Supplemental Material S1. Summary of the mathematical details of the latent class model and the procedures for construction of all model-derived conditional probabilities.

Ordinal Latent Class Model

Cumulative Logit Models for Ordinal Classes

A latent class model (LCM) relates a set of observed (manifest) categorical variables to a collection of L unobserved (latent) classes. In our setting, the manifest variables – the MBSImP component task scores – are ordinal categorical variables, and we assume the existence of a latent ordinal categorical variable which gives rise to the observed data. We define C as our class variable with L levels. We model cumulative logits for the first L-1 levels of C as

$$\ln[\text{Prob}(C \leq c)/\text{Prob}(C \geq c + 1)] = \alpha_c, \ \ c = 1, 2, \ldots, L\text{-}1, \quad (1)$$

where 'ln' indicates a natural logarithm. We constrain the intercept parameters as $\alpha_1 \leq \alpha_2 \leq \ldots \leq \alpha_{L-1}$, thereby imposing a natural ordering on the corresponding cumulative probabilities, and construct cumulative probabilities via the inverse logit function, so that

$$\text{Prob}(C \leq c) = 1/[1 + \exp(-\alpha_c)], \ \ c = 1, 2, \ldots, L\text{-}1.$$

Cumulative probabilities naturally yield class prevalences with

$$\text{Prob}(C = 1) = \text{Prob}(C \leq 1),$$

$$\text{Prob}(C = c) = \text{Prob}(C \leq c) - \text{Prob}(C \leq c - 1), \text{ for } c = 2, \ldots, L\text{-}1,$$

and

$$\text{Prob}(C = L) = 1 – \text{Prob}(C \leq L\text{-}1).$$

Item Response Probabilities

Consider the jth MBSImP component with ordinal scores ranging from 0 to $R(j)$. We use the notation '$R(j)$' to indicate that the maximum possible score is component dependent. (As previously noted, the maximum MBSImP score ranges from 2 to 4 depending on the specific component under consideration.) Let S_{jk} be the score for the kth swallowing task for component j, $k = 1, 2, \ldots, K(j)$. Here again, we adopt the notation '$K(j)$' to indicate the dependence of the number of swallowing tasks on the specific component under consideration. We construct cumulative logits for the $\{S_{jk}\}$ conditional on latent class c as follows:

$$\ln[\text{Prob}(S_{jk} \leq \ell | C = c)/\text{Prob}(S_{jk} \geq \ell{+}1 | C = c)] = \rho_{jk\ell|c}, \ \ \ell = 0, 1, \ldots, R(j)\text{-}1 \quad (2)$$

with the constraint that $\rho_{jk0|c} \leq \rho_{jk1|c} \leq \ldots \leq \rho_{jk,R(j)\text{-}1|c}$, and

$$\text{Prob}(S_{jk} \leq \ell | C = c) = 1/[1 + \exp(-\rho_{jk\ell|c})], \ \ \ell = 0, 1, \ldots, R(j)\text{-}1.$$

We construct the set of *item response probabilities*, that is, the collection of conditional score probabilities given latent class, based on differences of cumulative probabilities where

$$\text{Prob}(S_{jk} = 0|C = c) = \text{Prob}(S_{jk} \leq 0|C = c),$$

$$\text{Prob}(S_{jk} = \ell|C = c) = \text{Prob}(S_{jk} \leq \ell|C = c) - \text{Prob}(S_{jk} \leq \ell\text{-}1|C = c), \text{ for } \ell = 1, \ldots, R(j)\text{-}1,$$

and

$$\text{Prob}(S_{jk} = R(j)|C = c) = 1 - \text{Prob}(S_{jk} \leq R(j)\text{-}1|C = c).$$

Item response probabilities provide a probabilistic characterization of the component task scores for each latent class.

Log-Likelihood Construction

Construction of the log-likelihood is facilitated by assuming conditional independence. Within the context of MBSImP scoring, this assumption allows us to conclude that, once a subject's latent class is known, knowledge of specific swallow task scores provide no additional information about scores for any remaining task within the same domain. For subject i in class c, let S_{ijk} be the swallow task score for component j, task k with observed value s_{ijk}. Assuming scores follow a multinomial distribution, and appealing to the simplifying assumption of conditional independence, the ith subject's contribution to the log-likelihood is given by

$$L_i = \sum_j \sum_k \sum_\ell 1(s_{ijk} = \ell) \ln[\text{Prob}(S_{ijk} = \ell|C_i = c)],$$

where $1(s_{ijk} = \ell)$ is an indicator function which takes on a value of 1 if the observed swallowing task score, s_{ijk}, is ℓ, and takes on a value of 0 otherwise. The summation is over all components j for the specified domain (oral or pharyngeal), all tasks $k = 1, 2, \ldots K(j)$, and all scores $\ell = 0, 1, \ldots, R(j)$. The full log-likelihood is then constructed as the sum of the individual L_i across all subjects.

