Showa Univ J Med Sci **29**(1), 27～36, March 2017

Original

Prediction of Aspiration by Perceptual Evaluation of Pre-swallow Wet Voice and Wet Expiratory Sounds in Adults Diagnosed with Head and Neck Cancer

Michiyo YAMAKAWA*[1)], Kaoru YOKOYAMA[1)], Yoshiko TAKEI[1)], Kazuyoshi KAWABATA[2)], Koji TAKAHASHI[1)] and Michael E. GROHER[3)]

Abstract : This study investigated the relationships between the perceptual evaluation of "wetness" of both pre-swallow voice and expiratory sound and video fluoroscopic swallowing study (VFSS) findings. Pre-swallow phonation of the vowel "a" and expiratory sounds were recorded immediately before VFSS in 51 patients with head and neck cancer. During VFSS, subjects were requested to swallow 3 ml of a jelly-like, radiopaque test food. A total of 61 samples of "a" phonations and expiratory sounds were investigated in this study. These sound samples were randomized and presented to 12 examiners with various years of experience in dysphagia management. The examiners perceptually evaluated the wetness of sound samples using a 5-point "wetness" grade. VFSS findings were evaluated using the 8-point penetration aspiration (PA) scale. The relationships between the wetness of sound samples and VFSS findings were analyzed. Penetration / aspiration without materials ejected out of the airway can be predicted by the wetness of sound samples. In this study, both the pre-swallow wet voice and wet expiratory sounds were suitable for predicting penetration / aspiration after swallowing. High inter-rater and intra-rater reliabilities were verified in the high- and low-experience examiners, with no significant difference evident between these groups. These findings suggest that clinicians could predict penetration / aspiration in head and neck cancer patients by perceptually evaluating the wetness of pre-swallow voice and expiratory sounds regardless of clinical experience.

Key words : pre-swallow wet voice, pre-swallow wet expiratory sound, video fluoroscopic swallowing study, dysphagia prediction, 8-point penetration-aspiration scale

Introduction

The video fluoroscopic swallowing study (VFSS) is widely used in the field of dysphagia management[1)], although the patient's exposure to radiation could be of concern. In addition, invasive techniques such as VFSS are not suitable for patients who have their swallowing status

[1)] *Department of Special Needs Dentistry, Division of Oral Rehabilitation Medicine, Showa University School of Dentistry,* 2-1-1 Kitasenzoku, Ota-ku, Tokyo 145-8515, Japan.
[2)] *Department of Head and Neck, Cancer Institute Hospital.*
[3)] *Truesdail Center for Communicative Disorders, University of Redlands.*
* To whom corresponding should be addressed.

reviewed on a daily basis, including some head and neck cancer patients during the early post-operative stage, and acute and subacute stroke patients, because these patients often have unstable medical conditions and their dysphagic symptoms can change rapidly[2].

On the other hand, cervical auscultation (CA) is a noninvasive technique that can be used daily to evaluate oropharyngeal dysphagia by listening to swallow and respiratory sounds using a stethoscope placed on the patient's neck[3]. Borr *et al*[4] indicated that CA has 70% specificity and 94% sensitivity in detecting signs of dysphagia, while our previous study[5] and work by Sarraf Shirazi *et al*[6] also verified that the frequency characteristics of respiratory sounds after swallow detected from the neck provide an important clue in detecting oropharyngeal dysphagia. Gurgling or liquid vibrating expiratory sounds are consistent with the presence of material on / above the vocal folds or in the pharyngeal cavities, and vocal quality in this circumstance is often described as "wet"[7, 8]. Murugappan *et al*[9] found that phonation with liquid material on the vocal folds shows irregular and aperiodic characteristics, using ex vivo porcine larynges as a model, while Ryu *et al*[10] reported a significant change in the vocal quality of patients with penetration or aspiration after swallowing.

