VARIABLE SELECTION FOR LATENT TRANSITION ANALYSIS

A DISSERTATION

SUBMITTED ON THE SECOND DAY OF MARCH 2015

TO THE DEPARTMENT OF BIOSTATISTICS AND BIOINFORMATICS

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

OF THE SCHOOL OF PUBLIC HEALTH AND TROPICAL MEDICINE

OF TULANE UNIVERSITY

FOR THE DEGREE

OF

DOCTOR OF PHILOSOPHY

BY

CARA JOYCE

APPROVED:

Leann Myers, Ph.D. (Chair)

Larry Webber, Ph.D.

Marie A. Krousel-Wood, M.D., M.S.P.H.

Acknowledgements

I would like to thank my advisor and committee chair Dr. Leann Myers for her wonderful guidance and support during my time at Tulane, and her encouragement towards completing my dissertation. I am very grateful to Dr. “Tonette” Wood for introducing me to the amazing CoSMO research team, and the many opportunities she provided throughout. I would also like to express my deep gratitude to Dr. Larry Webber for his fantastic advice, inspiration, and mentorship.

Thanks to the faculty and colleagues at Tulane who impress and invigorate with their passion for Public Health, and to my friends for the good times and camaraderie. Finally, I owe it all to my family: my parents Michael and Deborah, my brothers Neal and Paul, and my partner Jorge Calderón.

TABLE OF CONTENTS

ABSTRACT

***Background*:** Latent Transition Analysis (LTA) is an effective tool used to investigate underlying groupings in multivariate categorical data over time. Variable selection is not a primary focus in LTA, however; researchers choose data elements *a priori* and include all variables in the model.

***Objective*:** To expand and evaluate variable search methods developed for Latent Class Analysis for application to the longitudinal LTA setting, and to apply these techniques to an expansive data set describing transitions in medication adherence.

***Methods*:** Simulated multivariate categorical data for two time periods consisted of a mixture of variables that separate statuses and noise variables. LTA models were specified via 1) a headlong search algorithm that identifies variables that account for significant between-group variation in probability, and 2) an item elimination method that sequentially removes variables that do not affect classification. Sample size, item response probabilities, number of statuses, and transition probabilities varied in simulations.

The Cohort Study of Medication Adherence in Older Adults (CoSMO), a prospective study of the barriers to antihypertensive medication adherence, served as an example of these procedures applied to real-world data. Each algorithm identified status-informative variables from a survey on mental and physical quality of life, and categorical variables measuring medication adherence, psychosocial factors, healthcare utilization, and clinical variables were profiled to measure and characterize changes in CoSMO participants' statuses over two years.

Results: Both algorithms performed well with fewer starting variables, large sample size relative to the number of variables included, and sizeable statuses across time. Simulations, as well as the application to CoSMO data, suggested that headlong search may select fewer variables and a simpler model whereas backwards elimination favors retention of more variables and a more complex model.

INTRODUCTION

Latent class analysis (LCA) identifies underlying groupings among subjects described by categorical variables. The latent variable is categorical, estimated from multiple categorical variables, and is not observed directly. LCA is an appealing approach for understanding complex interactions in multivariate categorical data. This method can identify subgroups of subjects or individuals that share common traits as measured by the collected variables used to specify the model (1). In psychology and the behavioral sciences, application of LCA methods to survey data distinguishes qualitative differences in individuals based on response patterns to survey questions. Some examples include modeling youth drug abuse and sexual risk behaviors (2, 3), depressive symptomatology (4, 5), and reasons for and patterns of medication non-adherence (6, 7).

For longitudinal data, an extension of LCA called latent transition analysis (LTA) includes both the latent class variable and probability of change in model specification (8). LTA is a person-centered approach that assumes individuals may transition from one group to another over time, with a latent status assigned at each time point. Estimation of transition probability is informative to public health researchers who seek to identify those susceptible to engagement in risky behaviors (9-11) or diminished levels of health or wellness over time (12, 13), for example.

The identification of the appropriate number of classes and statuses for LCA and LTA, respectively, is the focus of model selection techniques. Variable selection methods are not often discussed in the literature, however, and latent models are typically specified with all available variables (14). Inclusion of all variables is not recommendable in studies with more variables than observations or with broad data collection. An identifiable model may not exist when the multivariate contingency tables are complex. The inclusion of non-differentiating variables can lead to the selection of a final model with poorly separated groups or other problems of interpretability.

Dean and Raftery (14) and Bartolucci, Montanari, and Pandolfi (15, 16) each propose a new method for identifying appropriate variables for LCA models. In an extension of a previous technique for variable selection in clustering continuous variables (17), Dean and Raftery utilize a headlong search algorithm that identifies variables that explain a significant amount of variance in subgroups. At various steps in this procedure, variables are proposed for inclusion or exclusion by differences in the Bayesian information criterion (BIC) until convergence is achieved. This algorithm identified correct variables in simulation studies and provided improvements in classification. Databases with a large number of single nucleotide polymorphisms offered an example on real data, and the method detected similar class structures with far fewer variables than all available.

Bartolucci, Montanari, and Pandolfi present a variable reduction technique with the goal of identifying the smallest subset of variables possible that still fully describes the latent classes (15, 16). The procedure begins with all available variables in an LCA model, and the maximum posterior probabilities assign individuals to classes. Then, the algorithm removes variables one at a time based on the proportion of individuals that switch assigned class and the Kullback-Leibler distance. Bartolucci *et al* provide results of this technique applied to a large data set of survey items on the health of elderly patients in Italy. A step-by-step analysis of the algorithm shows that removing several variables will have no impact on classification and that the removal of nearly two-thirds of variables results in less than 5% change in classification. A shorter survey tool in this context reduces cost for future data collection and minimizes missing responses due to survey fatigue.

While there are no established procedures for conducting variable selection in the latent transition analysis setting, it is possible to extend and test the cross-sectional latent class methods discussed here on repeated measures data. Upon validation of these techniques, researchers conducting longitudinal studies will have methods for investigating change in individuals when full-variable models are not practical, parsimonious, or estimable. For example, key items selected from lengthy surveys can describe patterns of change when many candidate questions exist. Similarly, LTA models can identify principal risk factors for negative health outcomes when an extensive database of self-reported variables, clinical characteristics, medical records, and claims data are at hand.

LITERATURE REVIEW

Latent class analysis

Paul Lazarsfeld proposed the fundamentals of latent class theory in an assessment of survey data collected on American soldiers in 1950 (18, 19). Together with Neil Henry, he presented a full review of his work in a first textbook on latent class models (20); this 1968 text Latent Structure Analysis also introduced the theory of latent class models for repeated measures data, though it lacked a reliable approach for estimation of model parameters. Goodman furthered this work with a method for obtaining the maximum likelihood estimates for parameters in LCA models (21). This numerical method is similar to the expectation maximization (EM) algorithm that is an option in many statistical software packages used today (22-24).

Model notation varies in the literature; the parameters presented here are those of Collins and Lanza (1, 22). The variables for the latent class model are j = 1,…,J observed variables with response categories r_j= 1,…, R_j. The model is specified by considering the cross-tabulation of the J variables with W = $\prod_{j=1}^{J} R_j$ cells. If the model consists of survey items, for example, each cell of the cross-tabulation can be thought of as a full response pattern to the J questions. The latent variable L has categorical levels c=1,…,C and is error-free. Specifying models with two classes, three classes, and so on can determine the appropriate value of c; measures such as the likelihood ratio statistic, Akaike's Information

Criterion (AIC), and Bayesian Information Criterion (BIC) compared across models can be used to assess fit. The latent variable is not collected directly; it is measured from the J observed variables which are subject to error.

The overall probability of each level of the J variables is ρ_{j,r_j}, and probability within each class is $\rho_{j,r_j|c}$. The prevalence γ_c represents the proportion of subjects in latent class c, and $\sum_{c=1}^{C} \gamma_c = 1$. For a specific response pattern y with item responses y_j, the probability of the y response pattern as a function of γ and ρ is given as:

$$P(Y = y) = \sum_{c=1}^{C} \gamma_c \prod_{j=1}^{J} \prod_{r_{j=1}}^{R_j} \rho_{j,r_j|c}^{I(y_j=r_j)}$$

There is no closed-form solution to this formula, and the EM algorithm provides estimates of the maximum likelihood values of ρ and γ parameters (22). A specific γ_c in the above formula replaces the summation of γ_c to obtain the joint probability of membership in latent class c and response pattern y for an individual:

$$P(Y = y, L = c) = \gamma_c \prod_{j=1}^{J} \prod_{r_{j=1}}^{R_j} \rho_{j,r_j|c}^{I(y_j=r_j)}$$

A manipulation of these formulas provides a method to calculate the posterior probability of membership in the class c given a response pattern, with the help of Bayes rule:

$$P(L = c \mid Y = y) = \frac{\gamma_c \prod_{j=1}^{J} \prod_{r_j=1}^{R_j} \rho_{j,r_j|c}^{I(y_j=r_j)}}{\sum_{c=1}^{C} \gamma_c \prod_{j=1}^{J} \prod_{r_j=1}^{R_j} \rho_{j,r_j|c}^{I(y_j=r_j)}}$$

Researchers often assign class to individuals based on the maximum posterior probability observed across all possible classes.

Latent transition analysis

Latent transition analysis employs a latent Markov model to estimate the probability of transitions in class over time (25). Assuming the previous notation of LCA, relevant parameters and formulas for LTA are as follows. The J indicator variables are now collected at times t = 1,…,T and response categories within each J variable at a given time t are indexed as $r_{j,t}$ = 1,…,$R_{j,t}$. The latent variable L has S associated latent statuses (previously referred to as classes) labeled s_t.= 1, …, S at time t. The prevalence of the latent status s at time t is δ_{s_t}. The notation for item response probabilities expands to $\rho_{j,r_{j,t}|s_t}$ and corresponds to the probability of response $r_{j,t}$ to item j given latent status s_t. Transition probabilities from time t to time t + 1 are given as $\tau_{s_{t+1}|s_t}$, visualized in matrix form as:

$$\begin{pmatrix} \tau_{1_{t+1}|1_t} & \tau_{2_{t+1}|1_t} & \cdots & \tau_{S_{t+1}|1_t} \\ \tau_{1_{t+1}|2_t} & \tau_{2_{t+1}|2_t} & \cdots & \tau_{S_{t+1}|2_t} \\ \cdots & \cdots & \cdots & \cdots \\ \tau_{1_{t+1}|S_t} & \tau_{2_{t+1}|S_t} & \cdots & \tau_{S_{t+1}|S_t} \end{pmatrix}$$

The probability of Y, a vector of responses over all times is then given by:

$$P(Y = y) = \sum_{s_1=1}^{S} \cdots \sum_{s_T=1}^{S} \delta_{s_1} \tau_{s_{2|s_1}} \cdots \tau_{s_{T|s_{T-1}}} \prod_{t=1}^{T} \prod_{j=1}^{J} \prod_{r_{j,t}=1}^{R_j} \rho_{j,r_{j,t}|s_t}^{I(y_{j,t}=r_{j,t})}$$

The EM algorithm provides estimates of latent status prevalence at time 1, transition probabilities, and item-response probabilities at each time point. The estimated prevalence of each status at time 1 and transition probabilities allow for calculation of status prevalence at later times. In some circumstances, it may be desirable to constrain the item-response probabilities to be equal across times to reduce the number of parameters estimated.

LTA is gaining popularity as a method to understand stage-sequential processes in a variety of psychology and public health studies. Lanza and Collins assessed transitions in dating and sexual risk behavior in young adults (26), whereas Roberts and Ward introduced LTA for nursing research with an example that models changes in attitudinal barriers to pain management in cancer patients (27). Researchers employed LTA to study substance abuse patterns in a variety of populations, including college-aged students (28), Hispanics in the US and

Puerto Rico (29), and high-risk women (30). Relevant to the CoSMO study of older adults, Collins proposed LTA for aging research (31).

Headlong search algorithm

Badsberg first presented a headlong search algorithm for modeling contingency table data in 1992 (32). His method begins by identifying a minimal model with as few variables as possible, then allows variables to enter and exit the model with weak acceptance and rejection criteria, respectively, through an iterative process. Dean and Raftery adapted these methods to a variable selection technique in model-based clustering to choose the number of clusters and appropriate variables for clustering (17), and they later applied this methodology to the latent class setting (14).

In the headlong search algorithm, the first step is to determine the identifiable model with the greatest number of classes specified on all variables. The variable list is rank-ordered by the sums of the variances of probability across classes, because higher between-group variation in probability may indicate utility in clustering. The smallest number of variables required for a $c \geq 2$ class model functions as the initial clustering list. The user may define upper and lower limits of change in BIC (0 and -100, for example), and the algorithm begins.

> *Inclusion step:* Consider variables in the ranked variable list one at a time. Specify models with and without the new variable for a range of classes,

and choose from each based on minimum values of BIC. If the decline in BIC is below the lower limit of change, remove the variable from future consideration. If the improvement in BIC is above the upper limit of change, include the variable in the clustering list, and if between the upper and lower limit, return the variable to the end of the ranked variable list for future consideration. The inclusion step ends after testing each of the ranked variables or including one variable, whichever comes first.

Exclusion step: Consider each of the clustering variables one at a time. Compare BIC in models with and without each variable. Again, if the change is below the lower limit, remove the variable entirely from consideration. Exclude the variable and return it to the end of the ranked variable list if the change in BIC is between the lower and upper limit. Retain the variable otherwise.

Convergence of the algorithm occurs when the clustering list remains the same after successive inclusion and exclusion steps.

Dean and Raftery applied this algorithm to two simulated data sets and two real data sets. In a simulation of two data sets, they generated variables for two and three class models, some with different item response prevalence in each class (informative to LCA) and others with constant prevalence (non-informative). They found that the algorithm selected the correct variables for these simulated data, and the classification rates were similar to or better than those achieved using full-variable models.

Two real data sets provided further examples of this methodology. The Hungarian heart disease data set consists of five categorical variables thought to be relevant to the diagnosis of heart disease into categories of > 50% diameter or < 50% diameter narrowing in a major vessel. Variable selection led to a three-variable model with two classes. A clinical designation served as a comparison to this model. The correct classification rate, sensitivity, and specificity of the simpler model were similar to the five-variable model.

The HapMap project data set of single nucleotide polymorphisms (SNPs) also demonstrated performance of the algorithm. With the goal of defining ethnic/regional groups, over six hundred SNPs were available for latent class models. Over one third of these were unnecessary in the specification of a three-class model, though classification was again similar to the model with all variables.

Backwards elimination for item selection

Bartolucci, Montanari, and Pandolfi present criteria for item selection in extended LCA models (15) and in LCA multidimensional item response theory models (16). The objective of the item selection method is to reduce lengthy, validated surveys to a smaller number of categorical questionnaire items to identify latent classes. The assumption of validated surveys allows one to consider a LCA model specified on all items as representative of “true” latent classes and to assign LCA membership based on posterior probability.

First, the researcher chooses the number of latent classes with standard measures such as AIC and BIC and, perhaps, knowledge of the dimension of the latent construct of interest. This model includes all survey items, and it results in calculations of posterior probability of membership in each class for each individual. Then, the true latent class is the class with highest probability of membership for the purpose of the algorithm. Thresholds for class assignment are optional as well. For example, it may be reasonable to discard subjects with low posterior probabilities of membership across all classes, either due to atypical response patterns or for skipping many questions. The authors suggest discarding subjects with less than 0.95 maximum posterior probability.

To begin elimination of variables, consider a set of variables *A* and a candidate variable for elimination j. At the start of the algorithm, *A* is the full list of categorical variables. The algorithm begins with all items in the LCA model, eliminating one variable at a time based on one or two statistics:

1. $\mathbf{F}^{\boldsymbol{A}\setminus\mathbf{j}}$, the proportion of the sample that changes membership from the true latent class to a new class once variable j is removed from the set of existing items *A*

2. $\boldsymbol{D}^{A\backslash j} = \sum\sum \widehat{z_{\iota c}} \log \frac{\widehat{Z_{\iota c}}}{\widehat{z_{\iota c}^{A\backslash J}}}$, the Kullback-Liebler distance. $\widehat{Z_{\iota c}}$ is the posterior probability of membership in class c for subject i under the full item specification, and $\widehat{z_{\iota c}^{A\backslash J}}$ similarly but with all variables in *A* except variable j.

First, compute $\mathbf{F}^{A\backslash j}$ for every j ∈ *A*. A variable is eligible for removal if it uniquely has the lowest value of $\mathbf{F}^{A\backslash j}$ for all variables in *A*. The Kullback-Liebler distance (33) functions as a tiebreaker when two or more variables have the same values of $\mathbf{F}^{A\backslash j}$. From information theory, Kullback-Liebler is a measure of discrimination between two tests, or in this case, posterior probabilities from the model with *A* and in the model with *A* but not j. The stopping criterion is the choice of the researcher and depends on the tolerance for misclassification.

