

Transitions in Depressive Symptoms After 10 Years of Follow-up Using PROC LTA in SAS[®] and Mplus

Seungyoung Hwang, Johns Hopkins University Bloomberg School of Public Health

ABSTRACT

PROC LTA is the most popular and powerful SAS procedure for latent transition analysis used throughout a wide variety of scientific disciplines. However, PROC LTA does not provide standard errors of the parameter estimates and thus constructing 95% confidence intervals around estimates is not possible.

In this paper, the author shows how to examine transitions in latent statuses of depressive symptoms after 10 years of follow-up using PROC LTA in SAS. The author then examines whether clinical characteristics predict membership in the different statuses and transitions between latent statuses over time using both SAS and Mplus. Mplus programming code is provided to compute standard errors of the parameter estimates. The dataset used is based on the Baltimore Epidemiologic Catchment Area Study (1-3).

This paper gently guides SAS and Mplus users—even those with limited experience in statistics or who have never used these software—through a step-by-step approach to using SAS and Mplus for latent transition analysis, and gives advice on how to interpret the results. This paper is suited to students who are beginning their study of social and behavioral health sciences and to professors and research professionals who are conducting research in the fields in epidemiology, clinical psychology, or health services research.

CASE STUDY: BALTIMORE EPIDEMIOLOGIC CATCHMENT AREA FOLLOW-UP

The Baltimore Epidemiologic Catchment Area (ECA) Follow-up Study was a population-based longitudinal survey designed to examine the incidence and prevalence of psychiatric disorders over the adult life course (1-3).

Depressive symptoms were assessed for 1,023 adult household residents twice, in 1994 and 2004. The variables are as follows:

- dsij is the j^{th} depressive symptom (1 = dysphoria; 2 = anhedonia; 3 = changes in appetite; 4 = sleep disturbances; 5 = restlessness; 6 = tiredness; 7 = worthlessness; 8 = thinking problems; 9 = suicidal thoughts or behavior) at time i (1 = year of 1994; 2 = year of 2004), (1 = present; 2 = absent).
- heart 1 = lifetime heart disease; 0 = no lifetime heart disease.
- diabetes 1 = lifetime diabetes; 0 = no lifetime diabetes.
- cancer 1 = lifetime cancer; 0 = no lifetime cancer.

Table 1 provides depressive symptoms in 2004 by depressive symptoms assessed in 1994.

Table 1. Depressive symptoms in 1994 vs. 2004 (N = 1,023)				
Had depressive symptom in 1994?	Yes	Yes	No	No
Had depressive symptom in 2004?	Yes	No	Yes	No
Dysphoria	147 (48%)	161 (52)	115 (16)	597 (84)
Anhedonia	38 (38)	62 (62)	46 (5)	870 (95)
Appetite	133 (51)	126 (49)	191 (25)	571 (75)
Sleep	128 (50)	127 (50)	165 (22)	602 (78)
Slow or Restless	27 (26)	77 (74)	70 (8)	847 (92)
Tired	69 (45)	86 (55)	129 (15)	737 (85)
Worthless	49 (38)	79 (62)	74 (8)	818 (92)
Thinking Problems	46 (38)	75 (62)	83 (9)	816 (91)
Suicidal Thoughts or Behavior	121 (49)	127 (51)	103 (13)	670 (87)

Note: This table was adapted from Hwang. 2015 (4). Numbers in parentheses are row percentages, stratified by presence of a depressive symptom in 1994.

In general, depressive symptoms that were absent in 1994 were likely also to be absent in 2004. Of particular interest is that a substantial proportion of adults had persistent depressive symptoms over the 10-year period.

To study nine depressive symptoms as a pattern or set of latent variables in contrast to a focus on individual symptoms, I applied the latent transition model (5, 6). This model provides for simultaneous estimation: (i) latent status prevalences, (ii) item-response probabilities, and (iii) transition probabilities. In addition, latent transition analysis appears to be an informative model for examining predictors of transitions from one status of symptoms to another.