A pre-swallow wet voice and gurgling or liquid vibration sounds might be caused by aspiration or penetration and / or pharyngeal residue of saliva or nasal secretions. Thus, several investigations have been done concerning the predictability of penetration or aspiration using acoustical or auditory voice evaluation of such sounds[11–14]. Aspiration / penetration with materials ejected out of the airway could be judged as a safe swallow, whereas aspiration / penetration of materials not ejected out of the airway is unsafe. In addition, the wet voice and wet respiratory sound (gurgling or liquid vibration sound) have rarely been studied simultaneously.

The aim of this study was to investigate the relationships between VFSS findings and perceptual evaluation of the "wetness" of pre-swallow voice and expiratory sounds. In addition, variability in such perceptual evaluations with clinical experience was investigated. Inter- and intra-rater reliability of listeners' perception of vocal quality were also calculated. Effective perceptual evaluation of pre-swallow voice 'wetness' and expiratory sounds in detecting dysphagic conditions could be valuable in predicting such an outcome in patients without auscultation by stethoscope.

Subjects and methods

Subjects

The study comprised 51 head and neck cancer patients (38 males and 13 females : mean age, 64.9±12.8 years) with suspected dysphagia who received VFSS from May to September 2013. Table 1 details the patients' age, gender, and tumor site. Thirty-four patients were treated with surgery alone, five patients were treated with chemotherapy, radiotherapy, and surgery, five patients were treated with chemotherapy and radiotherapy, four patients were treated with radiotherapy and surgery, and three patients were treated with radiotherapy alone.

Of 61 sound samples in total (detected from 51 patients), 38 sound samples were detected

Table 1. Patient profile

Gender	Number of patients	Age (years±1SD)
Male	38	64.8±11.6
Female	13	65.4±16.0
Total	51	64.9±12.8

Tumor site		Number of samples
Tongue, oral floor	19	23
Maxilla	9	9
Hypopharynx	7	12
Mandible	6	7
Oropharynx	3	3
Larynx	2	2
Other	5	5
Total	51	61

from 30 patients receiving tube feeding at the time of the VFSS (range tube-feeding days, 11-606). Patients wearing a speaking valve after tracheotomy (n = 23) were included. Patients who could not follow auditory commands or could not phonate either with tracheal occlusion or non-occlusion were excluded. Based on these criteria, 3 post-tracheotomy patients were excluded, leaving 51 subjects for this investigation.

The study was approved by the ethics committee of the Showa University School of Dentistry (acceptance number : 2012-014) and the Cancer Institute Hospital (2013-1075). All subjects provided written informed consent.

Recording methods of voice and expiratory sounds

A microphone (COS-11D HWM, Sanken Microphone, Tokyo, Japan) was placed 3 cm laterally from the angle of the mouth without contacting the face. Patients then were asked to phonate the vowel "a" three times, and to exhale voluntarily three times for approximately 1.5 sec per each expiration attempt. The reason why the vowel "a" was chosen as the targeted vowel in this study is that "a" is the easiest vowel to pronounce, even for post-surgical oral-maxilla-pharyngeal cancer patients.

The detected vowel and voluntary expiration sounds were amplified, digitally converted with a 48 kHz sampling rate, and recorded to DVCAM tape (PDVM-41N, SONY) through a digital HD videotape recorder (GV-HD700, SONY).

During detection of vowel and expiration sounds, the tracheostomy speaking-valve opens and closes by respiration and extraneous sounds might be produced as crackles. To avoid these extraneous sounds the speaking valve was covered briefly with surgical tape without any adverse effects on subjects.

