BRANDY SINCO'S APPENDICES AND REFERENCES FOR:

"USING PATH ANALYSIS TO TEST A HYPOTHESIS ON THE THEORY OF CHANGE IN HEMOGLOBIN A1C (HBA1C) AMONG CLIENTS IN A CULTURALLY TAILORED DIABETES INTERVENTION FOR AFRICAN AMERICANS AND LATINOS"

PRESENTED TO APHA STATISTICS SECTION ON 10/30/2012

Table of Contents

Appendix A: SAS Code for SEM Power Calculation

```
/* Power calculation for REACH SEM paper */
ods html path="c:\tempfiles";
Proc IML;
alpha=.05; /* significance level */
rmsea0=.05; /* null hyp value */
rmseaa=.08; /* alt hyp value */
d=58;/* degrees of freedom */
n=188; /* sample size */
ncp0=(n-1)*d*(rmsea0**2);
ncpa=(n-1)*d*(rmseaa**2);

if rmsea0<rmseaa then
do;
 cval = cinv(1-alpha, d, ncp0);
 power = 1 - probchi(cval, d, ncpa);
end;

if rmsea0>rmseaa then
do;
 cval = cinv(alpha, d, ncp0);
 power = probchi(cval, d, ncpa);
end;

Print rmsea0 rmseaa alpha d n cval[Format=10.3] power[Format=5.3] ncp0 ncpa;
quit;
ods html close;
```

Appendix B: SAS Code for Assessment of Multivariate Normality

```
ods html path="c:\tempfiles";
ods graphics on;
/* Obtain Mahalanobis distance from principle components */
proc princomp data=SEMData std out=out_all outstat=outstat noprint;
   var Post_H1c Post_V110 Post_SMB Post_PAID Post_SelfEff_MH
       Pre_H1c Pre_V110 Pre_SMB Pre_PAID Pre_SelfEff_MH
       RaceInd GenderInd Age_BLIW Cohort2 Randomization2 HSGrad
       TotalClasses JrnyHlthGrp SelfMgtGrp FHADrVisitsBin;
run;

/* Compute squares of principle components */
data mahalanobis_all;
  set out_all;
  d2 = uss(of prin:);
 run;

/* Histogram and QQ Plot of Mahalanobis Distance */
Title 'Multivariate Diagnostics - All';
Proc Univariate Data=Mahalanobis_all;
 Var  d2;
 Label  mahalanobis_distance_to_mean="Squared Distance";
 Symbol1 V=Dot;

 Histogram / Gamma(Alpha=10 Sigma=2 Theta=0 Fill)
 CFill=Blue Name="Data W/Post";
 Inset Gamma(KSD KSDPval) / Header='Goodness of Fit Test'
 Position=(95,95) RefPoint=TR;

 QQPlot / Gamma(Alpha=10 Sigma=2 Theta=0 L=1)
 PctlMinor PCTLSCALE Name="Data W/Post";
Run;
/* p = 0.2500 Kolmogorov */
ods graphics off;
ods html close;
```

Appendix C: Derivation of Multivariate Skewness and Kurtosis Formulas for Bernoulli Random Variables

The skewness of a Bernoulli variable, X, with probability of event, π, can be computed from:

- $\mu_3 = E\left(X-\pi\right)^3 = \sum_{i=0}^{3}\binom{3}{i}E\left(X^i\right)\left(-\pi\right)^{3-i}.$
- $\mu_3 = -\binom{3}{0}\pi^3 + \binom{3}{1}E\left(X\right)\pi^2 - \binom{3}{2}E\left(X^2\right)\pi + \binom{3}{3}E\left(X^3\right).$
- $\mu_3 = -\pi^3 + 3\pi^3 - 3\pi^2 + \pi.$

✓ To standardize μ_3, divide by variance$^{3/2}$, $[\pi(1-\pi)]^{3/2}$, and the standardized skewness is $\frac{\left(1-2\pi\right)}{\sqrt{\pi\left(1-\pi\right)}}$.

The kurtosis of a binary variable, X, with probability of event, π, can be computed from

- $\mu_4 = E\left(X-\pi\right)^4 = \sum_{i=0}^{4}\binom{4}{i}E\left(X^i\right)\left(-\pi\right)^{4-i}.$
- $\mu_4 = \binom{4}{0}\pi^4 - \binom{4}{1}E\left(X\right)\pi^3 + \binom{4}{2}E\left(X^2\right)\pi^2 - \binom{4}{3}E\left(X^3\right)\pi + \binom{4}{4}E\left(X^4\right).$
- $\mu_4 = \pi^4 - 4\pi\pi^3 + 6\pi\pi^2 - 4\pi\pi + \pi = -3\pi^4 + 6\pi^3 - 4\pi^2 + \pi.$
- To standardize, divide by $[\pi(1-\pi)]^2$, subtract 3, and the standardized kurtosis is $\frac{\left(3\pi^2-3\pi+1\right)}{\pi\left(1-\pi\right)} - 3\frac{\pi\left(1-\pi\right)}{\pi\left(1-\pi\right)} = \frac{\left(3\pi^2-3\pi+1\right)+\left(3\pi^2-3\pi\right)}{\pi\left(1-\pi\right)} = \frac{1-6\pi\left(1-\pi\right)}{\pi\left(1-\pi\right)}.$

