Research Article

Classification of Physiologic Swallowing Impairment Severity: A Latent Class Analysis of Modified Barium Swallow Impairment Profile Scores

Jonathan Beall,^a Elizabeth G. Hill,^a Kent Armeson,^a Kendrea L. (Focht) Garand,^b Kate (Humphries) Davidson,^c and Bonnie Martin-Harris^d

Purpose: Our objectives were to (a) identify oral and pharyngeal physiologic swallowing impairment severity classes based on latent class analyses (LCAs) of the Modified Barium Swallow Impairment Profile (MBSImP) swallow task scores and (b) quantify the probability of severity class membership given composite MBSImP oral total (OT) and pharyngeal total (PT) scores.

Method: MBSImP scores were collected from a patient database of 319 consecutive modified barium swallow studies. Because of missing swallow task scores, LCA was performed using 25 multiply imputed data sets.

Results: LCA revealed a three-class structure for both oral and pharyngeal models. We identified OT and PT score intervals to assign subjects to oral and pharyngeal impairment latent severity classes, respectively, with high probability (probability of class membership ≥ 0.9 given OT

or PT scores within specified ranges) and high confidence (95% credible interval [CI] widths ≤ 0.24 for all total scores within specified ranges). OT scores ranging from 0 to 10 and from 14 to 18 yielded assignments in Oral Latent Classes 1 and 2, respectively, while OT = 22 was assigned to Oral Latent Class 3. PT scores ranging from 0 to 13 and from 18 to 24 yielded assignments in Pharyngeal Latent Classes 1 and 2, respectively, while PT = 26 was assigned to Pharyngeal Latent Class 3.

Conclusions: LCA of MBSImP task-level data revealed significant underlying oral and pharyngeal ordinal class structures representing increasingly severe gradations of physiologic swallow impairment. Clinically meaningful OT and PT score ranges were derived facilitating latent class assignment.

Supplemental Material: <https://doi.org/10.23641/asha.12315677>

^aDepartment of Public Health Sciences, Medical University of South Carolina, Charleston

^bDepartment of Speech Pathology and Audiology, University of South Alabama, Mobile

^cDepartment of Otolaryngology–Head & Neck Surgery, Medical University of South Carolina, Charleston

^dDepartment of Communication Sciences and Disorders, Northwestern University, Evanston, IL

Correspondence to Bonnie Martin-Harris: bonnie.martinharris@northwestern.edu

Editor-in-Chief: Julie Barkmeier-Kraemer

Editor: Georgia A. Malandraki

Received September 5, 2019

Revision received December 12, 2019

Accepted February 3, 2020

https://doi.org/10.1044/2020_AJSLP-19-00080

Publisher Note: This article is part of the Special Issue: Select Papers From the 2018 Charleston Swallowing Conference at Northwestern University.

The modified barium swallow study (MBSS) is the most commonly applied method for examining oropharyngeal swallowing physiology and impairment (Leonard & McKenzie, 2008; Logemann, 1997; Martin-Harris et al., 2008; Martin-Harris & Jones, 2008; Martin-Harris et al., 2000; Peladeau-Pigeon & Steele, 2013) and permits real-time dynamic visualization of bolus flow and structural movement throughout the upper aerodigestive tract using videofluoroscopic imaging. The MBSS further allows for detection of the presence, timing, depth, and patient response to airway invasion (penetration/aspiration) while assisting in the identification of the physiologic cause of swallowing impairment (Coyle, 2017; Rosenbek et al., 1996).

The formulation and documentation of impressions regarding the severity of physiologic swallowing impairment are routinely made in clinical practice as part of assessment reporting. These findings are communicated to other health care providers through electronic medical records

Disclosure: The authors have declared that no competing interests existed at the time of publication.

and through verbal communication to clinicians who will be responsible for providing swallowing therapy and routine patient care. Classifications of severity typically include terms such as “mild,” “moderate,” “severe,” or “profound.” At issue, however, is the lack of any empirical data to support the structure of these severity descriptors or impression labels applied by the clinician. What is missing in the literature and represents a constant call from swallowing clinicians is a method to extract physiologic swallowing impairment severity from videofluoroscopic images obtained during the MBSS. This current investigation seeks to fill this clinical practice gap. While physiologic severity classification does not sufficiently inform treatment decisions or management planning, severity estimates represent one common factor used in making multidimensional management decisions, despite the current, subjective nature of the judgment.

Modified Barium Swallow Impairment Profile Swallow Task Scores

The Modified Barium Swallow Impairment Profile (MBSImP) is a research-based, standardized, and validated approach for interpretation and quantitation of MBSS findings (Martin-Harris et al., 2008; Northern Speech Services, 2018). The MBSImP provides quantitative evaluation of 17 physiologic components of swallowing distributed across three functional domains: oral, pharyngeal, and esophageal. MBSImP components are scored on an ordinal scale from 0 (indicating no impairment) to a maximum of 2, 3, or 4, depending on the specific component. The MBSImP scoring instrument is composed of six measures of oral impairment (see Table 1), 10 measures of pharyngeal impairment (see Table 2), and one measure of esophageal clearance impairment in the upright position (Martin-Harris et al., 2008; Northern Speech Services, 2018). The MBSImP protocol requires administration of a variety of standardized bolus viscosities and volumes (referred to as swallowing tasks) in lateral and anterior–posterior (A-P) viewing planes, including thin (< 15 cps) barium (two trials of 5 ml via teaspoon, one cup sip [20 ml], and sequential swallows from cup [40 ml]), nectar (150–450 cps) barium (one trial of 5 ml via teaspoon, one cup sip [20 ml], and sequential swallows from cup [40 ml]), thin-honey (800–1,800 cps) barium (one trial of 5 ml via teaspoon), pudding (4,500–7,000 cps) barium (one trial of 5 ml via teaspoon), and a solid (one-half portion of a Lorna Doone shortbread cookie coated with 3-ml pudding barium; Varibar E-Z-EM, Inc.; 40% w/v ratio). When using the research protocol, all components are scored for all swallowing tasks in the lateral view, with the following exceptions: Tongue control during bolus hold (Component 2) is not scored during sequential swallows of thin liquid and nectar, pudding, and cookie tasks; bolus preparation/mastication (Component 3) is scored only for the cookie task; and the two components scored only in the A-P viewing plane—pharyngeal contraction (Component 13) and esophageal clearance (Component 17)—are

scored only for the 5-ml nectar and pudding tasks. Therefore, the full research protocol yields 127 scores per patient—13 components evaluated for nine swallow tasks, one component evaluated for five tasks, one component evaluated for a single task, and two components evaluated for two tasks.