A total of 61 samples of phonatory and expiratory sounds were fed to a personal computer

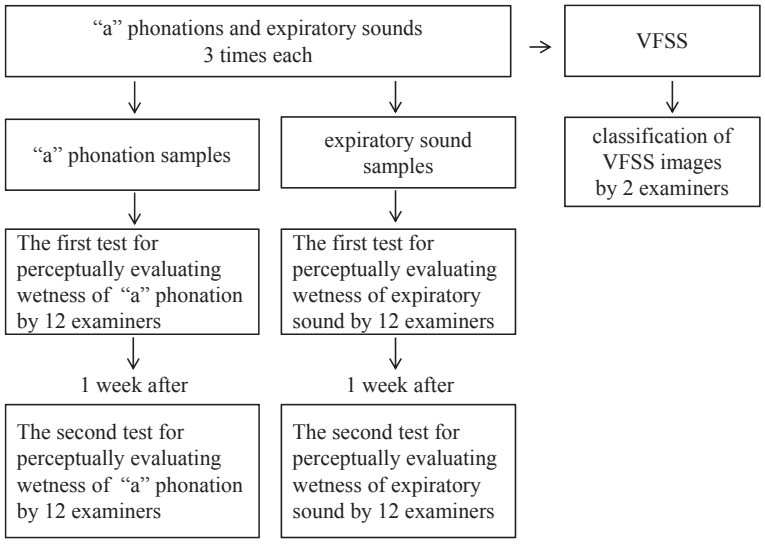

Fig. 1. Evaluation procedure.
VFSS, video fluoroscopic swallowing study

(Folio 9470m, HP) for sound editing.

Evaluation of the phonatory and expiratory sounds

Using editing software (EDIUS Pro 6.5, Grass Valley, CA), phonatory and expiratory sounds were separated and randomized. Using a headphone (ATH-M50, Audio-Technica, Tokyo), both sound samples were presented repeatedly to 12 examiners (ten dentists and two speech pathologists) with varying years of experience in dysphagia management. Examiners who had diagnosed over one thousand swallowing studies were assigned to the high-experience group (n = 5), and examiners who had diagnosed less than one thousand swallowing studies were assigned to the low-experience group (n = 7).

Prior to the main experiment, two experts in dysphagia diagnosis and management with 30 and 15 years' experience judged the sound samples for reliability. These samples were judged from 0 to 4; 0 was normal; 1 was mildly wet, close to normal; 2 was mildly wet; 3 was moderately wet; and 4 was severely wet. Agreement calculations were made for samples that were considered normal and those considered severely wet, and these two sound samples served as the reference samples for all judges. In the main experiment judges listened to all of the sound samples and assigned a number. Prior to listening to each sound, the two reference samples (0 and 4) were played to assist in the final scoring procedure. To evaluate intra-rater reliability, examiners re-evaluated the vowel and expiratory sounds one week after the first evaluation. All sound samples were then re-randomized and presented to the examiners under the same condition. The intra-rater reliability and the inter-rater reliability was calculated in both high- and low-experience groups.

A schema of this investigation is shown in Fig. 1.

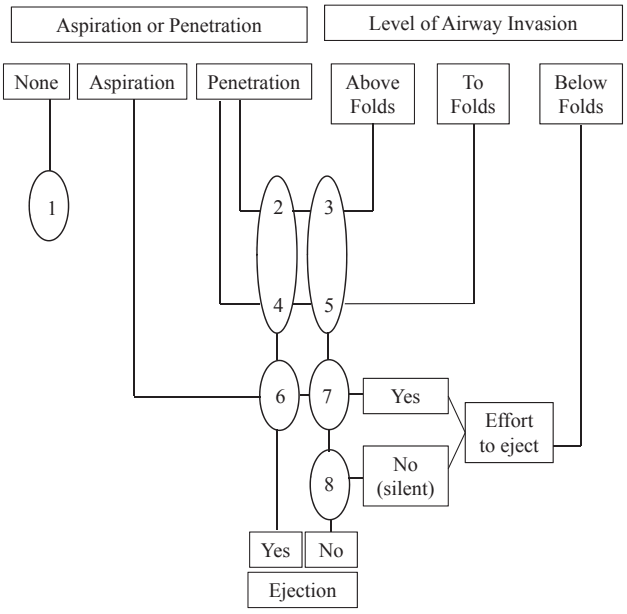

Fig. 2. Schematic representation of the penetration-aspiration scale (modified from reference 15)

Videofluoroscopic swallowing study (VFSS)

The VFSS was done immediately after the phonatory and expiratory sound recording. Subjects were requested to swallow 3 ml of a jelly-like radiopaque test food (thick liquid consistency; 50.0%v/v iopamidol containing 0.27% of agar or 33.3%v/v iopamidol containing 1.57% gelatin) provided by syringe.