Bartolucci *et al* developed this method with the ULISSE project ("Un Link Informatico sui Servizi Sanitari Esistenti per l'Anziano", or "a computerized network on health care services for older people") in mind. The ULISSE study involves extensive surveys of elderly Italian patients in nursing homes. The goal of LCA models for ULISSE was to acquire understanding of health care status and quality of services categories among older adults in Italy. 1,744 patients responded to 75 survey questions related to health conditions in this study, covering domains such as cognitive conditions, humor and behavioral disorders, and activities of daily living. This was a cross-sectional analysis of ULISSE baseline data, though it is a longitudinal study. Five latent classes describe

ULISSE patients at baseline, and separate applications of the algorithm treat missing values as both missing at random and not missing at random.

When treating missing observations as missing at random, no classifications change in the first 6 steps of the algorithm, and the removal of the first 28 items changes the class membership for less than 1% of patients. The first 50 steps of the algorithm lead to misclassification of 3.6% of patients. Then, imposing the 0.95 minimum posterior probabilities for inclusion reduces effective sample size by 5%. In this sample, the elimination of 25 items leads to no misclassification, and 50 steps of the algorithm results in only 2.3% of patients changing class assignment. These results show that a much shorter survey has potential for a reduction in costs and time invested with minimal loss in information.

DATA AND METHODS

Extension to longitudinal data

The headlong search algorithm and the backwards elimination method are easily modified to accommodate longitudinal data. For Dean and Raftery, the first step is to find the identifiable LCA model with the greatest number of classes specified on all variables, allowing multiple observations per individual. The sums of variances of probabilities across classes *and* across time are used to rank order the variable list. Then, the algorithm proceeds as usual, with changes in BIC calculated at each step to allow variables to enter and exit the LTA model specification.

Similarly, the backward elimination method of Bartolucci *et al* is modified to accommodate longitudinal data. The two statistics calculated at each step are adapted as follows:

1. $\mathbf{F_{lta}}^{A\backslash \mathbf{j}}$, the proportion of incorrect statuses across time once variable j is removed from the set of existing items *A*.
2. $\mathbf{D_{lta}}^{\mathbf{A\backslash j}} = \sum\sum\sum \widehat{\boldsymbol{z}_{\mathbf{ict}}} \, \mathbf{log} \frac{\widehat{\boldsymbol{Z}_{\mathbf{ict}}}}{\mathbf{z}_{\mathbf{ict}}^{\widehat{\mathbf{A\backslash j}}}}$, a modified Kullback-Liebler distance. $\widehat{\mathbf{Z_{ict}}}$ is defined as the posterior probability of membership in status c for subject i at time t under the full item specification, and $\boldsymbol{z}_{ict}^{\widehat{A\backslash J}}$ takes the same definition but with all variables in *A* except variable j.

Simulations

The number of variables and item response probabilities varied in simulations, with longitudinal, dichotomous variables and latent statuses simulated for two, four, and five statuses at two time points. Over half of the simulated data contained informative variables with strong separation and simple structure (34), i.e., a different set of items uniquely defined each status and each set of informative items had dissimilar item response probabilities in other statuses. As a two-class LCA example, a binary variable j that has item response probability $\rho_{j|1} = 0.8$ for class 1 and $\rho_{j|2} = 0.1$ for class 2 would show strong separation on variable j, whereas $\rho_{j|1} = 0.5$ in class 1 and $\rho_{j|2} = 0.6$ in class 2 would be relatively weaker. In a latent class model with simple structure, each item is either non-informative, or has high probability within a single class only. Latent class models with complex structures, however, may have items with high probabilities in more than one class. Half of the simulation settings for the four- and five-status data had complex structure with the purpose of examining the performance of these methods in cases with less well-defined separation. For all settings, half of the variables were non-informative and constant across all statuses.

Additionally, sample sizes, probabilities of status membership, and transition probabilities varied. For each condition, simulated data had either small (N=200) or large (N=1000) sample sizes. In settings with high transition probabilities, 36% of subjects changed to a random status from time 1 to time 2, and 12% of subjects changed to a random status in settings with low transition probabilities. Low and high transition probability specifications contrasted how these methods perform in scenarios with an expectation that individuals will maintain their latent status designation over time versus more volatile situations with greater anticipated changes. Table 1 lists these conditions.

Table 1: Setting for simulated longitudinal data

Sample Size			200, 1000	
Status proportions			Equal, Unequal	
Probability of Transition			High, Low	
Number of Replications			200	
			Tables with Settings*	
# of statuses	**# of variables**	**Model Structure**	**Item Response Probabilities***	**Transition Probabilities**
2	10	Simple	A1	A2
4	16	Simple	2 (left)	3
4	16	Complex	2 (right)	3
5	50	Simple	A3	A4
5	50	Complex	A5	A4

*"A" tables are in the appendix

As an example, table 2 represents settings for simple and complex four-status models. These conditions were tested on samples sizes of N=200 and N=1000, and with high and low transition probabilities for each sample size. In the simple model with equal groups, each status represents 25% of observations in the sample. The first status is characterized by high probabilities of affirmative responses to items 1 and 2, the second status has high probabilities of affirmative responses to items 3 and 4, the third status to items 5 and 6, and the fourth status to items 7 and 8. Items 9 through 16 are non-informative as probabilities are constant across statuses. Item response probabilities were the same for each status in the simple model with unequal groups, with status 1 through 4 representing 5%, 10%, 25%, and 60% of the sample, respectively (not shown).

For the complex structure with unequal groups, high probabilities of affirmative responses to items 1 through 8 characterize status 1. The second status has high probabilities of affirmative responses to items 1 through 4, and third status to items 5 through 8. The fourth status has low item response probabilities for items 1 through 8. The complex structure with equal groups has the same item response probabilities as the complex structure with unequal groups of table 2, however, each status will represent 25% of the total sample (not shown).

Table 2: Four status model conditions for time 1

	Simple Structure, Equal Groups				**Complex Structure, Unequal Groups**			
	Status 1 (25%)	**Status 2 (25%)**	**Status 3 (25%)**	**Status 4 (25%)**	**Status 1 (5%)**	**Status 2 (10%)**	**Status 3 (25%)**	**Status 4 (60%)**
Item 1	0.80	0.05	0.05	0.05	0.80	0.80	0.05	0.05
Item 2	0.80	0.10	0.10	0.10	0.80	0.80	0.10	0.10
Item 3	0.05	0.80	0.05	0.05	0.80	0.80	0.05	0.05
Item 4	0.10	0.80	0.10	0.10	0.80	0.80	0.10	0.10
Item 5	0.05	0.05	0.80	0.05	0.80	0.05	0.80	0.05
Item 6	0.10	0.10	0.80	0.10	0.80	0.10	0.80	0.10
Item 7	0.05	0.05	0.05	0.80	0.80	0.05	0.80	0.05
Item 8	0.10	0.10	0.10	0.80	0.80	0.10	0.80	0.10
Item 9	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20
Item 10	0.30	0.30	0.30	0.30	0.30	0.30	0.30	0.30
Item 11	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.40
Item 12	0.50	0.50	0.50	0.50	0.50	0.50	0.50	0.50
Item 13	0.60	0.60	0.60	0.60	0.60	0.60	0.60	0.60
Item 14	0.70	0.70	0.70	0.70	0.70	0.70	0.70	0.70
Item 15	0.80	0.80	0.80	0.80	0.80	0.80	0.80	0.80
Item 16	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90

The complex structure of table 2 may be illustrative of survey tools that evaluate at-risk populations. If each item represents a question about a distinct risky behavior or characteristic, the first small status would be the highest risk group exhibiting high probabilities of most types of risky behavior. The fourth large status would be the lowest risk group. The second and third status represent individuals that share some risky behaviors with the first status, but may be at risk for different reasons (in practice, to be determined by a qualitative comparison of items 1 through 4 versus 5 through 8).

Transition probabilities varied in simulations as well. Continuing with the four-status example and presented in table 3, 64% of observations maintained status assignment from time 1 to time 2 in half of the settings and 88% maintained status assignment in the other half. In settings with unequal groups, this results in different status sizes, no longer guaranteeing that $\boldsymbol{\delta}_{s_1} = \boldsymbol{\delta}_{s_2}$.

Table 3: Transition probabilities for the four-status model

Equal Groups: Probability of Transition from Time 1 to Time 2		Time 2							
		High Transition Probability				Low Transition Probability			
		Status 1 (25%)	Status 2 (25%)	Status 3 (25%)	Status 4 (25%)	Status 1 (25%)	Status 2 (25%)	Status 3 (25%)	Status 4 (25%)
Time 1	Status 1 (25%)	64%	12%	12%	12%	88%	4%	4%	4%
	Status 2 (25%)	12%	64%	12%	12%	4%	88%	4%	4%
	Status 3 (25%)	12%	12%	64%	12%	4%	4%	88%	4%
	Status 4 (25%)	12%	12%	12%	64%	4%	4%	4%	88%

Unequal Groups: Probability of Transition from Time 1 to Time 2		Time 2							
		High Transition Probability				Low Transition Probability			
		Status 1 (14.6%)	Status 2 (17.2%)	Status 3 (25.0%)	Status 4 (43.2%)	Status 1 (8.2%)	Status 2 (12.4%)	Status 3 (25.0%)	Status 4 (54.4%)
Time 1	Status 1 (5%)	64%	12%	12%	12%	88%	4%	4%	4%
	Status 2 (10%)	12%	64%	12%	12%	4%	88%	4%	4%
	Status 3 (25%)	12%	12%	64%	12%	4%	4%	88%	4%
	Status 4 (60%)	12%	12%	12%	64%	4%	4%	4%	88%

Additional tables with item response probabilities and transition probabilities are listed in the appendix.

Evaluating Variable Selection Techniques

Both variable selection techniques determine the classifying variables as well as the number of statuses in the LTA model. As a first step, the percent of simulations that correctly select all the status-informative variables was calculated. Furthermore, the percent of simulations that select no non-informative variables was considered, as acceptable classification may be possible with fewer variables. These percentages were compared for the two techniques.

Minimal detectable differences in proportions by the Likelihood Ratio chi-square test are provided for 50%, 80%, and 90% power in table 4, for a range of potential proportions for the inferior variable selection technique. These values assume Type I error rate of $\alpha=0.05$ and N=200 simulations. At 80% power, the minimal detectable difference between methods ranges from a 13.9% difference if the inferior technique selects correct variables in 50% of simulations (a 28% improvement) up to 7.0% if the lower proportion method is 90% (an 8% improvement). These differences are applicable to each setting from table 1; smaller differences are detectable when combining settings to examine small versus large sample size, high versus low transition probability, simple versus complex structure, equal versus unequal status sizes, and the number of latent statuses.

Table 4: Power to detect differences in variable selection techniques

Method 1	**Method 2**	**Difference**	**Power**
50%	59.8%	9.8%	50%
50%	63.9%	**13.9%**	80%
50%	66.0%	16.0%	90%
60%	69.4%	9.4%	50%
60%	73.2%	**13.2%**	80%
60%	75.2%	15.2%	90%
70%	78.6%	8.6%	50%
70%	82.0%	**12.0%**	80%
70%	83.7%	13.7%	90%
80%	87.3%	7.3%	50%
80%	90.0%	**10.0%**	80%
80%	91.3%	11.3%	90%
90%	95.2%	5.2%	50%
90%	97.0%	**7.0%**	80%
90%	97.8%	7.8%	90%

Next, the issue of label switching in LCA and LTA is considered (35). As class labels are arbitrary for each of the c = 1,…,C classes, class 1 from one LCA model may be labeled as class 2 in another, for example. Labels may switch for different starting values of the EM algorithm, though the resulting model parameters and posterior probabilities will result as identical upon a reordering of the labels. Additionally, differences in the variables selected to specify each model can lead to different class labels irrespective of starting values of the EM algorithm. For each variable selection technique, the class labels were reordered to match the labels assigned in simulating the longitudinal data. An example in table 5 shows two LCA results that have identical accuracy in detecting latent class definitions but are subject to label-switching. Class prevalence and item response probabilities are simply reversed and should be reordered prior to evaluating concordance of subjects' assignment by the two models.

Table 5: Label switching example

Model 1			Model 2		
	Class 1 **(70%)**	**Class 2** **(30%)**		**Class 1** **(30%)**	**Class 2** **(70%)**
Item 1	0.70	0.10	**Item 1**	0.10	0.70
Item 2	0.80	0.20	**Item 2**	0.20	0.80
Item 3	0.90	0.30	**Item 3**	0.30	0.90
Item 4	0.15	0.80	**Item 4**	0.80	0.15
Item 5	0.25	0.80	**Item 5**	0.80	0.25
Item 6	0.10	0.10	**Item 6**	0.10	0.10
Item 7	0.30	0.30	**Item 7**	0.30	0.30
Item 8	0.50	0.50	**Item 8**	0.50	0.50
Item 9	0.70	0.70	**Item 9**	0.70	0.70
Item 10	0.90	0.90	**Item 10**	0.90	0.90

*shaded rows indicate variables included in LCA specification

In practice, an examination of item response probabilities for two or more LCA models reveals a solution to the label switching problem. Tuller, Drotar, and Lubke present a method for automating label switching for simulation studies (35). After first arranging true latent class (from simulations) and predicted latent class (from posterior probability assignment) into a matrix for each resulting technique separately, rearrangement of the predicted latent classes to form the strongest diagonal allows for comparisons of each of the two techniques to known class assignment. A class assignment criterion determines if this rearrangement is warranted and not attributable to chance alone. $\mathbf{CA_{crit}}$ is set as a value such that:

$$\frac{\mathbf{1}}{\mathbf{C}} < \mathbf{CA_{crit}} < \mathbf{1},$$

representing a value between random allocation to the C classes ($\frac{\mathbf{1}}{\mathbf{C}} \times \mathbf{100}\%$) and perfect concordance (100%). $\mathbf{CA_{crit}}$ is a minimum value to be compared to the proportion correctly classified per class, a requirement to ensure classification is better than chance alone. Tuller, Drotar, and Lubke suggests a 20% improvement in classification over random allocation in choosing $\mathbf{CA_{crit}}$. For example, in an LCA model with 5 classes, each diagonal element would need to represent at least $\mathbf{CA_{crit}} = \frac{\mathbf{1.2}}{\mathbf{5}} = \mathbf{24}$% of the true latent class, a 20% improvement over random allocation. Less than a 20% improvement in classification over random allocation may represent instances of a spurious correction to labels or poor assignment in general.

For each simulation, the optimal arrangement in terms of correctly classifying observations into statuses was selected and overall correct classification rates computed. Considering a lower reference method proportion at the minimum $\mathbf{CA_{crit}}$ level and select values of improvement over random allocation, minimum differences required for 50%, 80%, and 90% power are provided in table 6. This again assumes 200 simulations per setting and α=0.05, with greater power to detect differences in techniques when combining settings for main effect comparisons. The correct classification rates are presented overall and for each time point separately.

Table 6: Power to detect differences in correct classification

Statuses	% improvement over random allocation*	Reference Method: % Correctly Classified	Comparison Method: % Correctly Classified	Power to detect
2	20%	60%	66.0%	50%
			69.4%	80%
			73.2%	90%
	50%	75%	83.0%	50%
			86.1%	80%
			87.6%	90%
	80%	90%	95.2%	50%
			97.0%	80%
			97.8%	90%
4	20%	30%	39.4%	50%
			43.5%	80%
			45.7%	90%
	100%	40%	49.8%	50%
			54.0%	80%
			56.2%	90%
	200%	75%	83.0%	50%
			86.1%	80%
			87.6%	90%
5	20%	24%	32.9%	50%
			36.9%	80%
			39.0%	90%
	100%	40%	49.8%	50%
			54.0%	80%
			56.2%	90%
	400%	80%	87.3%	50%
			90.0%	80%
			91.3%	90%

*correct classification rates of 50%, 25%, and 20% for 2, 4, and 5-status models respectively

Finally, cross-tabulations of true versus estimated status membership were constructed along with Cohen's kappa as a measure of agreement (36). The adjusted Rand index is an additional measure of the similarity between two classifications (37). The adjusted Rand index ranges between -1 and 1, with -1 representing perfect discordance and 1 representing perfect concordance.

In addition to comparisons at each individual setting, results from simulations were combined to consider each main effect in isolation: small versus large sample size, high versus low transition probability, equal versus unequal status prevalence, and simple versus complex model structure. Smaller differences in techniques among these main effects are detectable when compared to individual simulation settings in tables 4 and 6. Table 7 provides the number of simulations for each of these main effects.