Overall Impression Scores

Whenever possible, administration of all swallowing tasks is essential to elucidate swallowing impairment relative to each consistency and task, because of the known effect of bolus variables on swallowing physiology and impairment (Hazelwood et al., 2017), but scoring each swallowing task individually can be burdensome in a busy clinical setting. To alleviate this burden, an abbreviated scoring protocol can be implemented in which, generally speaking, only tasks resulting in the worst (maximum) score for each component are scored, with the rationale that the maximum score reflects the greatest physiologic impairment and thereby warrants priority attention during clinical intervention (Martin-Harris et al., 2008; Northern Speech Services, 2018). These scores are referred to as overall impression (OI) scores. Observations are noted in the patient record regarding the influence of bolus variables and swallowing tasks on swallowing physiology, and thus, component scores could change relative to these influences. However, only baseline scores are included in deriving the OI score. Although most component OI scores are derived directly from the maximum score across all tasks, there are exceptions. For two oral components, lip closure (Component 1) and oral residue (Component 5) and two pharyngeal components, tongue base retraction (Component 15) and pharyngeal residue (Component 16), the maximum task scores of both 0 and 1 are assigned OI scores of 0 (Martin-Harris et al., 2008; Northern Speech Services, 2018). In addition, for the oral component, initiation of the pharyngeal swallow (Component 6), if all nonsolid swallow task scores are 0 and the cookie task score is either a 0 or 1, the OI score is 0 (Martin-Harris et al., 2008; Northern Speech Services, 2018). The nonzero task scores annotated in these exceptions reflect anatomically distinct but normal physiology and therefore are not scored as impaired. These modifications to OI score construction ensure that a score of 0 always reflects unimpaired swallowing physiology. Definitions and constructions of OI scores for all MBSImP oral and pharyngeal components are summarized in Tables 1 and 2, respectively. This abbreviated scoring algorithm reduces the number of scores from 127 to 17 (one OI score per component).

Oral and Pharyngeal Total Scores

Composite oral and pharyngeal scores are constructed by summing the OI scores, resulting in oral total (OT) and pharyngeal total (PT) scores. These summative measures provide quantitative characterizations of a patient’s overall impairment in the oral and pharyngeal domains. Despite reporting of both OI component scores and composite OT/

Table 1. Modified Barium Swallow Impairment Profile (MBSImP) component scores with overall impression (OI) and total scoring algorithms for oral domain components.

MBSImP component	Swallow tasks ^a	Score range	OI score construction	OI score values
1. Lip closure	TL5ml, TLCS, TLSS, N5ml, NCS, NSS, H5ml, P5ml, C	0–4	0 if maximum task score is 0 or 1; maximum task score otherwise	0, 2, 3, or 4
2. Tongue control during bolus hold	TL5ml, TLCS, N5ml, NCS, H5ml	0–3	Maximum task score	0, 1, 2, or 3
3. Bolus preparation/ mastication	C	0–3	Maximum task score	0, 1, 2, or 3
4. Bolus transport/ lingual motion	TL5ml, TLCS, TLSS, N5ml, NCS, NSS, H5ml, P5ml, C	0–4	Maximum task score	0, 1, 2, 3, or 4
5. Oral residue	TL5ml, TLCS, TLSS, N5ml, NCS, NSS, H5ml, P5ml, C	0–4	0 if maximum task score is 0 or 1; maximum task score otherwise	0, 2, 3, or 4
6. Initiation of pharyngeal swallow	TL5ml, TLCS, TLSS, N5ml, NCS, NSS, H5ml, P5ml, C	0–4	0 if all nonsolid swallow task scores are 0 AND the cookie swallow task score is 0 or 1; maximum task score, otherwise	0, 1, 2, 3, or 4
Total score			Oral total = sum of oral domain OI scores Score range = 0–22	

Note. TL5ml = thin liquid 5 ml; TLCS = thin liquid cup sip; TLSS = thin liquid sequential swallow; N5ml = nectar 5 ml; NCS = nectar cup sip; NSS = nectar sequential swallow; H5ml = honey 5 ml; P5ml = pudding 5ml; C = cookie.

^aThe initial thin liquid 5 ml swallow task is performed to allow subjects to acclimate to the taste and texture of barium and is not scored.

PT scores in clinical and research practice, these quantitative characterizations have not yet been translated to terms related to severity (e.g., mild, moderate, severe). Yet, terminology classifying severity is frequent within the clinical and research realms despite a lack of evidence to support the validation of such classifications. Given the lack of validated measures of physiologic impairment severity, we approach the problem of severity classification by assuming MBSImP scores are themselves manifestations of an underlying (latent) severity class structure. Therefore, the primary objectives of this study were to (a) identify oral and pharyngeal physiologic swallowing impairment severity classes based on latent class analyses of the MBSImP swallow task scores and (b) quantify the probability of severity class membership given MBSImP OT and PT scores.

Method

Participants and Primary Analysis Set

Data were retrospectively analyzed from the original MBSImP database of 345 adults (18 years of age or older) consecutively referred for routine MBSS as part of medical management between 2005 and 2008 at the Medical University of South Carolina in Charleston, South Carolina, and at Saint Joseph's Hospital in Atlanta, Georgia. Institutional review board approval was obtained at both sites. Since patients could complete more than one swallow study during the 4-year period and to ensure independence of the examination data, we restricted analysis to include data only from the initial MBSS for each patient. This

exclusion resulted in a final sample size of 319 MBS exams included for oral and pharyngeal domain latent class identification.

