INTRODUCTION

Dysphagia is one of the main complications associated with aspiration pneumonia. With increases in older adult populations, dysphagia causing aspiration pneumonia and nutritional impairment is a growing problem. With greater focus in modern medicine on quality of life, the importance of accurate swallowing assessment with appropriate evaluation tools is drawing greater attention. Videofluoroscopic swallowing study (VFSS) and fiberoptic endoscopic evaluation of swallowing (FEES) are well known and widely used diagnostic tools for dysphagia. VFSS provides a real-time view of oral, pharyngeal, and cervical esophageal phases of swallowing, and FEES is useful for direct visualization of the anatomy of the nasopharynx and supraglottic structures. As each study has its own strengths and plays a complementary role with the other, combining these two improves their sensitivity for detecting aspiration and residue, compared to use of only one.

Correlations between Aspiration and Pharyngeal Residue Scale Scores for Fiberoptic Endoscopic Evaluation and Videofluoroscopy

Jin A Yoon¹, Sang Hun Kim¹, Myung Hun Jang¹, Sung-Dong Kim², and Yong Beom Shin¹
¹Department of Rehabilitation Medicine, Pusan National University School of Medicine and Biomedical Research Institute, Pusan National University Hospital, Busan;
²Department of Otorhinolaryngology, Pusan National University School of Medicine and Biomedical Research Institute, Pusan National University Hospital, Busan, Korea.

Purpose: To examine correlations among rating scales and findings suggestive of tracheal aspiration and pharyngeal residue between fiberoptic endoscopic evaluation of swallowing (FEES) and videofluoroscopic swallowing study (VFSS) in dysphagia patients. Materials and Methods: We studied patients referred to our hospital for dysphagia assessment. Three raters judged the residue severity and laryngeal penetration or tracheal aspiration of FEES and VFSS. The raters applied the Penetration-Aspiration Scale (PAS) for tracheal aspiration and pixel-based circumscribed area ratio and Yale Pharyngeal Residue Severity for post-swallow residue in VFSS and FEES, respectively. Anatomy-physiologic findings during FEES associated with tracheal aspiration were also analyzed. Results: A total of 178 participants were enrolled in our study. In correlation analysis, PAS ($r=0.74$), vallecula retention ($r=0.76$), and piriform sinus retention ($r=0.78$) showed strong positive correlations between FEES and VFSS. Intra-rater agreement between VFSS and FEES was good for PAS ($κ=0.65$) and vallecula ($κ=0.65$) and piriform sinus retention ($κ=0.69$). Among 72 patients who showed subglottic shelf residue, a suspected finding of aspiration, in FEES, 68 had concomitant tracheal aspiration during VFSS. Both vocal fold hypomobility and glottic gap during phonation were significantly associated with findings suggestive of tracheal aspiration during FEES ($p<0.05$). Conclusion: Quantitative and reliable aspiration and post swallow residue rating scales showed strong positive correlations and good agreement between VFSS and FEES.

Key Words: Deglutition, deglutition disorders, vocal cord dysfunction, endoscopes, fluoroscopy

Received: May 29, 2019 Revised: October 11, 2019
Accepted: October 28, 2019

The authors have no potential conflicts of interest to disclose.
© Copyright: Yonsei University College of Medicine 2019
This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (https://creativecommons.org/licenses/by-nc/4.0/) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.
have been various studies comparing the results of these two examinations, there is still no evidence to support any one of these two exams as a gold standard or if their results are interchangeable.5,6

Determining the presence of tracheal aspiration and severity of residue during a swallowing study is a key factor for clinicians when making treatment plans. Previous studies comparing these two exams have found that their agreement for the presence or absence of tracheal aspiration is very high. However, FEES consistently appears to show worse aspiration and residue severity scores and to be more sensitive because it directly visualizes the laryngeal anatomy in both aspiration and residue scales.6-8 Nonetheless, there are no fundamental differences in the perceptions of swallowing materials between the two exams.