Table 7: Number of simulations for each setting

Sample Size	
Small	4,000
Large	4,000
Structure	
Simple	4,800
Complex	3,200
Status Prevalence	
Equal	4,000
Unequal	4,000
Transition Probability	
Low	4,000
High	4,000
Number of Statuses	
2	1,600
4	3,200
5	3,200

Simulations were implemented and evaluated through a series of SAS macros. A labelswitching macro, invoked in both variable selection techniques, optimized variable labels for any given “test” and “true” class or status assignment. Another macro determined the appropriate number of latent statuses for a given variable set through BIC comparisons, also invoked in both variable selection techniques. Finally, specific to the headlong search algorithm, macros for inclusion and exclusion steps test variable lists for potential variables to drop from models, add to models, or remove from future model consideration.

Master macro programs for each variable selection technique included call parameters for the number of statuses as well as all relevant setting information (status proportion, transition probability, structure, and sample size). Each macro began with a step to create new simulated LTA data in PROC IML. Then, the latent transition analysis procedure PROC LTA specified models based on all variables, and only status-informative variables. For headlong search, the ranked variable list was derived from the model with the greatest number of latent statuses, and inclusion and exclusion steps were repeated until the algorithm converged. For backwards elimination, variables were considered for removal one at a time, and classification rates were compared between the full and simpler model. Variables were removed one by one until the classification between full and current model fell below a given threshold. Once variable selection terminated for a given iteration, PROC LTA specified a final model based on the variable subset chosen. The labelswitching macro was invoked for model comparisons. Appended SAS datasets tracked classification statistics, variable selection results, and algorithm convergence at each iteration, and at each step within iteration. Both algorithms were coded and implemented in SAS version 9.3 (SAS Institute, Cary, NC) and PROC LTA Version 1.3.0.

The Cohort Study of Medication Adherence in Older Adults (CoSMO)

The Cohort Study of Medication Adherence in Older Adults (CoSMO) is a prospective study of the barriers to antihypertensive medication adherence in older adults; the design of the study and baseline characteristics were previously described (38). In CoSMO, adults ages 65 and older with essential hypertension were chosen at random from a roster of a large managed-care organization with a demographically diverse population in southeastern Louisiana. Trained surveyors contacted those selected to participate between August 21st, 2006 and September 30th, 2007, and those eligible who provided verbal informed consent were enrolled. The Ochsner Clinic Foundation's Institutional Review Board and the privacy board of the managed care organization approved CoSMO.

Participant surveys, the electronic medical record, and the administrative databases of the managed-care organization were the main data sources for measures such as demographics, blood pressure, and medication adherence. The survey included assessment of socio-demographic factors, clinical factors, health care system factors, antihypertensive medication treatment-related variables, and adherence to antihypertensive medication. The 8-item Moriskey Medication Adherence Scale was used to assess medication adherence. Scores on this self-report scale range from zero to eight, with MMAS-8 scores of <6, 6 to <8, and 8 reflecting low, medium and high adherence, respectively (39). A previous study showed that a two-point change in MMAS-8 may represent real change in antihypertensive medication adherence among CoSMO participants (40). Additionally, pharmacy fill data sets function as an objective measure of adherence by means of the Proportion of Days Covered (PDC) (41).

Factors such as gender, depressive symptoms, life events, marital status, and use of calcium channel blockers were associated with a decline of two or more points in MMAS-8 score in a pooled logistic regression model (42). A previous cross-sectional study identified a 4-item self-report tool to identify low adherers among those with uncontrolled blood pressure (43). CoSMO participants completed the RAND Medical Outcomes Study 36-item tool as a measure of quality of life (44). For this study, latent transition analysis provides a new longitudinal perspective to mental and physical quality of life in this cohort of older adults. Levels of adherence as exhibited in subsets of participants by quality of life status may aid in identification of at-risk patients.

Separate latent transition analysis models for CoSMO were defined by applying each algorithm to the quality of life survey. Posterior probabilities of latent status membership were calculated, and relevant demographic and clinical characteristics were profiled for each baseline status. Concordance in status assignments was compared along with a qualitative assessment of status similarity. The percent with low adherence as measured by self-report and pharmacy fill measures was calculated for each status in each model.

RESULTS

Two Status Simulations

Algorithms converged to final variable lists in all simulations with two statuses, in a total of sixteen hundred simulations for eight settings. Table 8 lists the percentages of simulations selecting *all* status-informative variables and selecting *only* status-informative variables (i.e., no non-informative variables). None of the headlong search simulations resulted in selection of all status-informative variables. Four of the five status-informative variables were retained in at least 74% of headlong search simulations for each setting. Across all headlong search settings, four correct variables were retained in 82.4% of settings, three correct variables were retained in 16.4% of settings, and two correct variables were retained in 1.2% of settings. Backwards elimination also resulted in nearly 0% of simulations selecting all status-informative variables in settings with low transition probability. For small sample sizes and high transition probability, 41.0% and 19.5% of simulations selected all status-informative variables for settings with equal and unequal status proportions, respectively. These percentages increased to 84.5% and 37.5% for equal and unequal status proportions, respectively, at a large sample size.

At least four variables were required for identifiable models. As such, non-informative variables were retained in cases where only two or three status-informative variables were selected in headlong search simulations, 17.6% of simulations across all settings. Headlong search was more likely to select only status-informative variables at a smaller sample size, 87.9% at N=200 observations versus 77.0% at N=1000 observations. Backwards elimination simulations, however, removed all non-informative variables in nearly all settings.

Table 8: Two-status variable selection statistics

Sample Size	Status proportions	Transition Probability	HS^: % select all status-informative variables	BE^^: % select all status-informative variables	HS^: % select only status-informative variables	BE^^: % select only status-informative variables
N=200	Equal	High	0.0%	41.0%	87.0%	100.0%
		Low	0.0%	0.0%	86.0%	100.0%
	Unequal	High	0.0%	19.5%	87.5%	99.5%
		Low	0.0%	0.5%	91.0%	98.5%
N=1000	Equal	High	0.0%	84.5%	77.5%	100.0%
		Low	0.0%	0.0%	81.0%	100.0%
	Unequal	High	0.0%	37.5%	74.0%	100.0%
		Low	0.0%	0.0%	75.5%	100.0%

^HS: headlong search algorithm of Dean and Raftery
^^BE: backward elimination method of Bartolucci, Montanari, and Pandolfi

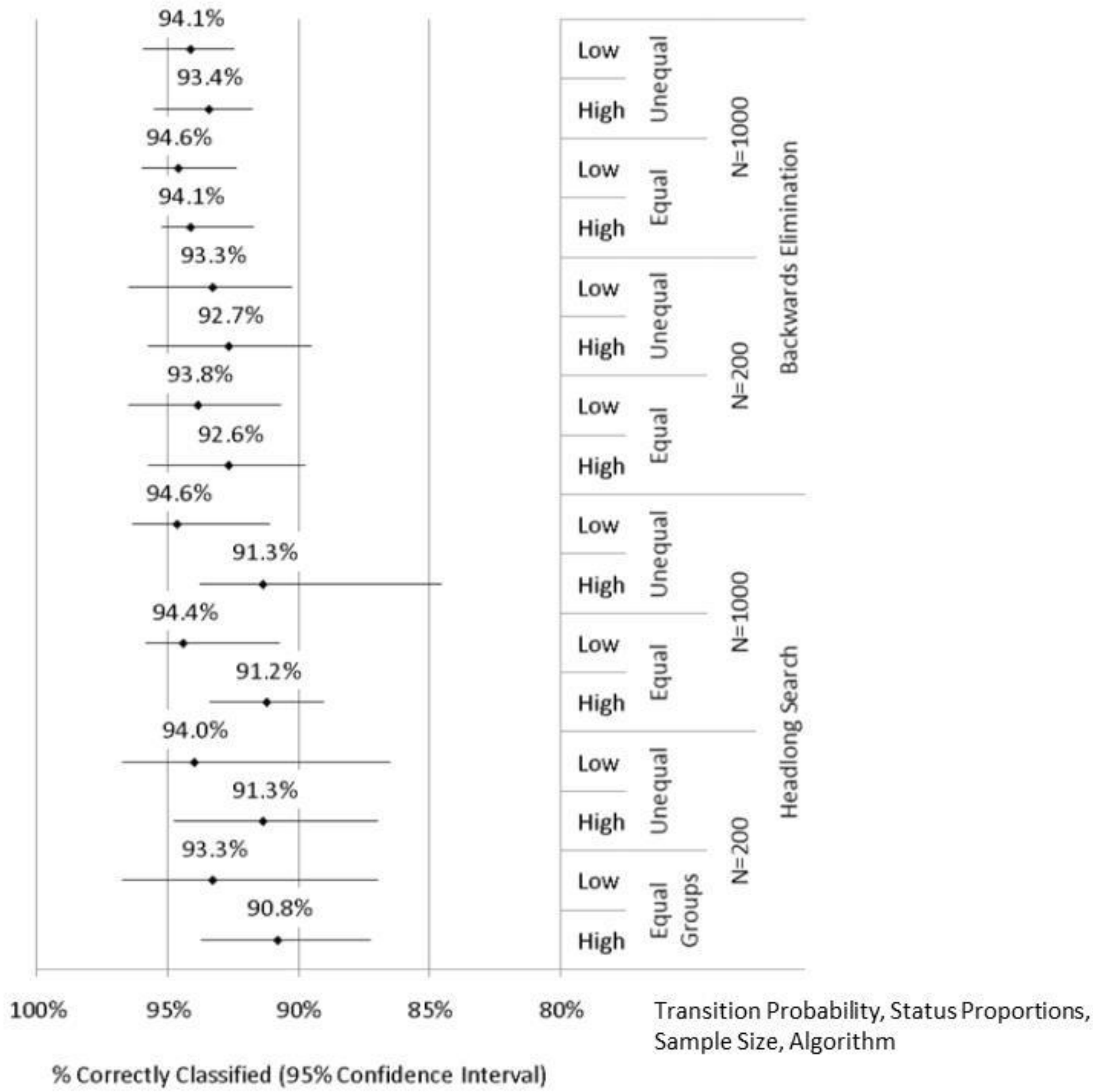
Two statuses were correctly identified as optimal by BIC in all simulations for both algorithms. Figure 1 contains the percent correctly classified across both time periods and 95% empirical confidence intervals for two-status simulation settings. While some trends are evident, all confidence intervals overlap within and across variable selection algorithms.

Correct classification rates for the headlong search algorithm overall ranged from 89.8% (in settings with small sample size, unequal status proportions, and high transition probabilities) to 94.6% (in settings with large sample size, unequal status proportions, and low transition probabilities). Among the headlong search settings, those with low transition probabilities had correct classification rates 3%-4% higher than those with high transition probabilities, regardless of sample size or status proportion.

Correct classification rates were similar or slightly higher for the backwards elimination algorithm and overall ranged from 92.6% (for small sample sizes, equal status proportions, and high transition probabilities) to 94.6% (at large sample sizes, equal status proportions, and low transition probabilities).

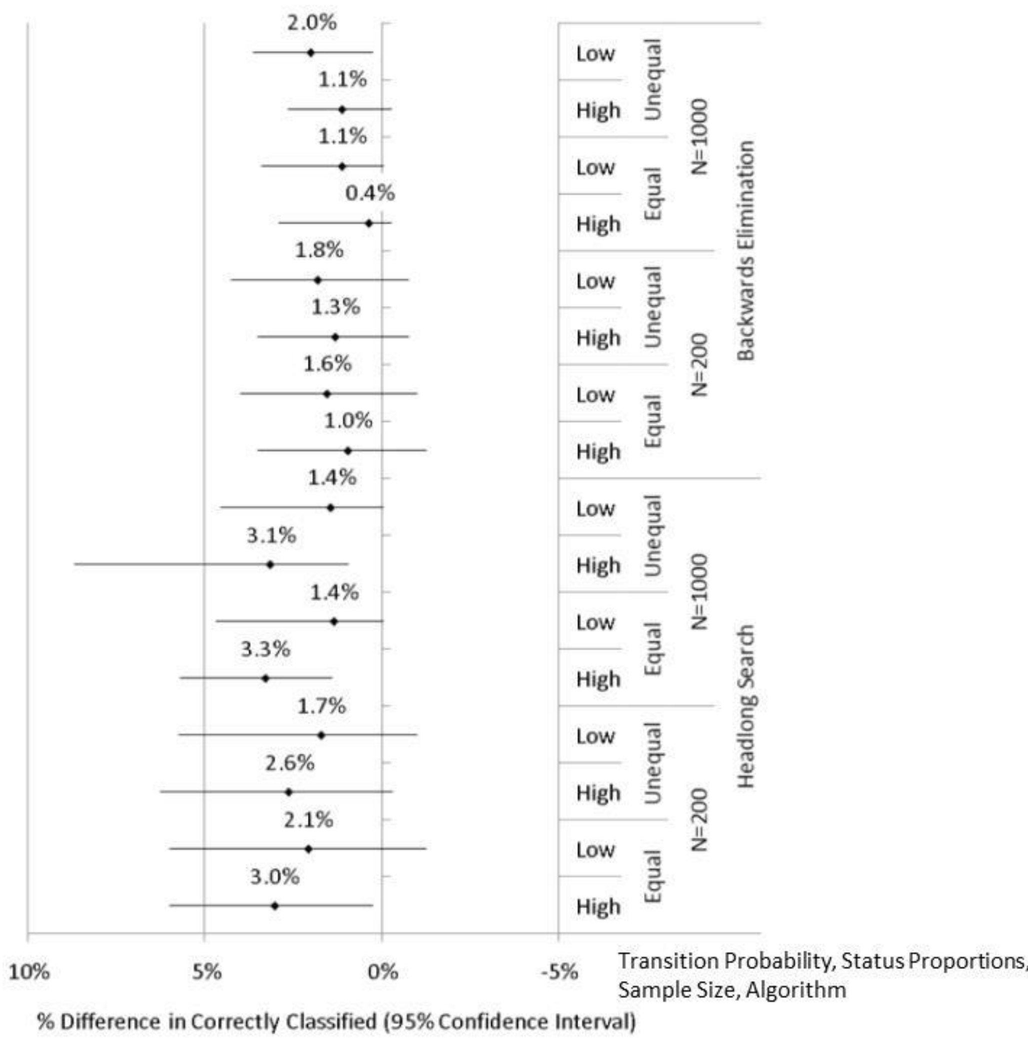
The confidence intervals of correct classification percentages were generally 1.5-2 times larger for small sample sizes, holding status proportion and transition probability constant within algorithm. For settings with high transition probability, the backwards search algorithm had a correct classification rate approximately two percentage points higher than the headlong search algorithm for each combination of sample size and status proportion. Correct classification rates, 95% confidence intervals, and directional trends were similar for time 1 and time 2 when considered separately (data not shown).

Figure 1: Percent correctly classified in two-status simulations by setting and variable selection method



To compare algorithm-based variable selection to full scale utilization, status assignment based on final selected variables was compared to status assignment based on full (ten) variables. After each iteration of each simulation, the correct classification proportion after variable selection was subtracted from correct classification proportion prior to variable selection. These differences were averaged across simulation setting and are shown in figure 2. Utilizing fewer, mostly informative variables resulted in a difference of 2% or less across all settings of the backwards elimination method and across low transition probability settings of the headlong search algorithm. Models from high transition probability settings of the headlong search algorithm had slightly higher differences in correct classification rates on selected variable specifications than on all variable specifications.

Figure 2: Differences between percent correctly classified with all variables and algorithm-selected variables for two-status simulations



Similarly, status assignment based on final variables selected was compared to status assignment based on *only* status-informative (five) variables, as misclassification due to assignment by maximum posterior probabilities may occur even when using only the appropriate classification variables. The differences between correct classification rates from status-informative variable models and final selected variable models were averaged across simulation settings and are displayed in figure 3. These differences are nearly identical to the differences in figure 2. In general, the simpler models, specified on a few informative variables, provided similar correct classification rates to models specified on all variables including non-informative variables, as well models using only status-informative variables.

