Latent Variable Mixture Modeling

A Flexible Statistical Approach for Identifying and Classifying Heterogeneity

Sarah J. Schmiege ▼ Paula Meek ▼ Angela D. Bryan ▼ Hans Petersen

Background: Latent variable mixture modeling is becoming increasingly popular in nursing research, in part due to the sophistication of the method in identifying relationships, patterns, and clusters in the data.

Objective: The aim of this study was to provide an overview of mixture modeling techniques, specifically as applied to nursing research, and to present examples from two studies to illustrate how these techniques may be used cross-sectionally and longitudinally.

Methods: The first data example demonstrates the use of latent profile analysis as applied to the St. George respiratory symptoms questionnaire in 2,232 smokers from the Lovelace Smokers Cohort. The second data example demonstrates growth mixture modeling as applied to condom use trajectories among 728 at-risk adolescents on probation.

Results: Three classes of symptoms emerged among the smokers cohort: those who were high on all symptoms, those who were low on all symptoms, and those who were high on cough and phlegm only. These classes were then distinguishable by participant gender and wood smoke exposure. In the second data example, four classes of condom use emerged. Only 59% of the sample indicated the previously reported decline in condom use over time; condom use remained stable or significantly increased for the remaining 41%.

Discussion: Both sets of results provide additional substantive information about patterns in the data that were not apparent from previously applied traditional methodological techniques. Considerations for the use of latent variable mixture modeling in nursing research are discussed.

Key Words: growth mixture modeling · latent class analysis · mixture modeling

The use of latent variable mixture modeling in nursing research has been increasing in popularity. This is, in part, because of the fact that these methods provide an innovative approach for answering a variety of substantive research questions that are frequently not possible with more traditional methods (e.g., analysis of variance [ANOVA], regression, cluster analysis, and factor analysis). Nurse researchers have worked to understand complex relationships and patterns that are part of improving health and well-being. Mixture modeling statistical methodology can assist nurse researchers in taking the next steps in understanding multifaceted relationships, intricate patterns, and clusters of symptoms that are needed to help guide interventions. The goal of this manuscript is to provide an overview of mixture modeling that will help nurse researchers to understand suitable uses of the methodology.

Overview of Methods
At a broad level, mixture modeling (i.e., latent class techniques) refers to a family of modeling techniques in which individuals are classified into subpopulations based on heterogeneity in the data (McLachlan & Peel, 2000; Muthe, 2002). Individuals who are similar along a dimension of interest (e.g., behavior and attitudes) are grouped into the same class, and those who are different are grouped into different classes. Mixture modeling is used to determine whether a categorical latent variable underlies the measured variables, wherein the category represents an individual’s projected assignment to a cluster or latent class (Gibson, 1959; Lazarsfeld & Henry, 1968). Individuals within the same class are expected to be similar to one another in regard to the variables of interest, but different from those in other classes. Mixture modeling can be distinguished from other analyses used to group individuals, such as cluster analysis, in that the classification of individuals into groups is done using a latent variable based on unobserved heterogeneity in the data (i.e., individual differences that are not directly measurable). Latent variables are those that are not observed directly but are inferred from measured or directly observed variables. (A familiar context for understanding latent variables is factor analysis, in which the commonalities among a variety of measured items are represented by

Sarah J. Schmiege, PhD, is Assistant Research Professor, Department of Biostatistics and Informatics, Colorado School of Public Health, University of Colorado-Denver, Aurora.

Paula Meek, PhD, RN, FAAN, is Professor, College of Nursing, University of Colorado-Denver, Aurora.

Angela D. Bryan, PhD, is Professor, Department of Psychology and Neuroscience, University of Colorado-Boulder.

Hans Petersen, MS, is Senior Research Associate, Lovelace Respiratory Research Institute, Albuquerque, New Mexico.

DOI: 10.1097/NNR.0b013e3182539f4c

Copyright © 2012 Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.
a single, in this case continuous, latent variable.) In mixture modeling, the latent variable is categorical because it represents likely class assignment based on the observed responses.