Although the original MBSImP study was conducted between 2005 and 2008, there have been no changes to the standardized scoring protocol in the intervening time that would jeopardize generalizability of the current analysis. Additionally, all clinicians for the original study demonstrated $\geq 80\%$ inter- and intrarater reliability as recommended for standardized MBSImP scoring (Martin-Harris et al., 2008). Oral domain latent classes were constructed using swallowing task scores from all oral MBSImP components (42 task scores in total). Pharyngeal domain latent class construction was based on swallowing task scores from all pharyngeal MBSImP components (81 task scores in total), excluding those derived from pharyngeal contraction (Component 13) because of missing data for components evaluated in the A-P view related to patient positioning obscuring the raters' view of the lateral pharyngeal walls, a requirement for accurate scoring for that specific component.

For the 42 oral component swallowing tasks, 139 subjects (44%) had complete data, 134 subjects (42%) were missing 10 or fewer task scores, and 46 subjects (14%) were missing 11 or more task scores. For the 81 pharyngeal component swallowing tasks, 188 subjects (59%) had complete data, 75 subjects (24%) were missing 10 or fewer task scores, and 56 subjects (18%) were missing 11 or more task scores. To mitigate the potential for bias and imprecision due to missing data, all analyses were performed using multiple imputation (further discussed in the Statistical Considerations section).

Table 2. Modified Barium Swallow Impairment Profile (MBSImP) component scores with overall impression (OI) and total scoring algorithms for pharyngeal domain components.

MBSImP component	Swallow tasks ^a	Score range	OI score construction	OI score values
7. Soft palate elevation	TL5ml, TLCS, TLSS, N5ml, NCS, NSS, H5ml, P5ml, C	0–4	Maximum task score	0, 1, 2, 3, or 4
8. Laryngeal elevation	TL5ml, TLCS, TLSS, N5ml, NCS, NSS, H5ml, P5ml, C	0–3	Maximum task score	0, 1, 2, or 3
9. Anterior hyoid excursion	TL5ml, TLCS, TLSS, N5ml, NCS, NSS, H5ml, P5ml, C	0–2	Maximum task score	0, 1, or 2
10. Epiglottic movement	TL5ml, TLCS, TLSS, N5ml, NCS, NSS, H5ml, P5ml, C	0–2	Maximum task score	0, 1, or 2
11. Laryngeal vestibular closure–height of swallow	TL5ml, TLCS, TLSS, N5ml, NCS, NSS, H5ml, P5ml, C	0–2	Maximum task score	0, 1, or 2
12. Pharyngeal stripping wave	TL5ml, TLCS, TLSS, N5ml, NCS, NSS, H5ml, P5ml, C	0–2	Maximum task score	0, 1, or 2
13. Pharyngeal contraction (A-P view only)	N5ml, P5ml	0–3	Maximum task score	0, 1, 2, or 3
14. Pharyngoesophageal segment opening	TL5ml, TLCS, TLSS, N5ml, NCS, NSS, H5ml, P5ml, C	0–3	Maximum task score	0, 1, 2, or 3
15. Tongue base retraction	TL5ml, TLCS, TLSS, N5ml, NCS, NSS, H5ml, P5ml, C	0–4	0 if maximum task score is 0 or 1; maximum task score otherwise	0, 2, 3, or 4
16. Pharyngeal residue	TL5ml, TLCS, TLSS, N5ml, NCS, NSS, H5ml, P5ml, C	0–4	0 if maximum task score is 0 or 1; maximum task score otherwise	0, 2, 3, or 4
Total score			Pharyngeal total = sum of pharyngeal domain OI scores, Including A-P View Component 13, Score range = 0–29, Excluding A-P View Component 13 ^b , Score range = 0–26	

Note. TL5ml = thin liquid 5 ml; TLCS = thin liquid cup sip; TLSS = thin liquid sequential swallow; N5ml = nectar 5 ml; NCS = nectar cup sip; NSS = nectar sequential swallow; H5ml = honey 5 ml; P5ml = pudding 5 ml; C = cookie; A-P = anterior–posterior.

^aThe initial thin liquid 5 ml swallow task is performed to allow subjects to acclimate to the taste and texture of barium and is not scored.

^bComponent 13 is frequently missing due to subject-level physical limitations that preclude A-P view evaluation.

Ordinal Latent Class Model

Latent class models (LCMs), first introduced in a mathematically rigorous framework by Lazarsfeld and Henry (1968), are used in settings in which the quantity of scientific interest is a classification variable, which is not directly measurable or observable. Rather, the investigator observes a set of imperfectly measured indicators (also referred to as “items” or “manifest variables”), and the latent class variable is defined based on its effect on these indicators. Specifically, the LCM is defined via the relationship between the latent class variable and the probability of response for the collection of indicators (items) within each latent class. A fundamental property of this relationship is the idea that the unobservable latent classes account for any association between the observed items. This assumption, commonly referred to as “conditional independence,” induces independence among items within the same class. We refer the interested reader to the book by Collins

and Lanza (2010) and references therein for an updated presentation of LCMs, a discussion of analysis and model-fitting techniques, and an overview of model extensions and applications.

In the LCM original presentation, the observed indicators are dichotomous, but generalizations of LCMs to ordinal manifest variables have been proposed (DeSantis et al., 2008). For our application, the manifest variables—the MBSImP swallowing task scores—are ordinal, and we further assume the latent class variable is ordinal, with levels reflecting increasing physiologic impairment severity. Following Agresti (2013, pp. 301–314), we model both the ordinal MBSImP swallowing task scores and the ordinal latent classes using cumulative logit models. Briefly, cumulative logits are defined in a manner similar to that in logistic regression, but where the ordinal variable is dichotomized by grouping values into two categories—those less than or equal to a given cut-point versus those exceeding the cut-point. The number of cumulative logits associated with

a given ordinal variable is always one fewer than the variable's total number of levels. Finally, our LCM relates class probabilities and score probabilities within class to the observed subject-level data via a stochastic (i.e., probabilistic) relationship, assuming the observed swallow task scores for each component arise from appropriately parameterized multinomial distributions.