A few studies have assessed the degrees of aspiration or retention by comparing severity through equivalent scaling and quantitative analysis of both exams.7 The scales used for post-swallow residue have had no definite classification standard and have not been validated.4 Therefore, it is not clear how clinicians ought to interpret and combine the results of these studies and clinically relevant tracheal aspiration and post-swallow residue severity. Therefore, additional investigations comparing the results of both studies through valid, reliable scales are necessary. The aim of this study was to examine correlations among the scoring and intuitive findings suggestive of aspiration and post-swallow residue in VFSS and FEES using the most valid, reliable rating scales.

MATERIALS AND METHODS

Subjects
In this prospective observational study, participants who were referred to Pusan National University Hospital for assessment of dysphagia between June 2017 and March 2018 were initially included if they met the following criteria: 1) patients who were referred to our hospital for the assessment of dysphagia; 2) patients who were able to take several bolus challenges of one consistency by mouth. Patients who were unable to sit upright during the exam, were uncooperative with both exams due to cognitive impairment, or who complained of severe discomfort during endoscope insertion were excluded from the study. Approval of this study was obtained from the institutional review board of Pusan National University Hospital (IRB Number: 1706-003-056). Patients were enrolled in the study after providing written informed consent.

Procedure
All patients underwent FEES and VFSS on the same day. A 3.0 mm flexible fiberoptic endoscope Pentax CP-1000, VN9-CP (both Pentax Japan Inc., Tokyo, Japan) was used. Mixed epinephrine and lidocaine HCl/chlorhexidine digluconate gel using a cotton swab were introduced to the nasal mucosa 3 min before scope insertion to enhance comfort during the examination. VFSS was performed using a Toshiba Ultimax-I DREX-UI80 (Toshiba America Medical System, Inc., Tustin, CA, USA) that provides digital storage of high-resolution images at a rate of 30 frames per second. We used different quantities and consistencies of boluses: barium impregnation using 1, 5, and 10 mL of thin liquid and two spoonfuls (3 mL) of puree (yogurt) in equal proportions. Barium powder 97% w/w (Solotop solution 70, Taejoon Pharm Co. Ltd., Yongin, Korea) was mixed for both VFSS and FEES to ensure identical viscosity during exams. The patients were allowed to spontaneously swallow for each type of food. At the time of swallowing, patients were allowed to swallow in whatever position was best for them, as recommend by Langmore10, after which they were repositioned to allow for viewing of the anatomy following the swallow attempt.

The exams were stopped when an event of massive aspiration was observed. For patients with a tracheostomy tube, FEES was performed with the cuff deflated, unless the patient was ventilator-dependent.

Scoring
In this study, one physician conducted a test for VFSS, while the two other physicians witnessed the test and scored the procedure at the same time. For FEES, the same physician conducted a test, and the two other physicians scored the procedure based on the recorded video of the entire test. Three raters judged the severity of residue and the presence of laryngeal penetration or tracheal aspiration; they were blinded to the participants’ details and refrained from pairing the FEES and VFSS recordings. The three raters were attending physicians specializing in rehabilitation medicine who had been performing VFSS for at least 7 years and FEES for at least 3 years with more than 200 cases a year. The raters used the Penetration-Aspiration Scale (PAS) to provide scores for VFSS and FEES. On an 8-point scoring system, scores of 6-8 were given when material entered the airway below the vocal folds, suggestive of tracheal aspiration. During the FEES exam, the final area of food residue after swallowing two types of diet, suggestive of penetration (false or true vocal fold) or aspiration (subglottic shelf) were described (Fig. 1A). The highest PAS during the examination was recorded as representative in both VFSS and FEES. In addition, he distribution of anatomic-physiologic findings suggestive of penetration or aspiration in FEES and VFSS and anatomy-physiological findings during FEES associated with findings suggestive of tracheal aspiration were analyzed.

Post-swallow residue during VFSS was measured based on pixel-based circumscribed area ratios and the ratio of residue relative to available vallecula and pyriform sinus space.11 The pixel-based circumscribed area ratio was utilized to minimize the limitations in quantifying the pharyngeal residue in two-
dimensional VFSS in the lateral view. The “freehand” tool in the picture archiving and communication system toolbar was used to divide the residue area by the area of the vallecula or pyriform sinuses, and this ratio was used to grade severity (Fig. 2). A total of five scaling measurements included no residue, trace (ratios from 0.01–0.05), mild (ratios from 0.05–0.25), moderate (ratios from 0.25–0.50), and severe (ratios over 0.50) were recorded (Fig. 2).