An advantage of mixture modeling techniques over traditional approaches (e.g., ANOVA and regression) is a more nuanced understanding of how similarities among individuals and differences between individuals can be captured and identified because classifications are not based solely on observed characteristics (e.g., gender and age). Mixture models also outperform traditional cluster analysis (Magidson & Vermunt, 2002), in that they allow for (a) variation in the types of variables that may be used as input data (e.g., continuous, categorical, counts, and censored), (b) greater flexibility in the types of models that can be estimated, (c) greater accuracy of classification, and (d) greater ability to include predictor variables to improve class distinctions.

There is great flexibility in how these techniques can be applied to data to answer research questions (Muthén, 2008; Table 1). One of the most common applications is latent class analysis (LCA) and latent profile analysis (LPA). Latent class analysis is used when the measured variables are categorical (e.g., yes or no for presence of a symptom or disease); LPA has the same goal as LCA but is used when the measured variables are continuous (e.g., body mass index). Both methods are used with cross-sectional data when the goal is to determine whether individuals can be clustered together on the same latent class based on their responses to a variety of measured indicators (e.g., multiple symptoms or indicators of health status), as well as whether there is only one class or multiple classes to characterize the sample. As an example, DeVon, Ryan, Rankin, and Cooper (2010) used LCA to classify patients based on their symptoms of acute coronary syndromes. In doing so, four symptom subgroups emerged: heavy symptom burden, chest pain only, sweating and weakness, and short of breath. These distinctions provide more information than what would be learned by traditional analyses used to evaluate simply level or type of symptoms.

Another modeling technique is to create latent classes based on the regression of one variable on another—regression mixture modeling. Standard regression analysis assumes a homogenous population, characterizing the relationship between independent and dependent variables using a single regression function (although observed group membership, such as gender, may be taken into account as a moderating or control variable). In contrast, regression mixture modeling is used to test for multiple regression functions related to unobserved heterogeneity within the population. In this approach, participants are clustered into latent classes based on similarity in the degree of relationship between the variables of interest (Ding, 2006; McLachlan & Peel, 2000). For example, to address inconsistencies in the literature of whether alcohol use has an impact on risky sexual behavior, Schmiege, Levin, and Bryan (2009) used regression mixture modeling to create latent classes characterizing the impact of alcohol use on condom use and frequency of intercourse. Three classes emerged: Alcohol use did not predict either outcome for a subset of participants, alcohol predicted condom use but not frequency of intercourse for another subset, and alcohol predicted both outcomes for the third subset. In these analyses, the commonly assumed relationship of alcohol use to risky sexual behavior applied to some, but not all, individuals.

Mixture modeling can also be extended to longitudinal data analysis, in which classes of participants are created based on their trajectories of responses over time. Standard growth curve modeling analysis estimates a common trajectory applied to all individuals (although there may be variability in the trajectory). Mixture modeling is useful in situations where the trajectory patterns are not the same for every individual and may be used to capture both normative and non-normative growth patterns (Muthén et al., 2002). For example, some individuals may demonstrate increased behavior over time, whereas others show decreased behavior; others may not change at all. Mixture modeling of longitudinal data may be done using latent class growth analysis (LCGA; Nagin, 1999; Roeder, Lynch, & Nagin, 1999) or growth mixture modeling (GMM; Muthén, 2004; Muthén & Asparouhov, 2009). In both cases, heterogeneity in the trajectory of an outcome over time is captured