Figure 1 is a graphical depiction of our ordinal LCM summarizing the relationships between the latent classes and score probabilities within class, with the observed subject-level data. Ellipses (called “nodes”) represent specific LCM components, and arrows indicate the direction of influence between nodes. Solid arrows reflect a stochastic relationship, while hollow arrows reflect a logical relationship in which the daughter (terminal) node is constructed from the parent (founder) node(s). To illustrate, for each subject, every MBSImP component task score is considered to be a realization from a multinomial distribution (bottom right of Figure 1). The subject-level score probabilities within

Figure 1. Graphical depiction of the ordinal latent class model. Ellipses (called “nodes”) represent specific latent class model (LCM) components, and arrows indicate the direction of influence between nodes. Solid arrows reflect a stochastic (probabilistic) relationship, while hollow arrows reflect a logical relationship in which the daughter node is constructed from the parent node(s). A detailed model description is provided in Supplemental Material S1. MBSImP = Modified Barium Swallow Impairment Profile.

class are depicted as having a direct influence on subjects' impairment scores via a stochastic relationship, as these probabilities are the multinomial parameters. Furthermore, these subject-level score probabilities are logically derived from each subject's assigned latent class and the model-based score probabilities within class. Full details pertaining to the LCM specification are provided in Supplemental Material S1.

LCM-Derived Probabilities

The ordinal LCM components—class prevalences and score probabilities within class—are the building blocks for construction of a number of model-derived probabilities. Specifically, score probabilities within class are used to construct OI score probabilities within class, which are in turn used to construct OT and PT score probabilities within class. From the latter, we construct the probabilities of greatest clinical relevance, those summarizing the probability of latent class membership given observed OT or PT scores. Since we assume latent ordinal classes reflect increasing levels of physiologic impairment severity, estimation of class membership probabilities given specific total scores allows us to quantify a patient's likelihood of a particular level of impairment severity given the patient's OT or PT score. These probabilities are constructed based on an application of Bayes' rule using the OT and PT score probabilities within class and the estimated class prevalences. Specifically, for an OT score of x ,

$$P(LC = c | OT = x) = [P(OT = x | LC = c) \cdot P(LC = c)] / P(OT = x), \quad (1)$$

where the notation “ $P(A | B)$ ” represents the probability of Event A conditional on Event B and $P(OT = x)$ in the denominator is derived using the law of total probability. The same approach is used to estimate $P(LC = c | PT = y)$. Details describing the full derivation of all LCM-derived probabilities are provided in Supplemental Material S1.

Statistical Considerations

Missing Data

Because of the potential for patients with more severe symptoms to have missing MBSImP swallowing task data, a data set composed of only subjects with complete MBSImP data would likely contain subjects with milder impairment than the general population of patients referred for MBSS. Indeed, for our data based on an analysis of deviance for nested simple logistic regression models, missing any oral swallowing task score was significantly associated with oral intake status ($p = .033$), and missing any pharyngeal swallowing task score was significantly associated with diagnostic category ($p = .014$), oral intake status ($p = .0002$), feeding tube status ($p = .0002$), and maximum penetration–aspiration scale (PAS) score ($p = .0047$). Therefore, in order to mitigate the potential for bias and imprecision due to missing data, we conducted all analyses using multiply

imputed data rather than data composed only of complete cases. Specifically, we used multivariate imputation by chained equations (MICE; van Buuren, 2007) to construct 25 oral and 25 pharyngeal multiply imputed data sets, where predictor variables used to impute data included all oral or pharyngeal swallow tasks, age (continuous), sex, diagnostic category (cardiothoracic, gastroenterology [GI], head and neck cancer, neurological, pulmonary, or other), oral intake status (no restrictions, some fluid or food restrictions, or NPO), and maximum PAS score (no penetration/aspiration [PAS = 1], penetration [PAS = 2–5], aspiration [PAS = 6–8]). Variables included in the imputation model were selected based on their potential clinical relevance and their observed association with oral or pharyngeal swallowing task score missingness. Additionally, in constructing the 25 multiply imputed data sets, missing values of clinical and demographic variables were also imputed, although these variables were not subsequently used to fit the LCMs. Rates of missingness for these variables were as follows: 4%, age; 4%, sex; 4%, diagnostic category; 1%, oral intake status; 1%, feeding tube status; and 12%, PAS.

Model Fitting and Inference

For each imputed data set, we fit our ordinal LCM using a Bayesian Markov chain Monte Carlo approach with vague Cauchy priors centered at 0 and scale equal to 10 for the cumulative logit intercept parameters (see Supplemental Material S1, Equations 1 and 2; Gelman et al., 2008). We fit models for three and four ordinal latent classes and evaluated fit based on a modification of the deviance information criteria (DIC) for finite mixture models, DIC3, as proposed by Celeux et al. (2006). We assessed convergence of individual model parameters using Geweke's convergence diagnostic statistic (Geweke, 1992), where a statistic value between -1.96 and 1.96 was considered to indicate convergence. We used a burn-in of 50,000 iterations with a post burn-in chain length of 25,000 thinned every 25 iterations to mitigate autocorrelation, resulting in a posterior chain of length of 1,000. To conduct inference, posterior chains from each multiple imputation data analysis were combined, resulting in a final chain length of 25,000 from which to conduct the pooled analysis across all imputations (Gelman et al., 2004, p. 520; Zhao & Reiter, 2010). We summarized posterior estimates of class prevalences and probabilities of interest using medians and 95% CIs, where the latter is an interval estimate obtained from a Bayesian model akin to a confidence interval. Subjects' latent class assignments were based on the class assigned with the greatest frequency across all iterations in the combined chain.

