

The Stability, Reliability, and Validity of Videofluoroscopy Measures for Patients with Head and Neck Cancer

Jacqui J. Frowen · Susan M. Cotton ·
Alison R. Perry

Published online: 6 February 2008
© Springer Science+Business Media, LLC 2008

Abstract The purpose of this study was to contrast the psychometric properties (stability, test-retest reliability, construct, and concurrent validity) of three different tools used for evaluating videofluoroscopy swallowing studies (VFSS): (1) rating the presence or absence of a swallowing disorder, (2) the Bethlehem Assessment Scale (BAS), and (3) biomechanical measures. These three tools were applied to the same three examinations of two different consistencies (liquid and semisolid), taken from 40 VFSSs of patients with head and neck (H&N) cancer. Stability of swallowing across three swallows was a concern for three measures with the liquid consistency and nine measures with the semisolid consistency. Test-retest reliability was found to vary considerably for the two consistencies (liquids, 0.53–1.00; semisolids, 0.45–1.00). Examination of construct validity of the BAS and biomechanical measures indicated that six factors represented swallowing function, but different factors represented swallowing under liquid and semisolid conditions. Concurrent validity of the presence/absence of disorder variables was less than adequate. These results are discussed in the following contexts: (1) psychometric properties of VFSS may not be adequate for clinical and research environments and (2) psychometric properties of VFSS measures appear to vary as a function of bolus consistency.

Keywords Dysphagia · Deglutition · Deglutition disorders · Videofluoroscopy · Head and neck cancer

Videofluoroscopy swallowing studies (VFSSs) remain the most common clinical method for instrumental assessment of swallowing (dys)function. Different ways of interpreting VFSSs have been described and include descriptions of dysfunction [1–3]; binary ratings, i.e., whether a swallowing movement is present or absent [4–7]; and/or the use of rating scales to describe the degree of impairment [8–10]. In addition to such methods, researchers have developed computer software to make detailed temporal, distance, and biomechanical measures of swallowing physiology from VFSSs [11–14].

The interpretations of VFSSs, and the clinical decisions that are based on such an assessment, remain subjective because we have no consensus or standardized criteria regarding abnormality and/or what would be considered an “unsafe” swallow. The differing observations and measurements that can be taken from VFSS images are limitations to its generalizability and usefulness in both clinical and research settings. Comparisons across studies may not be achievable when there are wide variations in measurement and interpretation. Furthermore, many measures are time-consuming to undertake so do not lend themselves to a clinical setting or, indeed, to many research environments.

In a recent article [15] we suggested the need to establish a core set of swallowing functions that are pertinent to assess when evaluating dysphagia. A minimum data set that can quickly and easily capture crucial information about swallowing function is desirable. Such measures need to be an *accurate* representation of the swallow,

J. J. Frowen (✉) · A. R. Perry
School of Human Communication Sciences,
Faculty of Health Sciences, La Trobe University,
Bundoora 3086, Victoria, Australia
e-mail: J.Frowen@latrobe.edu.au

S. M. Cotton
ORYGEN Youth Health, Department of Psychiatry,
University of Melbourne, Locked Bag 10, 35 Poplar Road,
Parkville 3052, Victoria, Australia

which means ensuring that the measures are stable, reliable, and valid.

Stability considers the level of “agreement” of two or more swallows during a VFSS, i.e., whether the true score remains unchanged across multiple swallows or whether there are systematic biases, resulting in significantly different values, for any/all of the swallows being measured.

Reliability of a tool considers both agreement (as described above) and consistency, i.e., whether the tool is “dependable” and will measure only the trait of interest. Reliability investigates the degree of random error (due to chance) while acknowledging any systematic error (which is predictable and consistent) [16].

An accurate measure of swallowing must also be *valid*, i.e., it must measure what it is intended to measure. *Construct validity* is the degree to which a theoretical construct is measured (e.g., swallowing), while *concurrent validity* is the degree to which outcomes from one test will correlate with the outcomes of a criterion test, given at the same time [16].

It is important to recognize that reliability and validity of measurement tools may vary, depending on the population of interest. For example, the psychometric properties of such measures may be distinct for individuals with no swallowing difficulties compared to a clinical sample, such as people with head and neck (H&N) cancer. In people with H&N cancer, both the disease itself and the subsequent treatment(s) are associated with (often severe) dysphagia [15, 17–20]. Nevertheless, there may be a significant difference in the degree of dysphagia, depending on factors such as the site and/or size of the primary tumor and/or the type or extent of treatment(s). Presenting dysphagia may change in the same individual across the care pathway. Problems experienced before treatment may be more or less severe—or simply different—when compared with those recorded immediately and/or six months after treatment. These variations, both across and within individuals, highlight the need for accurate measurement of swallowing function and the importance of selecting tools that allow stable, reliable, and valid measurements to be made. Thorough investigation of the psychometric properties of tools is the only way to ensure such accuracy. Thus, the first aim of this study was to compare the stability, reliability, and validity of three different types of measures used to analyze the same VFSSs taken from a sample of patients with H&N cancer. These measures were (1) the presence or absence of a swallow disorder, (2) the Bethlehem Assessment Scale (BAS) [21], and (3) detailed biomechanical measures.

Prior researchers have identified differences in swallowing physiology across different bolus consistencies, resulting in different measurement outcomes such as increased tongue base to pharyngeal wall contact, increased

oral and pharyngeal transit times, and increased cricopharyngeal opening for semisolids versus liquids [22–24]. A second aim was therefore to determine whether the psychometric properties of the tools differed when used to analyze swallowing of two bolus consistencies: liquids and semisolids. Using this approach, we aimed to identify key elements for inclusion when assessing dysphagia in a specific group of patients with H&N cancer.

Methods

Participants

Videofluoroscopy swallowing studies (VFSSs) were undertaken with 40 patients who had H&N cancer, three months after completing either radiotherapy ($n = 10$) or chemoradiotherapy ($n = 30$) for definitive treatment. These patients were sequentially recruited from the Peter MacCallum Cancer Centre (PMCC) and The Alfred Hospital, Melbourne, Australia. Thirty-two men and eight women were recruited (age = 42–80 years; mean = 59.5 years). Eighteen of the participants (45%) had been treated for primary tumors of the tonsil. Twelve (30%) had primary tumors within the base of tongue and ten participants (25%) had a tumor of the larynx. Just over half of the participants (52.5%) had large (T3–T4) [25] tumors, while 47.5% had smaller (T1–T2) tumors (Table 1). When examined against The Cancer Council Victoria’s statistics for H&N cancers in 2004 [26], this sample was considered representative of H&N cancer patients throughout the State, for both gender and age.

The variability in the site and size of tumors and the treatment received is due to the sequential nature of the participants’ recruitment. Participants ranged in feeding status, from being completely dependent on enteral nutrition to tolerating a normal diet with no restrictions (Table 1).

The study was approved by the Human Ethics Committees of PMCC, The Alfred Hospital, and La Trobe University, with written informed consent being obtained from all participants.

Procedure

VFSSs were conducted using a standard protocol. Participants were asked to complete three swallows each of measured 3 ml of barium liquid and 3 ml of semisolid (fruit puree) mixed with barium sulfate. All boluses were administered via a 5 ml teaspoon. A bolus volume of 3 ml was chosen because it is small enough to reduce the risk of aspiration but large enough to elicit a swallow reflex [28].

Table 1 Participant demographics

Characteristic	Number (%)
Gender	
Male	32 (80)
Female	8 (20)
Age	
40–49	9 (22.5)
50–59	14 (35)
60–69	9 (22.5)
70–80	8 (20)
Tumor site	
Tonsil	18 (45)
Base of tongue	12 (30)
Larynx	10 (25)
Tumor size ^a	
T1–2	19 (47.5)
T3–4	21 (52.5)
Treatment	
Radiotherapy only	10 (25)
Chemoradiotherapy	30 (75)
Feeding status	
Dependent on enteral nutrition ^b	12 (30)
Modified oral diet ^c	15 (37.5)
Normal oral diet ^d	13 (32.5)

^a Based on TNM staging [25]

^b Rated as 1 or 2 on the Swallowing: Activity Limitation scale of the Australian Therapy Outcome Measures (AusTOMs) [27]

^c Rated as 3 or 4 on the Swallowing: Activity Limitation scale of the AusTOMs

^d Rated as 5 on the Swallowing: Activity Limitation scale of the AusTOMs

Other researchers have demonstrated that a 3 ml volume is often one that patients can swallow throughout all stages of radiotherapy treatment [29–31].