### Table 1. Overview of Example Research Questions Matched to Common Mixture Modeling Methods and Data Characteristics

<table>
<thead>
<tr>
<th>Example Research Question</th>
<th>Data Characteristics or Conditions</th>
<th>Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Does a latent class that can cluster individuals based on their responses to key symptoms (pain, fatigue, sleep, shortness of breath) exist?</td>
<td>Cross-sectional</td>
<td>Latent Class Analysis</td>
</tr>
<tr>
<td>Can a latent class be used to capture variability in the effect of symptoms on quality of life?</td>
<td>Categorical or Continuous Longitudinal</td>
<td>Regression Mixture Modeling</td>
</tr>
<tr>
<td>Do different individuals show unique patterns of symptom trajectories over time?</td>
<td>Categorical or Continuous Longitudinal</td>
<td>Growth Mixture Modeling (within class variation can occur)</td>
</tr>
</tbody>
</table>

Mixture Modeling

Copyright © 2012 Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.
using a latent variable; the distinction is that no variation across individuals is allowed within classes in LCGA, whereas GMM allows within-class variation across individuals. Latent transition analysis (LTA; Collins & Wugalter, 1992; Lanza & Collins, 2008) represents another mixture modeling technique for analyzing longitudinal data, whereby latent class indicators are measured over time, but individuals may transition from one latent class to another across time points. For further reading on these topics, see Clogg (1995), Collins and Lanza (2010), and Hagenaars and McCutcheon (2002) for LCA; Muthén and Muthén (2000) and Muthén (2004) for GMM; and Muthén (2001, 2002) for latent variable mixture modeling more generally.

Mixture Modeling in Nursing Research

There are a variety of ways in which mixture modeling may be used in nursing research to drive a better understanding of patient burden and response to intervention. For example, LCA or LPA may be used to generate symptom clusters, LCGA or GMM could be applied to identify clusters based on illness trajectories, and a variety of mixture modeling techniques could be used to identify latent subgroups of individuals based on response to an intervention. Applications of mixture modeling in nursing research have increased in the past decade, likely because of increased recognition of the value of these techniques in contributing to substantive research questions and because advances in statistical computing have made these techniques more accessible to investigators.

In much of this work, LCA and LPA were used to identify symptom clusters; for example, among patients with acute coronary symptoms (DeVon et al., 2010), those with acute myocardial infarction (Ryan et al., 2007), childhood cancer survivors (Finnegan et al., 2009), and patients receiving biotherapy (Dodd et al., 2011). Other LCA and LPA applications in nursing have included identifying classes of care recipients and their caregivers based on patterns of community-based support service (Beeber, Thorpe, & Clipp, 2008), classifying harmful behaviors among nursing home residents (Draper, Brodaty, & Low, 2002), identifying patterns of service utilization among caregivers of older individuals (Hong, 2010), classifying influences on health behaviors (Kaiser, Brown, & Baumann, 2010), identifying patterns of visits to nursing homes to persons with dementia by family caregivers (Yamamoto-Mitani, Aneshensel, & Levy-Storms, 2002), creating typologies of residential care and assisted living (Park, Zimmerman, Sloane, Gruber-Baldini, & Eckert, 2006), and examining patterns of asthma control perception among adolescents (Rhee, Belyea, & Elward, 2008). An advantage of this approach is that it allows a more nuanced understanding of how individuals may be classified (e.g., the four distinct classes of symptoms among patients with acute coronary syndromes in DeVon et al., 2010), rather than simply assigning to them values of high or low on one dimension or outcome at a time.

Other nursing investigators have used GMM techniques to categorize individuals based on their trajectories of scores (e.g., symptom trajectories) over multiple time points. Chen, Wang, and Huang (2008) classified individuals based on their trajectory of functional status scores at four time points in the 6 months after hospitalization. Dunn, Cooper, et al. (2011) modeled latent subgroups of trajectories of symptoms 6 months after surgery for breast cancer, and Dunn, Aouizerat, et al. (2011) modeled trajectories of anxiety during and after radiation therapy among oncology patients and their family caregivers. Femia, Zerri, Stephens, and Greene (2007) used GMM to examine patterns of behavioral and psychological symptoms of dementia over time. Nies and Sun (2008) used GMM to examine level of responsiveness to a longitudinal intervention to increase walking. The unique classifications resulting from these methods have clear treatment and prevention implications.