The 3-status model is preferred over 2- and 4-status models based on the BIC ($BIC_3 = 3730.53$, $BIC_4 = 3753.95$, and $BIC_2 = 4091.71$) and model interpretability (7). Item-response probabilities given a specific latent status in 1994 and in 2004 from the 3-status model of depressive symptoms appear in the **Figure**. The y-axis represents the probability of having a depressive symptom, while the x-axis refers to indicator variables used for the LTA. The six lines represent the depressive symptom patterns for the three latent statuses in 1994 and in 2004. Since the lines were parallel to each other, it indicates three statuses with different degree (i.e., severe depression, mild depression, no depression), rather than type of depression or noticeable symptom. Adults in the “severe depression” category were likely to report that they have experienced all nine symptoms (probabilities were all greater than 50%). Adults in the “nondepressed” category were likely to report that they have never experienced depressive symptoms (probabilities were all less than 13%). Patterns of depressive symptoms in the “mild depression” category behaved somewhere between the two latent statuses above. The patterns of depressive symptoms were similar at the two time points.

Figure. Item-response probabilities

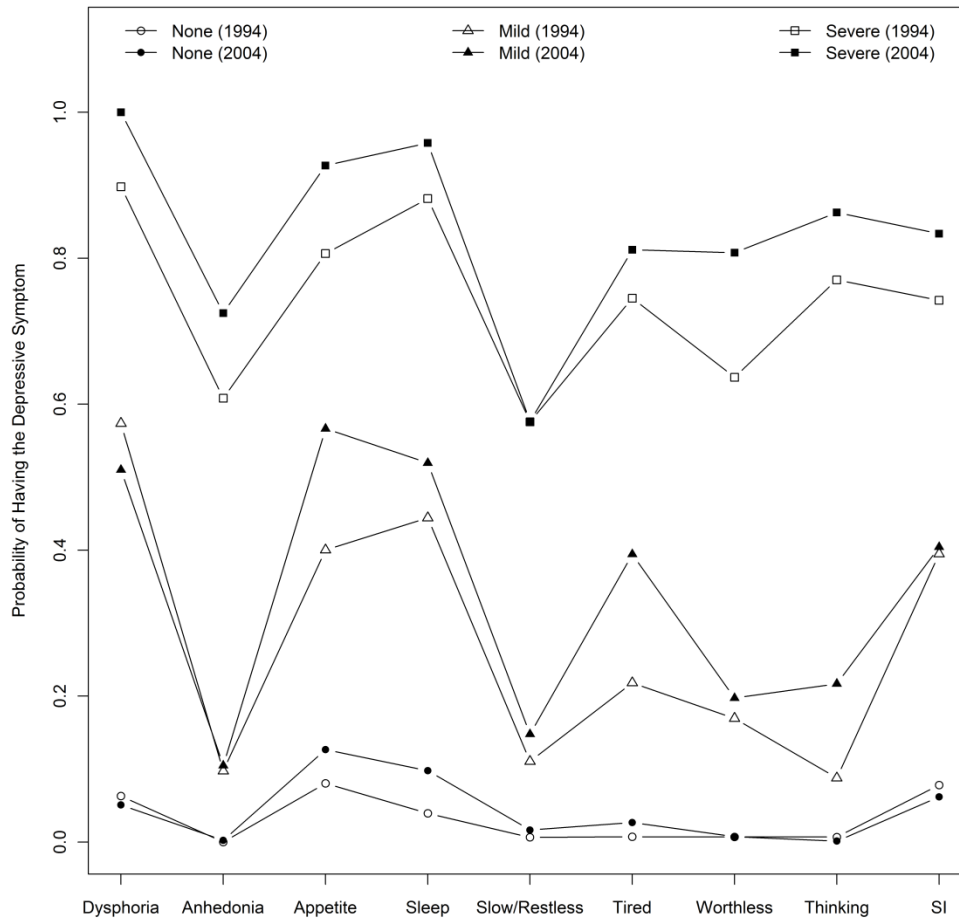


Table 2. Prevalence of latent statuses and transition probabilities in latent status membership.