Post-swallow residue during FEES was scored using the Yale Pharyngeal Residue Severity scale, which is the only proven, valid, and reliable vallecula and pyriform sinus severity rating scale. The five-point rating scale is measured using an anatomical, image-based rating scale. The ratings are no residue, trace (1–5%), mild (5–25%), moderate (25–50%), and severe (>50%) (Fig. 1B). The highest residue score during the examination was recorded as representative in both VFSS and FEES.

**Statistical analysis**

Values are given as mean±standard deviation (SD) or as percentages where appropriate. Spearman correlation was applied for analyses showing some degree of association between the variables. Weighted Kappas were calculated between raters and two exams. Inter-rater kappa was calculated using Fleiss kappa. Cohen’s linear weighted kappa was used for ordinal variables. Chi-Square test or Fisher’s exact test was used for categorical variables. *p* values less than 0.05 were considered to indicate statistical significance. All data were analyzed with the use of R 3.6.0. IRR package (R Foundation for Statistical Computing, Vienna, Austria) was used for weighted Kappa analysis.

**RESULTS**

A total of 199 participants were referred to our hospital for the
assessment of dysphagia between June 2017 and March 2018. Finally, 178 participants (55 female, 123 male; mean age 62.8 ±14.1 yr) were enrolled in our study. The reasons for dysphagia were stroke in 53 (29.8%), head and neck cancer in 38 (21.3%), cervical spinal cord injury in 36 (20.2%), vocal fold immobility after intubation in 8 (4.5%), and others, including general de-conditioning, facial fracture, and deep neck infection in 43 (24.2%) (Table 1).

Mean PAS scores were 4.44±2.90 for VFSS and 4.72±2.81 for FEES. Mean vallecula retention scale scores were 3.01±1.26 for VFSS and 2.84±1.15 for FEES; those for pyriform sinus retention were 2.32±1.35 for VFSS and 2.47±1.39 for FEES. Inter-rater reliability kappa statistics for aspiration and post-swallow residue rating in both exams are described in Table 2. Correlation analysis between FEES and VFSS findings are described in Table 3. In correlation analysis, PAS (r=0.74) and vallecula (r=0.76) and pyriform sinus retention (r=0.78) exhibited strong positive correlations between FEES and VFSS. Intra-rater agreement between VFSS and FEES was good for PAS (κ=0.65) and vallecula (κ=0.65) and pyriform sinus retention (κ=0.69).

Anatomic-physiologic findings suggestive of tracheal aspiration were compared between the two studies (Table 4). Among patients who had tracheal aspiration during VFSS (n=77), 53 (68.8%) had subglottic shelf residue, 16 (20.8%) had true vocal fold residue, 6 (7.8%) had false vocal fold residue, and 2 (2.6%) had no evidence of food residue inside laryngeal vestibule during FEES. Among 72 patients who showed subglottic shelf residue, a suspected finding of aspiration, in FEES, 53 had concomitant tracheal aspiration during VFSS. In addition, associations between vocal fold movement and tracheal aspiration during FEES was analyzed (Table 5). Both vocal fold hypomobility and glottic gap during phonation were significantly associated with tracheal aspiration (p<0.05).

**DISCUSSION**

The primary purpose of this study was to compare swallowing scales in VFSS and FEES through quantitative analysis. Corre-

### Table 1. Clinical and Pathological Characteristics of Patients (n=178)

| Characteristics             | Values |
|-----------------------------|--------|
| Age (yr)                    | 62.8±14.1 |
| Sex                         |        |
| Male                        | 123 (69.1) |
| Female                      | 55 (30.9)  |
| Reason for dysphagia        |         |
| Stroke                      | 53 (29.8)  |
| Head and neck cancer        | 38 (21.3)  |
| Cervical cord injury        | 36 (20.2)  |
| Vocal cord injury           | 8 (4.5)    |
| Others                      | 43 (24.2)  |

Values are presented as mean±standard deviation or n (%) unless otherwise indicated.