Guide to Model Estimation

As a first step, the statistical software program used for mixture modeling needs to be considered. At the time of this writing, mixture modeling techniques were not possible in standard statistical packages (e.g., SPSS and Stata), with the exception of user-contributed routines that run under SAS for LCA (Proc LCA; Lanza, Dziak, Huang, Xu, & Collins, 2011), latent transition analysis (Proc LTA; Lanza et al., 2011), and LCGA (Proc Traj; Jones, Nagin, & Roeder, 2001). Specialized software specific to carrying out LCA exists, including LatentGold, LEM, Panmark, and WinMira, to name a few. Mplus software (Muthén & Muthén, 1998/2010) is a latent variable modeling program (i.e., structural equation models and latent growth curve models) that is a powerful and flexible tool for estimating a variety of mixture models. Of note is that it is possible to estimate a variety of cross-sectional (LCA and LPA) and longitudinal (GMM, LCGA, and LTA) mixture models in Mplus. Furthermore, observed variables can be continuous, binary, ordered categorical, counts, censored, or combinations of these.

Estimation of mixture models can be considered an iterative process whereby the investigator estimates multiple models, each with an increasing number of classes, and then determines which model provides the best representation of the data. Mixture modeling techniques may be considered exploratory in the sense that the investigator is choosing the best-fitting model from several competing models. Confidence in the final solution is based on several statistical indices of fit, as well as the theoretical meaningfulness and conceptual interpretability of the class structure (Muthén, 2002; Muthén & Muthén, 2000).

Statistical measures of fit recommended by Nylund, Asparouhov, and Muthén (2007) include the Bayesian Information Criterion (BIC) Index, the Lo–Mendell–Rubin likelihood ratio test (LMR-LRT; Lo, Mendell, & Rubin, 2001), and a bootstrap likelihood ratio test (BLRT; McLachlan & Peel, 2000). Lower BIC numbers indicate a better fitting model; for both LMR-LRT and BLRT, fit is determined by a significance test comparing the log-likelihood ratio from the estimated model to that from a model with one fewer class, where a significant value indicates that the estimated model fits the data significantly better than a model with one fewer class.

Using statistical indices to determine the best-fitting model is a large part of deciding on the number of classes that exist in a data set, although interpretability and labeling of the classes by the investigator is also important. Once the statistical indices are examined to determine the best-fitting model, the next step is to examine the class structure to

Copyright © 2012 Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.
provide a conceptual interpretation of the observed structure and assign labels to the observed classes. This step is part of the decision-making process in that there may be instances where the best-fitting statistical solution is not considered viable because it does not make conceptual sense or the investigator must decide between two or more close-fitting models. As a result, there is a level of subjectivity in mixture modeling, and this step, in particular, relies on the substantive content knowledge of the investigative team.

As part of the modeling process, investigators also typically wish to determine how the observed class structure relates to other variables in the data set. In other words, there are a variety of measured variables in the data set, and it may be of interest to determine how these variables relate to the class structure (Muthén, 2002; Vermunt & Magidson, 2002). For example, perhaps class structure differs by gender (an observed variable), wherein women are more likely to end up in one class and men are more likely to end up in another. There are two broad ways in which the relationship of class structure to other variables of interest may be assessed. One common option (e.g., DeVon et al., 2010; Kaiser et al., 2010; Rhee et al., 2008; Ryan et al., 2007) is to take a stepwise approach by, first, generating each individual’s most likely class assignment using a specialized software program for mixture modeling. The likely class assignment variable is then merged back into the main data set, and then standard statistical techniques (ANOVA for continuous observed variables and chi-square for categorical observed variables) are used to determine whether class membership is related to differences in the observed variable. The second approach is to stay entirely within the mixture modeling framework (and software package) to estimate the impact of the predictor variables on most likely class membership. For example, in Mplus, the relationship of observed variables to the class structure can be estimated by treating the observed variable as a covariate to predict class membership using multinomial logistic regression. Of the two approaches, the latter approach is preferable because class membership is treated as the probability of class assignment, rather than as a final assignment variable, thereby reducing sampling variability, and because covariates may provide useful information about the class formation that could lead to changes in likely class membership (Muthén, 2004). Regardless of which of these two approaches is used, inclusion of covariates or explanatory variables should be theoretically driven and specified a priori, rather than carried out in an exploratory fashion.

