.2%; aspiration=100%; limitations of oral intake=97.8%) than the specificity (aspiration=61.0%, limitations of oral intake=68.6%, OD in general=82.1%). The highest sensitivity (100%) was detected for aspiration and the highest specificity (82%) for OD in general. The positive likelihood ratios for aspiration and limitations of oral intake were low (aspiration=2.6; limitations of oral intake=3.1). The best positive likelihood ratio was detected for OD in general, with a good value of 5.4.

Raw data, sensitivity, specificity, likelihood ratio and efficiency are presented in Table II for OD in general, aspiration and limitations of oral intake.

Item pool analysis. All four volumes contributed significantly to the final outcome of the WST (χ^2 >15.23; p<0.001) (see Table III).

The third swallow of 2 ml (S1c) was necessary in only 23 patients and did not contribute significantly to the final result of the WST (n=23; χ^2 >4.44; p=0.06). The same was true for S2c (n=4; χ^2 >.44; p=0.75).

Analysis of the three failure criteria demonstrated that wet voice before swallow only contributed significantly to the final outcome of the WST before the first 2 ml swallow (χ^2 =5.20; p<0.05. No failure due to this criterion was recorded at any other volume, no crosstabs with chi-square calculations could be performed. Of the two other failure criteria, that is wet voice/voice change after swallowing and cough/throat clearing, only cough/throat clearing consistently significantly contributed to the outcomes of the WST (Table IV).

Discussion

The presented study underlines the high risk of OD in patients after surgery for HNC. Besides the high prevalence of aspiration (49%), limitations of oral intake (56%) and OD in general (65%), we identified an unexpectedly high rate of

Table III. Comparison of the swallow attempts S1a-c (2 ml), S2a-c (5 ml), S3 (10 ml) and S4 (20 ml) with the final outcome of the water swallow test

Swallow attempt	Volume (ml)	N	χ^2
1a-c	2	80	18.53
2a-c	5	52	16.05
3	10	39	15.23
4	20	33	23.00

df=1 for all calculations; all p<0.001.

silent aspirators which is more often documented in patients following (chemo)radiotherapy and rather unusual for postsurgical patients (19). The known influence of tumor stage on the risk for OD is confirmed by our study. Aspiration and limitations of oral intake depend significantly on tumor stage (20, 21). However, patient age and tumor site had no significant effect.

A screening tool for OD needs to be validated with good test quality criteria besides good feasibility, reliability, and safety for the patient (22). The presented WST meets the majority of these demands: it is handy, easy-to-use with an administration time of less than 10 minutes; inter- and intrarater reliability is excellent and the progressive volume administration guarantees patient safety. The test quality criteria of our WST were noteworthy. One of the main focuses of a screening tool for OD is its high sensitivity to ensure detection of OD in this cohort of patients. On this regard, our WST is excellent, with sensitivity ranging from 96% to 100%. The specificity, although not as high as the sensitivity, is satisfactory, ranging from 61% to 82%. Together with the efficiency values, with numbers of correct decisions ranging from 80% to 91%, the WST seems to be a perfect screening tool for the detection of OD in general, or of aspiration or limitations of oral intake in patients after surgery for HNC. In fact, it is of interest how well the WST predicts the three reference complications. Whereas predictive values answer that question, they are vulnerable to shifts in disease due to their prevalence dependence (23), so that a transfer of test results to another population is risky. However, the positive likelihood ratio combines sensitivity and specificity and indicates, independent of prevalence, how more or less likely patients fulfilling one of the reference criteria are prone to have a positive test result than patients who do not fulfill the criteria (24). The positive likelihood ratio of our WST is 2.6 for aspiration and 3.1 for limitations of oral intake. Given the fact that a test with a positive likelihood ratio of 1 indicates the same probability for patients with and without the target condition for having a specific test result, our WST does not predict aspiration or limitations of oral intake satisfactorily. In

Table IV. Item reduction of failure criteria.

Volumes		Wet voice after	Cough/throat clearing
(ml)	N	χ^2	χ^2
2	75	10.99***	16.00***
5	52	8.75**	42.76***
10	39	1.83	15.23***
20	33	14.22**	13.94***

df=1 for all calculations; **p<0.01; ***p<0.001.

fact, a positive likelihood ratio value of between 2 and 5 generates only small changes in probability (25). However, the Patterson screening tool for OD yielded sensitivities between 67% and 94% for the identification of aspiration in patients with HNC following (chemo)radiotherapy, with a positive likelihood ratio of 3.4 pre-treatment and post-treatment values between 1.4 and 1.8 (13).