If the subject used a special prosthesis such as a palatal augmentation device, he or she was allowed to use it during the evaluation. The material was given to subjects 1–4 times depending on their VFSS findings. VFSS images were recorded to DVCAM tape (PDVM-41N, SONY) through a digital HD videotape recorder (GV-HD700, SONY).

Classification of VFSS images

A modified schematic representation of the penetration-aspiration (PA) scale developed by Rosenbek et al is shown in Fig. 2[15]. Using the PA scale, two dentists with 15 years of experience in dysphagia management evaluated the VFSS findings with frame-by-frame observation, and then classified the images by consensus into two groups: patients graded as 1, 2, 4, or 6 (group 1, material does not enter the airway or material enters the airway, remains above, contacts or passes the vocal folds, and is ejected from the airway) and patients graded as 3, 5, 7, or 8 (group 2, material enters the airway, remains above, contacts or passes the vocal folds, and is not ejected from the airway or the trachea; or material enters the airway, passes below the vocal folds, and no effort is made to eject, namely, silent aspiration). If a patient showed different VFSS findings across multiple swallow trials, the trial with the worst performance was used for classification.

Table 2. Number of sound samples for each group

Group No (PA scale)[a]	Number of sound samples
Group 1 (1, 2, 4, 6)[b]	39
Group 2 (3, 5, 7, 8)[c]	22

[a] PA scale = Penetration-aspiration scale
[b] Group 1 (1, 2, 4, 6) = Material does not enter the airway or material enters the airway, remains above, contacts or passes the vocal folds, and is ejected from the airway
[c] Group 2 (3, 5, 7, 8) = Material enters the airway, remains above, contacts or passes the vocal folds, and is not ejected from the airway or the trachea; or material enters the airway, passes below the vocal folds, and no effort is made to eject, namely, silent aspiration

Statistical Analysis

A statistical analysis was performed using the package IBM SPSS statistics 2.0. Differences between the degrees of wetness on phonatory and expiratory sounds and the PA score groups were analyzed statistically using the Mann-Whitney U test.

To analyze the intra-rater reliability, Kappa coefficients[16] were computed for each examiner on the first evaluation and on the second evaluation one week later. Agreement scores between two evaluations were analyzed in two conditions, those with a perfect match and those within one score difference (κ: in perfect match, κw: within one score difference). These computations were evaluated in the high- and low-experience group. Inter-rater reliability was evaluated the same way, based only on the first evaluation scores.

Sensitivity and specificity were analyzed for group 1 and group 2 by VFSS findings (PAS1, 2, 4, 6 and PAS3, 5, 7, 8). The degree of wetness 1–4 (1: mildly wet, close to normal, 2: mildly wet, 3: moderately wet, 4: severely wet) was evaluated as positive and the degree of wetness 0 (0: normal) was evaluated as negative.

Results

Classification of VFSS findings

Table 2 details VFSS classifications according to the PA scale. Thirty-three patients were PA scale 1, five patients were PA scale 2, eight patients were PA scale 3, four patients were PA scale 5, one patient was PA scale 6, ten patients were PA scale 8, and no patient was PA scale 4 or 7.

The degrees of wetness of "a" phonation and expiratory sound in each PA scale group

The degrees of wetness of "a" phonation and expiratory sound in each PA scale group are shown in Table 3.