Because of the complexity of mixture models, problems may occur during model estimation. A common problem is that of local maxima, where the estimation algorithm converges on a local best solution rather than the overall, global best solution. This issue may be addressed by running the algorithm many times using different start values each time, a process that is now automated in many software programs (e.g., Mplus). For example, the default in Mplus is to use 100 random starts in model estimation to ensure the consistency of model results, although this default can be overridden to include even more random starts (e.g., hundreds or thousands of start values) if error messages about local maxima persist. Another estimation issue is that of model identification. Although a complete discussion of this issue is beyond the scope of this article, researchers should be aware of the potential for problems during model estimation and should check their output carefully for error or warning messages.

Users of these techniques should also consider issues related to sample size and missing data prior to model estimation. These techniques are not appropriate for very small sample sizes. However, it is challenging to provide concrete guidelines around sample size, given that sample size requirements are specific to the type of model under investigation, the separation of the classes, and the characteristics of the data (Marsh, Lüdtke, Trautwein, & Morin, 2009). Power calculations are recommended for arriving at an appropriate sample size and Mplus provides a Monte Carlo simulation feature for performing power calculations for mixture modeling (Muthén & Muthén, 2002).

It is also rare that a data set includes complete data; some level of missing data is to be expected. Two state-of-the-art approaches for handling data that can at least be considered missing at random (Rubin, 1976) are full information maximum likelihood estimation (Enders & Bandalos, 2001) and multiple imputation (Schafer, 1997). Full information maximum likelihood is available in Mplus (and is currently the default setting) but is not available for mixture modeling techniques in other statistical packages. In this case, using multiple imputation prior to applying mixture modeling capabilities may be a good option.

Examples

Two examples are provided to demonstrate the application of these techniques to cross-sectional and longitudinal data. These examples are intended for illustrative purposes only, rather than to advance the substantive literature. Mplus version 6.0 was used for both examples (Muthén & Muthén, 1998/2010). The general approach for each example was to start by estimating a two-class model and then to use an iterative process to estimate models with one additional class until there remained no evidence from the data that information would be gained by estimating additional classes. Once the class structure was finalized, the impact of other explanatory variables was estimated directly within the Mplus software program; each variable was examined as a predictor of class membership using multinomial logistic regression.

Data Example 1: LPA of Respiratory Symptoms

The data for this example come from the Lovelace Smokers Cohort (Sood et al., 2010), a large cohort study of smokers in New Mexico. Participants were 2,232 individuals with at least 10 pack-years of smoking history. The sample was predominantly women (76.26%), 17.55% of participants were Hispanic, and at the time these data were collected participants ranged in age from 40 to 77 years (M = 56.27, SD = 9.42). Just over half of participants were current smokers (58.61%); a smaller percentage (10.56%) reported recent exposure to wood smoke.

There have been several publications using this cohort, for example, the examination of wood smoke exposure and gene promoter methylation as predictors of COPD risk (Sood et al., 2010). The goal of the present analyses was...
to apply LPA to determine whether participants could be grouped into latent classes based on their respiratory symptoms. A secondary aim was to determine how class membership was related to participant gender, current smoking status, and wood smoke exposure, all variables shown previously to be related to respiratory symptoms (Sood et al., 2010). Participants completed questionnaires to provide information about demographics, respiratory diseases, smoking history, and wood smoke exposure. Respiratory symptoms were assessed using the Saint George Respiratory Questionnaire (Jones, Quirk, Baveystock, & Littlejohns, 1992) symptom subscale. Six items were used for analysis: cough, phlegm, shortness of breath, wheezing, frequency of severe attacks, and number of good days. Items were maintained on the original 1–5 scale, rather than using the weighted values, and items were reverse-coded so that higher scores indicated greater symptomatology.