However, much research on OD screening tools using water swallowing screening was carried out on neurological patients, mainly those with acute and chronic stroke, using a wide range of bolus volumes and with a large variability in validity, reliability and practicability. In the systematic literature review of Daniels et al., only 16 articles out of more than 800 were identified as being eligible regarding a total of 14 study quality criteria such as blinded condition for screeners and for the instrumental examiners, representative patient groups and acceptable delay between clinical and instrumental diagnostics (8). Even in these studies, there is a wide range for sensitivity, specificity and positive likelihood ratio with the highest positive likelihood ratio value around 9 in a study of McCullough et al. (12) (9.5 for 3-oz thin liquid; 9.2 for 10 ml thin liquid, and 6.8 for 5 ml thin liquid) but with unacceptably low sensitivity of around 40% to 50%. However, the 3-oz swallow test in this study was not used if a patient had already demonstrated signs of moderate to severe swallowing impairment, which may explain the low sensitivity. Our test results regarding the predictability of aspiration, however, are comparable with the TOR-BSST of Martino et al. (26), who used, besides tongue movements, the 50-ml water test by Kidd et al. (27) with voice change before and after swallows. With the TOR-BSST, Martino et al. were able to identify aspiration in patients after stroke with a positive likelihood ratio of around 2.6 (2.65 for acute stroke (n=24): sensitivity=96.3%, specificity=63.6%; 2.5 for rehabilitation after stroke (n=35): sensitivity=80.0%, specificity=68.0%).

However, even if our WST as analyzed did not yield a high predictability for aspiration and for limitations of oral intake, it was able to identify OD with a need for further intervention, under a good positive likelihood ratio of 5.4.

Finally, the presented WST was easily and quickly administered. Nevertheless, whereas all of the four-volume levels were important for the final WST outcome, the item pool analysis revealed that the third swallow attempt at 2 ml and 5 ml did not contribute significantly to the WST outcome and can therefore be eliminated. Of the three failure criteria, wet voice before the first 2-ml swallow was important but was without significance for the rest of the WST. Of the two remaining criteria only cough/throat clearing contributed consistently with high significance to the final outcome of the WST, so that cough/throat clearing and wet voice before the beginning of the WST should be finally included in our WST.

Further research should validate the item-reduced WST and analyze its clinical utility for patients with HNC following (chemo)radiotherapy.

References

- Suarez-Cunqueiro MM, Schramm A, Schoen R, Seoane-Leston J, Otero-Cepeda XL, Bormann KH, Kokemueller H, Metzger M, Diz-Dios P and Gellrich NC: Speech and swallowing impairment after treatment for oral and oropharyngeal cancer. Arch Otolaryngol Head Neck Surg 134: 1299-1304, 2008.
- 2 Meyer A, Dietz A, Wollbruck D, Oeken J, Danker H, Meister EF, Sandner A, Volkel W, Brahler E and Singer S: Swallowing disorders after partial laryngectomy. Prevalence and predictors. HNO 60: 892-900, 2012.
- 3 Pauloski BR: Rehabilitation of dysphagia following head and neck cancer. Phys Med Rehabil Clin N Am 19: 889-928, 2008.
- 4 Gagnebin J, Jaques B and Pasche P: Reconstruction of the anterior mouth floor by surgical flap microanastomosis: Oncologic and functional results. Schweiz Med Wochenschr Suppl 116: 39S-42S, 2000.
- 5 Konsulov SS: Surgical treatment of anterolateral tongue carcinoma. Folia Med 47: 20-23, 2005.
- 6 Semenov YR, Starmer HM and Gourin CG: The effect of pneumonia on short-term outcomes and cost of care after head and neck cancer surgery. Laryngoscope 122: 1994-2004, 2012.
- 7 Kyrgidis A, Tzellos TG, Kechagias N, Patrikidou A, Xirou P, Kitikidou K, Bourlidou E, Vahtsevanos K and Antoniades K: Cutaneous squamous cell carcinoma (SCC) of the head and neck: Risk factors of overall and recurrence-free survival. Eur J Cancer 46: 1563-1572, 2010.
- 8 Daniels SK, Anderson JA and Willson PC: Valid items for screening dysphagia risk in patients with stroke: A systematic review. Stroke 43: 892-897, 2012.
- 9 Kaspar K and Ekberg O: Identifying vulnerable patients: Role of the EAT-10 and the multidisciplinary team for early intervention and comprehensive dysphagia care. Nestle Nutr Inst Workshop Ser 72: 19-31, 2012.
- 10 Wu MC, Chang YC, Wang TG and Lin LC: Evaluating swallowing dysfunction using a 100-ml water swallowing test. Dysphagia 19: 43-47, 2004.
- 11 Patterson JM, McColl E, Carding PN, Kelly C and Wilson JA: Swallowing performance in patients with head and neck cancer: A simple clinical test. Oral Oncol 45: 904-907, 2009.