Category of Depression in 1994 (Estimated Prevalence)	Category of Depression in 2004 (Estimated Prevalence)		
	None (57%)	Mild (33%)	Severe (9%)
None (62%)	0.77*	0.19	0.04
Mild (27%)	0.33	0.62	0.05
Severe (11%)	0.07	0.41	0.52

*Diagonal transition probabilities in bold to facilitate interpretation.

The prevalences of the three latent statuses at the two time points and transition probability estimates are shown in **Table 2**. Parameter restrictions were imposed so that the item-response probabilities are equal across the two times and the meaning of the latent statuses remains constant over time (6). In 1994, the nondepressed latent status (62%) was the most prevalent, followed by the mild and severe depression statuses (27% and 11%). The overall prevalences were similar in 2004.

The probability that adults were in the same status at follow-up as at baseline was 0.77 for “nondepressed,” 0.62 for “mild depression,” and 0.52 for “severe depression.” Adults in the severe depression latent status at baseline were the most likely to transition to another status across time (41% to mild depression and 7% to nondepressed status). Of particular interest is that 23% of the persons in the nondepressed latent status at baseline transitioned to the mild or severe depression statuses at follow-up.

Next, I examined whether clinical characteristics predicted membership in the different statuses at baseline. The three covariates were self-reported lifetime heart disease, lifetime diabetes, and lifetime cancer (results are shown below in **Table 3**).

Table 3. Odds ratio with 95% confidence intervals for predictors of latent status in 1994.

Covariate	Latent status in 1994		
	None	Mild	Severe
heart disease	Reference	1.83 (0.97 - 3.45)	1.83 (0.96 - 3.49)
diabetes	Reference	1.25 (0.69 - 2.25)	1.36 (0.68 - 2.75)
cancer	Reference	1.13 (0.52 - 2.49)	1.27 (0.50 - 3.19)

I found that adults who reported lifetime heart disease, diabetes, or cancer were 83%, 36%, and 27% more likely, respectively, to be in the severe depression latent status relative to the nondepressed status. Similarly, adults with lifetime heart disease or diabetes were 83% and 25% more likely to be in the mild depression latent status relative to the nondepressed status.

Since PROC LTA in SAS does not provide standard errors of the parameter estimates, 95% confidence intervals and associated p-values could not be computed. That is, I am not sure whether the three clinical characteristics above were significant or not. Mplus programming shows that all the six 95% confidence intervals included the null value of 1, indicating that lifetime heart disease, diabetes, and cancer were not significant predictors of 1994 latent status membership.

Additionally, I examined whether the clinical characteristics are significant predictors of transitions between latent statuses of depressive symptoms. Estimates of odds ratios for the transitions from the nondepressed latent status into the severe or mild depression statuses appear in **Table 4**.

Table 4. Odds ratios reflecting the effects of medical conditions on transition from nondepressed latent status in 1994 to severe or mild depression latent statuses in 2004 (N = 1,023).

Covariate	Latent status in 2004		
	None	Mild	Severe
heart disease	Reference	3.44 (1.51 - 7.87)	1.36 (0.26 - 7.19)
diabetes	Reference	1.46 (0.62 - 3.41)	0.52 (0.06 - 4.76)
cancer	Reference	2.22 (0.69 - 7.21)	1.04 (0.07 - 16.07)

The association of covariates with change in latent status between 1994 and 2004, as measured using the odds ratio, is given in **Table 4**. Lifetime diabetes and cancer were not significantly associated with transitions to the mild depression categories compared with transitions to the nondepressed category (the associated confidence intervals included null). However, adults who reported having lifetime heart disease were more than three times as likely to make a transition to the mild depression category as to the nondepressed category (odds ratio, 3.44; 95% confidence interval, 1.51-7.87). **Table 4** shows that there were no statistically significant associations of specific covariates with transition to the severe

depression category compared with transition to the nondepressed category. Once again, PROC LTA in SAS does not provide standard errors of the parameter estimates. 95% confidence intervals were computed using Mplus programming.