The statistical fit indices, including the BIC value, the log-likelihood value, the LMR-LRT, and the BLRT for two-, three-, and four-class solutions are presented in Table 2. On the basis of an initial evaluation of the statistical fit indices, the four-class model appeared to fit the data the best, as it had a lower BIC value and a significantly lower log-likelihood value compared to the three-class model. However, there were problems in the estimation of the four-class model. An error message showed that the best log-likelihood value was not replicated and that the solution may therefore not be trustworthy; furthermore, it was not possible to replicate the best log-likelihood value in 62% of the bootstrap draws. According to the error message, the problem may have been due to problems with model identification or with local maxima. The number of random starts was increased to 10,000 and still encountered problems in the estimation of the four-class model. To identify whether the source of the problem was identification or problems with local maxima, a five-class model was estimated; the same error message was received, suggesting the problem was likely due to model identification.

On the basis of the statistical concerns with the four-class model, the observed class structure of the two-, three-, and four-class solutions were examined carefully to guide decision making for the final model. The mean values for each of the symptoms for each class are presented in Figure 1, with the two-class solution presented in Figure 1a, the three-class in Figure 1b, and the four-class in Figure 1c. As shown,

The three- and four-class solutions were similar in that they both included a class of individuals with low scores on all symptoms, a class with high scores on all symptoms, and a

<table>
<thead>
<tr>
<th>Class</th>
<th>BIC</th>
<th>Log-likelihood</th>
<th>LMR-LRT</th>
<th>BLRT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Two-class solution</td>
<td>41,304.51</td>
<td>−20,579.01</td>
<td>4,016.98***</td>
<td>4,091.40***</td>
</tr>
<tr>
<td>Three-class solution</td>
<td>39,223.17</td>
<td>−19,511.35</td>
<td>2,096.47***</td>
<td>2,135.31***</td>
</tr>
<tr>
<td>Four-class solution*</td>
<td>37,776.51</td>
<td>−18,761.04</td>
<td>1,473.34***</td>
<td>1,500.63***</td>
</tr>
</tbody>
</table>

Notes. BIC = Bayesian Information Criteria; LMR-LRT = Lo-Mendell-Rubin likelihood ratio test; BLRT = bootstrap likelihood ratio test.

*Error message obtained during model estimation stating that the best log-likelihood value was not replicated and that there may be problems with model identification; this error message maintained even after increasing the number of random starts considerably.

***p < .001.
class with high scores on the cough and phlegm symptoms only, with particularly low scores on number of attacks and good days. The difference is that the four-class solution contained an additional class that showed the same pattern of high scores on cough and phlegm only as observed in the three-class solution, although severity of attacks and good days scores were slightly lower in this class. The distinction between the three- and four-class solutions was not seen as substantively useful.

The three-class solution was maintained as the final model, based on the estimation problems encountered in the four- and five-class models, and because the additional class did not contribute to model interpretability. The first class included 14% of the sample and evidenced high endorsement of all the symptom items. This class was thus labeled the High Symptoms class. The second class contained 41% of the sample, evidenced low endorsement of all the symptom items, and was labeled the Low Symptoms class. The third class contained 45% of the sample and was unique in that individuals in this class strongly endorsed the cough and phlegm items but showed less endorsement of the other symptoms. This class was labeled the Dominant Airway Clearance Symptoms class.

Participant gender, current smoking status, and wood smoke exposure were examined as predictors of class membership. Gender distinguished membership in the Low Symptoms class relative to the Dominant Airway Clearance Symptoms class, where being male predicted membership in the Dominant Airway Clearance Symptoms class (\(B = .26\) (SE = .12), \(p < .05\)). Wood smoke exposure distinguished membership in the High Symptoms class relative to the Dominant Airway Clearance Symptoms class, where greater wood smoke exposure predicted membership in the High Symptoms class (\(B = .42\) (SE = .21), \(p < .05\)). Current smoking status did not relate to class membership.