The degree of wetness of "a" phonation was significantly higher in group 2 than in group 1 at both the first and second evaluation test ($P < 0.05$), while the degree of wetness of expiratory sound was significantly higher in group 2 than in group 1 only at the first evaluation test ($P < 0.05$).

Table 3. The degree of wetness of "a" phonation and expiratory sound

Group No (PA scale)[a]	The degree of wetness of "a" phonation (mean±1SD)	The degree of wetness of expiratory sound (mean±1SD)
First evaluation test		
Group 1 (1, 2, 4, 6)[b]	1.2±0.9 *	1.6±1.0 *
Group 2 (3, 5, 7, 8)[c]	2.0±1.2	2.1±0.8
Second evaluation test		
Group 1 (1, 2, 4, 6)[b]	1.3±0.9 *	1.7±1.0
Group 2 (3, 5, 7, 8)[c]	2.0±1.2	2.1±0.9

[a] PA scale = Penetration-aspiration scale
[b] Group 1 (1, 2, 4, 6) = Material does not enter the airway or material enters the airway, remains above, contacts or passes the vocal folds, and is ejected from the airway
[c] Group 2 (3, 5, 7, 8) = Material enters the airway, remains above, contacts or passes the vocal folds, and is not ejected from the airway or the trachea; or material enters the airway, passes below the vocal folds, and no effort is made to eject, namely, silent aspiration
* $P < 0.05$ (Mann-Whitney U test)

In summary, patients in group 2 (PA scale 3, 5, 7, 8; penetration / aspiration without materials ejected out of the airway) could be judged effectively by wetness of the phonatory and expiratory sounds. Moreover, phonatory wetness was more easily detected than expiratory wetness.

Intra-rater reliability

Kappa coefficients of phonatory and expiratory sounds for intra-rater reliability are shown in Table 4. Agreement scores were obtained using the means of Kappa coefficients (κ: in perfect match, κw: in one score difference). Moderate agreement scores were achieved for data with a perfect match between investigators ($\kappa = 0.40$ to 0.43), while high agreement scores were shown for data within one score difference ($\kappa w = 0.82$ to 0.94).

On phonatory and expiratory sounds, high agreement was shown for data within one score difference. There was no difference between the high- and low-experience clinician groups. There was no significant difference in reliability between the high-experience group and low-experience group, on phonatory or expiratory sound scoring.

Inter-rater reliability

Kappa coefficients of phonatory and expiratory sounds for inter-rater reliability are shown in Table 5. Low agreement scores were shown for data with a perfect match between investigators ($\kappa = 0.16$ to 0.26), while high agreement scores were shown for data within one score difference ($\kappa w = 0.62$ to 0.78).

On phonatory and expiratory sounds, high agreement was shown for data within one score difference. There was no significant difference in reliability between the high-experience group and low-experience group.

Table 4. Intra-rater reliability for phonatory and expiratory sounds

	κ	κw
"a" phonation		
High-experience group	0.42	0.90
Low-experience group	0.44	0.88
Total	0.43	0.89
Expiratory sound		
High-experience group	0.40	0.94
Low-experience group	0.41	0.82
Total	0.41	0.87

κ = Kappa coefficients in perfect match
κw = Kappa coefficients within one score difference

Table 5. Inter-rater reliability for phonatory and expiratory sounds

	κ	κw
"a" phonation		
High-experience group	0.26	0.78
Low-experience group	0.24	0.63
Total	0.25	0.68
Expiratory sound		
High-experience group	0.16	0.66
Low-experience group	0.23	0.62
Total	0.21	0.63

κ = Kappa coefficients in perfect match
κw = Kappa coefficients within one score difference

Sensitivity and specificity

Sensitivity and specificity was 0.73 and 0.51, respectively, for the "a" phonation, and 0.91 and 0.31, respectively for the expiratory sound, showing high sensitivities of both sound scores.