Participants were seated upright, viewed in the lateral plane, and the fluoroscope was focused on the lips anteriorly, the soft palate superiorly, the cervical vertebrae posteriorly, and the lower end of the cervical esophagus inferiorly. A 5-cent Australian coin was taped under the participant's chin to correct for magnification, allowing for calibration during post-hoc analyses. Images were recorded on high-quality videotape, using either the analog Kay Swallowing Workstation (Kay Elemetrics, Lincoln Park, NJ) or a Sony (Model SVO-9500MDP) VHS videocassette recorder (Sony, Australia), depending at which hospital the procedure was conducted. Images were then digitized into .avi format using the video capture software "Virtual Dub" (GNU General Public License) to allow all post-hoc measurements to be made on a personal computer.

Measurements

Three types of measures were completed on all six swallows (three liquid, three semisolid) from each of the 40 participants: (1) the presence or absence of swallow disorders, (2) the BAS, and (3) detailed temporal, distance, and biomechanical measures (collectively referred to as "biomechanical measures").

Presence or Absence of Swallowing Disorders

Numerous swallowing "disorders" are documented in the H&N dysphagia literature and 12 of these were selected for this study on the basis of their reported prevalence [4, 5, 7, 14, 30, 32–34]. These were judged as being "present" or "absent" during each swallow and were operationalized from the literature [2–4, 28, 32] as follows:

- *Poor bolus formation*: material spreads around oral cavity and/or part of the bolus prematurely spills into the pharynx.
- *Prolonged oral transit*: more than 1 second between initiation of the oral swallow (first posterior movement of the bolus from the hold position) and the bolus passing through the faucial arches and/or repeated tongue-pumping motion.
- *Reduced velopharyngeal closure*: velopharyngeal closure is incomplete and/or material enters the nasal cavity and nasal regurgitation seen.
- *Delayed onset of swallow reflex*: the head of the bolus is beyond the point where the lower edge of the mandible crosses the tongue base before the swallow is initiated (the first frame showing laryngeal elevation).
- *Base of tongue (BOT) and/or posterior pharyngeal wall (PPW) weakness*: reduced posterior movement of the BOT with reduced/incomplete contact to the PPW.
- *Reduced laryngeal elevation*: limited superior and/or anterior movement of the larynx during the swallow.
- *Reduced epiglottic inversion*: absent or incomplete (remains horizontal/does not completely close off the laryngeal vestibule) tilting of the epiglottis during the swallow.
- *Reduced laryngeal vestibule closure*: incomplete contact of the arytenoid to epiglottic base.
- *Pharyngeal residue*: any portion of the bolus (more than trace) remains in the valleculae and/or pyriforms and/or BOT and/or PPW postswallow.
- *Cricopharyngeal muscle dysfunction*: delayed opening when the bolus reaches the cricopharyngeus and/or residue in the pyriform sinuses after the swallow.
- *Laryngeal penetration*: part of the bolus enters the larynx and remains at or above the level of the vocal folds.

- *Aspiration (including silent aspiration)*: part of the bolus enters the larynx and passes below the vocal folds into the subglottis.

The Bethlehem Assessment Scale (BAS)

The BAS has 11 domains, each of which are rated on a four-point ordinal rating scale, where 1 represents “normal” and 4 represents a “severe dysfunction” [21]. The 11 domains include lip function, tongue function, jaw function, soft palate elevation, swallow reflex, hyoid elevation, residue in valleculae, residue in pyriform sinuses, aspiration, pharyngeal wall function, and cricopharyngeal function. The BAS has been validated using an Australian population of people with motor neuron disease and has good interrater reliability [35]. Because measures from the BAS are separately rated over each of the 11 different swallowing domains (rather than one rating made of the swallow function as a whole), this was a suitable scale for use in the current study.

Biomechanical Measures

There have been many biomechanical measures used in analysis of VFSSs [11, 12, 28]. In this study we chose those

measures most frequently identified in the H&N cancer literature as being a problem following radiotherapy and where the methods for analyzing them had been clearly described [12, 14, 36, 37]. Temporal measures were completed using frame-by-frame analysis of each swallow to identify the first and last frames that showed each movement of interest. Distance and biomechanical measures were completed using the public domain software program “IMAGE” developed at the National Institutes of Health (Washington, DC). This is a public domain program available at <http://rsb.info.nih.gov/nih-image>. The methods for completing measures from VFSS images using the IMAGE program have been previously described [12, 13]. Table 2 lists the biomechanical measures used in this study, references for how to measure them, and an indication of the direction of *better* scores.

Statistical Analysis

The statistical analyses of data address three aspects of measurement: *stability* (whether the values obtained for the three swallows were significantly different), *reliability* (both agreement and consistency of observed swallows), and *validity* (specifically construct validity), across liquid and semisolid boluses.

Table 2 Biomechanical measures

Variable	Units of measure	Reference (for description of measures)
<i>Temporal measures:</i>		
Oral transit time	seconds	28
Pharyngeal transit time	seconds	28
Pharyngeal delay time	seconds	28
Duration of tongue base to posterior pharyngeal wall contact	seconds	14
Duration of laryngeal vestibule closure	seconds	14, 35
Duration of laryngeal elevation	seconds	14, 35
Duration of cricopharyngeal opening	seconds	14
<i>Distance and anatomical measures:</i>		
Extent of laryngeal excursion ^a	millimeters	12
Extent of hyoid excursion ^a	millimeters	12
Maximum posterior movement of tongue base ^a	millimeters	14, 34
Maximum anterior movement of posterior pharyngeal wall ^a	millimeters	14, 34
Total pharyngeal area at maximum pharyngeal constriction	centimeters ²	12
Extent of laryngeal vestibule closure	millimeters	30
Maximum width of cricopharyngeal opening ^a	millimeters	12
<i>Additional observational measures:</i>		
Approximate percentage of pharyngeal residue	Percentage	2
Number of swallows required to clear the bolus from the pharynx	count	2
Approximate percentage of penetration/aspiration	Percentage	2
Penetration/aspiration scale	ordinal scale	8

^a Higher scores for these measures represent better function. Scores were therefore transformed for these variables prior to data analysis. For all other variables, lower scores represent better function

Stability

Stability was analyzed for the categorical data (presence/absence of swallow disorders) using the Cochran's Q test statistic. Stability across the three swallows was analyzed with all other data (BAS and biomechanical measures), using a series of one-way repeated-measures analysis of variance (ANOVA). The assumption of sphericity was examined using Mauchly's test of sphericity and, when violated ($p < 0.05$), a Greenhouse Geisser test was used to adjust the degrees of freedom. Post-hoc analyses, using the least significant difference (LSD) test, were conducted on variables (from BAS and biomechanical measures) with significant F values from the ANOVA models to enable identification of which of the three swallows significantly differed.

Reliability

For the categorical data (presence/absence of swallow disorders), the overall percentage of agreement was calculated for each variable by assigning "agreement" (i.e., the disorder is present/absent in all three swallows) or "nonagreement" (the disorder is present in some swallows but absent in others) to each of the 40 participants' VFSSs. Reliability was considered to be acceptable across the three swallows for a variable if the percentage of agreement was 75% or above, indicating "good" agreement [16]. To assess both the level of agreement and the degree of correlation (reliability) for the three swallows for the BAS and biomechanical measures, intraclass correlation coefficient (ICC 3,1) was used. Portney and Watkins [16] state that values of 0.75 or above indicate "good" reliability, so this value was chosen to represent "adequate" reliability in this study.