- 12 McCullough GH, Rosenbek JC, Wertz RT, McCoy S, Mann G and McCullough K: Utility of clinical swallowing examination measures for detecting aspiration post-stroke. J Speech Lang Hear Res 48: 1280-1293, 2005.
- 13 Patterson JM, Hildreth A, McColl E, Carding PN, Hamilton D and Wilson JA: The clinical application of the 100 ml water swallow test in head and neck cancer. Oral Oncol 47: 180-184, 2011.
- 14 Martino R, Pron G and Diamant N: Screening for oropharyngeal dysphagia in stroke: Insufficient evidence for guidelines. Dysphagia 15: 19-30, 2000.
- 15 Langmore SE, Schatz K and Olsen N: Fiberoptic endoscopic examination of swallowing safety: A new procedure. Dysphagia 2: 216-219, 1988.
- 16 Langmore SE: Endoscopic Evaluation and Treatment of Swallowing Disorders. Stuttgart: Thieme pp. 101-143, 2001.
- 17 Rosenbek JC, Robbins JA, Roecker EB, Coyle JL and Wood JL: A penetration-aspiration scale. Dysphagia 11: 93-98, 1996.
- 18 Crary MA, Mann GD and Groher ME: Initial psychometric assessment of a functional oral intake scale for dysphagia in stroke patients. Arch Phys Med Rehabil 86: 1516-1520, 2005.
- 19 Nguyen NP, Frank C, Moltz CC, Vos P, Millard C, Smith HJ, Dutta S, Lee H, Martinez T, Karlsson U, Nguyen LM and Sallah S: Dysphagia severity and aspiration following postoperative radiation for locally advanced oropharyngeal cancer. Anticancer Res 28: 431-434, 2008.
- 20 Zuydam AC, Rogers SN, Brown JS, Vaughan ED and Magennis P: Swallowing rehabilitation after oro-pharyngeal resection for squamous cell carcinoma. Br J Oral Maxillofac Surg 38: 513-518, 2000.
- 21 Nicoletti G, Soutar DS, Jackson MS, Wrench AA and Robertson G: Chewing and swallowing after surgical treatment for oral cancer: Functional evaluation in 196 selected cases. Plast Reconstr Surg 114: 329-338, 2004.
- 22 Martino R: Screening and clinical assessment of oropharyngeal dysphagia. Nestle Nutr Inst Workshop Ser 72: 53-56, 2012.
- 23 Westover MB, Eiseman NA, Cash SS and Bianchi MT: Information theoretic quantification of diagnostic uncertainty. Open Med Inform J 6: 36-50, 2012.
- 24 Deeks JJ and Altman DG: Diagnostic tests 4: Likelihood ratios. BMJ 329: 168-169, 2004.
- 25 Jaeschke R, Guyatt GH and Sackett DL: Users' guides to the medical literature. III. 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 271: 703-707, 1994.
- 26 Martino R, Silver F, Teasell R, Bayley M, Nicholson G, Streiner DL and Diamant NE: The Toronto Bedside Swallowing Screening Test (TOR-BSST): Development and validation of a dysphagia screening tool for patients with stroke. Stroke 40: 555-561, 2009.
- 27 Kidd D, Lawson J, Nesbitt R and MacMahon J: Aspiration in acute stroke: A clinical study with videofluoroscopy. Q J Med 86: 825-829, 1993.

Received May 30, 2013 Revised July 9, 2013 Accepted July 10, 2013