DISCUSSION

Recent decades have seen tremendous applications of latent transition analysis in various academic fields. However, few have reported step-by-step instructions to perform each technique in both SAS and Mplus at the same time. In this paper, I examined whether clinical characteristics were more or less likely to predict membership in the different statuses of depressive symptoms and predict transitions between latent statuses over time using SAS (PROC LTA procedure) and Mplus. This paper provides the example SAS and Mplus programming codes for exactly the same procedures. The emphasis is on statistical tools and model interpretations which are easily applicable to social and behavioral health studies.

STATISTICAL ANALYSIS USING SAS

Here is the code for analyzing the latent transition models with PROC LTA for the Baltimore ECA data:

```
* ===== Figure ===== *;
PROC LTA DATA=SASData;
  nstatus 3;
  ntimes 2;
  items    ds11-ds19
           ds21-ds29;
  categories 2 2 2 2 2 2 2 2 2;
  *measurement times;
  seed 592667;
RUN;
```

```
* ===== Table 2 ===== *;
PROC LTA DATA=SASData;
  nstatus 3;
  ntimes 2;
  items    ds11-ds19
           ds21-ds29;
  categories 2 2 2 2 2 2 2 2 2;
  measurement times;
  seed 592667;
RUN;
```

```
* ===== Table 3 ===== *;
PROC LTA DATA=SASData;
  nstatus 3;
  ntimes 2;
  items    ds11-ds19
           ds21-ds29;
  categories 2 2 2 2 2 2 2 2 2;
  covariates1 heart;
  *covariates1 diabetes;
  *covariates1 cancer;
  referencel 1; /* reference = nondepressed */
```

```

    measurement times;
    seed 592667;
RUN;

* ===== Table 4 ===== *;
PROC LTA DATA=SASData;
    nstatus 3;
    ntimes 2;
    items    ds11-ds19
            ds21-ds29;
    categories 2 2 2 2 2 2 2 2 2;
    covariates1 heart;
    *covariates1 diabetes;
    *covariates1 cancer;
    reference1 1; /* reference = nondepressed */
    covariates2 heart;
    *covariates2 diabetes;
    *covariates2 cancer;
    reference2 1; /* reference = nondepressed */
    measurement times;
    beta prior = 1;
    seed 592667;
RUN;

```

The BETA PRIOR statement was used in **Table 4** to invoke a stabilizing prior distribution on the β parameters so that it solves most sparseness-related estimation problems (8).

STATISTICAL ANALYSIS USING MPLUS

Here is the code for analyzing the latent transition models with Mplus for the Baltimore ECA data:

```

* ===== Figure ===== *;
TITLE: Baltimore ECA Study - LTA of Depressive Symptoms
DATA: File = MplusData.txt;
VARIABLE: Names = ds11-ds19 ds21-ds29 heart diabetes cancer;
Usevariables = ds11-ds19 ds21-ds29;
Classes = C1(3) C2(3);
CATEGORICAL = ds11-ds19 ds21-ds29;
MISSING = ALL(-99);
ANALYSIS:
    TYPE = mixture;
    STARTS = 1000 250;
Model:
    %OVERALL%
    C2 ON C1;
MODEL C1:
    %C1#1%
    [ds11$1-ds19$1];
    %C1#2%
    [ds11$1-ds19$1];
    %C1#3%
    [ds11$1-ds19$1];
MODEL C2:
    %C2#1%
    [ds21$1-ds29$1];