Validity

Factor analysis was used as both a data reduction technique and as a measure of construct validity of the BAS and biomechanical measures. Principal axis factoring with oblique rotation was conducted on swallow 2 for both liquids and semisolids to formulate a pattern matrix and thereby identify the factor loadings for all variables across the two measurement scales. Extraction of factors was based on the eigenvalue greater than 1 rule [38]. Factor loadings greater than 0.4 were considered important to the interpretation of the factor structure, and, therefore, variables that had a poor loading on the factor (i.e., < 0.4) were removed from the analysis.

Variables that clustered together (i.e., were "related" in some way) to form a factor were then analyzed using Cronbach's alpha to establish the degree of correlation between these variables (internal consistency of each of the factors). Cronbach's alpha > 0.75 was considered indicative of a strong relationship between variables loading on that factor (i.e., good internal consistency) [16].

The variables within each factor were also analyzed using Pearson's product moment correlation coefficient (r) to establish which variables were highly correlated with each other (suggesting that they essentially measure the same aspect of the underlying trait) and which were poorly correlated with each other (suggesting that they are "related" but measure different aspects of the underlying trait). Each factor was then assigned a clinical "label" according to the common theme or theoretical construct that characterized that group of variables.

To assess the concurrent validity of the presence/absence variables, point-biserial correlations (r_{pb}) with the factor scores derived from the factor analyses for liquid and semisolid factors were calculated. Only the presence/absence variables that were considered to be clinically related to each factor were analyzed. The purpose of this approach was to capture the relationship between the categorical variables and the other (BAS and biomechanical) measures, and the strength of this relationship was considered "good" to "excellent" for correlation coefficient values above 0.75 [16].

Intra- and interrater reliability for all measures was examined, respectively, by 10% of all measures being repeated by the same examiner six weeks later and independently by a second examiner. The percentage agreement between each measure was calculated for the presence/absence measures and the BAS (BAS measures were considered to be in agreement if the scores were within one rating of each other). Intra- and interrater reliability was at least 96% and 92% (for the presence/absence measures) and 100% and 92% in agreement (for the BAS), respectively. For the biomechanical measures, using ICC, intrarater reliability was at least 83%, and the interrater reliability was at least 80%, except for the variable *maximum PPW movement*, which had a reliability of 72%.

Results

Stability

The descriptive statistics for all variables and swallows are shown in Tables 3 and 4 for liquid and semisolid consistencies, respectively. Two domains of the BAS—*lip function* and *jaw function*—were excluded from all subsequent analyses, because all swallows (for all participants

Table 3 Descriptive statistics and reliability of the three swallows—liquids

Variable	Statistic	Swallow 1	Swallow 2	Swallow 3	Reliability estimate
Presence of swallowing disorder					% agreement ^a
Poor bolus formation	% (n)	25 (10)	27.5 (11)	23.1 (9)	85.0
Prolonged oral transit		7.5 (3)	7.5 (3)	5.1 (2)	100.0
Reduced velopharyngeal closure		2.6 (1)	2.6 (1)	0 (0)	97.4
Delayed swallow reflex		62.5 (25)	67.5 (27)	66.7 (26)	75.0
BOT/PPW weakness		57.5 (23)	60.0 (24)	59.0 (23)	95.0
Reduced laryngeal elevation		32.5 (13)	35.0 (14)	38.5 (15)	87.5
Reduced epiglottic inversion		32.5 (13)	30.0 (12)	25.6 (10)	95.0
Reduced laryngeal vestibule closure		42.5 (17)	37.5 (15)	38.5 (15)	90.0
Pharyngeal residue		55.0 (22)	67.5 (27)	61.5 (24)	82.5
Cricopharyngeal dysfunction		15.0 (6)	17.5 (7)	12.8 (5)	97.5
Penetration		40.0 (16)	42.5 (17)	33.3 (13)	80.0
Aspiration		12.5 (5)	10.0 (4)	12.8 (5)	85.0
Bethlehem Assessment Scale					ICC (3,1)
Lip function	M (SD) ^b	1 (0)	1 (0)	1 (0)	na
Tongue function		1.38 (0.54)	1.38 (.54)	1.44 (0.60)	0.92
Jaw function		1 (0)	1 (0)	1 (0)	na
Soft palate elevation		1.63 (0.82)	1.47 (0.76)	1.53 (0.76)	0.79
Swallow reflex		1.95 (0.92)	2.08 (0.84)	1.97 (0.87)	0.78
Hyoid elevation		2.87 (0.57)	2.87 (0.57)	2.79 (0.70)	0.89
Residue in valleculae		2.21 (0.73)	2.10 (0.75)	2.21 (0.83)	0.83
Residue in pyriform sinuses		1.69 (0.73)	1.77 (0.63)	1.77 (0.74)	0.81
Aspiration		1.85 (1.16)	1.92 (1.18)	1.79 (1.17)	0.84
Pharyngeal wall function*		2.31 (0.95)	2.10 (0.91)	2.10 (0.91)	0.80
Cricopharyngeal function		1.28 (0.69)	1.28 (0.61)	1.26 (0.60)	0.96
Biomechanical measures					ICC (3,1)
Oral transit time	M (SD)	0.35 (0.25)	0.30 (0.19)	0.31 (0.23)	0.53
Pharyngeal transit time		0.86 (0.22)	0.90 (0.25)	0.86 (0.22)	0.73
Pharyngeal delay time		0.15 (0.21)	0.17 (0.23)	0.14 (0.20)	0.62
Duration BOT-PPW contact		0.47 (0.19)	0.48 (0.20)	0.45 (0.17)	0.80
Duration vestibule closure		0.64 (0.26)	0.63 (0.23)	0.60 (0.17)	0.79
Duration larynx elevation		1.30 (0.45)	1.28 (0.42)	1.21 (0.36)	0.63
Duration cricopharyngeal opening*		0.50 (0.15)	0.53 (0.14)	0.52 (0.13)	0.84
Extent larynx excursion		2.50 (0.63)	2.40 (0.71)	2.36 (0.62)	0.78
Extent hyoid excursion		1.56 (0.53)	1.53 (0.51)	1.51 (0.47)	0.73
Maximum BOT movement		0.55 (0.33)	0.54 (0.32)	0.49 (0.28)	0.80
Maximum PPW movement		0.37 (0.22)	0.41 (0.23)	0.37 (0.22)	0.84
Area at maximum pharyngeal constriction		1.13 (0.97)	1.23 (1.09)	1.24 (1.15)	0.94
Extent vestibule closure		0.05 (0.10)	0.05 (0.10)	0.05 (0.09)	0.91
Maximum width cricopharyngeal opening		0.53 (0.18)	0.56 (0.15)	0.56 (0.19)	0.78
Percent pharyngeal residue		16.41 (16.02)	17.56 (15.38)	16.41 (15.77)	0.87
Number of swallows to clear bolus*		2.46 (1.12)	2.67 (1.03)	2.67 (1.01)	0.91
Percent penetrated/aspirated		2.74 (3.83)	2.67 (4.62)	3.21 (6.35)	0.80
Penetration-aspiration scale		2.44 (2.27)	2.26 (2.23)	2.31 (2.39)	0.85

BOT = base of tongue; PPW = posterior pharyngeal wall; na = not applicable

* Significant difference between the three swallows, $p < 0.05$

^a Percentage agreement was determined when a patient had data on at least two swallows

^b Means and standard deviations are based on cases where data for all three swallows has been recorded. Patients who did not have data for all three swallows were excluded