### Table 2. Inter-Rater Reliability Kappa Statistics for VFSS and FEES Findings

| Kappa  | 95% CI               |
|--------|----------------------|
|        | Lower    | Upper    |
| VFSS   |          |          |
| PAS    | 0.89      | 0.82     | 0.95     |
| Vallec. | 0.91      | 0.85     | 0.96     |
| Pyriform sinus | 0.88      | 0.82     | 0.94     |
| FEES   |          |          |
| PAS    | 0.76      | 0.70     | 0.83     |
| Vallec. | 0.79      | 0.72     | 0.86     |
| Pyriform sinus | 0.71      | 0.65     | 0.78     |

VFSS, videofluoroscopic swallowing study; FEES, fiberoptic endoscopic evaluation of swallowing; PAS, Penetration-Aspiration Scale.

### Table 3. Distribution of Anatomic-Physiologic Findings Suggestive of Penetration or Aspiration in FEES and VFSS

| FEES findings | VFSS findings |
|---------------|---------------|
| None          | 35 (72.9)     | 13 (25.0) | 2 (2.6)  |
| False VF residue | 3 (6.3) | 10 (19.2) | 6 (7.8)  |
| True VF residue   | 4 (8.3) | 16 (30.8) | 16 (20.8) |
| Subglottic shelf residue | 6 (12.5) | 13 (25.0) | 53 (68.8) |
| Total           | 48 (100)     | 52 (100)  | 77 (100) |

FEES, fiberoptic endoscopic evaluation of swallowing; VFSS, videofluoroscopic swallowing study; VF, vocal fold.

Values are presented as n (%).

### Table 4. Association between VF Movement and Tracheal Aspiration during FEES

| PAS (FEES) | 95% CI |
|------------|--------|
| >6         | 0.002  |
| ≤6         |        |
| VF hypomobility |       |
| Yes        | 16 (9.0) | 6 (3.4)  |
| No         | 54 (30.3) | 102 (57.3)|
| VF contact | <0.001† |
| Yes        | 59 (33.1) | 105 (59.0)|
| No         | 13 (7.3)  | 1 (0.6)   |

VF, vocal fold; FEES, fiberoptic endoscopic evaluation of swallowing; PAS, Penetration-Aspiration Scale.

Values are presented as n (%) unless otherwise indicated.

*Significant difference (p<0.05); †Chi-square test or Fisher’s exact test for categorical variable.
lation analysis revealed strong positive correlation and good agreement between FEES and VFSS. Furthermore, when we compared findings suggestive of tracheal aspiration in both studies, all except 2 patients who showed tracheal aspiration during VFSS showed intuitive findings suggestive of penetration or aspiration in FEES. In addition, there was a significant association with tracheal aspiration and vocal fold movement. Both FEES and VFSS provide valuable information in the assessment of dysphagia; therefore, their selection should be clinically mandated. VFSS provides comprehensive information about structures from the lips to the stomach and about the flow of food boluses in the sagittal plane. Meanwhile, FEES focuses mostly on imaging of the pharyngeal region from the nasopharynx to the hypopharynx in the axial plane. There is no gold standard, as neither tool has been shown to be superior in detecting abnormal findings during swallowing. As both modalities yield completely different views, fundamental differences in the perception of findings make it difficult for clinicians to combine the results of the two exams and to draw unified prescription. As both modalities have their respective strengths and they have complementary roles, VFSS and FEES should be used to balance each other. Although the results of both studies should not be similarly interpreted and are not completely interchangeable, efforts are required to precisely analyze and compare patterns when the two tests are performed by the same examiners in a sufficient number of patients in order to determine how to best combine the results of these two tests and, ultimately, to achieve an optimal treatment plan for the patient.

Herein, PAS showed strong positive correlation and good agreement between the two exams. PAS is a reliable measure in both VFSS and FEES, although Colodny reported that FEES was more reliable for assessing penetration than VFSS and that VFSS was more reliable for the assessment of the various severities of aspiration. In previous studies comparing PAS between these two examinations, FEES was reported to consistently score 1 point higher, although these studies had a limited sample size. In our study, the mean PAS score was also higher in FEES, but only by 0.3 points. High positive correlation and good inter-rater agreement were observed for the quantitative measurement of post-swallow residue, rather than perceptual grading used in previous studies. Previous studies comparing the results of the two exams found that FEES consistently shows worse scores in the residue scale, compared to VFSS. However, the standards for defining residue severity are ambiguous and not generally quantitative. In our study, using Yale Pharyngeal Residue Severity for FEES showed no statistically significant difference between pixel-based scoring in VFSS. Since this study was the first to use this same quantitative scale for VFSS and FEES and to analyze the two tests not by equivalent scaling but by quantitative analysis, we consider that it has enough significance to supplement the limitations of previous studies.