Figure 3: Differences between percent correctly classified with status-informative variables and algorithm-selected variables for two-status simulations

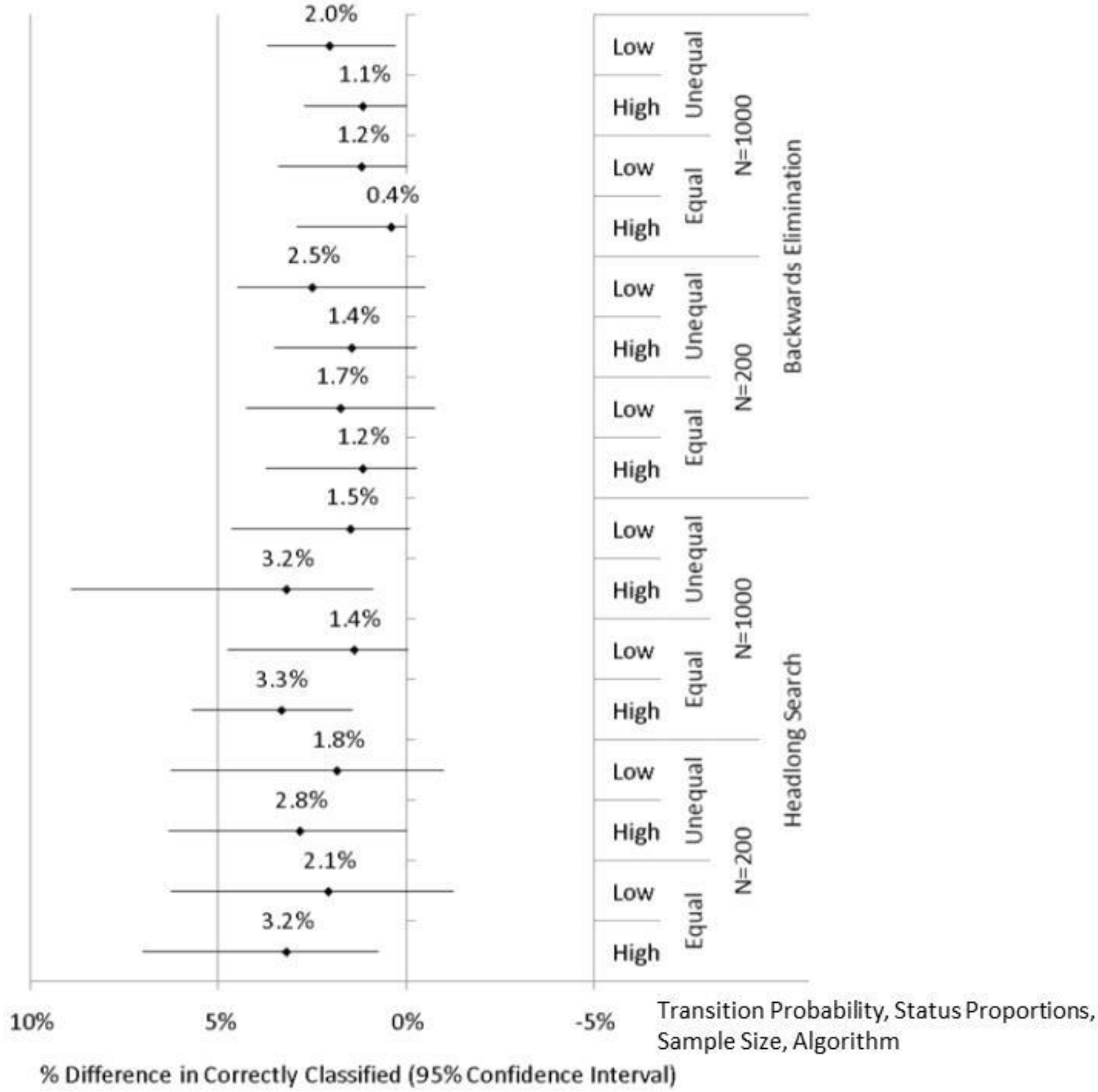


Table 9 contains correct classification rates comparing simulated and predicted status for two-status models aggregated across both time periods. Results were similar for each time period separately (data not shown). Status 1, representing 50% of observations for simulations with equal status proportions and > 50% for other settings (55.6%-70%, depending on time period and transition probability), attained similar or higher percentages correctly classified compared to status 2.

Table 9: Percent correctly classified for each status in two-status simulations

			Headlong Search^	
Sample Size	Status proportions	Transition Probability	Status 1 (95% CI)	Status 2 (95% CI)
N=200	Equal	High	90.9% (82.1%, 96.8%)	90.5% (81.7%, 97.5%)
		Low	93.0% (77.6%, 98.0%)	93.5% (84.9%, 98.5%)
	Unequal	High	93.6% (84.3%, 98.5%)	87.5% (76.1%, 95.8%)
		Low	95.4% (85.3%, 98.9%)	90.9% (81.5%, 96.8%)
N=1000	Equal	High	91.0% (86.6%, 94.6%)	91.3% (85.0%, 95.0%)
		Low	94.3% (88.7%, 97.0%)	94.4% (89.6%, 96.6%)
	Unequal	High	93.8% (87.0%, 97.2%)	87.1% (77.7%, 92.2%)
		Low	96.3% (92.9%, 98.0%)	91.0% (82.9%, 94.8%)
			Backward Elimination^^	
			Status 1 (95% CI)	Status 2 (95% CI)
N=200	Equal	High	92.8% (85.4%, 97.2%)	92.4% (84.2%, 97.1%)
		Low	94.0% (88.3%, 97.9%)	93.5% (87.6%, 97.8%)
	Unequal	High	94.8% (88.5%, 98.5%)	88.9% (79.0%, 96.6%)
		Low	95.2% (90.1%, 98.9%)	89.1% (78.0%, 97.1%)
N=1000	Equal	High	94.1% (90.0%, 95.7%)	94.2% (90.2%, 95.6%)
		Low	94.4% (90.2%, 96.6%)	94.7% (91.7%, 96.7%)
	Unequal	High	95.4% (93.2%, 97.4%)	90.0% (85.0%, 95.0%)
		Low	96.5% (94.6%, 97.7%)	89.1% (82.7%, 94.1%)

^headlong search algorithm of Dean and Raftery
^^backward elimination method of Bartolucci, Montanari, and Pandolfi
CI – confidence interval

As indicated by high correct classification rates, kappa and adjusted Rand statistics were high across all simulations (see table 10). While confidence intervals again overlapped across all settings, the backwards elimination agreement statistics were mostly equivalent or slightly higher than those of the headlong search algorithm.

Table 10: Kappa and adjusted Rand statistics for two-status simulations

			Headlong Search^	
Sample Size	Status Proportions	Transition Probability	Kappa (95% CI)	Adjusted Rand (95% CI)
N=200	Equal	High	0.81 (0.74, 0.87)	0.67 (0.55, 0.77)
		Low	0.86 (0.74, 0.93)	0.76 (0.55, 0.87)
	Unequal	High	0.78 (0.03, 0.89)	0.66 (0.00, 0.81)
		Low	0.86 (0.70, 0.93)	0.78 (0.53, 0.87)
N=1000	Equal	High	0.82 (0.78, 0.87)	0.69 (0.61, 0.75)
		Low	0.89 (0.81, 0.92)	0.79 (0.66, 0.84)
	Unequal	High	0.81 (0.67, 0.86)	0.69 (0.47, 0.76)
		Low	0.88 (0.80, 0.92)	0.79 (0.67, 0.86)
			Backwards Elimination^^	
Sample Size	Status Proportions	Transition Probability	Kappa (95% CI)	Adjusted Rand (95% CI)
N=200	Equal	High	0.85 (0.80, 0.91)	0.73 (0.64, 0.84)
		Low	0.88 (0.81, 0.93)	0.77 (0.66, 0.86)
	Unequal	High	0.84 (0.77, 0.91)	0.73 (0.62, 0.84)
		Low	0.85 (0.78, 0.92)	0.76 (0.64, 0.86)
N=1000	Equal	High	0.88 (0.83, 0.90)	0.78 (0.70, 0.82)
		Low	0.89 (0.85, 0.92)	0.80 (0.72, 0.85)
	Unequal	High	0.86 (0.82, 0.90)	0.75 (0.70, 0.83)
		Low	0.86 (0.82, 0.91)	0.77 (0.72, 0.84)

^headlong search algorithm of Dean and Raftery
^^backward elimination method of Bartolucci, Montanari, and Pandolfi
CI – confidence interval

Four Status Simulations

The performance of variable selection algorithms was less stable in some simulation settings with four statuses. As with two-status simulations, the headlong search algorithm omitted at least one of the eight status-informative variables in nearly every simulation and setting (see table 11). The headlong search algorithm successfully limited to only status-informative variables for most settings, however; ≥98% for all settings with complex structure as well as simple structure settings at large sample sizes.

Table 11: Four-status variable selection statistics

Structure	Sample Size	Status proportions	Transition Probability	HS^: % select all status-informative variables	BE^^: % select all status-informative variables	HS^: % select only status-informative variables	BE^^: % select only status-informative variables
Simple	N=200	Equal	High	0.0%	87.0%	96.5%	3.0%
			Low	0.5%	64.0%	97.5%	3.5%
		Unequal	High	0.0%	16.0%	82.5%	77.5%
			Low	0.0%	6.5%	69.0%	92.0%
	N=1000	Equal	High	0.0%	12.5%	100.0%	63.0%
			Low	0.0%	4.5%	100.0%	66.0%
		Unequal	High	0.0%	85.0%	99.0%	1.0%
			Low	0.0%	76.5%	98.0%	11.5%
Complex	N=200	Equal	High	0.0%	14.5%	98.5%	95.5%
			Low	0.0%	17.0%	100.0%	96.0%
		Unequal	High	0.0%	11.0%	98.5%	84.0%
			Low	0.0%	2.0%	100.0%	86.0%
	N=1000	Equal	High	0.0%	4.0%	100.0%	18.0%
			Low	0.0%	11.0%	100.0%	28.0%
		Unequal	High	0.0%	28.5%	100.0%	52.0%
			Low	0.0%	19.0%	100.0%	69.5%

^HS: headlong search algorithm of Dean and Raftery
^^Backward Elimination method of Bartolucci, Montanari, and Pandolfi

There was greater variability in both the percent selecting all status-informative variables and the percent selecting only status-informative variables for the backwards search algorithm. This is largely due to incorrect identification of the number of statuses (by BIC) from the full list of sixteen variables. Table 12 shows that for small sample sizes, only those settings with equal status proportions, or N=50 per status, successfully detected all four statuses. Other small sample size settings, which contained statuses between 5%-17.2% of sample, mostly simplified to two or three status models and likewise had low values for the percent selecting all status-informative variables. At a larger sample size, the correct number of statuses was more likely to be identified based on full-variable specification. Across all complex settings of large sample size, 88.6% correctly identified four status models (84.5%-90.5% across different status proportions and transition probabilities). For simple structure settings, 99.3% of simulations with equal status proportions and 60.0% with unequal proportions correctly identified four status models.

Table 12: Number of statuses identified in four-status variable selection settings

Headlong Search^								
				Final Number of Statuses Identified (% of Simulations)				
Structure	Sample Size	Status Proportions	Transition Probability	2	3	4	5	6
Simple	N=200	Equal	High	68.0%	28.5%	3.5%	0.0%	0.0%
			Low	49.5%	32.5%	18.0%	0.0%	0.0%
		Unequal	High	73.5%	26.5%	0.0%	0.0%	0.0%
			Low	73.5%	26.5%	0.0%	0.0%	0.0%
	N=1000	Equal	High	0.5%	80.0%	14.5%	4.0%	1.0%
			Low	0.0%	68.5%	20.0%	11.0%	0.5%
		Unequal	High	13.0%	84.0%	3.0%	0.0%	0.0%
			Low	9.0%	86.0%	4.5%	0.5%	0.0%
Complex	N=200	Equal	High	79.0%	21.0%	0.0%	0.0%	0.0%
			Low	80.0%	20.0%	0.0%	0.0%	0.0%
		Unequal	High	85.5%	14.5%	0.0%	0.0%	0.0%
			Low	83.5%	16.0%	0.5%	0.0%	0.0%
	N=1000	Equal	High	88.0%	9.0%	0.0%	3.0%	0.0%
			Low	13.0%	48.0%	27.0%	12.0%	0.0%
		Unequal	High	73.0%	27.0%	0.0%	0.0%	0.0%
			Low	24.5%	48.0%	26.0%	1.5%	0.0%

Table 12 (continued): Number of statuses identified in four-status variable selection settings

Backwards Elimination^^								
				Final Number of Statuses Identified (% of Simulations)				
Structure	Sample Size	Status proportions	Transition Probability	2	3	4	5	6
Simple	N=200	Equal	High	0.0%	17.0%	82.5%	0.5%	0.0%
			Low	0.0%	17.0%	81.5%	1.5%	0.0%
		Unequal	High	83.0%	16.0%	1.0%	0.0%	0.0%
			Low	91.0%	7.5%	1.5%	0.0%	0.0%
	N=1000	Equal	High	0.0%	0.0%	100.0%	0.0%	0.0%
			Low	0.0%	0.0%	98.5%	0.5%	1.0%
		Unequal	High	0.0%	9.5%	73.0%	9.0%	8.5%
			Low	0.0%	22.5%	47.0%	18.0%	12.5%
Complex	N=200	Equal	High	0.0%	99.5%	0.5%	0.0%	0.0%
			Low	0.0%	100.0%	0.0%	0.0%	0.0%
		Unequal	High	81.0%	19.0%	0.0%	0.0%	0.0%
			Low	61.0%	38.5%	0.5%	0.0%	0.0%
	N=1000	Equal	High	0.0%	9.5%	89.0%	1.5%	0.0%
			Low	0.0%	0.5%	90.5%	9.0%	0.0%
		Unequal	High	0.0%	0.5%	90.5%	9.0%	0.0%
			Low	0.0%	6.5%	84.5%	9.0%	0.0%

^headlong search algorithm of Dean and Raftery
^^backward elimination method of Bartolucci, Montanari, and Pandolfi

The adjusted Rand statistic serves as the primary measure of agreement as many settings failed to identify the correct number of statuses in the final specifications and kappa could not be calculated. In table 13, the effect of misidentified variables and/or incorrect status number is evident in lower statistics than previously exhibited by two-status models. At large sample sizes and equal status proportions, the backwards elimination methods perform significantly better than headlong search. Combined across structure type and transition probability, the adjusted Rand (95% confidence interval) for backwards elimination is 0.808 (0.748, 0.870), significantly higher than that of headlong search 0.485 (0.377, 0.654). Most other adjusted Rand statistics are equivalent or better in the backwards elimination method within and across settings, though not significantly so.

Table 13: Adjusted Rand statistics for four-status simulations

	Sample Size	Status proportions	Transition Probability	HS: Adjusted Rand (95% CI)	BE: Adjusted Rand (95% CI)
Simple Structure	N=200	Equal	High	0.30 (0.17, 0.52)	0.66 (0.28, 0.85)
			Low	0.41 (0.19, 0.75)	0.73 (0.33, 0.90)
		Unequal	High	0.50 (0.00, 0.72)	0.54 (0.33, 0.75)
			Low	0.64 (0.00, 0.82)	0.64 (0.52, 0.86)
	N=1000	Equal	High	0.45 (0.23, 0.59)	0.80 (0.76, 0.85)*
			Low	0.54 (0.26, 0.75)	0.85 (0.81, 0.89)*
		Unequal	High	0.61 (0.18, 0.75)	0.61 (0.35, 0.84)
			Low	0.74 (0.37, 0.84)	0.66 (0.36, 0.90)
Complex Structure	N=200	Equal	High	0.42 (0.29, 0.54)	0.59 (0.50, 0.66)
			Low	0.47 (0.41, 0.59)	0.62 (0.53, 0.68)
		Unequal	High	0.42 (0.27, 0.67)	0.41 (0.30, 0.70)
			Low	0.43 (0.23, 0.78)	0.52 (0.28, 0.84)
	N=1000	Equal	High	0.43 (0.39, 0.57)	0.76 (0.57, 0.81)*
			Low	0.52 (0.37, 0.69)	0.83 (0.77, 0.88)*
		Unequal	High	0.45 (0.27, 0.66)	0.73 (0.45, 0.82)
			Low	0.56 (0.23, 0.78)	0.80 (0.52, 0.87)

^HS: headlong search algorithm of Dean and Raftery
^^BE: backward elimination method of Bartolucci, Montanari, and Pandolfi
*BE adjusted Rand statistic significantly higher than HS

Five Status Simulations

Five status simulations demonstrated limited success. With headlong search, only four variables were retained in most simulations across all settings. The four selected were status-identifying variables in nearly 100% of simulations, though only two- or three-status models were identified by BIC based on these few variables.

For backwards elimination, only simulations with simple structure and large sample size correctly identified five-status models based on all 50 variables, regardless of transition probability or status proportions (91.4% of these simulations). BIC was optimized for three- or four-status models at all other settings. In 57.0% of all backwards elimination simulations, no non-informative variables were eliminated, and in nearly 0% of simulations were only status-informative variables retained. More non-informative variables were eliminated for larger sample sizes.

Application to CoSMO

The Cohort Study of Medication Adherence in Older Adults (CoSMO) included data collection from a wide range of sources including many surveys with categorical predictors suitable for latent transition analysis. A total of 2,194 CoSMO participants were recruited and completed the baseline survey between August 2006 and September 2007. Follow-up surveys were completed one and two years after the baseline survey. The 1,956 patients who completed the 2nd follow-up survey, subsequently referred to only as the "follow-up" survey, serve as the analytic sample for the CoSMO application.