Latent Class Model-Derived Probabilities

The various components of the ordinal latent class model – class prevalences, cumulative probabilities, and item response probabilities – form the building blocks for construction of a number of probabilities of interest that allow for further characterization of latent classes.

OI Score Probabilities Given Class

As previously mentioned, an abbreviated scoring algorithm based only on OI scores is frequently used in clinical practice. Therefore, it is of clinical interest to evaluate the probability that an OI score is equal to a specific value given latent class membership. Specifically, let $\boldsymbol{OI}_j$ be the OI score for the jth MBSImP component. We are interested in evaluating the conditional probability of OI given latent class, $\text{Prob}(\boldsymbol{OI}_j = \ell|C = c)$, $\ell = 0, 1, \ldots, R(j)$. Recall that, with the exception of components 1, 5, 6, 15 and 16, $\boldsymbol{OI}_j$ is equal to the maximum task score, which we denote as M_j. To derive the probabilities of interest, we begin as in previous constructions with cumulative probabilities, with an initial focus on maximum scores, given their relationship to OI scores. Specifically, we start with construction of $\text{Prob}(M_j \leq \ell|C = c)$, and appeal to the wellknown fact that the maximum among a collection of scores is at most ℓ if and only if all scores are no more than ℓ. Specifically,

$$\text{Prob}(M_\text{j} \leq \ell | C = c) = \text{Prob}(S_{\text{j}1} \leq \ell, S_{\text{j}2} \leq \ell, \ldots, S_{\text{j,K(j)}} \leq \ell | C = c).$$

Under our assumption of conditional independence, we simplify the joint probability as a product of marginal probabilities, and write

$$\text{Prob}(M_\text{j} \leq \ell | C = c) = \text{Prob}(S_{\text{j}1} \leq \ell | C = c) \text{ x } \text{Prob}(S_{\text{j}2} \leq \ell | C = c) \text{ x } \ldots \text{ x } \text{Prob}(S_{\text{j,K(j)}} \leq \ell | C = c).$$

The probabilities of interest pertaining to the OI scores are then derived from these cumulative probabilities as follows:

$$\text{Prob}(\boldsymbol{OI}_\text{j} = 0 | C = c) = \text{Prob}(M_\text{j} = 0 | C = c) = \text{Prob}(M_\text{j} \leq 0 | C = c),$$

$$\text{Prob}(\boldsymbol{OI}_\text{j} = \ell | C = c) = \text{Prob}(M_\text{j} = \ell | C = c) =$$

$$\text{Prob}(M_\text{j} \leq \ell | C = c) - \text{Prob}(M_\text{j} \leq \ell\text{-}1 | C = c), \ell = 1, \ldots, R(j)\text{-}1,$$

and

$$\text{Prob}(\boldsymbol{OI}_\text{j} = R(j) | C = c) = \text{Prob}(M_\text{j} = R(j) | C = c) = 1 - \text{Prob}(M_\text{j} \leq R(j)\text{-}1 | C = c).$$

For components $j = 1$, 5, 15 and 16, OI scores are limited to values of 0, 2, 3 or 4 since maximum task scores of either 0 or 1 yield an OI score of 0. For these components,

$$\text{Prob}(\boldsymbol{OI}_\text{j} = 0 | C = c) = \text{Prob}(M_\text{j} = 0 | C = c) + \text{Prob}(M_\text{j} = 1 | C = c),$$

$$\text{Prob}(\boldsymbol{OI}_\text{j} = 1 | C = c) = 0,$$

and

$$\text{Prob}(\boldsymbol{OI}_\text{j} = \ell | C = c) = \text{Prob}(M_\text{j} = \ell | C = c), \ell = 2, 3 \text{ or } 4,$$

where $\text{Prob}(M_\text{j} = \ell | C = c)$ is constructed as described.