To evaluate the association between latent class assignment and demographic and clinical variables of interest, we fit univariate proportional odds models (Agresti, 2013, pp. 301–307) for each multiply imputed data set, with latent class assignment as the dependent variable. Analyses across the 25 imputed data sets were then pooled as described by Little and Rubin (1987, pp. 255–259) to conduct inference.

All analyses were conducted using R Version 3.6.0 (R Core Team, 2019) and SAS Version 9.4. Multiple imputation was performed using the MICE package in R (van Buuren & Groothuis-Oudshoorn, 2011). LCMs were constructed using the NIMBLE library in R (de Valpine et al., 2017).

Results

Model Fit

Although we observed smaller DIC3 values for both the oral and pharyngeal four-class models, DIC is prone to overfitting (Robert & Titterton, 2002) and should therefore be interpreted with caution. Additionally, we noted a slightly higher proportion of LCM parameters converging for three- versus four-class models. Across the 25 multiple imputation data analyses, average LCM parameter convergence rates for the three-class oral and pharyngeal models were 93.2% and 92.0%, respectively; for four-class oral and pharyngeal LCMs, convergence rates were 90.5% and 90.4%, respectively. Finally, the four-class model demonstrated reduced separation between latent classes when compared to the three-class model based on an evaluation of the posterior probabilities of class given OT and PT scores. Specifically, for the four-class model, all posterior probabilities of membership in the highest latent class given total scores were either extremely small or only of moderate size with wide corresponding CIs. While these results for the four-class model allowed us to identify which subjects should be excluded from the highest class, they provided equivocal information about total scores that would offer a conclusion of membership in the highest class. In comparison, separation between all classes was excellent for the three-class models. Given these observations—concern for overfitting, higher overall convergence rates, and improved interpretability—we elected to base inference on the three-class models. For the three-class oral model, 153 subjects (48%) were assigned to Class 1 (none/mild), 113 (35%) were assigned to Class 2 (moderate), and 53 (17%) were assigned to Class 3 (severe). For the three-class pharyngeal model, 178 subjects (56%) were assigned to Class 1 (none/mild), 75 subjects (24%) were assigned to Class 2 (moderate), and 66 subjects (21%) were assigned to Class 3 (severe).

Class Associations With Demographic and Clinical Variables

Tables 3 and 4 summarize the associations between oral and pharyngeal ordinal latent classes and demographic and clinical variables. Diagnostic category, oral intake status, feeding tube status, and maximum PAS were all found to be significantly associated with latent class. Most notably, there was a significant increase in the odds of being assigned to a higher class: for NPO restrictions relative to no restrictions ($p < .0001$, both oral and pharyngeal models), if a feeding tube was present versus not present ($p < .0001$, both models), and if the maximum PAS indicated aspiration

Table 3. Associations between oral ordinal latent classes and demographic/clinical variables.

Variable	Class 1 N = 153 (48%)	Class 2 N = 113 (35%)	Class 3 N = 53 (17%)	p ^a
Age (years)				
< 40	9 (2.4)	8 (2.7)	10 (4.2)	Ref.
40–59	33 (3.8)	20 (3.9)	22 (5.7)	.074
60+	58 (4.1)	72 (4.4)	68 (6.4)	.17
Sex				
Female	56 (4.1)	58 (4.7)	66 (6.6)	Ref.
Male	44 (4.1)	42 (4.7)	34 (6.6)	.31
Diagnostic category ^b				
Cardiothoracic	7 (2.1)	11 (3.0)	9 (4.0)	.40
Gastroenterology	20 (3.2)	4 (1.9)	8 (3.7)	.0004
Head and neck cancer	23 (3.5)	11 (3.0)	30 (6.3)	.66
Neurological	11 (2.5)	29 (4.3)	21 (5.6)	.0062
Other	21 (3.3)	12 (3.2)	10 (4.2)	.025
Pulmonary	18 (3.1)	32 (4.5)	23 (5.7)	.095
Oral intake status				
No restrictions	59 (4.0)	22 (3.9)	13 (4.7)	Ref.
Fluid/food restrictions	33 (3.8)	55 (4.7)	47 (6.9)	.13
NPO	8 (2.2)	23 (4.0)	40 (6.7)	< .0001
Feeding tube status				
Absent	85 (2.9)	69 (4.4)	45 (6.8)	Ref.
Present	15 (2.9)	31 (4.4)	55 (6.8)	< .0001
Max PAS				
No penetration/aspiration (PAS = 1)	35 (3.9)	21 (4.2)	3 (2.6)	Ref.
Penetration (PAS = 2–5)	50 (4.1)	43 (4.9)	17 (5.2)	.0049
Aspiration (PAS = 6–8)	15 (2.9)	36 (4.9)	80 (5.6)	< .0001

Note. Summary measures are column percents and standard errors based on a pooled analysis of multiply imputed data. PAS = penetration–aspiration scale.

^ap values are based on a pooled analysis of parameter estimates obtained from proportional odds models fit for each multiply imputed data set (see text). ^bBecause there was no reasonable reference category for Diagnostic Category, six individual proportional odds models were fit, each with Diagnostic Category dichotomized as subjects with the specified diagnosis versus those without the diagnosis.

relative to no penetration or aspiration ($p < .0001$, both models). Additionally, there was a significant increase in the odds of assignment to a lower class for a GI primary diagnosis relative to no GI diagnosis ($p = .0004$, oral; $p = .026$, pharyngeal) and for other diagnoses ($p = .025$, oral; $p = .026$, pharyngeal). Finally, there was a significant increase in the odds of assignment to a higher oral class for a neurological primary diagnosis relative to no neurological diagnosis ($p = .0062$) and a significant increase in the odds of assignment to higher pharyngeal class for a head and neck cancer primary diagnosis ($p = .028$).