In FEES, there is often image loss at the time of swallowing due to pharyngeal structures of the swallowed material covering the tip of the endoscope. Among the symptoms of swallowing-related aspiration, only about 7% are reported to occur during actual swallowing. Nevertheless, according to Langmore, most aspirations occur in the ‘during swallow’ period. Therefore, transglottic tracheal aspiration with a PAS higher than 6 should be inferred based on findings after swallowing. On the other hand, in cases of spillage via a posterior laryngeal route, subglottic shelf residue may not be observed, and as such, the absence of aspiration cannot be concluded with certainty, which is one limitation of FEES. Thus, in the same context, 2 (2.6%) of patients with tracheal aspiration in VFSS had no evidence of aspiration during FEES; however, the lack of residue in the vocal fold or subglottic shelf cannot be used directly to conclude absence of aspiration. Additional reports will be needed on the timing of aspiration during swallowing in FEES. Nonetheless, as it is possible to miss small amounts of aspiration through fluoroscopy as well, combining the results of the two tests could allow for the best sensitivity.

With vocal fold mobility, which could be additionally confirmed during FEES, we conducted an additional analysis to examine the usefulness of FEES as a diagnostic tool along with VFSS. During FEES, anatomic-physiologic assessment using a speaking or breathing task should be performed before swallowing evaluation. When we checked the correlation between laryngeal function and tracheal aspiration, subjects with vocal fold hypomobility or incomplete glottic contact on phonation showed a significantly higher rate of tracheal aspiration. Therefore, when performing FEES, it is important to focus on whether vocal cord hypomobility and complete glottic contact are achieved during phonation: both are important anatomic-physiologic assessment categories before swallowing evaluation to predict tracheal aspiration during FEES.

The limitation of this study is that VFSS and FEES could not be performed simultaneously. However, both studies were performed on the same day and the difference between PAS scores in the two examinations was smaller in our study than seen in a previous study using simultaneous examinations, although the vallecula residue grading was higher in VFSS. Thus, we believe that, even when the above considerations are taken into account, our results are reliable due to the large sample size. In addition, the purpose of our study was to compare scaling itself in both studies for making clinical decisions: both studies are usually not performed at the same time in the clinical field. We also compared the highest and worst PAS and pharyngeal residue scale scores for the two tests, but did not divide and analyze scores according to diet and volume. Nevertheless, since PAS is still used in the clinical field and since the highest score, which refers to the highest level of aspiration or residue, represents a patient’s swallowing function, we believe comparing these results would be still meaningful.

Quantitative and reliable aspiration and post swallow resi-
due rating scales showed strong positive correlation and good agreement between VFSS and FEES. Both studies can work synergistically and may be of use as complementary studies of swallowing evaluation.

AUTHOR CONTRIBUTIONS

Conceptualization: Jin A Yoon and Yong Beom Shin. Data curation: Jin A Yoon, Sang Hun Kim, and Myung Hun Jang. Formal analysis: Jin A Yoon and Yong Beom Shin. Investigation: Jin A Yoon, Sang Hun Kim, and Myung Hun Jang. Methodology: Jin A Yoon and Sung-Dong Kim. Project administration: Jin A Yoon and Yong Beom Shin. Resources: Sang Hun Kim and Sung-Dong Kim. Supervision: Yong Beom Shin. Validation: Jin A Yoon, Sang Hun Kim, and Myung Hun Jang. Visualization: Yong Beom Shin. Writing—original draft: Jin A Yoon. Writing—review & editing: Yong Beom Shin.