Patient characteristics including demographics, social determinants, clinical, behavioral, and healthcare variables at baseline and at follow-up are presented in table 14. Questions related to demographics, hypertension duration, clinic visits, reduction of medication due to cost, smoking, alcohol use, lifestyle modifications, and marital status were obtained from survey questions. Medical claims data were used to obtain patients' history of stroke, diabetes, myocardial infarction, and heart failure. Classes of antihypertensive medication were determined from pharmacy fill records. A score of ≥ 16 using the Center for Epidemiological Studies Depression Scale qualified participants as having depressive symptoms(45). Health food, herbal supplements, or relaxation techniques to improve blood pressure control were considered uses of Complementary and Alternative medicine (46). Healthcare satisfaction variables were dichotomized with those reporting poor or fair satisfaction as "low" (47). The bottom tertile of mental and physical quality of life was defined as "low" for the RAND Medical Outcomes Study 36-item tool (48). Low social support was defined by the lowest tertile of the RAND Medical Outcomes Study Social Support survey (49). The John Henry Active Coping scale score was dichotomized with scores below the median considered "low" (50). A median cut point was also applied to the Perceived Stress Scale (51).

At baseline, 47.9% of CoSMO participants were ≥ 75 years old, 58.2% were female, 29.9% were black, and 80% had a high school education or greater. Over 60% had been diagnosed with hypertension ≥ 10 years and over half had ≥ 4 clinic visits per year. 31.7% and 32.8% had low baseline physical and mental quality of life, respectively; the percent with low physical quality of life increased to 38.7% at follow-up while the percent with low mental quality of life decreased slightly to 28.1%.

*Table 14: Characteristics of CoSMO participants at baseline and follow-up**

Participant Characteristics, n (%)	Baseline	Follow-Up
Demographics		
Age ≥ 75 years	937 (47.9)	1,198 (61.3)
Female	1,139 (58.2)	--
Black race	584 (29.9)	--
High school education or greater	1,564 (80.0)	--
Clinical Variables		
Hypertension duration ≥ 10 Years	1,209 (62.0)	--
Body mass index: ≥ 30 kg/m^2	1,498 (76.7)	1,426 (74.5)
History of Stroke	227 (11.6)	464 (23.7)
History of Diabetes	797 (40.8)	995 (50.9)
History of Myocardial Infarction	211 (10.8)	479 (24.5)
History of Congestive Heart Failure	137 (7.0)	609 (31.1)
Healthcare Variables		
≥4 clinic visits per year	1,038 (53.2)	1,238 (63.4)
3+ classes of antihypertensive medication^	830 (43.2)	817 (48.7)
Reduced medication due to cost	65 (3.3)	47 (2.4)
Uses Complementary and Alternative Medicine	524 (26.8)	519 (26.5)
Low Overall Satisfaction with Healthcare	90 (4.6)	91 (4.8)
Low Satisfaction with Access	82 (4.2)	76 (4.0)
Low Satisfaction with Communication	199 (10.2)	157 (8.3)
Behavioral Variables		
Never a smoker	957 (49.3)	954 (49.6)
<2 alcoholic drinks per week	1,529 (78.5)	1,503 (78.7)
Increasing fruits and vegetables	1,336 (68.3)	1,469 (75.1)
Reducing salt	1,570 (80.3)	1,545 (79.0)
Quality of Life		
Low Physical Quality of Life	618 (31.7)	754 (38.7)
Low Mental Quality of Life	640 (32.8)	548 (28.1)
Social Determinants		
Depressive symptoms	244 (12.5)	267 (13.7)
Low Coping	925 (47.3)	1,087 (55.9)
High Stress	659 (33.7)	544 (28.6)
Low Social Support	653 (33.4)	701 (35.8)
Married	1,127 (57.6)	1,052 (53.8)
Antihypertensive Medication Adherence		
Low MMAS-8^^	279 (14.3)	247 (12.8)
Low PDC ^^^	468 (28.5)	393 (23.9)

*N=1,956 completed baseline and 2nd follow-up survey
^in the year prior
^^MMAS-8 – Morisky Medication Adherence Scale
^^^PDC = Proportion of days covered

Physical and mental quality of life (QOL) questions had response categories of yes/no, and 3-, 5-, and 6-point Likert scales. The positive or negative implication of possible responses and distribution of response categories suggested natural cut points for dichotomous variables representing quality of life. “Positive” quality of life responses were reference categories and “negative” quality of life responses were response categories of interest estimated by $\rho_{j,r_j,t|s_t}$ in latent transition models. Table 15 displays the percent with the negative quality of life category at baseline and follow-up. At least 98% of participants responded to both items at baseline and follow-up, save for questions on vigorous activity (96.4%) climbing several flights of stairs (97.4%), and walking more than one mile (95.8%). These three questions on physical functioning of a more strenuous nature may have represented ambiguous necessity for adults in this age group. Across both time periods, 1,645 participants (84.5%) completed all survey questions. Several items of physical quality of life displayed significant changes over two years, particularly within the physical functioning and physical health subscales. Increases in the percent not full of pep (26.8% at baseline versus 29.4% at follow-up, p=0.029) and the percent with physical/emotional interfering with social activity most or all of the time (35.7% at baseline versus 38.2% at follow-up, p=0.049) were two sole examples of decline in mental QOL. Participants were less likely to report that they were limited in work or other activities due to emotional health (14.8% versus 12.7%, p=0.020) or that their health now was somewhat or much worse than one year ago (15.2% versus 13.1%, p=0.041) at follow-up than they were at baseline.

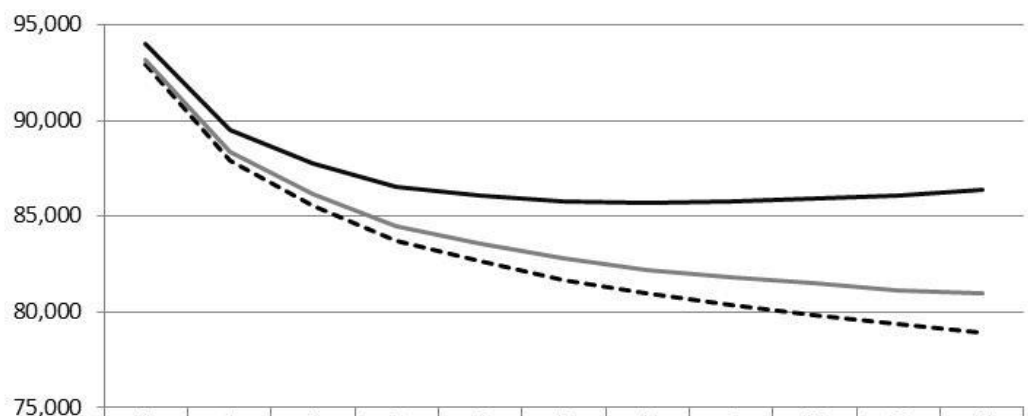
Table 15: Quality of life responses at baseline and follow-up*

Response*	N (%)*	% at Baseline	% at Follow-Up	p-value**
Physical Functioning				
Limited a little/a lot in:				
PF1: Vigorous activities	1,885 (96.4)	82.6	86.7	<0.001
PF2: Moderate activities	1,932 (98.8)	43.9	55.8	<0.001
PF3: Lifting or carrying groceries	1,942 (99.3)	37.0	45.1	<0.001
PF4: Climbing several flights of stairs	1,906 (97.4)	69.4	71.6	0.036
PF5: Climbing one flight of stairs	1,939 (99.1)	42.2	46.0	0.001
PF6: Bending, kneeling, stooping	1,947 (99.5)	60.3	68.8	<0.001
PF7: Walking more than one mile	1,873 (95.8)	63.5	70.7	<0.001
PF8: Walking several blocks	1,926 (98.5)	46.3	55.2	<0.001
PF9: Walking one block	1,950 (99.7)	27.8	33.1	<0.001
PF10: Bathing or dressing self	1,952 (99.8)	12.2	13.3	0.190
Physical Health				
Problems due to physical health:				
PH1: Cut down time on work/other activities	1,949 (99.6)	29.1	32.4	0.008
PH2: Accomplish less than liked	1,944 (99.4)	46.6	49.4	0.028
PH3: Limited in work or other activities	1,945 (99.4)	46.5	51.4	<0.001
PH4: Difficulty performing work/activities	1,941 (99.2)	42.4	48.1	<0.001
Body Pain				
BP1: Has body pain	1,949 (99.6)	74.6	77.3	0.009
BP2: Pain interfered with normal work	1,944 (99.4)	52.1	53.9	0.151
General Health				
GH1: Fair / poor general health	1,953 (99.8)	28.3	25.6	0.011
GH2: Health worse than 1 year ago	1,950 (99.7)	15.2	13.1	0.041
GH3: Get sick easier than other people	1,950 (99.7)	8.7	8.5	0.836
GH4: Not as healthy as anyone known	1,950 (99.7)	32.1	31.3	0.485
GH5: Expects health to get worse	1,947 (99.5)	27.4	27.0	0.702
GH6: Health is not excellent	1,948 (99.6)	41.2	43.5	0.039
Emotional Health				
Problems due to emotional health:				
EH1: Cut down time on work/other activities	1,945 (99.4)	14.8	13.9	0.350
EH2: Accomplish less than liked	1,935 (98.9)	20.8	18.8	0.060
EH3: Limited in work or other activities	1,939 (99.1)	14.8	12.7	0.020
Mental Health				
At least a little of the time feels:				
MH1: Very nervous	1,950 (99.7)	40.4	39.9	0.663
MH2: Nothing will cheer up	1,948 (99.6)	25.6	24.7	0.556
MH3: Down hearted and blue	1,948 (99.6)	43.5	43.3	0.827
At least a good bit of the time feels:				
MH4: Is a not happy person	1,949 (99.6)	26.7	23.1	0.147
MH5: Is not calm and peaceful	1,952 (99.8)	29.1	27.7	0.247
Vitality				
At least a little of the time feels:				
V1: Not full of pep	1,941 (99.2)	26.8	29.4	0.029
V2: Worn out	1,952 (99.8)	52.2	50.9	0.310
V3: Tired	1,950 (99.7)	62.0	63.4	0.235
V4: Not a lot of energy	1,942 (99.3)	25.1	26.9	0.097
Social Functioning				
Physical/emotional interferes with social:				
SF1: Most or all of the time	1,942 (99.3)	35.7	38.2	0.049
SF2: To at least a slight extent	1,948 (99.6)	32.1	33.5	0.258

*questions labeled with domain/number, referenced in table 16
**p-value for McNemar's test

Each variable selection algorithm was applied to the 36 dichotomized items separately. Figure 4 displays the likelihood ratio statistic G-squared, Akaike information criteria (AIC), and Bayesian information criteria (BIC) for latent transition analysis models specified on all items. BIC is minimized for an eight-status model. Neither G-squared nor AIC have reached a minimum value at twelve statuses, though the rate of decline begins to wane from eight statuses and on.

Figure 4: Fit statistics for 36-item latent transition analysis models by number of statuses



	2	3	4	5	6	7	8	9	10	11	12
G-squared	92,904	87,888	85,516	83,703	82,623	81,669	80,940	80,377	79,856	79,338	78,947
AIC	93,198	88,336	86,122	84,471	83,557	82,773	82,218	81,833	81,494	81,162	80,961
BIC	93,993	89,547	87,760	86,546	86,081	85,757	85,672	85,768	85,921	86,092	86,405

Number of Latent Statuses

For the headlong search algorithm, the variable list was rank-ordered by the sums of variances of probability across statuses and across time from the eight-status model. The smallest number of variables required for an identifiable model was four, and the initial clustering list was: limited in walking several blocks

(PF8), limited in work or other activities due to physical health (PH3), difficulty performing work or other activities due to health (PH4), and limited in walking one block due to health (PF9). Appendix table A6 contains the rank-ordering of the full list.

Headlong search selected variables through four steps (see table 16). BIC was minimized to 1085.06 for a two-status model utilizing the initial four variables. The third ranked variable, having accomplished less than liked due to physical health (PH2), was included in step 1 after difficulty climbing one flights of stairs (PF5) and feeling downhearted and blue (MH3) did not significantly lower BIC. The BIC of this model was minimized to 662.10 for three statuses as opposed to two statuses on the initial variable set. After the inclusion of PH2, PF9 was excluded as BIC was lowered with its removal, again with a three-status model in Step 2. Finally, each remaining variable was tested for inclusion with none lowering BIC below 127.44 (119.34 from step 2, plus 8.10 BIC for a one-class model with a single variable). The next best potential variables in step 3 were both from the mental quality of life subset of questions; these were questions about feeling very nervous (question MH1 with BIC = 158.92 for a three-status model) and not feeling calm and peaceful (question MH5 with BIC = 163.51 for a three-status model). Finally, in step 4, no variables were eligible for exclusion, and the algorithm terminated.

Table 16: Steps in headlong search variable selection

Step	Variable(s) Proposed	Step Type and Result	Current BIC^	Test BIC	Difference	Final Statuses
0	PF8, PH3, PH4, PF9	Initialize	1085.06	----	----	2
1	PF5, MH3, PH2*	Inclusion, Successful	1093.16	662.10	431.06	3
2	PF9*	Exclusion, Successful	670.20	119.34	550.86	3
3	All not in current model	Inclusion, Unsuccessful	127.44	----	----	3
4	PF8, PH3, PH4, PH2	Exclusion, Unsuccessful	127.44	----	----	3

*starred items are the final variables chosen for inclusion/exclusion in non-terminal steps
^8.10 added to previous test BIC to account for additional variable

Latent status was assigned at baseline and follow-up based on maximum predicted probability of status membership. Status 1, representing 35.3% of CoSMO participants, had the highest physical and mental quality of life across all items (see table 17). Very few participants in this status had trouble walking several blocks (5.1%) or were limited in their physical health (13.5%, 6.5%, and 3.8% for PH2, PH3, and PH4, respectively). Their mental quality of life was similarly better than average, with low percentages reporting problems due to emotional health and lowest percentages with negative mental health, vitality, and social functioning responses. Status 3, 44.1% of CoSMO participants, had poorest mental and physical quality of life. A majority reported trouble walking one or several blocks, lifting groceries, and over 80% had problems due to physical health per questions PH2-PH4. Status 2, 20.6% of participants, fell in between the first in third in terms of both quality of life metrics.

Table 17: Quality of life responses at baseline for the headlong search three-status model

Response, %	Status 1 (35.3%)	Status 2 (20.6%)	Status 3 (44.1%)
Physical Functioning			
Limited a little/a lot in:			
PF1: Vigorous activities	64.6	84.4	95.4
PF2: Moderate activities	15.7	38.2	69.4
PF3: Lifting or carrying groceries	10.6	31.3	60.8
PF4: Climbing several flights of stairs	42.6	72.6	89.7
PF5: Climbing one flight of stairs	12.5	41.0	66.6
PF6: Bending, kneeling, stooping	33.0	57.2	83.5
PF7: Walking more than one mile	27.5	72.6	87.7
PF8: Walking several blocks	5.1	58.5	73.7
PF9: Walking one block	1.0	26.2	50.2
PF10: Bathing or dressing self	1.2	6.9	23.4
Physical Health			
Problems due to physical health:			
PH1: Cut down time on work / other activities	4.9	13.2	55.9
PH2: Accomplish less than liked	13.5	23.9	83.7
PH3: Limited in work or other activities	6.5	11.0	95.1
PH4: Difficulty performing work or other activities	3.8	14.4	86.5
Body Pain			
BP1: Has body pain	60.4	70.7	87.7
BP2: Pain interfered with normal work	24.7	41.4	79.1
General Health			
GH1: Fair / poor general health	11.9	21.8	44.4
GH2: Health worse than 1 year ago	5.9	7.4	26.3
GH3: Get sick easier than other people	3.3	7.4	13.5
GH4: Not as healthy as anyone known	16.5	27.9	46.8
GH5: Expects health to get worse	20.8	23.6	34.5
GH6: Health is not excellent	18.4	30.5	64.5
Emotional Health			
Problems due to emotional health:			
EH1: Cut down time on work / other activities	3.9	12.4	24.7
EH2: Accomplish less than liked	7.4	18.7	32.8
EH3: Limited in work or other activities	2.8	9.9	26.9
Mental Health			
At least a little of the time feels:			
MH1: Very nervous	28.2	37.6	51.5
MH2: Down in the dumps, nothing will cheer up	13.1	23.3	36.1
MH3: Down hearted and blue	30.3	41.4	54.9
At least a good bit of the time feels:			
MH4: Is a not happy person	12.6	21.5	35.9
MH5: Is not calm and peaceful	17.3	23.6	41.2
Vitality			
At least a little of the time feels:			
V1: Not full of pep	10.9	20.2	42.8
V2: Worn out	29.4	39.1	66.5
V3: Tired	44.6	57.6	77.8
V4: Not a lot of energy	6.8	16.1	44.0
Social Functioning			
Physical/emotional interferes with social activity:			
SF1: Most or all of the time	11.0	21.0	54.0
SF2: To at least a slight extent	14.5	26.7	57.1

Overall, 26.5% of CoSMO participants changed status from baseline to follow-up based on this three-status model. As may be expected in CoSMO's elderly population, transition probabilities demonstrated a decline in quality of life, as Status 1 decreased in size and Status 3 increased in size from baseline to follow-up. 80.9% of participants who had highest quality of life in Status 1 at baseline remained in this category at follow-up; most that transitioned did so to Status 3, 18.7%. Nearly half of those in the middle Status 2 had declined to Status 3 over the follow-up time period. 76.6% of those in the lowest quality of life category remained in this Status 3 at follow-up. 5.3% of Status 3 participants migrated to Status 1, and 18.1% migrated to Status 2.