```

```
%C2#2%
[ds21$1-ds29$1];
%C2#3%
[ds21$1-ds29$1];
OUTPUT: tech1 tech8;
```

```
* ===== Table 2 ===== *
```

```
TITLE: Baltimore ECA Study - LTA of Depressive Symptoms
DATA: File = MplusData.txt;
VARIABLE: Names = ds11-ds19 ds21-ds29 heart diabetes cancer;
Usevariables = ds11-ds19 ds21-ds29;
Classes = C1(3) C2(3);
CATEGORICAL = ds11-ds19 ds21-ds29;
MISSING = ALL(-99);
ANALYSIS:
  TYPE = mixture;
  STARTS = 1000 250;
```

```
Model:
```

```
%OVERALL%
C2 ON C1;
MODEL C1:
  %C1#1%
  [ds11$1-ds19$1](1-9);
  %C1#2%
  [ds11$1-ds19$1](10-18);
  %C1#3%
  [ds11$1-ds19$1](19-27);
MODEL C2:
  %C2#1%
  [ds21$1-ds29$1](1-9);
  %C2#2%
  [ds21$1-ds29$1](10-18);
  %C2#3%
  [ds21$1-ds29$1](19-27);
OUTPUT: tech1 tech8;
```

```
* ===== Table 3 ===== *
```

```
TITLE: Baltimore ECA Study - LTA of Depressive Symptoms
DATA: File = MplusData.txt;
VARIABLE: Names = ds11-ds19 ds21-ds29 heart diabetes cancer;
Usevariables = ds11-ds19 ds21-ds29 heart;
!Usevariables = ds11-ds19 ds21-ds29 diabetes;
!Usevariables = ds11-ds19 ds21-ds29 cancer;
Classes = C1(3) C2(3);
CATEGORICAL = ds11-ds19 ds21-ds29;
MISSING = ALL(-99);
ANALYSIS:
  TYPE = mixture;
  STARTS = 1000 250;
```

```
Model:
```

```
%OVERALL%
C2 ON C1;
C1 ON heart;
!C1 ON diabetes;
!C1 ON cancer;
```

```

MODEL C1:
  %C1#1%
  [ds11$1-ds19$1](1-9);
  %C1#2%
  [ds11$1-ds19$1](10-18);
  %C1#3%
  [ds11$1-ds19$1](19-27);

```

```

MODEL C2:
  %C2#1%
  [ds21$1-ds29$1](1-9);
  %C2#2%
  [ds21$1-ds29$1](10-18);
  %C2#3%
  [ds21$1-ds29$1](19-27);

```

```

OUTPUT: tech1 tech8;

```

```

* ===== Table 4 ===== *;

```

```

TITLE: Baltimore ECA Study - LTA of Depressive Symptoms

```

```

DATA: FILE = MplusData.txt;

```

```

VARIABLE: Names = ds11-ds19 ds21-ds29 heart diabetes cancer;

```

```

Usevariables = ds11-ds19 ds21-ds29 heart;

```

```

!Usevariables = ds11-ds19 ds21-ds29 diabetes;

```

```

!Usevariables = ds11-ds19 ds21-ds29 cancer;

```

```

Classes = C1(3) C2(3);

```

```

CATEGORICAL = ds11-ds19 ds21-ds29;

```

```

MISSING = all(-99);

```

```

ANALYSIS:

```

```

TYPE = mixture;

```

```

STARTS = 1000 250;

```

```

Model:

```

```

%OVERALL%
[C2#1] (a1);
[C2#2] (a2);
C2#1 ON C1#1 (b11);
C2#1 ON C1#2 (b12);
C2#2 ON C1#1 (b21);
C2#2 ON C1#2 (b22);
C1 ON heart;
!C1 ON diabetes;
!C1 ON cancer;

```

```

MODEL C1:
  %C1#1%
  [ds11$1-ds19$1*1](1-9);
  C2#1 ON heart (g11);
  C2#2 ON heart (g21);
  !C2#1 ON diabetes (g11);
  !C2#2 ON diabetes (g21);
  !C2#1 ON cancer (g11);
  !C2#2 ON cancer (g21);
  %C1#2%
  [ds11$1-ds19$1*0](10-18);
  C2#1 ON heart (g12);
  C2#2 ON heart (g22);
  !C2#1 ON diabetes (g12);
  !C2#2 ON diabetes (g22);
  !C2#1 ON cancer (g12);