Discussion

Wet voice can be caused by aspiration or penetration of saliva and / or nasal secretions. Some studies have shown that wet voice during or after swallow may give the clinician an auditory clue that there is threat to the upper airway with subsequent aspiration. Detection of sounds can be done by placing a stethoscope on the lateral neck. This technique is non-invasive and can easily be a routine part of the clinical examination using test swallow items.

In the present investigation, a jelly-like radiopaque test food liquid was chosen as the VFSS test material because it is typically very safe for dysphagic patients to swallow. Using varying test food consistencies, Warms and Richards[13] found no association between the presence of a wet voice and penetration or aspiration of prandial material after swallowing, suggesting that a wet voice could be more predictive of saliva and / or mucoid secretions in the airway rather than indicative of prandial material in the airway.

There are two different types of swallowing in humans, namely the volitional and spontaneous swallow. Spontaneous swallows usually occur as the result of accumulated saliva and / or mucoid secretions while awake or during sleep[17], and are assessed by dripping water directly into the oropharynx[18]. Such an evaluation method is quite invasive, whereas the perceptual evaluation of phonatory or expiratory wetness prior to a swallow is a non-invasive technique that potentially could detect dysphagia or the risk of dysphagia.

Several investigations have studied the predictability of penetration or aspiration using a voice evaluation[11–14]. For instance, Linden *et al*[11, 12] correlated clinical signs of wet phonation, abnormal phonatory quality, harsh phonation, and breathy phonation with subglottic penetration ; however, approximately 1 / 3 of subglottic penetrations were not predicted by overall clinical

indicators including posture, palatal gag, pharyngeal gag, and tongue atrophy. Hey *et al*[14] also investigated the predictability of aspiration and oral intake in 80 post-operative head and neck cancer patients using various kinds of boluses (thin liquid to particulate solid). They reported that wet voice had high specificities (aspiration = 90%, limitation of oral intake = 94%), and low sensitivities (aspiration = 41%, limitation of oral intake = 40%). None of these previous studies clarified whether ejection after aspiration / penetration is related to pre-swallow wet voice. In contrast, the present study revealed that penetration or aspiration without materials ejected out of the airway (PA scale 3, 5, 7, 8) could be predicted from evaluating pre-swallow phonatory and expiratory sounds.

Overall, our results revealed a slight difference between wetness scores on phonatory and expiratory sounds, suggesting that both tasks should be evaluated to effectively predict penetration / aspiration after swallow, especially when materials are not ejected from the airway. It is important to note that in patients with recurrent nerve paralysis, hoarseness may complicate the evaluation of wet voice, and evaluating the wet, voluntary expiratory sounds may be more useful in such cases than the wet phonatory sounds in predicting airway competence.

Clinicians with varying years of experience in dysphagia management performed the perceptual evaluations of wetness for pre-swallow voice and expiratory sounds. Comparisons among these evaluators based on clinical experience showed high inter-rater and intra-rater reliabilities when analyzing the data within one score difference.

Results from prior studies and the present study suggest that clinicians are able to predict penetration / aspiration without materials ejected out of the airway in a small bolus of thick liquid by evaluating wetness of pre-swallow voice and expiratory sound in post-surgical head and neck cancer patients. Therefore it may be important to include this evaluation as a routine part of the clinical examination to detect dysphagia. The technique of using auditory information for detecting dysphagia in post-surgical head and neck cancer patients might also be useful for families and caregivers of these patients as a way to monitor their swallow safety at home. Future research should verify the effectiveness of this procedure for detecting dysphagia in a variety of patients beyond those with head and neck cancer.

Conclusion

The main findings from this investigation were that penetration / aspiration could be predicted from the wetness of pre-swallow voice and expiratory sounds when materials were not ejected from the airway during the VFSS without the use of a stethoscope. In addition, clinical judgments of phonatory and expiratory wetness were not affected by clinical dysphagia management experience. We suggest that the perceptual evaluation of vocal and expiratory wetness be a routine part of the clinical evaluation of suspected oropharyngeal dysphagia.