Finally, for component $j = 6$, an OI score of 0 occurs when all non-solid task scores are 0 and the cookie task score is either 0 or 1. Therefore $\text{Prob}(\boldsymbol{OI}_6 = 0 | C = c)$ is constructed as

$$\text{Prob}(S_{61} = 0, S_{62} = 0, \ldots, S_{68} = 0, S_{69} \leq 1 | C = c) =$$

$$\text{Prob}(S_{61} = 0 | C = c) \text{ x } \text{Prob}(S_{62} = 0 | C = c) \text{ x } \ldots \text{ x } \text{Prob}(S_{68} = 0 | C = c) \text{ x } \text{Prob}(S_{69} \leq 1 | C = c),$$

where tasks $k = 1$ through 8 are the non-solid tasks, task $k = 9$ is the cookie task, and we appeal to conditional independence to simplify the joint probability as a product of marginal probabilities. For the remaining OI scores for component 6, we have

$$\text{Prob}(\boldsymbol{OI}_6 = 1|C = c) = \text{Prob}(M_6 = 1|C = c) - \text{Prob}(S_{61} = 0, S_{62} = 0, \ldots, S_{68} = 0, S_{69} = 1|C = c),$$

and

$$\text{Prob}(\boldsymbol{OI}_6 = \ell|C = c) = \text{Prob}(M_6 = \ell|C = c),\ \ell = 2, 3, 4,$$

where $\text{Prob}(M_6 = \ell|C = c)$ is constructed as previously described.

Class Probabilities Given Oral and Pharyngeal Total Scores

Using the conditional OI score probabilities for a given latent class, we then construct the probabilities of primary interest to our study, namely, $\text{Prob}(C = c|\boldsymbol{OT} = x)$ and $\text{Prob}(C = c|\boldsymbol{PT} = y)$, where $\boldsymbol{OT}$ and $\boldsymbol{PT}$ are the oral and pharyngeal total score variables, and x and y are their observed values, respectively. Let $(\boldsymbol{OI}_1, \boldsymbol{OI}_2, \ldots, \boldsymbol{OI}_6)$ and $(\boldsymbol{OI}_7, \boldsymbol{OI}_8, \ldots, \boldsymbol{OI}_{16})$ be the overall impression score vectors for the oral and pharyngeal components so that $x = \boldsymbol{OI}_1 + \boldsymbol{OI}_2 + \ldots + \boldsymbol{OI}_6$ and $y = \boldsymbol{OI}_7 + \boldsymbol{OI}_8 + \ldots + \boldsymbol{OI}_{16}$. Furthermore, let $\varphi(x)$ and $\varphi(y)$ be the set of all OI score vectors that yield an oral total score of x and a pharyngeal total score of y, respectively. For example, for an oral total score of 1, $\varphi(x = 1) = \{(0, 1, 0, 0, 0, 0), (0, 0, 1, 0, 0, 0), (0, 0, 0, 1, 0, 0), (0, 0, 0, 0, 0, 1)\}$. (Recall that neither component 1 nor component 5 allows an OI score of 1, so $\varphi(x = 1)$ enumerates the only four ways in which an OT = 1 is possible across the six oral components.) It follows that

$$\text{Prob}(\boldsymbol{OT} = 1|C = c) = \sum_{\varphi(x = 1)} \text{Prob}(\boldsymbol{OI}_1, \boldsymbol{OI}_2, \ldots, \boldsymbol{OI}_6|C = c) =$$

$$\sum_{\varphi(x = 1)} [\text{Prob}(\boldsymbol{OI}_1|C = c) \text{ x } \text{Prob}(\boldsymbol{OI}_2|C = c) \text{ x } \ldots \text{ x } \text{Prob}(\boldsymbol{OI}_6|C = c)],$$

where we assume OI scores are conditionally independent given class. We use the same approach for other OT and PT scores. From these probabilities, we construct the probabilities of primary interest, $\text{Prob}(C = c|\boldsymbol{OT} = x)$ and $\text{Prob}(C = c|\boldsymbol{PT} = y)$, using an application of Bayes' rule. Specifically,

$$\text{Prob}(C = c|\boldsymbol{OT} = x) = [\text{Prob}(\boldsymbol{OT} = x|C = c) \text{ x } \text{Prob}(C = c)]/\text{Prob}(\boldsymbol{OT} = x)$$

and

$$\text{Prob}(C = c|\boldsymbol{PT} = y) = [\text{Prob}(\boldsymbol{PT} = y|C = c) \text{ x } \text{Prob}(C = c)]/\text{Prob}(\boldsymbol{PT} = y),$$

where, using the law of total probability, the probabilities in the denominators are given by

$$\text{Prob}(\boldsymbol{OT} = x) = \text{Prob}(\boldsymbol{OT} = x|C = 1) \text{ x } \text{Prob}(C = 1) + \ldots + \text{Prob}(\boldsymbol{OT} = x|C = L) \text{ x } \text{Prob}(C = L),$$

and

$$\text{Prob}(\boldsymbol{PT} = y) = \text{Prob}(\boldsymbol{PT} = y|C = 1) \text{ x } \text{Prob}(C = 1) + \ldots + \text{Prob}(\boldsymbol{PT} = y|C = L) \text{ x } \text{Prob}(C = L).$$