The analysis uncovered a symptom pattern that was not clear with classic techniques using the mean of individual symptom items or a subscale score. The Dominant Airway Clearance Symptoms class, although commonly reported in the clinical setting, was not distinguished as a pattern response to standardized items on a questionnaire, and in this case, nearly half of the sample was classified into this group. Identification of these groups and the Dominant Airway Clearance Symptoms class, in particular, can advance care by focusing on these symptom phenotypes and more targeted interventions.

**Data Example 2: GMM of Risky Sexual Behavior Among Adolescents**

The data for this example come from a longitudinal study assessing sexual risk and substance use among youth (\(n = 728\)) recruited from juvenile probation offices in the Denver area. Participants were surveyed every 6 months for 2 years to yield 5 data points (0, 6, 12, 18, and 24 months). Retention rates ranged from 78.6% to 90.1% across follow-ups. Participants (485 men and 243 women) were ethnically diverse (40.9% Hispanic, 24.5% African American, and 15.7% White) and had a mean age of 15.71 years (\(SD = 1.05\) years). Most participants (87.52%) were experienced sexually at baseline.

Analyses testing the main hypotheses of this study are presented elsewhere (Bryan, Schmiege, & Magnan, in press), wherein a significant decrease in condom use was observed over the 2-year period and greater marijuana use was associated with a steeper decline in condom use. The goal was to apply GMM to provide a more nuanced understanding of the primary results by determining whether this growth trajectory was consistent across all participants or if there were variations in the condom use trajectories that could be captured by latent classes. Frequency of condom use was measured on a 5-point scale from never to always at each time point ("In the past 6 months, how much of the time did you use condoms when you had sexual intercourse"). This variable was treated as continuous for analytical purposes.

The two-, three-, four-, and five-class solutions were tested as possible fits to the data (Table 3). Collectively, each statistical measure of model fit supported the four-class model. The BIC value was the lowest for the four-class model, and both the LMR-LRT and the BLRT were significant for the two-, three-, and four-class models, indicating that each model fits the data better than a model with one fewer class (i.e., the two-class model fits better than a one-class model, and the three-class model fits better than a two-class model). However, these values were not significant for the five-class model, indicating that a five-class model did not fit the data significantly better than the four-class model. The four-class model was retained as the final model.

**TABLE 3. Model Estimation Results for Longitudinal Data Example**

<table>
<thead>
<tr>
<th>Class</th>
<th>BIC</th>
<th>Log-likelihood</th>
<th>LMR-LRT</th>
<th>BLRT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Two-class solution</td>
<td>7,843.50</td>
<td>-3,882.60</td>
<td>114.72***</td>
<td>120.58***</td>
</tr>
<tr>
<td>Three-class solution</td>
<td>7,780.53</td>
<td>-3,841.33</td>
<td>78.53***</td>
<td>82.54***</td>
</tr>
<tr>
<td>Four-class solution</td>
<td>7,574.38</td>
<td>-3,728.47</td>
<td>214.75**</td>
<td>225.73***</td>
</tr>
<tr>
<td>Five-class solution</td>
<td>7,590.98</td>
<td>-3,726.98</td>
<td>2.832</td>
<td>2.976</td>
</tr>
</tbody>
</table>

Notes. BIC = Bayesian Information Criteria; LMR-LRT = Lo-Mendell-Rubin likelihood ratio test; BLRT = bootstrap likelihood ratio test.