Acknowledgements

This study was supported by a Grant-in-Aid for Scientific Research (C) from the Japan Society for the Promotion of Science, KAKENHI (25463018).

Conflict of interest disclosure

None of the authors have any conflict of interest to declare.

References

1) Martin-Harris B, Jones B. The videofluorographic swallowing study. *Phys Med Rehabil Clin N Am*. 2008;**19**:769–785.
2) Farrell Z, Murphy E. A comment on "the natural history of dysphagia following a stroke" (Dysphagia 12:188–193, 1997) *Dysphagia*. 1998;**13**:230–231.
3) Zenner PM, Losinski DM, Mills RH. Using cervical ausclutation in the clinical dysphagia examination in long-term care. *Dysphagia*. 1995;**10**:27–31.
4) Borr C, Hielscher-Fastabend M, Lucking A. Reliability and validity of cervical auscultation. *Dysphagia*. 2007;**22**:225–234.
5) Hirano K, Takahashi K, Uyama R, *et al*. Diagnostic criteria of cervical auscultation for assessing sounds during swallowing: the relationships of auditory assessment with VF images, and acoustic characteristics of expiratory sounds immediately after swallowing. *J Jpn Stomatol Soc*. 2001;**50**:82–89. (in Japanese).
6) Sarraf Shirazi S, Buchel C, Daun R, *et al*. Detection of swallows with silent aspiration using swallowing and breath sound analysis. *Med Biol Eng Comput*. 2012;**50**:1261–1268.
7) Schulze-Delrieu KS, Miller RM. Clinical assessment of dysphagia. In *Perlman A, Schulze-Delrieu KA, eds. Deglution and its Disorders: Anatomy, Physiology, Clinical Diagnosis, and Management*. San Diego: Singular Publishing Group; 1997. pp125–152.
8) Goodrich SJ, Walker AI. Clinical swallow evaluation. In *Leonard R, Kendall K, eds. Dysphagia Assessment and Treatment Planning: A Team Approach*. San Diego: Singular Publishing Group; 1997. pp59–72.
9) Murugappan S, Boyce S, Khosla S. Acoustic characteristics of phonation in "wet voice" conditions. *J Acoust Soc Am*. 2010;**127**:2578–2589.
10) Ryu JS, Park SR, Choi KH. Prediction of laryngeal aspiration using voice analysis. *Am J Phys Med Rehabil*. 2004;**83**:753–757.
11) Linden P, Siebens AA. Dysphagia: predicting laryngeal penetration. *Arch Phys Med Rehabil*. 1983;**64**:281–284.
12) Linden P, Kuhlemeier KV, Patterson C. The probability of correctly predicting subglottic penetration from clinical observations. *Dysphagia*. 1993;**8**:170–179.
13) Warms T, Richards J. "Wet voice" as a predictor of penetration and aspiration in oropharyngeal dysphagia. *Dysphagia*. 2000;**15**:84–88.
14) Hey C, Lange BP, Aere C, *et al*. Predictability of oral and laryngopharyngeal function for aspiration and limitation of oral intake in patients after surgery for head and neck cancer. *Anticancer Res*. 2013;**33**:3347–3353.
15) Rosenbek JC, Robbins JA, Roecker EB, *et al*. A penetration-aspiration scale. *Dysphagia*. 1996;**11**:93–98.
16) Feinstein AR. Clinimetrics. New Haven: Yale Univ Press; 1987.
17) Ertekin C. Voluntary versus spontaneous swallowing in man. *Dysphagia*. 2011;**26**:183–192.
18) Yahagi R, Okuda-Akabane K, Fukami H, *et al*. Facilitation of voluntary swallowing by chemical stimulation of the posterior tongue and pharyngeal region in humans. *Neurosci Lett*. 2008;**448**:139–142.

[Received April 26, 2016 : Accepted November 7, 2016]