Table 4 Descriptive statistics and reliability of the three swallows—semisolids

Variable	Statistic	Swallow 1	Swallow 2	Swallow 3	Reliability estimate
Presence of swallowing disorder					% agreement ^a
Poor bolus formation*	% (n)	18.4 (7)	2.7 (1)	2.7 (1)	89.2
Prolonged oral transit**		47.4 (18)	24.3 (9)	24.3 (9)	81.1
Reduced velopharyngeal closure		2.7 (1)	2.7 (1)	0 (0)	94.6
Delayed swallow reflex		44.7 (17)	51.4 (19)	54.1 (20)	54.1
BOT/PPW weakness		52.6 (20)	45.9 (17)	45.9 (17)	94.6
Reduced laryngeal elevation		42.1 (16)	37.8 (14)	35.1 (13)	91.9
Reduced epiglottic inversion		26.3 (10)	18.9 (7)	18.9 (7)	97.3
Reduced laryngeal vestibule closure		15.8 (6)	10.8 (4)	13.5 (5)	82.5
Pharyngeal residue**		78.9 (30)	73.0 (27)	59.5 (22)	81.1
Cricopharyngeal dysfunction		15.8 (6)	13.5 (5)	10.8 (4)	97.3
Penetration		2.6 (1)	2.7 (1)	2.7 (1)	100.0
Aspiration		0 (0)	0 (0)	0 (0)	100.0
Bethlehem Assessment Scale					ICC (3,1)
Lip function	M (SD) ^b	1 (0)	1 (0)	1 (0)	na
Tongue function		1.40 (0.55)	1.29 (0.52)	1.26 (0.51)	0.69
Jaw function		1 (0)	1 (0)	1 (0)	na
Soft palate elevation		1.58 (0.75)	1.48 (0.62)	1.52 (0.67)	0.74
Swallow reflex		1.63 (0.77)	1.6 (0.69)	1.69 (0.68)	0.66
Hyoid elevation		2.86 (0.55)	2.89 (0.53)	2.83 (0.62)	0.94
Residue in valleculae*		2.66 (1.03)	2.54 (1.07)	2.37 (1.0)	0.82
Residue in pyriform sinuses		1.29 (0.52)	1.23 (0.43)	1.14 (0.43)	0.69
Aspiration		1.10 (0.37)	1.06 (0.34)	1.06 (0.34)	0.92
Pharyngeal wall function**		2.20 (0.96)	2.03 (0.99)	1.97 (0.95)	0.90
Cricopharyngeal function		1.17 (0.57)	1.11 (0.40)	1.11 (0.53)	0.89
Biomechanical measures					ICC (3,1)
Oral transit time	M (SD)	0.77 (0.57)	0.64 (0.51)	0.68 (0.55)	0.70
Pharyngeal transit time		0.92 (.36)	0.91 (0.34)	0.86 (0.32)	0.54
Pharyngeal delay time		0.20 (0.33)	0.20 (0.37)	0.15 (0.26)	0.56
Duration BOT-PPW contact		0.48 (0.18)	0.48 (0.12)	0.56 (0.12)	0.45
Duration vestibule closure		0.57 (0.17)	0.60 (0.11)	0.58 (0.13)	0.45
Duration larynx elevation		1.19 (0.36)	1.16 (0.33)	1.14 (0.32)	0.79
Duration cricopharyngeal opening*		0.39 (0.10)	0.42 (0.09)	0.45 (0.12)	0.65
Extent larynx excursion		2.42 (0.56)	2.49 (0.61)	2.51 (0.60)	0.75
Extent hyoid excursion		1.45 (0.40)	1.54 (0.47)	1.48 (0.42)	0.62
Maximum BOT movement		0.63 (0.36)	0.63 (0.37)	0.61 (0.30)	0.82
Maximum PPW movement		0.33 (0.20)	0.34 (0.19)	0.37 (0.22)	0.76
Area at maximum pharyngeal constriction*		1.21 (0.96)	1.15 (0.92)	1.02 (0.78)	0.89
Extent vestibule closure		0.03 (0.10)	0.01 (0.02)	0.01 (0.06)	0.48
Max. width cricopharyngeal opening**		0.46 (0.12)	0.52 (0.15)	0.52 (0.14)	0.75
Percent pharyngeal residue**		30.14 (26.02)	23.57 (21.65)	21.71 (20.40)	0.85
Number of swallows to clear bolus		2.94 (1.21)	2.83 (1.29)	2.86 (1.22)	0.91
Percent penetrated/aspirated		0.29 (1.69)	0.14 (0.85)	0.14 (0.85)	0.83
Penetration-aspiration scale		1.11 (0.68)	1.03 (0.17)	1.09 (0.51)	0.73

BOT = base of tongue; PPW = posterior pharyngeal wall; na = not applicable

* Significant difference between the three swallows, $p < 0.05$

** $p < 0.01$

^a Percentage agreement was determined when a patient had data on at least two swallows

^b Means and standard deviations are based on cases where data for all three swallows has been recorded. Patients who did not have data for all three swallows were excluded

and both consistencies) received a value of 1 (i.e., normal) for these factors, indicating zero variance. There were three variables for the liquid consistency and six variables for the semisolid consistency, where significant differences in mean values existed across the three swallows.

Liquids

For the liquid consistency, the following variables were significantly different across the three swallows: (1) *pharyngeal wall function* [BAS; $F(2, 76) = 3.40, p = 0.039$]; (2) *duration of cricopharyngeal opening* [biomechanical measure; $F(2, 76) = 4.08, p = 0.021$]; and (3) *number of swallows required to clear the bolus* [biomechanical measure; $F(1.31, 49.59) = 6.02, p = 0.011$]. Post-hoc analysis revealed that for all three variables, swallow 1 was significantly different than swallows 2 (all $p < 0.05$) and 3 (all $p < 0.05$). Higher (i.e., worse) scores were rated on swallow 1 for *pharyngeal wall function*, indicating that this variable improved as the swallow study progressed, whereas lower (i.e., better) scores were rated on swallow 1 for the *duration of cricopharyngeal opening* and *number of swallows required to clear the bolus*, indicating that these variables became worse as the swallow study progressed.

Semisolids

The following variables were significantly different across the three swallows for the semisolid consistency: (1) *poor bolus formation* [presence/absence measure; Cochran's Q (2) = 6.0, $p = 0.05$]; (2) *prolonged oral transit* [presence/absence measure; Cochran's Q (2) = 10.29, $p = 0.006$]; (3) *pharyngeal residue* [presence/absence measure; Cochran's Q (2) = 9.33, $p = 0.009$]; (4) *residue in valleculae* [BAS; $F(1.67, 56.84) = 4.14, p = 0.027$]; (5) *pharyngeal wall function* [BAS; $F(2, 68) = 5.93, p = 0.004$]; (6) *duration of cricopharyngeal opening* [biomechanical measure; $F(2, 68) = 12.38, p < 0.001$]; (7) *pharyngeal area at maximum constriction* [biomechanical measure; $F(1.641, 55.79) = 3.73, p = 0.038$]; (8) *maximum width of cricopharyngeal opening* [biomechanical measure; $F(2, 68) = 11.49, p < 0.001$]; and (9) *percent of pharyngeal residue* [biomechanical measure; $F(1.523, 51.78) = 10.83, p < 0.001$].

For the variables *presence of poor bolus formation* and *prolonged oral transit*, there were more incidences of oral stage disorders for swallow 1 compared with swallows 2 and 3, indicating that these variables improved as the study progressed. For the variable *presence of pharyngeal residue*, there was less pharyngeal residue present for swallow 3 than for swallows 1 and 2, indicating less residue as the

study progressed. Similarly, post-hoc analysis revealed that for *residue in valleculae* (BAS), swallows 1 and 3 were significantly different ($p = 0.023$), again with scores indicating less residue as the study progressed. Swallow 3 had significantly less *pharyngeal area at maximum constriction* than did swallows 1 ($p = 0.033$) and 2 ($p = 0.031$), and for the *percent of pharyngeal residue*, swallow 1 had significantly higher residue than did swallows 2 ($p = 0.003$) and 3 ($p = 0.001$), indicating improved performance during the swallow study for these variables. For *pharyngeal wall function* (BAS) and *maximum width of cricopharyngeal opening*, there were significantly higher mean scores for swallow 1 compared with swallows 2 ($p = 0.012, p < 0.01$) and 3 ($p = 0.003, p < 0.01$). For *duration of cricopharyngeal opening*, all three swallows differed at the $p < 0.05$ level.