Conditional Class Probabilities Given OT and PT Scores

The conditional probabilities of class membership given OT and PT scores with corresponding 95% CIs are given in Tables 5 and 6, respectively. Both the oral and pharyngeal models yield clear class separations based on total scores. For the oral model, OT scores ranging from 0 to 10, ranging from 14 to 18, and equal to 22 yield assignments in Latent Classes 1, 2, and 3, respectively, with high probability and high confidence (i.e., narrow CI). For intermediate scores (11–13 and 19–21), we consider probabilities and CIs to be such that classification is equivocal. For example, given

an OT of 11, the probability of Class 1 membership is .93 (95% CI [.68, .98]) and the probability of Class 2 membership is .07 (95% CI [.28, .32]). Although these point estimates would favor membership in Class 1, the CI widths indicate a nonnegligible probability of Class 2 membership. Indeed, when the OT score is 11, the data are consistent with a Class 2 assignment probability as high as 32%. In general, we consider class assignment given total score to be equivocal if the lower bound of the 95% CI for the class assigned with the highest probability is less than .7. We therefore consider class assignment to be either 1 or 2 when OT is 11, 12, or 13 and either 2 or 3 when OT is 19, 20, or 21.

Using similar criteria for assigning class based on PT scores, total scores ranging from 0 to 13, ranging from 18 to 24, and equal to 26 yield assignments in Latent Classes 1, 2, and 3, respectively, with high probability and high confidence. Intermediate scores ranging from 14 to 17 are assigned to either Class 1 or 2, and a score of 25 is assigned to either Class 2 or 3.

Discussion

In this study, we used LCMs to identify an underlying ordinal three-class structure determined by MBSImP

Table 4. Associations between pharyngeal ordinal latent classes and demographic/clinical variables.

Variable	Class 1 N = 178 (56%)	Class 2 N = 75 (24%)	Class 3 N = 66 (21%)	p ^a
Age (years)				
< 40	10 (2.4)	6 (2.9)	10 (3.9)	Ref.
40–59	32 (3.5)	16 (4.3)	21 (5.0)	.085
60+	58 (3.8)	78 (4.9)	68 (5.8)	.068
Sex				
Female	54 (3.8)	65 (5.6)	63 (6.0)	Ref.
Male	46 (3.8)	35 (5.6)	37 (6.0)	.0859
Diagnostic category ^b				
Cardiothoracic	7 (1.9)	13 (3.9)	9 (3.6)	.28
Gastroenterology	17 (2.8)	7 (2.9)	8 (3.6)	.026
Head and neck cancer	17 (2.9)	16 (4.3)	33 (5.9)	.028
Neurological	19 (3.0)	25 (5.0)	14 (4.3)	.78
Other	19 (3.0)	7 (2.9)	12 (4.1)	.026
Pulmonary	21 (3.1)	31 (5.4)	24 (5.3)	.29
Oral intake status				
No restrictions	55 (3.7)	15 (4.1)	20 (4.9)	Ref.
Fluid/food restrictions	34 (3.6)	64 (5.6)	45 (6.1)	.065
NPO	10 (2.3)	21 (4.8)	35 (5.9)	< .0001
Feeding tube status				
Absent	85 (2.7)	60 (5.7)	53 (6.1)	Ref.
Present	15 (2.7)	40 (5.7)	47 (6.1)	< .0001
Max PAS				
No penetration/aspiration (PAS = 1)	37 (3.7)	7 (3.3)	19 (6.1)	Ref.
Penetration PAS = 2–5)	52 (3.8)	41 (5.8)	12 (4.3)	.0003
Aspiration (PAS = 6–8)	11 (2.4)	51 (5.9)	69 (7.0)	< .0001

Note. Summary measures are column percents and standard errors based on a pooled analysis of multiply imputed data. PAS = penetration–aspiration scale.

^aP values are based on a pooled analysis of parameter estimates obtained from proportional odds models fit for each multiply imputed data set (see text). ^bBecause there was no reasonable reference category for Diagnostic Category, six individual proportional odds models were fit, each with Diagnostic Category dichotomized as subjects with the specified diagnosis versus those without the diagnosis.

Table 5. Probability of class given oral total score, corresponding 95% credible intervals, and class assignment given total score.

Total Score	Class 1	Class 2	Class 3	Class assignment given total score
0	1 (1, 1)	0 (0, 0)	0 (0, 0)	1
1	1 (1, 1)	0 (0, 0)	0 (0, 0)	1
2	1 (1, 1)	0 (0, 0)	0 (0, 0)	1
3	1 (1, 1)	0 (0, 0)	0 (0, 0)	1
4	1 (1, 1)	0 (0, 0)	0 (0, 0)	1
5	1 (1, 1)	0 (0, 0)	0 (0, 0)	1
6	1 (1, 1)	0 (0, 0)	0 (0, 0)	1
7	1 (1, 1)	0 (0, 0)	0 (0, 0)	1
8	1 (1, 1)	0 (0, 0)	0 (0, 0)	1
9	1 (0.99, 1)	0 (0, 0.01)	0 (0, 0)	1
10	0.99 (0.94, 1)	0.01 (0, 0.06)	0 (0, 0)	1
11	0.93 (0.68, 0.98)	0.07 (0.02, 0.32)	0 (0, 0)	1 or 2
12	0.67 (0.28, 0.86)	0.33 (0.14, 0.72)	0 (0, 0)	1 or 2
13	0.3 (0.08, 0.56)	0.7 (0.44, 0.92)	0 (0, 0)	1 or 2
14	0.1 (0.02, 0.24)	0.9 (0.76, 0.98)	0 (0, 0)	2
15	0.03 (0.01, 0.08)	0.97 (0.92, 0.99)	0 (0, 0)	2
16	0.01 (0, 0.02)	0.99 (0.98, 1)	0 (0, 0)	2
17	0 (0, 0.01)	1 (0.99, 1)	0 (0, 0)	2
18	0 (0, 0)	1 (0.98, 1)	0 (0, 0.02)	2
19	0 (0, 0)	0.78 (0.45, 0.96)	0.22 (0.04, 0.55)	2 or 3
20	0 (0, 0)	0.26 (0.09, 0.57)	0.74 (0.43, 0.91)	2 or 3
21	0 (0, 0)	0.24 (0.02, 0.77)	0.76 (0.23, 0.98)	2 or 3
22	0 (0, 0)	0 (0, 0.01)	1 (0.99, 1)	3

Table 6. Probability of class given pharyngeal total score, corresponding 95% credible intervals, and class assignment given total score.