Appendix D: SAS Code for Missing Data Analysis

```
*** Step 1: Impute with Proc MI ***;
*** Include WithFlag (Withdrawal as covariate) ***;
Proc MI Data=TestMissing NImpute=20 seed=11292011 out=ImputeSEM;
Var WithFlag Age_BLIW Genderind RaceInd Cohort2 Randomization2 HSGrad
TotalClasses FHADrVisitsBinJrnyHlthGrp SelfMgtGrp
Pre_V110 Pre_SMB Pre_PAID Pre_SelfEff_MH
Post_V110 Post_SMB Post_PAID Post_SelfEff_MH
Pre_H1cPost_H1c ;
Run;

*** Step 2: Combine estimates from Proc CALIS ***;
Proc Calis Covariance Method=ML VARDEF=N Data=ImputeSEM;
LINEQS … list line equations;
VARIANCE List variances;
VAR List variables;
by _IMPUTATION_;
ODS Output LINEQSEQ=LINEQSEQ;
Run;
```

Appendix E: SAS Proc CALIS Code for SEM Analysis

```
ods html newfile=proc path="c:\tempfiles";
ods graphics on;
Proc Calis Covariance Residual Modification Kurtosis Method=FIML plots=residuals
Data=SEMData;
LINEQS
Post_V110 = P_PostV110_PreV110 Pre_V110 +
  P_PostV110_TotalClasses TotalClasses + P_PostV110_FHADrVisitsBin FHADrVisitsBin +
  P_PostV110_JrnyHlthGrp JrnyHlthGrp + P_PostV110_SelfMgtGrp SelfMgtGrp +
  E_PostV110,

Post_PAID = P_PostPAID_PrePAID Pre_PAID +
  P_PostPAID_TotalClasses TotalClasses + P_PostPAID_FHADrVisitsBin FHADrVisitsBin +
  P_PostPAID_JrnyHlthGrp JrnyHlthGrp + P_PostPAID_SelfMgtGrp SelfMgtGrp +
  E_PostPAID,

Post_SelfEff_MH = P_PostSelfEff_MH_PreSelfEff_MH Pre_SelfEff_MH +
  P_PostSelfEff_MH_TotalClasses TotalClasses + P_PostSelfEff_FHADrVisitsBin
FHADrVisitsBin +
  P_PostSelfEff_JrnyHlthGrp JrnyHlthGrp + P_PostSelfEff_SelfMgtGrp SelfMgtGrp +
  E_PostSelfEff_MH,

Post_SMB = P_PostSMB_PreSMB Pre_SMB + P_PostSMB_PostV110 Post_V110 +
  P_PostSMB_PostPAID Post_PAID + P_PostSMB_PostSelfEff Post_SelfEff_MH +
  E_PostSMB,

Post_H1c = P_PostH1c_PreH1c Pre_H1c + P_PostH1c_PostSMB Post_SMB +
  P_PostH1c_Age Age_BLIW + P_PostH1c_Gender Genderind +
  P_PostH1c_RaceInd RaceInd + P_PostH1c_Cohort2 Cohort2 +
  P_PostH1c_Rand Randomization2 + P_PostH1c_HSGrad HSGrad +
   E_Post_H1c;

VARIANCE
E_PostV110 = Var_PostV110,
E_PostSMB = Var_PostSMB,
E_PostPAID = Var_PostPAID,
E_PostSelfEff_MH = Var_PostSelfEff_MH,
E_Post_H1c = Var_PostH1c,
Pre_H1c = Var_PreH1c,
Pre_V110 = Var_PreV110,
Pre_SMB = Var_PreSMB,
Pre_PAID = Var_PrePAID,
Pre_SelfEff_MH = Var_PreSelfEff_MH,
Age_BLIW = Var_AgeBLIW,
GenderInd = Var_GenderInd,
```

```
RaceInd = Var_RaceInd,
HSGrad = Var_HSGrad,
Cohort2 = Var_Cohort2,
Randomization2 = Var_Randomization2,
JrnyHlthGrp = Var_JrnyHlthGrp,
SelfMgtGrp = Var_SelfMgtGrp,
FHADrVisitsBin = Var_FHADrVisitsBin,
TotalClasses = Var_TotalClasses;

VAR
Pre_H1c Pre_V110 Pre_SMB Pre_PAID Pre_SelfEff_MH
Age_BLIW Genderind RaceInd Cohort2 Randomization2 HSGrad
TotalClasses FHADrVisitsBin JrnyHlthGrp SelfMgtGrp
Post_H1c Post_V110 Post_SMB Post_PAID Post_SelfEff_MH;
ODS Output LINEQSEQ=LINEQSEQ SQMultCorr=SQMultCorr Fit=Fit;
Run;
ods graphics off;
ods html close;
```

REFERENCES

1. Center for Statistical Consultation And Research, University of Michigan. Applied structural equation modeling, may 10 - 14. 2010.