For both liquid and semisolid consistencies, in all instances where a statistically significant difference existed across the swallows, either swallow 1 or swallow 3 was the source of this difference, with swallow 2 being in between. Because of this difference, it was decided that for all variables and for both consistencies, data from swallow 2 only would be used for the subsequent analyses of validity.

Reliability

Liquids

For the liquid consistency, all categorical data (presence/absence of swallow disorders) reached a percentage agreement of 75% or higher, indicating a good level of agreement across the three swallows. All BAS measures demonstrated good reliability (i.e., ICCs > 0.75), with values ranging from 0.78 (BAS swallow reflex) to 0.96 (BAS cricopharyngeal function) (Table 3). There were five biomechanical measures that did not reach the predetermined reliability criterion set in this study. These included *oral transit time* (0.53), *pharyngeal transit time* (0.73), *pharyngeal delay time* (0.62), *duration of laryngeal elevation* (0.63), and *extent of hyoid excursion* (0.73).

Semisolids

There were a number of swallowing measures obtained from the semisolid swallows that did not reach a percentage agreement of 75% nor an ICC value of 0.75 (Table 4). For the presence/absence measures, the *presence of delayed swallow reflex* reached only 54.1% agreement across the three swallows.

On the BAS scale, measures that did not reach an ICC value of 0.75 were *tongue function* (0.69), *soft palate*

elevation (0.74), swallow reflex (0.66); and residue in pyriform sinuses (0.69). For the biomechanical measures, variables that did not reach 0.75 were oral transit time (0.70), pharyngeal transit time (0.54), pharyngeal delay time (0.56), duration of BOT-PPW contact (0.45), duration laryngeal vestibule closure (0.45), duration of cricopharyngeal opening (0.65), extent of hyoid excursion (0.62), extent of vestibule closure (0.48), and penetration-aspiration (0.73).

Validity

Factor analysis revealed the presence of six factors each for liquids and semisolids, with factor loadings above 0.4 and eigenvalues above 1.0. The factors comprised 22 variables for liquids and 18 variables for semisolids. The factor

“structures” underlying each condition (liquids and semisolids) was distinct and is discussed separately. The variables, factor loadings and Cronbach’s alpha values for each of the six factors are shown in Tables 5 and 6 for liquids and semisolids, respectively. The maximum and minimum correlations for the variables within each factor are shown in Tables 7 and 8 for liquids and semisolids, respectively.

Liquids

Six identified factors explained 66.72% of the total variance, with Cronbach’s alpha values ranging from 0.71 to 0.92. The most important factor, explaining 31.76% of the variability, was labeled Pharyngeal Motility, encompassing variables that represent residue of the bolus within the

Table 5 Factor analysis: pattern matrix showing factor loadings for each variable within the factor, clinical labels, and correlations—liquid consistency

Variables ^a	Factor and clinical label					
	1 Pharyngeal Motility	2 Duration of Pharyngeal/Laryngeal Functions	3 Extent of Pharyngeal/Laryngeal Functions	4 Swallow Initiation and Timing	5 Airway Protection	6 Duration of Pharyngeal Transit
Residue in valleculae (BAS)	0.835					
Number of swallows to clear bolus	0.835					
Pharyngeal wall function (BAS)	0.759					
Percent pharyngeal residue	0.695					
Maximum BOT movement	0.678					
Area at maximum pharyngeal constriction	0.594				−0.457	
Residue in pyriform sinuses (BAS)	0.494					
Duration BOT-PPW contact		0.789				
Duration vestibule closure		0.779				
Duration larynx elevation		0.49				
Maximum width of CPO			0.734			
Extent larynx excursion			0.588		−0.407	
Extent hyoid excursion			0.459			
Maximum PPW movement			0.449			
Pharyngeal delay time				−0.914		
Swallow reflex (BAS)				−0.591		
Pharyngeal transit time				−0.528		−0.423
Extent vestibule closure					−0.713	
Cricopharyngeal function (BAS)					−0.660	
Percent penetrated/aspirated					−0.594	
Aspiration (BAS)					−0.534	
Duration CPO						−0.878
% of Variance	31.757	15.451	6.705	4.581	4.377	3.853
Cronbach’s alpha	0.918	0.740	0.705	0.806	0.857	0.759

^a The 22 variables in this factor analysis were the only variables included because they all had loadings greater than 0.4

Table 6 Factor analysis: pattern matrix showing factor loadings for each variable within the factor, clinical labels, and correlations—semisolid consistency

Variables ^a	Factor and clinical label					
	1 Pharyngeal Motility	2 Swallow Initiation and Timing	3 Lower Pharyngeal Function	4 Extent of Pharyngeal/Laryngeal Functions	5 Duration of Pharyngeal/Laryngeal Functions	6 Temporal Measures
Residue in valleculae (BAS)	0.962					
Number of swallows to clear bolus	0.919					
Pharyngeal wall function (BAS)	0.91					
Percent pharyngeal residue	0.881					
Area at maximum pharyngeal constriction	0.805					
Maximum BOT movement	0.479					
Pharyngeal delay time		0.959				
Pharyngeal transit time		0.835				
Swallow reflex (BAS)		0.598				
Cricopharyngeal function (BAS)			0.905			
Aspiration (BAS)			0.876			
Residue in pyriform sinuses (BAS)			0.494			
Extent larynx excursion				0.878		
Extent hyoid excursion				0.854		
Duration vestibule closure					0.842	
Duration BOT-PPW contact					0.743	
Oral transit time						0.626
Duration larynx elevation						0.544
% of Variance	27.447	14.753	11.71	7.489	5.395	3.677
Cronbach's alpha	0.933	0.858	0.717	0.755	0.802	0.466

^a The 18 variables in this factor analysis were the only variables included because they all had loadings greater than 0.4

pharynx and movement of pharyngeal structures to propel the bolus through the pharynx. Specifically, this factor included seven variables: *residue in valleculae* (BAS), *number of swallows to clear bolus*, *pharyngeal wall function* (BAS), *maximum BOT movement*; *percent pharyngeal residue*, *area at maximum pharyngeal constriction*, and *residue in pyriform sinuses* (BAS). The other five factors were assigned the following clinical labels: Duration of Pharyngeal/Laryngeal Functions, Extent of Pharyngeal/Laryngeal Functions, Swallow Initiation and Timing, Airway Protection, and Duration of Pharyngeal Transit. All six factors were then correlated with specific variables from the scale “presence/absence of swallow disorders” that were considered to be clinically related (Table 9). Point-biserial correlations ranged from 0.044 to -0.718, with the strongest relationship observed between Factor 5 (Airway Protection) and the *presence of cricopharyngeal dysfunction* ($r_{pb} = -0.718$). Correlations were low overall, with almost all values being well below the predetermined level of 0.75.

Semisolids

The six factors identified for semisolids explained 70.47% of the total variance, with Cronbach's alpha values ranging from 0.47 to 0.93. Again, the most important factor related to the clinical concept of Pharyngeal Motility (explaining 27.45% of the variability) and included the same variables as liquids, excluding *residue in pyriform sinuses* (BAS). Clinical labels, as used for liquids, were assigned to three of the factors for semisolids: Duration of Pharyngeal/Laryngeal Functions, Extent of Pharyngeal/Laryngeal Functions, and Swallow Initiation and Timing. Factor 6 was assigned the clinical label Temporal Measures, though the clinical association of the variables *oral transit time* and *duration of laryngeal elevation* is unclear (Table 6). Cronbach's alpha for this factor was only 0.466. All six factors were correlated with clinically related variables from the scale “presence/absence of swallow disorders” (Table 9). Again, point-biserial correlations were low overall, ranging from -0.067 to 0.864.