**p < .01.

***p < .001.
The next step was to interpret the four classes. The mean condom use values over time for each class are presented in Figure 2. The first class included 19% of the sample; the condom use mean was moderate (M = 2.99 on a 5-point scale), and the condom use trajectory was nonsignificant in this class, indicating no change in condom use over time. This class was thus labeled the Moderate Baseline/No Change Over Time class. The second class comprised 25% of the sample; mean condom use was moderate to high at baseline (M = 4.00), and condom use significantly decreased over time. This class was labeled the Decrease Over Time class. The third class comprised 22% of the sample. Here mean condom use was low at baseline (M = 1.47), and condom use significantly increased over time. This class was labeled the Low Baseline/Increase Over Time class. The fourth class comprised 34% of the sample; mean condom use was at the upper end of the scale in this class (M = 5.00), and condom use in this class significantly decreased over time. This class was labeled the High Baseline/Decrease Over Time class.

These results provide a more complete understanding of the condom use trajectories observed in this population than the traditional growth modeling applied in Bryan et al. (in press). The previous conclusion was that condom use significantly decreased over time. The results of the GMM analyses demonstrate that this pattern only holds for 59% of the sample, the Decrease Over Time and High Baseline/Decrease Over Time classes. These two classes are distinguished further by the fact that 34% of the sample was at the upper limit of the condom use measure at baseline, and their scores could, therefore, only decrease or remain stable over time. The remaining 41% of participants whose condom use did not decrease over time were divided into those who showed a consistent, moderate pattern of condom use over time and those whose condom use scores were low at baseline and thus significantly increased over time.

**Discussion**

This presentation serves as a brief introduction to latent variable mixture modeling, a method that is used increasingly in nursing research as statistical computing becomes more powerful and nurse researchers become interested in how these methods may be applied to provide unique answers to substantive research questions. Two data applications of the method were illustrated: a cross-sectional, LPA of respiratory symptoms among individuals in the Lovelace Smoking Cohort, and a longitudinal, GMM of condom use over a 2-year period among at-risk adolescents. A three-class solution emerged in the LPA, wherein one class showed low scores on all the symptoms, another showed high scores on all the symptoms, and the third class, over 40% of the sample, showed high levels of cough and phlegm only. A four-class solution emerged in the GMMs, wherein only two classes showed the expected decrease in condom use over time and the two other classes showed moderate initial condom use and no change over time and low initial condom use and increasing scores over time. The substantive interpretation of these findings is beyond the scope of this study, because these examples are used for illustrative purposes only. What is notable about these findings, however, is that, in both cases, the latent classes added to the substantive understanding beyond what has been reported using traditional methods. In both examples, classes emerged that showed unique patterns in the data not identified by traditional methods (where the goal is to identify the average association among the variables, and patterns that deviate from the average association are treated simply as noise in the data).

Also provided here is an overview of the methodology used to estimate and interpret these models. In brief, the general procedure is to estimate a series of competing models with increasing numbers of classes and then to choose the most appropriate model based on statistical fit indices and conceptual interpretability. As illustrated by the first example, decisions in choosing a final model are not always straightforward. In this example, the statistical fit indices supported a four-class model over the three-class model; however, there were warning messages in the output that cautioned against trusting the results of the four-class model, and the additional class did not appear to add to the substantive interpretation of the findings. For these reasons, the three-class model was maintained. This result underscores the importance of viewing output carefully for error or warning messages. It also shows that, although model decisions should be made primarily based on statistical fit,
there is some level of subjectivity in arriving at a final solution. Having a substantive expert is crucial for the interpretation of the statistical solution. Model replication is similarly important when using these methods; it is possible that a different investigator would have decided that an alternative solution to the data could be considered more reasonable. These caveats around final model selection are important to understand when estimating models and to be a discerning consumer of mixture modeling methodology in other published studies.

In summary, mixture modeling approaches are useful for identifying and classifying heterogeneity in the data that is not observed directly. In particular, for identifying latent classes of variables that will help nurses to refine existing treatments and create new interventions for various conditions and disease processes. Given the complex nature of nursing research, mixture modeling techniques present an advantage over traditional approaches by allowing nuanced understanding of variability among individuals beyond patterns of simple high-versus-low behavior or symptoms or of simply the normal trajectory pattern over time. The unique classifications resulting from these methods have clear treatment and prevention implications for nursing research.

References