```



```

!C2#2 ON cancer (g22);
%C1#3%
[ds11$1-ds19$1*-1](19-27);
C2#1 ON heart (g13);
C2#2 ON heart (g23);
!C2#1 ON diabetes (g13);
!C2#2 ON diabetes (g23);
!C2#1 ON cancer (g13);
!C2#2 ON cancer (g23);
MODEL C2:
%C2#1%
[ds21$1-ds29$1*1](1-9);
%C2#2%
[ds21$1-ds29$1*0](10-18);
%C2#3%
[ds21$1-ds29$1*-1](19-27);
MODEL CONSTRAINT:
new(log11_x0 log12_x0 log21_x0 log22_x0 log31_x0 log32_x0
p11_x0 p12_x0 p13_x0
p21_x0 p22_x0 p23_x0
p31_x0 p32_x0 p33_x0
log11_x1 log12_x1 log21_x1 log22_x1 log31_x1 log32_x1
p11_x1 p12_x1 p13_x1
p21_x1 p22_x1 p23_x1
p31_x1 p32_x1 p33_x1
lo_smn_s lo_smn_m
lo_snm_s lo_snm_m
lo_msn_s lo_msn_m
lo_mns_s lo_mns_m
lo_nsm_s lo_nsm_m
lo_nms_s lo_nms_m);
log11_x0 = a1 + b11;
log12_x0 = a2 + b21;
log21_x0 = a1 + b12;
log22_x0 = a2 + b22;
log31_x0 = a1;
log32_x0 = a2;
p11_x0 = exp(log11_x0) / (exp(log11_x0) + exp(log12_x0) + 1);
p12_x0 = exp(log12_x0) / (exp(log11_x0) + exp(log12_x0) + 1);
p13_x0 = 1 / (exp(log11_x0) + exp(log12_x0) + 1);
p21_x0 = exp(log21_x0) / (exp(log21_x0) + exp(log22_x0) + 1);
p22_x0 = exp(log22_x0) / (exp(log21_x0) + exp(log22_x0) + 1);
p23_x0 = 1 / (exp(log21_x0) + exp(log22_x0) + 1);
p31_x0 = exp(log31_x0) / (exp(log31_x0) + exp(log32_x0) + 1);
p32_x0 = exp(log32_x0) / (exp(log31_x0) + exp(log32_x0) + 1);
p33_x0 = 1 / (exp(log31_x0) + exp(log32_x0) + 1);
log11_x1 = a1 + b11 + g11;
log12_x1 = a2 + b21 + g21;
log21_x1 = a1 + b12 + g12;
log22_x1 = a2 + b22 + g22;
log31_x1 = a1 + g13;
log32_x1 = a2 + g23;
p11_x1 = exp(log11_x1) / (exp(log11_x1) + exp(log12_x1) + 1);
p12_x1 = exp(log12_x1) / (exp(log11_x1) + exp(log12_x1) + 1);
p13_x1 = 1 / (exp(log11_x1) + exp(log12_x1) + 1);
p21_x1 = exp(log21_x1) / (exp(log21_x1) + exp(log22_x1) + 1);
p22_x1 = exp(log22_x1) / (exp(log21_x1) + exp(log22_x1) + 1);