Table 7 Intercorrelations of variables for each factor—liquid consistency

Factor	Correlation	
	Minimum	Maximum
Factor 1	0.302	0.855**
	Maximum BOT movement and percent pharyngeal residue	Number of swallows to clear bolus and percent pharyngeal residue
Factor 2	0.312	0.741**
	Duration BOT-PPW contact and duration larynx elevation	Duration BOT-PPW contact and duration vestibule closure
Factor 3	0.174	0.710**
	Maximum PPW movement and extent hyoid excursion	Extent larynx excursion and extent hyoid excursion
Factor 4	0.400*	0.732**
	Swallow reflex (BAS) and pharyngeal transit time	Pharyngeal delay time and pharyngeal transit time
Factor 5	0.230	0.749**
	Cricopharyngeal function (BAS) and extent larynx excursion	Aspiration (BAS) and percent penetrated/aspirated
Factor 6	0.612**	(factor includes 2 variables only)
	Pharyngeal transit time and duration CPO	

* $p < 0.05$ ** $p < 0.01$

Additional information regarding correlations among the variables within each factor is available upon request

Table 8 Intercorrelations of variables for each factor—semisolid consistency

Factor	Correlation	
	Minimum	Maximum
Factor 1	0.469**	0.872**
	Number of swallows to clear bolus and maximum BOT movement	Pharyngeal wall function (BAS) and percent pharyngeal residue
Factor 2	0.527**	0.908**
	Pharyngeal transit time and swallow reflex (BAS)	Pharyngeal transit time and pharyngeal delay time
Factor 3	0.294	0.813**
	Aspiration (BAS) and residue in pyriform sinuses (BAS)	Aspiration (BAS) and cricopharyngeal function (BAS)
Factor 4	0.701**	(factor includes 2 variables only)
	Extent larynx excursion and extent hyoid excursion	
Factor 5	0.670**	(factor includes 2 variables only)
	Duration vestibule closure and duration BOT-PPW contact	
Factor 6	0.304	(factor includes 2 variables only)
	Oral transit time and duration larynx elevation	

* $p < 0.05$ ** $p < 0.01$

Additional information regarding correlations among the variables within each factor is available upon request

Discussion

This is the first study in which researchers have used comprehensive psychometric techniques (i.e., stability, reliability, and validity) to investigate different swallowing measures taken from VFSSs. Using this approach, we highlight issues of measurement of swallowing in a sample of H&N cancer patients with oropharyngeal and laryngeal tumors treated with (chemo)radiotherapy.

Stability

Stability varied between bolus consistencies, with more measures having poorer stability for semisolids than for liquids. There were particular differences identified between swallows 1 and 3, highlighting the variable nature of swallowing from trial to trial.

Results from the semisolid consistency indicated that the swallow function improved as a VFSS progressed. This

Table 9 Correlations of BAS and biomechanical measures factors with presence/absence of swallow disorder variables

Factor	Presence/absence of swallow disorder variable	r_s
Liquids		
1. Pharyngeal motility	BOT/PPW weakness	0.707**
1. Pharyngeal motility	Pharyngeal residue	0.648**
2. Duration of pharyngeal/laryngeal functions	Reduced laryngeal elevation	0.082
2. Duration of pharyngeal/laryngeal functions	BOT/PPW weakness	0.083
2. Duration of pharyngeal/laryngeal functions	Reduced epiglottic inversion	0.083
2. Duration of pharyngeal/laryngeal functions	Reduced laryngeal vestibule closure	0.044
3. Extent of pharyngeal/laryngeal functions	BOT/PPW weakness	0.127
3. Extent of pharyngeal/laryngeal functions	Reduced laryngeal elevation	0.383*
3. Extent of pharyngeal/laryngeal functions	Pharyngeal residue	0.102
3. Extent of pharyngeal/laryngeal functions	Cricopharyngeal dysfunction	0.357*
4. Swallow initiation and timing	Delayed swallow reflex	-0.563**
4. Swallow initiation and timing	Penetration	-0.222
4. Swallow initiation and timing	Aspiration	-0.078
5. Airway protection	BOT/PPW weakness	-0.482**
5. Airway protection	Reduced laryngeal elevation	-0.425**
5. Airway protection	Reduced epiglottic inversion	-0.665**
5. Airway protection	Reduced laryngeal vestibule closure	-0.711*
5. Airway protection	Pharyngeal residue	-0.379*
5. Airway protection	Cricopharyngeal dysfunction	-0.718**
5. Airway protection	Penetration	-0.620**
5. Airway protection	Aspiration	-0.675**
6. Duration of pharyngeal transit	Reduced laryngeal elevation	0.189
6. Duration of pharyngeal transit	Cricopharyngeal dysfunction	0.116
Semisolids		
1. Pharyngeal motility	BOT/PPW weakness	0.727**
1. Pharyngeal motility	Pharyngeal residue	0.743**
2. Swallow initiation and timing	Delayed swallow reflex	0.498*
3. Lower pharyngeal function	BOT/PPW weakness	-0.067
3. Lower pharyngeal function	Pharyngeal residue	0.201
3. Lower pharyngeal function	Cricopharyngeal dysfunction	0.702**
3. Lower pharyngeal function	Penetration	0.864**
4. Extent of pharyngeal/laryngeal functions	BOT/PPW weakness	0.226
4. Extent of pharyngeal/laryngeal functions	Reduced laryngeal elevation	0.422**
5. Duration of pharyngeal/laryngeal functions	BOT/PPW weakness	0.137
5. Duration of pharyngeal/laryngeal functions	Reduced epiglottic inversion	0.087
5. Duration of pharyngeal/laryngeal functions	Reduced laryngeal vestibule closure	-0.162
6. Temporal measures	Poor bolus formation	0.392*
6. Temporal measures	Prolonged oral transit	0.446**
6. Temporal measures	Reduced laryngeal elevation	-0.292

* $p < 0.05$

** $p < 0.01$ (2-tailed)

applied to the variables *presence of poor bolus formation*, *presence of prolonged oral transit*, *presence of pharyngeal residue*, *residue in valleculae* (BAS), *pharyngeal area at maximum constriction*, *percent of pharyngeal residue*, and *pharyngeal wall function* (BAS). This suggests a practice effect (with improvement in pharyngeal constriction) and/or increased lubrication within the pharynx resulting in less pharyngeal residue of the semisolid bolus as the study progressed.

For the liquid consistency, analysis of the *pharyngeal wall function* (BAS) again indicated an improvement in function. However, there was an increase in the *number of swallows required to clear the liquid bolus* over the trial (indicating worsening function as the VFSS progressed). This may be an artifact of the coating effect of liquid barium. Thus, for swallows 2 and 3, more swallows were required to clear the already-coated anatomic structures than was the case for swallow 1. These results challenge

our commonly accepted practices where the mean of multiple swallows is often taken to represent overall swallow function [5, 14, 23, 31, 39]. The legitimacy of this can be assured only in instances where the three recorded swallows are very similar, which in this study was not always the case, particularly for the variables *pharyngeal wall function* (BAS) and *duration of cricopharyngeal opening*, where significant differences were found across the three swallows for both consistencies. Classic test theory states that an observed score (X) is a function of the true score (T) and an error component (E) [16]. As such, taking the mean of three swallows provides the mean (and standard error) of the observed scores but does not take into account the error component of each individual swallow. Using results from swallow 2 provides a measure that more accurately represents the true score \pm the error, such as potential practice (as seen in this study) or fatigue effects.

Reliability

Low ICCs were obtained for both liquid and semisolid consistencies for the following four variables: *oral transit time*, *pharyngeal transit time*, *pharyngeal delay time*, and *extent of hyoid excursion*. These findings suggest poor reliability of the measure itself, i.e., measurement error. According to Portney and Watkins [16], measurement error (i.e., the difference between the observed score and the true score) may be due to (1) rater (examiner) error, (2) the measuring instrument, or (3) variability in the trait or subject being measured. For example, for the variable *oral transit time*, brief oral “holding” of a bolus prior to the swallow or changes in oral lubrication may contribute to measurement error.