Table 18: Status transitions from baseline to follow-up for the headlong search model

		Follow-up		
	N **Total % / Row %**	Status 1 604 (30.9%)	Status 2 378 (19.3%)	Status 3 974 (49.8%)
Baseline	Status 1 690 (35.3%)	558 28.5 / 80.9	3 0.2 / 0.4	129 6.6 / 18.7
	Status 2 404 (20.6%)	0 0.0 / 0.0	219 11.2 / 54.2	185 9.5 / 45.8
	Status 3 862 (44.1%)	46 2.4 / 5.3	156 8.0 / 18.1	660 33.7 / 76.6

Table 19 contains participant characteristics by the three statuses, with Cochran Armitage trend tests conducted for the ordinal statuses. Those with worse quality of life (Status 3) were older, more likely to be female, have high BMI, and multiple past cardiovascular events. They had more clinic visits per year though with less satisfaction with healthcare, access, and communication with their doctors. Social determinants were worse for this status as well, as this group exhibited more depressive symptoms, low coping skills, high stress, and low social support. Finally, those in Status 3 filled more classes of antihypertensive medication in spite of lower adherence as measured by both self-report (MMAS-8) and pharmacy fill (PDC).

Table 19: Characteristics of CoSMO participants at baseline for the headlong search three-status model

Participant Characteristics, %	**Status 1 (35.3%)**	**Status 2 (20.6%)**	**Status 3 (44.1%)**	**p-value***
Demographics				
Age ≥ 75 years	41.2	52.0	51.4	<0.001
Female	51.5	61.6	62.1	<0.001
Black race	28.1	32.9	29.8	0.516
High school education or greater	85.2	80.0	75.8	<0.001
Clinical Variables				
Hypertension duration ≥ 10 Years	59.2	61.4	64.5	0.031
Body mass index: ≥ 30 kg/m²	70.1	77.2	81.7	<0.001
History of Stroke	7.7	12.1	14.5	<0.001
History of Diabetes	31.5	44.8	46.3	<0.001
History of Myocardial Infarction	6.5	11.6	13.8	<0.001
History of Congestive Heart Failure	6.1	6.4	8.0	0.135
Healthcare Variables				
≥4 clinic visits per year	40.4	54.7	62.7	<0.001
3+ classes of antihypertensive medication^	36.5	41.2	49.5	<0.001
Reduced medication due to cost	1.3	4.0	4.7	<0.001
Uses Complementary and Alternative Medicine	25.8	24.0	28.9	0.151
Low Overall Satisfaction with Healthcare	1.7	5.0	6.7	<0.001
Low Satisfaction with Access	2.5	3.0	6.2	<0.001
Low Satisfaction with Communication	6.4	8.9	13.9	<0.001
Behavioral Variables				
Never a smoker	47.9	54.0	48.2	0.992
<2 alcoholic drinks per week	72.3	81.1	82.1	<0.001
Increasing fruits and vegetables	66.5	68.8	69.5	0.217
Reducing salt	78.8	77.7	82.6	0.055
Quality of Life				
Low Physical Quality of Life	1.6	8.2	66.8	<0.001
Low Mental Quality of Life	20.9	33.3	42.1	<0.001
Social Determinants				
Depressive symptoms	3.2	10.6	20.8	<0.001
Low Coping	40.8	43.3	54.4	<0.001
High Stress	22.5	31.2	43.9	<0.001
Low Social Support	28.4	33.4	37.4	<0.001
Married	60.9	54.0	56.7	0.120
Antihypertensive Medication Adherence				
Low MMAS-8^^	13.2	10.6	16.8	0.032
Low PDC ^^^	24.9	26.8	33.2	<0.001

^in the year prior
^^MMAS-8 – Morisky Medication Adherence Scale; ^^^PDC = Proportion of days covered
*p-value for Cochran Armitage trend test

Next, application of the backwards elimination approach selected variables for an eight-status model. The first six variables could be removed with > 90% remaining in the same status; removing a further six variables as well resulted in > 80% remaining in the same status. Table 20 lists items eliminated along with $1 - \mathrm{F}^{A\backslash \mathrm{j}}$, $\mathrm{D}^{A\backslash \mathrm{j}}$ at each step through the 28th elimination, as models did not converge for fewer than eight variables.

Table 20: Steps in backwards elimination variable selection

Step	Item	$1 - \mathrm{F}^{A\backslash \mathrm{j}}$	$\mathrm{D}^{A\backslash \mathrm{j}}$	Step	Item	$1 - \mathrm{F}^{A\backslash \mathrm{j}}$	$\mathrm{D}^{A\backslash \mathrm{j}}$
1	SF1	0.924	265.7	15	V2	0.670	218.7
2	GH3	0.954	209.4	16	EH2	0.754	356.8
3	BP1	0.937	168.8	17	V3	0.724	345.0
4	MH4	0.929	158.7	18	PF9	0.740	254.1
5	GH5	0.926	162.9	19	MH2	0.741	261.3
6	PF1	0.916	170.7	20	GH6	0.705	284.0
7	PF6	0.899	150.0	21	BP2	0.694	275.1
8	PH1	0.878	159.8	22	SF2	0.664	229.3
9	GH4	0.874	106.7	23	PF5	0.675	102.1
10	PF3	0.847	137.3	24	MH1	0.638	182.3
11	PF7	0.833	231.5	25	PF10	0.638	111.1
12	GH2	0.835	142.9	26	V4	0.638	213.2
13	PH2	0.691	158.0	27	MH3	0.617	137.2
14	MH5	0.703	167.0	28	V1	0.578	268.8

The removal of 10 variables resulted in $1 - \mathrm{F}^{A\backslash \mathrm{j}}$ = 0.847, or 85% maintaining the same status assignment as the full-variable model. 85% classification is nearly seven times better than random assignment to eight statuses (or 12.5% correct classification), and the characteristics of the eight statuses were maintained in the simpler model (data not shown). No subscale was completely omitted due to variable elimination.

Baseline quality of life responses are provided in table 21. In keeping with the order of the three-status model from headlong search, statuses in table 21 are ordered from lowest percent in the bottom tertile of physical quality of life to highest percent in the bottom tertile of physical quality of life. Status 1, 21.0% of CoSMO participants, had best physical and mental quality of life across all 36 dichotomized items. Status 2 also had better physical quality of life than average, but these participants were more likely to respond in the negative to mental health questions, more often to be down hearted and blue, nervous, and down in the dumps. Status 3 participants responded that they were limited in physical functioning at rates similar or slightly above average compared to the overall sample; however, they were much less likely to report problems with physical, mental, and emotional health. Status 4 similarly had better emotional health, but worse physical health and social functioning than Status 3.

Status 5 exhibited slightly poorer physical functioning, and well below-average quality of life with respect to the emotional and mental health subscales. Status 6 had the lowest mental quality of life in terms of emotional health, mental health, and social functioning. Status 7 is characterized by poor physical quality of life, but better-than-average emotional and mental health. Finally, the 8.1% of participants in Status 8 had poorest physical functioning and general health.

Table 21: Quality of life responses at baseline for the backwards elimination eight-status model

Response, %	Status 1 (21.0%)	Status 2 (10.2%)	Status 3 (8.8%)	Status 4 (14.8%)
Physical Functioning				
Limited a little/a lot in:				
PF1: Vigorous activities	54.2	69.4	90.3	89.2
PF2: Moderate activities	7.8	10.5	41.2	29.2
PF3: Lifting or carrying groceries	4.2	8.5	34.5	20.8
PF4: Climbing several flights of stairs	23.6	49.0	91.7	67.3
PF5: Climbing one flight of stairs	2.7	12.0	53.5	21.3
PF6: Bending, kneeling, stooping	25.4	25.1	67.8	56.4
PF7: Walking more than one mile	17.6	16.6	98.2	52.8
PF8: Walking several blocks	2.5	3.5	92.4	7.6
PF9: Walking one block	0.0	0.0	40.4	0.0
PF10: Bathing or dressing self	0.0	0.0	9.9	2.8
Physical Health				
Problems due to physical health:				
PH1: Cut down time on work / other activities	2.2	6.5	7.0	22.1
PH2: Accomplish less than liked	8.5	12.6	9.3	51.0
PH3: Limited in work or other activities	1.2	6.5	8.8	58.0
PH4: Difficulty performing work or other activities	1.7	2.5	6.5	44.9
Body Pain				
BP1: Has body pain	48.4	65.8	69.0	77.8
BP2: Pain interfered with normal work	10.0	27.5	31.6	57.8
General Health				
GH1: Fair / poor general health	3.9	12.5	20.9	18.0
GH2: Health worse than 1 year ago	1.7	7.5	4.1	12.1
GH3: Get sick easier than other people	1.0	6.0	3.5	4.8
GH4: Not as healthy as anyone known	9.7	19.0	26.7	23.9
GH5: Expects health to get worse	17.1	26.0	19.8	25.7
GH6: Health is not excellent	6.1	21.5	26.9	41.5
Emotional Health				
Problems due to emotional health:				
EH1: Cut down time on work / other activities	0.0	6.5	2.3	2.4
EH2: Accomplish less than liked	1.5	11.6	5.2	9.0
EH3: Limited in work or other activities	0.5	0.5	1.7	4.2
Mental Health				
At least a little of the time feels:				
MH1: Very nervous	8.8	69.3	18.0	24.7
MH2: Down in the dumps, nothing will cheer up	0.0	44.7	5.2	4.5
MH3: Down hearted and blue	7.1	83.0	23.8	14.2
At least a good bit of the time feels:				
MH4: Is a not happy person	2.7	30.7	10.5	11.1
MH5: Is not calm and peaceful	3.2	41.7	5.8	12.1
Vitality				
At least a little of the time feels:				
V1: Not full of pep	7.1	6.5	17.4	20.1
V2: Worn out	15.8	49.5	29.7	39.1
V3: Tired	28.0	61.6	56.4	56.1
V4: Not a lot of energy	2.2	3.5	7.6	16.0
Social Functioning				
Physical/emotional interferes with social activity:				
SF1: Most or all of the time	2.9	18.6	11.0	17.6
SF2: To at least a slight extent	1.7	24.0	7.0	20.1

Table 21 (continued): Quality of life responses at baseline for the backwards elimination eight-status model

Response, %	Status 5 (12.8%)	Status 6 (10.0%)	Status 7 (14.3%)	Status 8 (8.1%)
Physical Functioning				
Limited a little/a lot in:				
PF1: Vigorous activities	89.1	94.8	96.4	98.7
PF2: Moderate activities	48.4	86.5	75.0	96.8
PF3: Lifting or carrying groceries	42.8	82.9	57.8	88.1
PF4: Climbing several flights of stairs	75.7	98.4	97.5	100.0
PF5: Climbing one flight of stairs	37.6	88.1	78.0	96.2
PF6: Bending, kneeling, stooping	69.2	91.2	86.7	94.3
PF7: Walking more than one mile	71.8	97.4	99.6	100.0
PF8: Walking several blocks	43.0	94.8	92.5	98.7
PF9: Walking one block	13.9	78.4	53.4	88.1
PF10: Bathing or dressing self	10.4	44.1	12.2	42.1
Physical Health				
Problems due to physical health:				
PH1: Cut down time on work / other activities	33.6	77.9	43.2	72.2
PH2: Accomplish less than liked	63.1	89.2	76.3	88.7
PH3: Limited in work or other activities	50.4	95.9	87.1	95.6
PH4: Difficulty performing work/other activities	51.4	88.7	80.1	93.7
Body Pain				
BP1: Has body pain	86.1	93.8	84.5	93.7
BP2: Pain interfered with normal work	70.8	90.8	72.7	91.1
General Health				
GH1: Fair / poor general health	31.9	69.2	32.0	74.8
GH2: Health worse than 1 year ago	18.3	39.7	16.1	40.9
GH3: Get sick easier than other people	13.5	27.8	5.0	19.5
GH4: Not as healthy as anyone known	32.5	60.8	42.7	74.2
GH5: Expects health to get worse	25.9	38.7	32.7	46.5
GH6: Health is not excellent	46.6	78.5	59.5	85.5
Emotional Health				
Problems due to emotional health:				
EH1: Cut down time on work / other activities	38.4	75.8	4.3	6.3
EH2: Accomplish less than liked	50.6	93.3	7.9	8.4
EH3: Limited in work or other activities	33.2	89.1	3.6	4.4
Mental Health				
At least a little of the time feels:				
MH1: Very nervous	78.5	86.7	22.2	54.1
MH2: Down in the dumps, nothing will cheer up	59.0	78.5	7.9	38.4
MH3: Down hearted and blue	90.0	92.8	26.4	57.2
At least a good bit of the time feels:				
MH4: Is a not happy person	49.0	67.5	15.4	40.3
MH5: Is not calm and peaceful	62.5	76.9	16.1	47.8
Vitality				
At least a little of the time feels:				
V1: Not full of pep	30.8	60.8	25.4	80.5
V2: Worn out	56.2	85.6	59.9	82.4
V3: Tired	74.1	92.8	72.8	90.6
V4: Not a lot of energy	33.2	63.2	26.6	85.4
Social Functioning				
Physical/emotional interferes with social activity:				
SF1: Most or all of the time	51.8	86.7	35.1	69.9
SF2: To at least a slight extent	67.3	89.2	39.0	78.3

Across all eight statuses, 59.8% changed status from baseline to follow-up. Rank-ordered by the percent with low physical quality of life, 71.4% maintained or worsened in status. The least stable statuses were 3, 4, and 7; of note, these three all had mostly positive responses on emotional and mental health questions, and differentiation between the three is due to vitality, social function, and physical health subscales. Statuses 1 and 8 had the highest proportions of participants that remained in the same category over time, 73.0% and 64.8% respectively.

Table 22: Status transitions from baseline to follow-up for the backwards elimination model

		Follow-up							
	N % of Baseline	Status 1 448 (22.9%)	Status 2 158 (8.1%)	Status 3 268 (13.7%)	Status 4 241 (12.3%)	Status 5 214 (10.9%)	Status 6 165 (8.4%)	Status 7 205 (10.5%)	Status 8 257 (13.1%)
Baseline	Status 1 411 (21.0%)	300 73.0%	17 4.1%	16 3.9%	31 7.5%	3 0.7%	0 0.0%	37 9.0%	7 1.7%
	Status 2 200 (10.2%)	41 20.5%	89 44.5%	15 7.5%	16 8.0%	23 11.5%	3 1.5%	11 5.5%	2 1.0%
	Status 3 172 (8.8%)	30 17.4%	5 2.9%	39 22.7%	76 44.2%	11 6.4%	1 0.6%	5 2.9%	5 2.9%
	Status 4 289 (14.8%)	49 17.0%	17 5.9%	39 13.5%	47 16.3%	31 10.7%	3 1.0%	99 34.3%	4 1.4%
	Status 5 251 (12.8%)	20 8.0%	30 12.0%	23 9.2%	5 2.0%	101 40.2%	26 10.4%	22 8.8%	24 9.6%
	Status 6 195 (10.0%)	3 1.5%	0 0.0%	10 5.1%	8 4.1%	33 16.9%	88 45.1%	9 4.6%	44 22.6%
	Status 7 279 (14.3%)	4 1.4%	0 0.0%	109 39.1%	54 19.4%	9 3.2%	15 5.4%	20 7.2%	68 24.4%
	Status 8 159 (8.1%)	1 0.6%	0 0.0%	17 10.7%	4 2.5%	3 1.9%	29 18.2%	2 1.3%	103 64.8%

Table 23 contains profiles of participant characteristics by the eight statuses. Statuses 1, 3, and 4 were least likely to have both low physical and mental quality of life, while Statuses 6 and 8 were most likely to have both low physical and mental quality of life. Status 7 exhibits lower physical quality of life but not mental quality of life, while Statuses 2 and 5 have the reverse with lower mental but not physical quality of life.

Statuses 5 and 6, with the greatest percent having low mental quality of life, also represented higher rates with depressive symptoms, stress, dissatisfaction with health care, and were most likely to have reduced their medication due to cost. These two statuses had the greatest proportion with low self-reported adherence, and low PDC in the case of Status 6. Statuses 1 and 3 had the best adherence by both self-report and pharmacy fill.