Total Score	Class 1	Class 2	Class 3	Class assignment given total score
0	1 (1, 1)	0 (0, 0)	0 (0, 0)	1
1	1 (1, 1)	0 (0, 0)	0 (0, 0)	1
2	1 (1, 1)	0 (0, 0)	0 (0, 0)	1
3	1 (1, 1)	0 (0, 0)	0 (0, 0)	1
4	1 (1, 1)	0 (0, 0)	0 (0, 0)	1
5	1 (1, 1)	0 (0, 0)	0 (0, 0)	1
6	1 (1, 1)	0 (0, 0)	0 (0, 0)	1
7	1 (1, 1)	0 (0, 0)	0 (0, 0)	1
8	1 (1, 1)	0 (0, 0)	0 (0, 0)	1
9	1 (1, 1)	0 (0, 0)	0 (0, 0)	1
10	1 (1, 1)	0 (0, 0)	0 (0, 0)	1
11	1 (1, 1)	0 (0, 0)	0 (0, 0)	1
12	1 (0.98, 1)	0 (0, 0.02)	0 (0, 0)	1
13	1 (0.82, 1)	0 (0, 0.18)	0 (0, 0)	1
14	0.98 (0.37, 1)	0.02 (0, 0.63)	0 (0, 0)	1 or 2
15	0.79 (0.08, 1)	0.21 (0, 0.92)	0 (0, 0)	1 or 2
16	0.26 (0.01, 0.99)	0.74 (0.01, 0.99)	0 (0, 0)	1 or 2
17	0.04 (0, 0.78)	0.96 (0.22, 1)	0 (0, 0)	1 or 2
18	0 (0, 0.05)	1 (0.85, 1)	0 (0, 0)	2
19	0 (0, 0.01)	1 (0.99, 1)	0 (0, 0)	2
20	0 (0, 0)	1 (1, 1)	0 (0, 0)	2
21	0 (0, 0)	1 (1, 1)	0 (0, 0)	2
22	0 (0, 0)	1 (1, 1)	0 (0, 0)	2
23	0 (0, 0)	1 (1, 1)	0 (0, 0)	2
24	0 (0, 0)	1 (0.96, 1)	0 (0, 0.04)	2
25	0 (0, 0)	0.87 (0.13, 0.99)	0.13 (0.01, 0.87)	2 or 3
26	0 (0, 0)	0.02 (0, 0.24)	0.98 (0.76, 1)	3

component scores across various swallowing tasks. The association of higher class with larger composite scores and observed significant associations between latent class and diagnostic category, oral intake status, feeding tube status, and maximum PAS score for both the oral and pharyngeal models provide strong evidence that the identified class structures represent increasingly severe gradations of oral and pharyngeal physiologic swallowing impairment.

Using the LCM score probabilities within class and class prevalence estimates, we were able to construct estimates of the conditional probability of class assignment given OT and PT scores. Specifically, we successfully identified ranges of OT and PT scores within which subjects are assigned to one of three oral or pharyngeal swallowing impairment severity classes with both high probability and high confidence. If confirmed, total scores in the specified ranges would be useful for categorizing patients in clinical trials and for demonstrating improvement in dysphagia classification following intervention.

We note that, for both domains, there are total score ranges within which class assignment is equivocal. While this may seem initially unsatisfying, our work further supports that appropriate clinical management of a patient with a score in one of these intervals would require additional information beyond that contained in the total score. For example, adding a measure of swallowing safety (e.g., PAS score), indices of cognitive-communication status,

or quality clinical measures is necessary to further inform on appropriate intervention strategies and plans of care.

Limitations

Because of concerns about the potential bias induced by missing data, we performed analysis and conducted subsequent inference using multiply imputed data rather than using a complete case analysis. An assumption underlying multiple imputation is that data are missing at random (Little & Rubin, 1987, pp. 14–17). The missing at random assumption states that a variable's propensity for missing values is associated with some other variable(s) but is not due to the missing data value itself. Although our imputation model included clinical variables found to be significantly associated with missingness, it is possible that other unmeasured variables influence the missingness potential for MBSImP swallowing task scores. Although we observed consistency comparing results from our multiple imputation analysis to a complete case analysis, with similar total score cut-points discriminating between classes (complete case analysis results not shown), we recommend validation of the current study using an independent data set of patients referred for MBSS, but with lower rates of missingness.

Additionally, we did not include pharyngeal contraction (Component 13), a score obtained in the A-P viewing

plane, in the construction of our pharyngeal LCM. Therefore, reported PT score ranges yielding class assignments are interpretable only for components scored in the lateral viewing plane. Furthermore, like Component 13, esophageal clearance scores (Component 17) were also not included in the model due to many missing observations of swallowing in the A-P viewing plane. For readers unfamiliar with the MBSImP components, it may seem unusual to have two surrogate measures of impairment, oral residue (Component 5) and pharyngeal residue (Component 16), that do not in themselves represent physiology. These two components are included in the tool because of their high score correlations with mechanistic impairment in our original and ongoing studies (Martin-Harris et al., 2008).

This swallowing impairment severity analysis classified patients into severity groupings based on the widely applied method of using composite OT and PT scores. We recognize the value of evaluating the severity class represented by the score ranges within each of the 17 swallowing components on the MBSImP, and this research is ongoing in large data sets.

Conclusions

Latent class analysis using MBSImP reveals significant underlying oral and pharyngeal ordinal class structure representing increasingly severe gradations of physiologic swallow impairment. Furthermore, clinically meaningful OT and PT score ranges were derived facilitating latent class assignment consistent with none/mild, moderate, and severe impairment. These findings support an evidence-based three-class structure system of severity using quantitative measures of swallow physiology in patients with dysphagia.