There were a number of variables for which a low-percentage agreement or ICC was calculated for one bolus consistency, but an adequate percentage agreement or ICC was calculated for the other consistency. For liquids, this finding applied only to *duration of laryngeal elevation*. For semisolids, this applied to *presence of delayed swallow reflex*, *tongue function* (BAS), *soft palate elevation* (BAS), *swallow reflex* (BAS), *residue in pyriform sinuses* (BAS), *duration of BOT-PPW contact*, *duration of laryngeal vestibule closure*, *duration of cricopharyngeal opening*, *extent of vestibule closure*, and *penetration-aspiration scale*. These results suggest that the measure may not be sensitive or accurate enough to capture that particular variable (or trait) during the swallow of that particular consistency. When examining *duration of BOT-PPW contact*, for example, some patients may have anticipated difficulties with a semisolid bolus and therefore performed a more effortful swallow, thus altering the dynamics of pharyngeal motility during all or some of their semisolid bolus

swallows. These same patients may not have anticipated or experienced difficulties with liquids, which may have contributed to the more reliable results in measuring *duration of BOT-PPW contact* for liquids.

There were a number of variables for both consistencies where stability was poor, yet reliability for that same variable was high (Tables 3 and 4). For example, the variable *percent of pharyngeal residue* for semisolids had poor stability, so the true score significantly changed (i.e., there was less residue) as the VFSS progressed. However, reliability for this variable was high (0.85), indicating that despite the poor agreement, there is good consistency and the variable can be considered an accurate representation of the trait of interest (i.e., the percentage of residue within the pharynx).

Validity

The validity section of this study was divided into two components: (1) construct validity of the BAS and biomechanical measures and (2) concurrent validity of the presence/absence measures. To our knowledge, using factor analytic techniques to understand the psychometric properties of swallowing measures is a novel approach. Several key findings can be drawn from the two separate factor analyses for liquid and for semisolid boluses.

First, the amount of variation explained in both models was large, 66.72% and 70.47% for liquids and semisolids, respectively. With these results we can be confident that the six factors identified for each consistency accurately capture the construct of swallowing. It should be noted, however, that the clinical significance of Factor 6 for semisolids (Temporal Measures) was unclear.

Second, Pharyngeal Motility was the factor that explained the majority of variation in both liquids and semisolids. This is clearly an essential aspect of swallowing that needs to be measured in this population, and the results of the factor analyses identified seven variables for liquids and six variables for semisolids that could be considered suitable for capturing the movement of pharyngeal structures during the swallow and/or pharyngeal residue after the swallow.

Third, there were three factors, in addition to Pharyngeal Motility, that clearly captured swallowing for both liquid and semisolid consistencies: Duration of Pharyngeal/Laryngeal Functions, Extent of Pharyngeal/Laryngeal Functions, and Swallow Initiation and Timing. However, two factors were revealed for each consistency (Factors 5 and 6 for liquids and Factors 3 and 6 for semisolids) that were exclusive to that consistency only, indicating that slightly different characteristics are observed when each consistency is swallowed. These differences across consistencies are essential to

consider when measuring swallowing from VFSSs, i.e., a tool that captures specific aspects of a liquid swallow cannot be assumed to be an equally valid tool for capturing a semisolid swallow.

The presence/absence measures correlated poorly with the factors for both semisolid and liquid consistencies, indicating poor concurrent validity of these dichotomous measures. This suggests that either much variability was unaccounted for (i.e., more than 49% in all cases) or that these variables are measuring a different construct, which seems unlikely. For example, Factor 2 for semisolids (Swallow Initiation and Timing), which included the variables *pharyngeal delay time*, *pharyngeal transit time*, and *swallow reflex* (BAS), did not correlate highly with the presence/absence measure *delayed swallow reflex* ($r_{pb} = 0.527$), although clinically we would anticipate that these measures should represent the same construct. Because these dichotomous variables (e.g., a disorder being present or absent) had little association with either the BAS or the biomechanical measures, this raises questions about their validity when being used to measure swallowing in H&N cancer patients. We acknowledge that the dichotomous variables are different in that they measure absolute abnormality, whereas the BAS or biomechanical measures measure the degree or severity of problems. These differences raise further questions about whether such dichotomous variables are sensitive to change over time, i.e., whether they can capture small but clinically meaningful changes from one VFSS to another, when they indicate only absolute abnormality.

Differences in Bolus Consistencies

This study identified a number of significant differences between the two bolus consistencies. While previous researchers have found that different bolus consistencies yield different results in terms of the variable being measured (e.g., pharyngeal transit time, duration of cricopharyngeal opening), the results of this study go further, identifying potential problems with the reliability and validity of measurement variables when applied to different consistencies. Swallowing different bolus consistencies needs to be conceptualized separately; not only do specific variables differ for different consistencies, but the swallow as a whole differs. Therefore, using different measurements for each consistency may be a more accurate way to measure swallowing and monitor changes over time.

Recommendations

Six underlying factors of importance for swallowing in H&N cancer patients were identified, but these factors

were slightly different across the two bolus consistencies, indicating the need to measure slightly different aspects of swallowing for each consistency to be accurately represented.

Intercorrelations among the variables within each factor can be used to identify which variables are the most apposite to measure. Variables that have a high loading on a factor (and are therefore a good representation of that factor) but low intercorrelations with each other (and therefore measure slightly different traits) are best. For example, Factor 1 for semisolids (Pharyngeal Motility) includes the variable *number of swallows required to clear the bolus*, which has a high loading on the factor (0.919) but a low intercorrelation with other variables (0.469 with the variable *maximum BOT movement*). We suggest that within each factor one or two measures are taken, selected on the basis of having a high loading on a factor, but a low intercorrelation with other variables, while taking into account the reliability of the measure itself.

Results of the factor analyses indicate that by selecting a minimum set of measures in this manner, adequate construct validity would be achieved, and the remaining variables within each factor are then not essential. For example, when considering the Duration of Pharyngeal/Laryngeal Functions (Factor 2) of liquids, if the variable *duration of BOT-PPW contact* was measured, then the variables *duration of vestibule closure* and *duration of laryngeal elevation* do not need to be measured. Because *duration of laryngeal elevation* was found to have poor reliability for liquids (Table 3), selecting an alternative measure to represent the Duration of Pharyngeal/Laryngeal Functions is appropriate.

Further research is needed to investigate the stability, reliability, and validity of swallowing measures that are being used with different populations. The issues identified in this article remain to be investigated with other measurement tools that are currently used in research and clinical settings, such as FEES, pharyngeal manometry, and oral tongue pressure measurements.

Conclusions

This study relates to specific data taken from a H&N cancer population that was treated with (chemo)radiotherapy for oropharyngeal or laryngeal tumors. However, the results may be applied to other populations, including normals, and to the examination of other swallowing tools.

The outcomes of this study raise questions about the current methods that are used to measure swallowing from VFSSs. We have found that different measures may be important for assessing swallowing of different consistencies, and that data taken from the mean of three swallows

may not be an accurate representation of swallowing because the stability of swallow sequences needs to be considered.

We have identified key elements that need to be included to comprehensively and accurately assess swallowing function in this H&N cancer population. We have not specified what the *best* measures are, but we suggest specific elements to capture the key characteristics of swallowing different bolus consistencies.

Ongoing investigation into the psychometric properties of our clinical tools will further contribute to our understanding of accurate measurement, thereby improving dysphagia research and practice.