Table 23: Characteristics of CoSMO participants at baseline for the backwards elimination eight-status model

Participant Characteristics, %	**Status 1 (21.0%)**	**Status 2 (10.2%)**	**Status 3 (8.8%)**	**Status 4 (14.8%)**
Demographics				
Age ≥ 75 years	39.9	44.5	49.4	48.8
Female	45.5	58.0	62.8	53.6
Black race	26.3	27.0	32.0	31.1
High school education or greater	88.8	84.5	79.1	82.0
Clinical Variables				
Hypertension duration ≥ 10 Years	58.0	61.1	62.2	61.2
Body mass index: ≥ 30 kg/m²	69.8	67.0	82.5	76.5
History of Stroke	7.3	7.5	11.6	11.8
History of Diabetes	30.4	27.0	45.3	40.5
History of Myocardial Infarction	6.6	6.5	9.3	9.3
History of Congestive Heart Failure	5.4	8.0	4.7	7.6
Healthcare Variables				
≥4 clinic visits per year	36.6	42.2	55.0	52.6
3+ classes of antihypertensive medication^	35.6	35.7	44.7	40.6
Reduced medication due to cost	1.2	3.0	2.3	1.4
Uses Complementary/Alternative Medicine	24.3	28.5	21.5	24.6
Low Overall Satisfaction with Healthcare	1.0	4.5	1.8	2.8
Low Satisfaction with Access	0.7	4.0	1.7	3.5
Low Satisfaction with Communication	4.6	9.0	4.7	9.1
Behavioral Variables				
Never a smoker	53.3	46.7	50.0	47.0
<2 alcoholic drinks per week	72.0	69.4	81.4	78.5
Increasing fruits and vegetables	67.6	65.5	68.0	66.8
Reducing salt	78.3	74.5	79.7	83.0
Quality of Life				
Low Physical Quality of Life	0.0	0.0	8.1	14.5
Low Mental Quality of Life	1.5	57.0	5.8	11.8
Social Determinants				
Depressive symptoms	0.0	9.0	2.9	2.8
Low Coping	35.3	46.2	40.1	47.1
High Stress	11.2	43.0	17.4	22.1
Low Social Support	20.9	38.0	25.0	28.7
Married	64.7	54.0	57.0	59.5
Antihypertensive Medication Adherence				
Low MMAS-8^^	10.2	14.5	10.5	13.5
Low PDC^^^	22.2	29.4	23.2	30.6

^in the year prior
^^MMAS-8 – Morisky Medication Adherence Scale; ^^^PDC = Proportion of days covered

Table 23 (continued): Characteristics of CoSMO participants at baseline for the backwards elimination eight-status model

Participant Characteristics, (%)	**Status 5 (12.8%)**	**Status 6 (10.0%)**	**Status 7 (14.3%)**	**Status 8 (8.1%)**
Demographics				
Age ≥ 75 years	48.6	55.9	52.7	50.3
Female	66.1	69.2	60.9	64.2
Black race	32.7	40.0	26.5	27.0
High school education or greater	78.1	65.5	77.1	74.8
Clinical Variables				
Hypertension duration ≥ 10 Years	59.8	58.2	71.3	66.0
Body mass index: ≥ 30 kg/m²	73.3	86.7	84.9	79.2
History of Stroke	12.4	15.9	13.6	17.6
History of Diabetes	43.8	51.3	44.1	56.6
History of Myocardial Infarction	11.6	14.9	14.3	18.9
History of Congestive Heart Failure	8.0	5.6	7.9	10.1
Healthcare Variables				
≥4 clinic visits per year	56.4	67.5	62.2	71.1
3+ classes of antihypertensive medication^	39.3	53.1	48.5	59.5
Reduced medication due to cost	6.8	8.7	1.8	4.4
Uses Complementary/Alternative Medicine	27.5	34.4	29.0	26.4
Low Overall Satisfaction with Healthcare	7.6	11.3	3.6	9.4
Low Satisfaction with Access	6.4	10.8	2.5	8.8
Low Satisfaction with Communication	15.1	19.0	8.3	19.0
Behavioral Variables				
Never a smoker	50.2	53.1	46.0	44.9
<2 alcoholic drinks per week	78.8	88.2	80.6	86.8
Increasing fruits and vegetables	72.9	75.4	69.5	58.5
Reducing salt	79.3	89.2	81.7	76.1
Quality of Life				
Low Physical Quality of Life	19.6	72.2	78.4	97.5
Low Mental Quality of Life	80.0	99.0	7.9	39.0
Social Determinants				
Depressive symptoms	24.7	55.4	3.2	21.4
Low Coping	51.8	63.1	49.1	58.5
High Stress	60.6	68.2	27.2	45.3
Low Social Support	47.8	51.3	31.5	35.8
Married	56.2	45.1	56.6	60.4
Antihypertensive Medication Adherence				
Low MMAS-8^^	19.1	22.6	12.9	14.5
Low PDC ^^^	29.6	44.4	29.3	28.7

^in the year prior
^^MMAS-8 – Morisky Medication Adherence Scale; ^^^PDC = Proportion of days covered

Latent status assignment for both headlong search variables and backwards elimination variables were compared at baseline to assess model differences (see table 24). Over 90% of Statuses 6, 7, and 8 of the backwards elimination model were classified in Status 3 of headlong search, representing the lowest levels of physical, and to a lesser extent mental, quality of life. Likewise, Statuses 1 and 2 of the backwards elimination model were most likely to be classified in Status 1 of the headlong search model, with better physical and mental quality of life metrics.

Table 24: Cross-tabulation of baseline backwards elimination status with baseline headlong search status

		Headlong Search		
	N % of Total / % of Row	Status 1 690 (35.3%)	Status 2 404 (20.7%)	Status 3 862 (44.1%)
Backwards Elimination	Status 1 411 (21.0%)	361 18.5 / 87.8	47 2.4 / 11.4	3 0.2 / 0.7
	Status 2 200 (10.2%)	155 7.9 / 77.5	43 2.2 / 21.5	2 0.1 / 1.0
	Status 3 172 (8.8%)	31 1.6 / 18.0	134 6.9 / 77.9	7 0.4 / 4.1
	Status 4 289 (14.8%)	103 5.3 / 35.6	60 3.1 / 20.8	126 6.4 / 43.6
	Status 5 251 (12.8%)	40 2.0 / 15.9	84 4.3 / 3.5	127 6.5 / 50.6
	Status 6 195 (10.0%)	0 0.0 / 0.0	6 0.3 / 3.1	189 9.7 / 96.9
	Status 7 279 (14.3%)	0 0.0 / 0.0	27 1.4 / 9.7	252 12.9 / 90.3
	Status 8 159 (8.1%)	0 0.0 / 0.0	3 0.2 / 1.9	156 8.0 / 98.1

As a measure of external validity to the models identified by variable selection techniques, the variables selected from these methods were compared to a previously developed “even-shorter” quality of life tool, the SF-12 (44). The SF-12 is a subset of the original 36 questions developed to estimate the eight domains on a single-page health survey. The questions included in the SF-12 are shown in table 25.

Table 25: SF-12 Quality of life survey questions compared to headlong search and backwards elimination

	Headlong Search	Backwards Elimination
Physical Functioning		
Limited a little/a lot in:		
PF2: Moderate activities		✓
PF4: Climbing several flights of stairs		✓
Physical Health		
Problems due to physical health:		
PH2: Accomplish less than liked	✓	✓
PH3: Limited in work or other activities	✓	✓
Body Pain		
BP2: Pain interfered with normal work		✓
General Health		
GH6: Health is not excellent		✓
Emotional Health		
Problems due to emotional health:		
EH2: Accomplish less than liked		✓
EH3: Did work or other activities less carefully*		✓
Mental Health		
At least a little of the time feels:		
MH3: Down hearted and blue		✓
At least a good bit of the time feels:		
MH5: Is not calm and peaceful		✓
Vitality		
At least a little of the time feels:		
V4: Not a lot of energy		✓
Social Functioning		
Physical/emotional interferes with social activity:		
SF1: Most or all of the time		

Check marks represent variables selected by each technique
*Wording was modified for CoSMO: "Limited in the kind of work or other activities"

Two of the four items selected from headlong search are also included in the SF-12. The SF-12 uses two physical health questions of the original four, omitting PH4 chosen by headlong search. The question related to limitations in walking several blocks (PF8) is not on the SF-12, though questions on moderate activities (PF2) and climbing several flights of stairs (PF4) are included. These variables had reasonable concordance of responses among CoSMO participants (74.6% between PF8 and PF2; 72.6% between PF8 and PF4; data not shown). No variables from the mental health domains, body pain, or general health were selected by the headlong search algorithm.

Of the ten items removed through backwards elimination, only SF1, the amount of time that physical or emotional problems interfered with social activity, is present on the SF-12. Backwards elimination retained the second social functioning question, the extent affected rather than the amount of time affected by physical or emotional interference. These items similarly exhibited high concordance among CoSMO participants (79.1%; data not shown).

DISCUSSION

The aphorism that “all models are wrong but some are useful” is a good place to start in considering variable selection for latent class or transition models. Each technique represents a practical, logical approach to finding acceptable subsets of variables for models when full-variable modeling is not possible or recommended. Headlong search begins with the notion of separation of classes due to between-class differences in item response probabilities, a desirable property in latent class modeling. Backwards elimination arises from the belief that valid scales provide true class definitions, and noise variables may be removed so long as class definition is maintained. Provided that statuses are sizeable and potential variable lists are not unwieldy, both methods are equally suitable for variable selection as observed in two-status simulations. The simpler models provide similar agreement and correct classification rates to models based on all variables, as well as model based on only informative variables.

Statuses with small sample size at any time period can result in failures to identify correct variables, incorrect inclusion of non-informative variables, incorrect identification of the number of latent statuses, and ultimately incorrect classification. Variable selection techniques applied to studies with long surveys but limited participation may not yield informative or accurate status definitions. LCA and LTA methods, often categorized alongside cluster and factor analysis within the family of multivariate techniques for handling high-dimensional data, are better employed in big-data applications where sample size will trump class complexity. While four- and five-status simulations point to pitfalls of both techniques, in practice, the researcher has options to overcome potential problems. For example, to the extent possible, sample size planning should account for the greatest number of classes one may hypothesize as reasonable for the study population, and a healthy ratio of observations to variables ensured prior to data collection. Furthermore, some variables may be removed from consideration before variable selection techniques are applied, when high levels of multicollinearity are present or when the items responses probabilities are too close to zero or one.

The choice of headlong search versus backwards elimination as a variable selection technique may hinge upon one or several considerations. Simulation results suggest that headlong search may select fewer variables and simpler models whereas backwards elimination may point to inclusion of more variables and more complex models. The three-status CoSMO model that arises from headlong search provides a straightforward categorization of three levels of physical quality of life at baseline and estimation of change over time. The eight statuses identified in the backwards elimination technique become a bit cumbersome to describe and compare, though the role and interaction of both physical and mental quality of life subscales are better represented by this model. Lacking a gold standard definition of class or status, one may choose either headlong search or backwards elimination models bearing in mind the interpretation and application of modeling results. As an example, more complex models may be useful in descriptive analyses to compare subgroups fully described by the scale or variables. Utilizing statuses identified from these models in behavioral interventions may not be feasible, and simpler models may be favored. Additionally, prior knowledge of and research into variables or scales may support one variable selection technique or the other. The psychometric properties including scale reliability should be determined prior to utilizing backwards elimination, for example, as “true” status definition is an assumption. The headlong search approach may be preferred if the ratio of variables to sample size is too high to have confidence that the correct number of statuses can be identified by the full-variable model.

STRENGTHS, LIMITATIONS, AND FUTURE DIRECTIONS

There are limitations to the two variable selection techniques discussed here, as well as to the scope of the assessment of these two tools. The method of Dean and Raftery requires an increasingly large sample size depending on the number of item response parameters specified in the initial variable set. Additionally, they propose strategies for defining reduced variable lists when models are not identifiable, though these require more user-intervention than possible for this simulation study. Dean and Raftery do not address missing data either, which is often a concern with survey data for LCA. The item reduction of Bartolucci *et al* poses some potential drawbacks as well. The authors recommend that this technique be applied to validated surveys to identify true latent class membership, which may not be available. Measurement error and a variety of sources of response bias may be a limitation as well, even when using recognized, validated surveys.

This analysis extended these two variable selection techniques from the cross-sectional LCA setting to longitudinal LTA. Expanded simulation is a strength of this study, as there are limited data from simulations in these published cross-sectional cases. Dean and Raftery simulated two-class and three-class data sets a single time each to test their methodology, and this study provides larger trends in correct identification of the number of classes as well as class membership not available in the initial proposal of the methodology. Bartolucci *et al* did not test their method on simulated data but evaluated its application on the ULISSE data set alone. Finally, CoSMO is the first application of these variable selection techniques to a real longitudinal data set. The quality of life scale is a validated survey tool, thus meeting the criterion of Bartolucci *et al* for the backwards variable elimination. CoSMO sample size and the number of survey items available are similar to the ULISSE data set as well. The LTA models from both techniques provide characterizations of quality of life and change over time for CoSMO participants. Differentiation of medication adherence by status membership illustrates the interplay of quality of life and compliance to prescribed medications in this cohort.

Variable selection for latent transition analysis can lead to future longitudinal CoSMO studies capitalizing on the broad data collection available across many patient characteristics. Variable selection performed on questions about reasons for non-adherence may suggest behavioral interventions for this population, and ways to identify those at risk for transitioning from good to poor adherers. To varying degrees, both quality of life models defined through variable selection on CoSMO demonstrated subgroups of participants with lower quality of life and lower medication adherence. Utility of these new models can be further assessed through an examination of baseline status and change in status to predict longer term health outcomes including blood pressure control and cardiovascular events such as myocardial infarction, stroke, heart failure, and CV death.

Aspects of each variable selection technique can be established through further studies and simulations. For headlong search, the high and low difference in BIC values for inclusion and exclusion of variables can be adjusted to perhaps allow for more complex models with added variables. Other fit statistics than BIC, including AIC and adjusted BIC, can be examined also. For backwards elimination, the stopping criteria, or correct classification value at which to terminate the variable elimination steps, should be defined given sample size, status size, and number of statuses present. Recommendations on defining the number of “true” classes or statuses based on all variables, and their assignment through posterior probabilities, should be considered as well; simulations showed that use of BIC alone did not always identify the correct number of statuses given four or five statuses and a high number of non-informative variables.

Additionally, future studies should explore cross-sectional LCA performance in simulations. Alternative settings with different numbers of classes, sample sizes, and item response probabilities should be tested for both LCA and LTA. Four and five status settings may exhibit improved variable selection and status concordance provided larger sample sizes; several thousand observations to represent a large study, or several million observations to represent data mining of a large database.