Acknowledgments

This work was supported in part by National Institute on Deafness and Other Communication Disorders Awards K23DC005764 (PI: Martin-Harris) and K24DC12801 (PI: Martin-Harris) and by the Biostatistics Shared Resource, Hollings Cancer Center, Medical University of South Carolina (P30 CA138313). The authors would like to acknowledge Julie Blair and R. Jordan Hazelwood for their time and collaborative efforts.

References

- Agresti, A. (2013). *Categorical data analysis* (3rd ed.). Wiley.
- Celeux, G., Forbes, F., Robert, C. P., & Titterton, D. M. (2006). Deviance information criteria for missing data models. *Bayesian Analysis*, 1(4), 651–674. <https://doi.org/10.1214/06-BA122>
- Collins, L. M., & Lanza, S. T. (2010). *Latent class and latent transition analysis: With applications in the social, behavioral, and health sciences*. Wiley.
- Coyle, J. L. (2017). Scaling the swallow: The penetration-aspiration scale quantifies how—and where—a patient's swallow goes awry. *The ASHA Leader*, 22(5), 36–38. <https://doi.org/10.1044/leader.OTP.22052017.36>
- de Valpine, P., Turek, D., Paciorek, C. J., Anderson-Bergman, C., Temple Lang, D., & Bodik, R. (2017). Programming with models: Writing statistical algorithms for general model structures with NIMBLE. *Journal of Computational and Graphical Statistics*, 26(2), 403–413. <https://doi.org/10.1080/10618600.2016.1172487>
- DeSantis, S. M., Houseman, E. A., Coull, B. A., Stemmer-Rachamimov, A., & Betensky, R. A. (2008). A penalized latent class model for ordinal data. *Biostatistics*, 9(2), 249–262. <https://doi.org/10.1093/biostatistics/kxm026>
- Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. (2004). *Bayesian data analysis*. Chapman & Hall.
- Gelman, A., Jakulin, A., Pittau, M. G., & Su, Y.-S. (2008). A weakly informative default prior distribution for logistic and other regression models. *The Annals of Applied Statistics*, 2(4), 1360–1383. <https://doi.org/10.1214/08-AOAS191>
- Geweke, J. (1992). Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. In J. O. Berger, J. M. Bernardo, A. P. Dawid, & A. F. M. Smith (Eds.), *Bayesian statistics* (Vol. 4, pp. 169–194). Oxford University Press. <https://doi.org/10.21034/sr.148>
- Hazelwood, R. J., Armeson, K. E., Hill, E. G., Bonilha, H. S., & Martin-Harris, B. (2017). Identification of swallowing tasks from MBSS that optimize the detection of physiologic impairment. *Journal of Speech, Language, and Hearing Research*, 60(7), 1855–1863. https://doi.org/10.1044/2017_JSLHR-S-16-0117
- Lazarsfeld, P. F., & Henry, N. W. (1968). *Latent structure analysis*. Houghton Mifflin.
- Leonard, R., & McKenzie, S. (2008). Dynamic swallow study: Instrumentation and measurement techniques. In R. Leonard & K. Kendall (Eds.), *Dysphagia assessment and treatment planning: A team approach* (2nd ed., 127–160). Singular.
- Little, R. J. A., & Rubin, D. B. (1987). *Statistical analysis with missing data*. Wiley.
- Logemann, J. A. (1997). Role of the modified barium swallow in management of patients with dysphagia. *Otolaryngology—Head & Neck Surgery*, 116(3), 335–338. [https://doi.org/10.1016/S0194-5998\(97\)70269-9](https://doi.org/10.1016/S0194-5998(97)70269-9)
- Martin-Harris, B., Brodsky, M. B., Michel, Y., Castell, D. O., Schleicher, M., Sandidge, J., Maxwell, R., & Blair, J. (2008). MBS measurement tool for swallow impairment—MBSImP: Establishing a standard. *Dysphagia*, 23(4), 392–405. <https://doi.org/10.1007/s00455-008-9185-9>
- Martin-Harris, B., & Jones, B. (2008). The videofluorographic swallowing study. *Physical Medicine and Rehabilitation Clinics of North America*, 19(4), 769–785. <https://doi.org/10.1016/j.pmr.2008.06.004>
- Martin-Harris, B., Logemann, J. A., McMahon, S., Schleicher, M., & Sandidge, J. (2000). Clinical utility of the modified barium swallow. *Dysphagia*, 15(3), 136–141. <https://doi.org/10.1007/s004550010015>
- Northern Speech Services. (2018). *Modified Barium Swallow Impairment Profile*. <https://www.mbsimp.com>
- Peladeau-Pigeon, M., & Steele, C. M. (2013). Technical aspects of a videofluoroscopic swallowing study. *Canadian Journal of Speech-Language Pathology and Audiology*, 37(3), 216–226.
- R Core Team. (2019). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing (Version 3.6.0). <https://www.R-project.org/>
- Robert, C. P., & Titterton, D. M. (2002). Discussion of “Bayesian measures model complexity and fit” by Spiegelhalter, D. J., Best, N. G., Carlin, B. P., & van der Linde, A. *Journal of the Royal Statistical Society, Series B*, 64(4), 583–639. <https://doi.org/10.1111/1467-9868.00353>

-
- Rosenbek, J. C., Robbins, J. A., Roecker, E. B., Coyle, J. L., & Wood, J. L. (1996). A penetration–aspiration scale. *Dysphagia*, 11(2), 93–98. <https://doi.org/10.1007/BF00417897>
- van Buuren, S. (2007). Multiple imputation of discrete and continuous data by fully conditional specification. *Statistical Methods in Medical Research*, 16(3), 219–242. <https://doi.org/10.1177/0962280206074463>
- van Buuren, S., & Groothuis-Oudshoorn, K. (2011). MICE: Multi-variate imputation by chained equations in R. *Journal of Statistical Software*, 45(3), 1–67. <https://doi.org/10.18637/jss.v045.i03>
- Zhao, X., & Reiter, J. P. (2010). A note on Bayesian inference after multiple imputation. *The American Statistician*, 64(2), 159–163. <https://doi.org/10.1198/tast.2010.09109>