```

```

p23_x1 = 1 / (exp(log21_x1) + exp(log22_x1) + 1);
p31_x1 = exp(log31_x1) / (exp(log31_x1) + exp(log32_x1) + 1);
p32_x1 = exp(log32_x1) / (exp(log31_x1) + exp(log32_x1) + 1);
p33_x1 = 1 / (exp(log31_x1) + exp(log32_x1) + 1);
lo_smn_s = log((p31_x1 / p33_x1) / (p31_x0 / p33_x0));
lo_smn_m = log((p32_x1 / p33_x1) / (p32_x0 / p33_x0));
lo_snm_s = log((p21_x1 / p22_x1) / (p21_x0 / p22_x0));
lo_snm_m = log((p23_x1 / p22_x1) / (p23_x0 / p22_x0));
lo_msn_s = log((p32_x1 / p33_x1) / (p32_x0 / p33_x0));
lo_msn_m = log((p31_x1 / p33_x1) / (p31_x0 / p33_x0));
lo_mns_s = log((p23_x1 / p22_x1) / (p23_x0 / p22_x0));
lo_mns_m = log((p21_x1 / p22_x1) / (p21_x0 / p22_x0));
lo_nsm_s = log((p12_x1 / p11_x1) / (p12_x0 / p11_x0));
lo_nsm_m = log((p13_x1 / p11_x1) / (p13_x0 / p11_x0));
lo_nms_s = log((p13_x1 / p11_x1) / (p13_x0 / p11_x0));
lo_nms_m = log((p12_x1 / p11_x1) / (p12_x0 / p11_x0));

```

OUTPUT: sampstat tech1;

The order of the statuses is arbitrary. Thus all 6 possibilities (i.e., severe-mild-none, severe-none-mild, mild-severe-none, mild-none-severe, none-severe-mild, and none-mild-severe) were considered in **Table 4** to calculate log odds ratios of transition from nondepressed latent status in 1994 to severe or mild depression latent statuses in 2004.

REFERENCES

1. Eaton WW, Regier DA, Locke BZ, et al: The Epidemiologic Catchment Area Program of the National Institute of Mental Health. Public health reports 1981; 96:319-325
2. Regier DA, Myers JK, Kramer M, et al: The NIMH Epidemiologic Catchment Area program. Historical context, major objectives, and study population characteristics. Archives of general psychiatry 1984; 41:934-941
3. Eaton WW, Kalaydjian A, Scharfstein DO, et al: Prevalence and incidence of depressive disorder: the Baltimore ECA follow-up, 1981-2004. Acta psychiatrica Scandinavica 2007; 116:182-188
4. Hwang S: Transitions in Depressive Symptoms After 10 Years of Follow-up Using PROC LTA, Orlando, FL, PharmaSUG, 2015
5. Lanza ST, Collins LM: A new SAS procedure for latent transition analysis: transitions in dating and sexual risk behavior. Developmental psychology 2008; 44:446-456
6. Collins LM, Lanza ST: Latent Class and Latent Transition Analysis: With Applications in the Social, Behavioral, and Health Sciences, Hoboken, NJ, John Wiley & Sons, 2010
7. Schwarz G: Estimating Dimension of a Model. Ann Stat 1978; 6:461-464
8. Lanza ST, Dziak JJ, Huang L, et al: PROC LCA & PROC LTA User's Guide Version 1.3.2, University Park, PA, The Pennsylvania State University, The Methodology Center 2015

ACKNOWLEDGMENTS

I took the opportunity to use data from the Baltimore Epidemiologic Catchment Area (ECA) study. The Baltimore ECA is supported by a grant from the National Institute on Drug Abuse, United States (PI: William W. Eaton, Ph.D.; R01 DA009897).

This work is dedicated to my parents, Jungja Han and Okgil Hwang. All I have and will accomplish are only possible due to their love and sacrifices. As always, most special thanks to Yun Kyoung Ryu for her advice and encouragement in getting me to finish this paper. She has been incredibly generous with her time in reviewing the draft and making many helpful suggestions. All remaining errors, omissions, and weaknesses are my sole responsibility.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Seungyoung Hwang, MS, MSE, GStat®

Biostatistician, Department of Mental Health,

DrPH Student, Department of Health Policy and Management

Bloomberg School of Public Health

Johns Hopkins University

624 North Broadway, Hampton House 880

Baltimore, MD 21205

Phone: 410-440-2040

Email: shwang25@jhu.edu

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration. Other brand and product names are trademarks of their respective companies.