APPENDIX

Table A1: Two-status item response probabilities for unequal groups

	Status 1 (70%)	Status 2 (30%)
Item 1	0.70	0.10
Item 2	0.80	0.20
Item 3	0.90	0.30
Item 4	0.15	0.80
Item 5	0.25	0.80
Item 6	0.10	0.10
Item 7	0.30	0.30
Item 8	0.50	0.50
Item 9	0.70	0.70
Item 10	0.90	0.90

Table A2: Two-status transition probabilities

Equal Groups: Probability of Transition from Time 1 to Time 2		Time 2			
		High Transition Probability		Low Transition Probability	
		Status 1 (50%)	Status 2 (50%)	Status 1 (50%)	Status 2 (50%)
Time 1	Status 1 (50%)	64%	36%	88%	12%
	Status 2 (50%)	36%	64%	12%	88%

Unequal Groups: Probability of Transition from Time 1 to Time 2		Time 2			
		High Transition Probability		Low Transition Probability	
		Status 1 (55.6%)	Status 2 (44.4%)	Status 1 (65.2%)	Status 2 (34.8%)
Time 1	Status 1 (70%)	64%	36%	88%	12%
	Status 2 (30%)	36%	64%	12%	88%

Table A3: Five status item response probabilities for simple structure, unequal groups

	Status 1 (5%)	**Status 2 (10%)**	**Status 3 (15%)**	**Status 4 (20%)**	**Status 5 (50%)**
Item 1	0.70	0.10	0.10	0.10	0.10
Item 2	0.75	0.15	0.15	0.15	0.15
Item 3	0.80	0.20	0.20	0.20	0.20
Item 4	0.85	0.25	0.25	0.25	0.25
Item 5	0.90	0.30	0.30	0.30	0.30
Item 6	0.10	0.70	0.10	0.10	0.10
Item 7	0.15	0.75	0.15	0.15	0.15
Item 8	0.20	0.80	0.20	0.20	0.20
Item 9	0.25	0.85	0.25	0.25	0.25
Item 10	0.30	0.90	0.30	0.30	0.30
Item 11	0.10	0.10	0.70	0.10	0.10
Item 12	0.15	0.15	0.75	0.15	0.15
Item 13	0.20	0.20	0.80	0.20	0.20
Item 14	0.25	0.25	0.85	0.25	0.25
Item 15	0.30	0.30	0.90	0.30	0.30
Item 16	0.10	0.10	0.10	0.70	0.10
Item 17	0.15	0.15	0.15	0.75	0.15
Item 18	0.20	0.20	0.20	0.80	0.20
Item 19	0.25	0.25	0.25	0.85	0.25
Item 20	0.30	0.30	0.30	0.90	0.30
Item 21	0.10	0.10	0.10	0.10	0.70
Item 22	0.15	0.15	0.15	0.15	0.75
Item 23	0.20	0.20	0.20	0.20	0.80
Item 24	0.25	0.25	0.25	0.25	0.85
Item 25	0.30	0.30	0.30	0.30	0.90
Item 26	0.10	0.10	0.10	0.10	0.10
Item 27	0.20	0.20	0.20	0.20	0.20
Item 28	0.30	0.30	0.30	0.30	0.30
……					
Item 33	0.80	0.80	0.80	0.80	0.80
Item 34	0.90	0.90	0.90	0.90	0.90
Item 35	0.10	0.10	0.10	0.10	0.10
Item 36	0.20	0.20	0.20	0.20	0.20
……					
Item 50	0.70	0.70	0.70	0.70	0.70

Table A4: Five-status transition probabilities

Equal Groups: Probability of Transition from Time 1 to Time 2		Time 2									
		High Transition Probability					Low Transition Probability				
		Status 1 (20.0%)	Status 2 (20.0%)	Status 3 (20.0%)	Status 4 (20.0%)	Status 5 (20.0%)	Status 1 (20.0%)	Status 2 (20.0%)	Status 3 (20.0%)	Status 4 (20.0%)	Status 5 (20.0%)
Time 1	Status 1 (20.0%)	64%	9%	9%	9%	9%	88%	3%	3%	3%	3%
	Status 2 (20.0%)	9%	64%	9%	9%	9%	3%	88%	3%	3%	3%
	Status 3 (20.0%)	9%	9%	64%	9%	9%	3%	3%	88%	3%	3%
	Status 4 (20.0%)	9%	9%	9%	64%	9%	3%	3%	3%	88%	3%
	Status 5 (20.0%)	9%	9%	9%	9%	64%	3%	3%	3%	3%	88%

Unequal Groups: Probability of Transition from Time 1 to Time 2		Time 2									
		High Transition Probability					Low Transition Probability				
		Status1 (12.2%)	Status2 (15.0%)	Status 3 (17.7%)	Status 4 (23.2%)	Status 5 (37.0%)	Status 1 (7.4%)	Status 2 (11.7%)	Status 3 (15.9%)	Status 4 (24.4%)	Status 5 (45.7%)
Time 1	Status 1 (5.0%)	64%	9%	9%	9%	9%	88%	3%	3%	3%	3%
	Status 2 (10.0%)	9%	64%	9%	9%	9%	3%	88%	3%	3%	3%
	Status 3 (15.0%)	9%	9%	64%	9%	9%	3%	3%	88%	3%	3%
	Status 4 (20.0%)	9%	9%	9%	64%	9%	3%	3%	3%	88%	3%
	Status 5 (50.0%)	9%	9%	9%	9%	64%	3%	3%	3%	3%	88%

Table A5: Five status item response probabilities for complex structure, unequal groups

	Status 1 (5%)	**Status 2** (10%)	**Status 3** (15%)	**Status 4** (20%)	**Status 5** (50%)
Item 1	0.50	0.50	0.10	0.10	0.10
Item 2	0.60	0.60	0.15	0.15	0.15
Item 3	0.70	0.70	0.20	0.20	0.20
Item 4	0.80	0.80	0.25	0.25	0.25
Item 5	0.90	0.90	0.30	0.30	0.30
Item 6	0.50	0.50	0.10	0.10	0.10
Item 7	0.60	0.60	0.15	0.15	0.15
Item 8	0.70	0.70	0.20	0.20	0.20
Item 9	0.80	0.80	0.25	0.25	0.25
Item 10	0.90	0.90	0.30	0.30	0.30
Item 11	0.50	0.50	0.10	0.50	0.10
Item 12	0.60	0.60	0.15	0.60	0.15
Item 13	0.70	0.70	0.20	0.70	0.20
Item 14	0.80	0.80	0.25	0.80	0.25
Item 15	0.90	0.90	0.30	0.90	0.30
Item 16	0.50	0.10	0.50	0.10	0.10
Item 17	0.60	0.15	0.60	0.15	0.15
Item 18	0.70	0.20	0.70	0.20	0.20
Item 19	0.80	0.25	0.80	0.25	0.25
Item 20	0.90	0.30	0.90	0.30	0.30
Item 21	0.50	0.10	0.50	0.10	0.10
Item 22	0.60	0.15	0.60	0.15	0.15
Item 23	0.70	0.20	0.70	0.20	0.20
Item 24	0.80	0.25	0.80	0.25	0.25
Item 25	0.90	0.30	0.90	0.30	0.30
Item 26	0.10	0.10	0.10	0.10	0.10
Item 27	0.20	0.20	0.20	0.20	0.20
Item 28	0.30	0.30	0.30	0.30	0.30
……					
Item 33	0.80	0.80	0.80	0.80	0.80
Item 34	0.90	0.90	0.90	0.90	0.90
Item 35	0.10	0.10	0.10	0.10	0.10
Item 36	0.20	0.20	0.20	0.20	0.20
……					
Item 50	0.70	0.70	0.70	0.70	0.70

Table A6: Quality of life questions ranked by variance across status (1 is highest variance)

	Rank
Physical Functioning	
Limited a little/a lot in:	
PF1: Vigorous activities	31
PF2: Moderate activities	9
PF3: Lifting or carrying groceries	10
PF4: Climbing several flights of stairs	16
PF5: Climbing one flight of stairs	5
PF6: Bending, kneeling, stooping	18
PF7: Walking more than one mile	8
PF8: Walking several blocks	**1**
PF9: Walking one block	**4**
PF10: Bathing or dressing self	32
Physical Health	
Problems due to physical health:	
PH1: Cut down time on work / other activities	12
PH2: Accomplish less than liked	7
PH3: Limited in work or other activities	**2**
PH4: Difficulty performing work or other activities	**3**
Body Pain	
BP1: Has body pain	33
BP2: Pain interfered with normal work	17
General Health	
GH1: Fair / poor general health	26
GH2: Health worse than 1 year ago	34
GH3: Get sick easier than other people	35
GH4: Not as healthy as anyone known	30
GH5: Expects health to get worse	36
GH6: Health is not excellent	21
Emotional Health	
Problems due to emotional health:	
EH1: Cut down time on work / other activities	24
EH2: Accomplish less than liked	20
EH3: Limited in work or other activities	22
Mental Health	
At least a little of the time feels:	
MH1: Very nervous	14
MH2: Down in the dumps, nothing will cheer up	15
MH3: Down hearted and blue	6
At least a good bit of the time feels:	
MH4: Is a not happy person	25
MH5: Is not calm and peaceful	18
Vitality	
At least a little of the time feels:	
V1: Not full of pep	28
V2: Worn out	27
V3: Tired	29
V4: Not a lot of energy	23
Social Functioning	
Physical/emotional interferes with social activity:	
SF1: Most or all of the time	11
SF2: To at least a slight extent	13

BIBLIOGRAPHY

1. Collins LM, Lanza ST. Latent class and latent transition analysis : with applications in the social behavioral, and health sciences. Hoboken, N.J.: Wiley; 2010.

2. Auerbach KJ, Collins LM. A multidimensional developmental model of alcohol use during emerging adulthood. J Stud Alcohol 2006;67(6):917-25.

3. Haydon AA, Herring AH, Prinstein MJ, Halpern CT. Beyond age at first sex: patterns of emerging sexual behavior in adolescence and young adulthood. The Journal of adolescent health : official publication of the Society for Adolescent Medicine 2012;50(5):456-63.

4. Mezuk B, Kendler KS. Examining variation in depressive symptoms over the life course: a latent class analysis. Psychological Medicine 2012:1-10.

5. Tisminetzky M, Bray BC, Miozzo R, Aupont O, McLaughlin TJ. Identifying symptom profiles of depression and anxiety in patients with an acute coronary syndrome using latent class and latent transition analysis. Int J Psychiatry Med 2011;42(2):195-210.

6. Jaeger S, Pfiffner C, Weiser P, Kilian R, Becker T, Langle G, et al. Adherence styles of schizophrenia patients identified by a latent class analysis of the Medication Adherence Rating Scale (MARS): A six-month follow-up study. Psychiatry Research 2012.

7. Ahn J, McCombs JS, Jung C, Croudace TJ, McDonnell D, Ascher-Svanum H, et al. Classifying patients by antipsychotic adherence patterns using

latent class analysis: characteristics of nonadherent groups in the California Medicaid (Medi-Cal) program. Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research 2008;11(1):48-56.

8. Velicer WF, Martin RA, Collins LM. Latent transition analysis for longitudinal data. Addiction 1996;91 Suppl:S197-209.

9. Bray BC, Lanza ST, Collins LM. Modeling Relations Among Discrete Developmental Processes: A General Approach to Associative Latent Transition Analysis. Struct Equ Modeling 2010;17(4):541-569.

10. Chung H, Park Y, Lanza ST. Latent transition analysis with covariates: pubertal timing and substance use behaviours in adolescent females. Stat Med 2005;24(18):2895-910.

11. Evers KE, Harlow LL, Redding CA, LaForge RG. Longitudinal changes in stages of change for condom use in women. Am J Health Promot 1998;13(1):19-25.

12. Cain AS, Epler AJ, Steinley D, Sher KJ. Stability and change in patterns of concerns related to eating, weight, and shape in young adult women: a latent transition analysis. J Abnorm Psychol 2010;119(2):255-67.

13. Windle M, Grunbaum JA, Elliott M, Tortolero SR, Berry S, Gilliland J, et al. Healthy passages. A multilevel, multimethod longitudinal study of adolescent health. Am J Prev Med 2004;27(2):164-72.

14. Dean N, Raftery AE. Latent Class Analysis Variable Selection. Annals of the Institute of Statistical Mathematics 2010;62(1):11-35.

15. Bartolucci FGE, Montanari. Pandolfi, Silvia. Item Selection by an extended Latent Class model: An application to nursing homes evaluation. Munich Personal RePEc Archive 2012.

16. Bartolucci F, Montanari GE, Pandolfi S. Dimensionality of the Latent Structure and Item Selection Via Latent Class Multidimensional IRT Models. Psychometrika 2012;77(4):782-802.

17. Raftery AE, Dean N. Variable Selection for Model-Based Clustering. Journal of the American Statistical Association 2006;101(473):168-178.

18. Lazarsfeld PF. The Logical and Mathematical Foundations of Latent Structure Analysis. Chapter 10 in Stouffer (1950). 1950.

19. Lazarsfeld PF. Some Latent Structures. 1950.

20. Lazarsfeld PF, Henry NW. Latent structure analysis. New York: Houghton, Mifflin; 1968.

21. Goodman LA. The analysis of systems of qualitative variables when some of the variables are unobservable : Part I. A modified latent structure approach. Chicago: University of Chicago; 1974.

22. Lanza ST, Collins LM, Lemmon DR, Schafer JL. PROC LCA: A SAS procedure for latent class analysis. Structural Equation Modeling-a Multidisciplinary Journal 2007;14(4):671-694.

23. Muthen LK, Muthen BO. Mplus User’s Guide Sixth Edition. 1998-2010.

24. Vermunt JK, Magidson J. Latent GOLD 4.0 User's Guide. In. Belmont, Massachusetts: Statistical Innovations Inc; 2005.

25. Langeheine R. Latent Variables Markov Models. In: Von Eye A, Clogg CC, editors. Latent variables analysis: Sage Publications, Inc; 1994.

26. Lanza ST, Collins LM. A new SAS procedure for latent transition analysis: transitions in dating and sexual risk behavior. Developmental Psychology 2008;44(2):446-56.

27. Roberts TJ, Ward SE. Using latent transition analysis in nursing research to explore change over time. Nursing Research 2011;60(1):73-9.

28. Lanza ST, Patrick ME, Maggs JL. Latent Transition Analysis: Benefits of a Latent Variable Approach to Modeling Transitions in Substance Use. J Drug Issues 2010;40(1):93-120.

29. Maldonado-Molina MM, Collins LM, Lanza ST, Prado G, Ramirez R, Canino G. Patterns of substance use onset among Hispanics in Puerto Rico and the United States. Addict Behav 2007;32(10):2432-7.

30. Lanza ST, Bray BC. Transitions in Drug Use among High-Risk Women: An Application of Latent Clsas and Latent Transition Analysis. Adv Appl Stat Sci 2010;3(2):203-235.

31. Collins LM. The measurement of dynamic latent variables in longitudinal aging research: quantifying adult development. Exp Aging Res 1991;17(1):13-20.

32. Badsberg JH. Model search in contingency tables by CoCo. Computational Statistics & Data Analysis 1992;1:251-256.

33. Kullback S, Leibler RA. On information and sufficiency. The Annals of Mathematical Statistics 1951;22(1):79-86.

34. Nylund KL MB, Asparouhov T. Deciding on the Number of Classes in Latent Class Analysis: A Monte Carlo Simulation Study. Structural Equation Modeling: An Interdisciplinary Journal 2007;14:535-569.

35. Tueller SJ, Drotar S, Lubke GH. Addressing the Problem of Switched Class Labels in Latent Variable Mixture Model Simulation Studies. Structural Equation Modeling: A Multidisciplinary Journal 2011;18(1):110-131.

36. Cohen J. A Coefficient of Agreement for Nominal Scales. Educational and Psychological Measurement 1960;20(1):37-46.

37. Hubert L, Arabie P. Comparing partitions. Journal of Classification 1985;2(1):193-218.

38. Krousel-Wood MA, Muntner P, Islam T, Morisky DE, Webber LS. Barriers to and determinants of medication adherence in hypertension management: perspective of the cohort study of medication adherence among older adults. Med Clin North Am 2009;93(3):753-769.

39. Morisky DE, Ang A, Krousel-Wood M, Ward HJ. Predictive validity of a medication adherence measure in an outpatient setting. J Clin Hypertens (Greenwich) 2008;10(5):348-54.

40. Muntner P, Joyce C, Holt E, He J, Morisky D, Webber LS, et al. Defining the minimal detectable change in scores on the eight-item Morisky Medication Adherence Scale. Ann Pharmacother 2011;45(5):569-75.

41. Choudhry NK, Shrank WH, Levin RL, Lee JL, Jan SA, Brookhart MA, et al. Measuring concurrent adherence to multiple related medications. Am J Manag Care 2009;15(7):457-64.

42. Krousel-Wood M, Joyce C, Holt E, Muntner P, Webber LS, Morisky DE, et al. Predictors of decline in medication adherence: results from the cohort study of medication adherence among older adults. Hypertension. 2011 Nov;58(5):804-10. Epub 2011 Oct 3.

43. Krousel-Wood M, Joyce C, Holt EW, Levitan EB, Dornelles A, Webber LS, et al. Development and evaluation of a self-report tool to predict low pharmacy refill adherence in elderly patients with uncontrolled hypertension. Pharmacotherapy 2013;33(8):798-811.

44. SF-36.org: A community for measuring health outcomes using SF tools. In; 2015.

45. Knight RG, Williams S, McGee R, Olaman S. Psychometric properties of the Centre for Epidemiologic Studies Depression Scale (CES-D) in a sample of women in middle life. Behav Res Ther 1997;35(4):373-380.

46. Krousel-Wood MA, Muntner P, Joyce CJ, Islam T, Stanley E, Holt EW, et al. Adverse effects of complementary and alternative medicine on antihypertensive medication adherence: findings from the cohort study of medication adherence among older adults. J Am Geriatr Soc 2010;58(1):54-61.

47. Krousel-Wood M, Joyce C, Holt E, Muntner P, Webber LS, Morisky DE, et al. Predictors of decline in medication adherence: results from the cohort study of medication adherence among older adults. Hypertension 2011;58(5):804-810.

48. Holt EW, Muntner P, Joyce CJ, Webber L, Krousel-Wood MA. Health-related quality of life and antihypertensive medication adherence among older adults. Age Ageing 2010;39(4):481-487.

49. Sherbourne CD, Stewart AL. The MOS social support survey. Soc Sci Med 1991;32(6):705-714.

50. Fernander AF, Duran RE, Saab PG, Llabre MM, Schneiderman N. Assessing the reliability and validity of the John Henry Active Coping Scale in an urban sample of African Americans and white Americans. Ethn Health 2003;8(2):147-161.

51. Cohen S, Kamarck T, Mermelstein R. A global measure of perceived stress. J Health Soc Behav 1983;24(4):385-396.