References

- Ekberg O, Nylander G. Pharyngeal dysfunction after treatment for pharyngeal cancer with surgery and radiotherapy. *Gastrointest Radiol* 1983;8:97–104.
- Logemann J. Evaluation and treatment of swallowing disorders 2nd ed. Austin TX: Pro-Ed, 1998.
- Ekberg O. Radiologic evaluation of swallowing. In: Groher ME (ed), *Dysphagia: Diagnosis and management*, 3rd ed. Boston: Butterworth-Heinemann, 1997, pp 191–222.
- Eisbruch A, Lyden T, Bradford CR, Dawson LA, Haxer MJ, Miller AE, Teknos TN, Chepeha DB, Hogikyan ND, Terrell JE, Wolf GT. Objective assessment of swallowing dysfunction and aspiration after radiation concurrent with chemotherapy for head-and-neck cancer. *Int J Radiat Oncol Biol Phys* 2002;53:23–28.
- Hughes PJ, Scott PMJ, Kew J, Cheung DMC, Leung SF, Ahuja AT, van Hasselt A. Dysphagia in treated nasopharyngeal cancer. *Head Neck* 2000;22:393–397.
- Chang Y-C, Chen S-Y, Lui LT, Wang T-G, Wang T-C, Hsiao T-Y, Li Y-W, Lien I-N. Dysphagia in patients with nasopharyngeal cancer after radiation therapy: A videofluoroscopic swallowing study. *Dysphagia* 2003;18:135–143.
- Connor NP, Cohen SB, Kammer RE, Sullivan PA, Brewer KA, Hong TS, Chappell RJ, Harari PM. Impact of conventional radiotherapy on health-related quality of life and critical functions of the head and neck. *Int J Radiat Oncol Biol Phys* 2006;65:1051–1062.
- Rosenbek JC, Robbins J, Roecker EB, Coyle JL, Wood JL. A penetration-aspiration scale. *Dysphagia* 1996;11:93–98.
- Alexander MS, Edelman GM, Man K, Randall CJ. Dysphagia after fast neutron therapy to the head and neck. *Br J Radiol* 1989;62:724–728.
- Price G, Jones C, Charlton R, Allen C. A combined approach to the assessment of neurological dysphagia. *Clin Otolaryngol Allied Sci* 1987;12:197–201.
- Logemann JA, Kahrilas PJ, Begelman J, Dodds W, Pauloski BR. Interactive computer program for biomechanical analysis of videoradiographic studies of swallowing. *AJR Am J Roentgenol* 1989;153:277–280.
- Leonard R, McKenzie S. Dynamic swallow study: Instrumentation, measurement techniques. In: Leonard R, Kendall K (eds), *Dysphagia assessment and treatment planning: A team approach*. San Diego: Singular Publishing Group, 1997, pp 127–160.
- Kendall KA, McKenzie SW, Leonard RJ, Jones C. Structural mobility in deglutition after single modality treatment of head and neck carcinomas with radiotherapy. *Head Neck* 1998;20:720–725.
- Lazarus CL, Logemann JA, Pauloski BR, Colangelo LA, Kahrilas PJ, Mittal BB, Pierce M. Swallowing disorders in head and neck cancer patients treated with radiotherapy and adjuvant chemotherapy. *Laryngoscope* 1996;106:1157–1166.
- Frowen JJ, Perry AR. Swallowing outcomes after radiotherapy for head and neck cancer: A systematic review. *Head Neck* 2006;28:932–944.
- Portney LG, Watkins MP. Foundations of clinical research: Applications to practice 2nd ed. Upper Saddle River NJ: Prentice Hall Health, 2000.
- Pauloski BR, Logemann JA, Rademaker AW, McConnel FMS, Heiser MA, Cardinale S, Shedd D, Lewin J, Baker SR, Graner D, Cook B, Milianti F, Collins S, Baker T. Speech and swallowing function after anterior tongue and floor of mouth resection with distal flap reconstruction. *J Speech Hear Res* 1993;36:267–276.
- Nicoletti G, Soutar D, Jackson M, Wrench AA, Robertson G. Chewing and swallowing after surgical treatment for oral cancer: Functional evaluation in 196 selected cases. *Plast Reconstr Surg* 2004;114:329–338.
- Pauloski BR, Rademaker AW, Logemann JA, Stein D, Beery Q, Newman L, Hanchett C, Tusant S, MacCracken E. Pretreatment swallowing function in patients with head and neck cancer. *Head Neck* 2000;22:474–482.
- Stenson KM, MacCracken E, List M, Haraf DJ, Brockstein B, Weichselbaum R, Vokes EE. Swallowing function in patients with head and neck cancer prior to treatment. *Arch Otolaryngol Head Neck Surg* 2000;126:371–377.
- Scott AG. The development of a scale to assess swallowing function in motor neuron disease using videofluoroscopic techniques. Victoria, Australia: School of Human Communication Sciences, Faculty of Health Sciences, La Trobe University, 1999.
- Dantas RO, Kern MK, Massey BT, Dodds WJ, Kahrilas PJ, Brasseur JG, Cook IJ, Lang IM. Effect of swallowed bolus variables on oral and pharyngeal phases of swallowing. *Am J Physiol* 1990;258(5 Pt 1):G675–G681.
- Lazarus CL, Logemann JA, Rademaker AW, Kahrilas PJ, Pajak TF, Lazar R, Halper A. Effects of bolus volume, viscosity, and repeated swallows in nonstroke subjects and stroke patients. *Arch Phys Med Rehabil* 1993;74:1066–1070.
- Stachler RJ, Hamlet SL, Mathog RH, Jones L, Heilbrun LK, Manov J, O'Campo JM. Swallowing of bolus types by postsurgical head and neck cancer patients. *Head Neck* 1994;16:413–419.
- Sobin L, Wittekind C. TNM Classification of Malignant Tumours, 6th ed. New York: Wiley, 2002.
- Thursfield V, Farrugia H, Giles G. Canstat: Cancer in Victoria 2004. Melbourne: The Cancer Council Victoria, 2006.
- Perry A, Skeat J. AusTOMs for speech pathology. Melbourne: La Trobe University, 2004.
- Logemann JA. Manual for the videofluorographic study of swallowing. 2nd ed. Austin TX: Pro-Ed, 1993.
- Lazarus CL, Logemann JA, Kahrilas PJ, Mittal BB. Swallow recovery in an oral cancer patient following surgery, radiotherapy, and hyperthermia. *Head Neck* 1994;16:259–265.
- Kotz T, Costello R, Li Y, Posner MR. Swallowing dysfunction after chemoradiation for advanced squamous cell carcinoma of the head and neck. *Head Neck* 2004;26:365–372.
- Lazarus CL, Logemann JA, Pauloski BR, Rademaker AW, Larson CR, Mittal BB, Pierce M. Swallowing and tongue function following treatment for oral and oropharyngeal cancer. *J Speech Lang Hear Res* 2000;43:1011–1023.
- Lazarus CL. Effects of radiation therapy and voluntary maneuvers on swallow function in head and neck cancer patients. *Clin Commun Disord* 1993;3:11–20.
- Smith RV, Kotz T, Beitler JJ, Wadler S. Long-term swallowing problems after organ preservation therapy with concomitant

- radiation therapy and intravenous hydroxyurea: initial results. *Arch Otolaryngol Head Neck Surg* 2000;126:384–389.
34. Carrara-de Angelis E, Feher O, Barros PB, Nishimoto IN, Kowalski LP. Voice and swallowing in patients enrolled in a larynx preservation trial. *Arch Otolaryngol Head Neck Surg* 2003;129:733–738.
35. Scott A, Perry A, Bench J. A study of interrater reliability when using videofluoroscopy as an assessment of swallowing. *Dysphagia* 1998;13:223–227.
36. Pauloski BR, Logemann JA. Impact of tongue base and posterior pharyngeal wall biomechanics on pharyngeal clearance in irradiated postsurgical oral and oropharyngeal cancer patients. *Head Neck* 2000;22:120–131.
37. Newman LA, Robbins KT, Logemann JA, Rademaker AW, Lazarus CL, Hamner A, Tusan S, Huang CF. Swallowing and speech ability after treatment for head and neck cancer with targeted intraarterial versus intravenous chemoradiation. *Head Neck* 2002;24:68–77.
38. Tabachnick BG, Fidell LS. *Using multivariate statistics*, 4th ed. Boston: Allyn and Bacon, 2001.
39. Kendall KA, Leonard RJ, McKenzie SW, Jones CU. Timing of swallowing events after single-modality treatment of head and neck carcinomas with radiotherapy. *Ann Otol Rhinol Laryngol* 2000;109:767–775.

Jacqui J. Frowen BSpPath (Hons)

Susan M. Cotton PhD

Alison R. Perry PhD, MRCSLT