ORCID iDs

Jin A Yoon https://orcid.org/0000-0001-5762-0559
Sang Hun Kim https://orcid.org/0000-0003-4849-5228
Myung Hun Jang https://orcid.org/0000-0002-9694-5409
Sung-Dong Kim https://orcid.org/0000-0002-8436-5722
Yong Beom Shin https://orcid.org/0000-0001-5026-1696

REFERENCES

1. Martin-Harris B, Jones B. The videofluorographic swallowing study. Phys Med Rehabil Clin N Am 2008;19:769-85.
2. Langmore SE, Schatz K, Olsen N. Fiberoptic endoscopic examination of swallowing safety: a new procedure. Dysphagia 1988;2: 216-9.
3. Kim JY, Koh ES, Kim HR, Chun SM, Lee SU, Jung SH. The diagnostic usefulness of the fiberoptic endoscopic evaluation of swallowing. J Korean Acad Rehab Med 2011;35:14-22.
4. Kelly AM, Leslie P, Beale T, Payten C, Drinnan MJ. Fibreoptic endoscopic evaluation of swallowing and videofluoroscopy: does examination type influence perception of pharyngeal residue severity? Clin Otolaryngol 2006;31:425-32.
5. ECRI Health Technology Assessment Group. Diagnosis and treatment of swallowing disorders (dysphagia) in acute-care stroke patients. Evid Rep Technol Assess (Summ) 1999;(8):1-6.
6. Kelly AM, Drinnan MJ, Leslie P. Assessing penetration and aspiration: how do videofluoroscopy and fiberoptic endoscopic evaluation of swallowing compare? Laryngoscope 2007;117:1723-7.
7. Rao N, Brady SL, Chaudhari G, Donzelli JJ, Wesling MW. Gold-standard? Analysis of the videofluoroscopic and fiberoptic endoscopic swallowing examinations. J Appl Res 2003;3:89-96.
8. Pisegna JM, Langmore SE. Parameters of instrumental swallowing evaluations: describing a diagnostic dilemma. Dysphagia 2016;31: 462-72.
9. Wu CH, Hsiao TY, Chen JC, Chang YC, Lee SY. Evaluation of swallowing safety with fiberoptic endoscope: comparison with videofluoroscopic technique. Laryngoscope 1997;107:396-401.
10. Langmore S. Endoscopic evaluation and treatment of swallowing disorders. 2nd ed. New York (NY): Thieme; 2011.
11. Rosenbek JC, Robbins JA, Roecker EB, Coyle JL, Wood JL. A penetration-aspiration scale. Dysphagia 1996;11:383-8.
12. Pearson WG Jr, Molfenter SM, Smith ZM, Steele CM. Image-based measurement of post-swallow residue: the normalized residue ratio scale. Dysphagia 2013;28:167-77.
13. Neubauer PD, Rademaker AW, Leder SB. The Yale Pharyngeal Residue Severity Rating Scale: an anatomically defined and image-based tool. Dysphagia 2015;30:521-8.
14. Neubauer PD, Hersey DP, Leder SB. Pharyngeal Residue Severity Rating Scales based on fiberoptic endoscopic evaluation of swallowing: a systematic review. Dysphagia 2016;31:352-9.
15. Langmore SE, Schatz K, Olsen N. Endoscopic and videofluoroscopic evaluations of swallowing and aspiration. Ann Otol Rhinol Laryngol 1991;100:678-81.
16. Langmore SE. Evaluation of oropharyngeal dysphagia: which diagnostic tool is superior? Curr Opin Otolaryngol Head Neck Surg 2003;11:485-9.
17. Hiss SG, Postma GN. Fiberoptic endoscopic evaluation of swallowing. Laryngoscope 2003;113:1386-93.
18. McCullough GH, Rosenbek JC, Robbins JA, Coyle JL, Wood JL. Ordinality and intervals of a penetration-aspiration scale. J Med Speech Lang Pathol 1998;6:65-72.
19. Colodny N. Interjudge and intrajudge reliabilities in fiberoptic endoscopic evaluation of swallowing (Fees(R)) using the penetration-aspiration scale: a replication study. Dysphagia 2002;17:308-15.
20. Smith CH, Logemann JA, Colangelo LA, Rademaker AW, Pauloski BR. Incidence and patient characteristics associated with silent aspiration in the acute care setting. Dysphagia 1999;14:1-7.