2. Kline RB. *Principles and Practice of Structural Equation Modeling.* 3rd ed. New York: The Guilford Press; 2011.

3. Johnson RA, Wichern DW. *Applied Multivariate Statistical Analysis.* 5th ed. New Jersey: Prentice Hall; 2002.

4. Kenny DA. Mediation. http://davidakenny.net/cm/mediate.htm. Accessed 1/6/2012.

5. Bollen KA. *Structural Equations with Latent Variables.* New York, NY: John Wiley & Sons, Inc.; 1989.

6. Joreskog KG, Sorbom D. Recent developments in structural equation modeling. *Journal of Marketing Research*. 1982;19(4):404-416.

7. SAS Institute. The CALIS procedure. http://support.sas.com/documentation/cdl/en/statug/63347/HTML/default/viewer.htm#calis_toc.htm. Accessed 7/26, 2011.

8. Gao S, Mokhtarian PL, Johnston RA. Non-normality of data in structural equation models. *Transportation Research Board's 87th Annual Meeting.*

9. Joreskog KG, Sorbom D. LISREL: Structural equation modeling. *Scientific Software International.* Chicago, IL:1996;8.

10. Hatcher L. *A Step-by-Step Guide to using SAS for Factor Analysis and Structural Equation Modeling.* Cary, NC: SAS Institute; 2003.

11. Suhr DD. Exploratory or confirmatory factor analysis? *SAS Global Forum Proceedings.* 2006.

12. Estep D. Course notes for stat 560, Colorado State University. 2011;1:224-248.

13. West BT, Welch KB, Galecki AT. *Linear Mixed Models: A Practical Guide using Statistical Software* . Boca Raton, FL: Chapman Hall / CRC Press; 2007.

14. Bentler PM. Comparative fit indices in structural models. *Psychological Bulletin*. 1990;107(2):238-246.

15. Steiger JH. Structural model evaluation and modification: An interval estimation approach. *Multivariate Behavioral Research*. 1990;25(2):173-180.

16. Hu L, Bentler PM. Cutoff criteria for fit indices in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*. 1999;6:1-55.

17. Loehlin JC. *Latent Variable Models: An Introduction to Factor, Path, and Structural Analysis.* 2nd ed. Hillsdale, NJ: Lawrence Erlbaum Associates; 1992.

18. MacCullum RC, Browne MW, Sugawara HM. Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods*. 1996;1:130-149.

19. Fitzgerald JT, Davis WK, Connell CM, Hess GE, Funnell MM, Hiss RG. Development and validation of the diabetes care profile. *Eval Health Prof.* 1996;19:208-230.

20. Williams GC, Freedman ZR, Deci EL. Supporting autonomy to motivate patients with diabetes for glucose control. *Diabetes Care*. 1998;21:1644-1651.

21. Polonsky WH, Anderson BJ, Lohrer PA, et al. Assessment of diabetes-related distress. *Diabetes Care*. 1995;18(6):754-760.

22. Polonsky WH, Welch G. Listening to our patients' concerns: Understanding and addressing diabetes-specific emotional distress. *Diabetes Care*. 1996;9:8-11.

23. Toobert DJ, Hampson SE, Glasgow RE. The summary of diabetes self-care activities measure. *Diabetes Care*. 2000;23(7):943-950.

24. Agresti A, Mehta CR, Patel NR. Exact inference for contingency tables with ordered categories. *JASA*. 1990;85:453-458.

25. Anupama N, Watts D. Exact methods in the NPAR1WAY procedure. *SAS Users Group International.* 1996.

26. Balanda KP, MacGillivray HL. Kurtosis: A critical review. *The American Statistician*. 1988;42(2):111-119.

27. Mardia KV. Measures of multivariate skewness and kurtosis with applications. *Biometrika*. 1970;57(3):519-530.

28. SAS Institute. Multivariate analysis concepts. support.sas.com/publishing/pubcat/chaps/56903.pdf. Accessed 2/27/2012.

29. Kolmogorov A. Sulla determinazione empirica di una legge di distribuzione. *Giorn. 1st. Ital. Attuari*. 1933(4):83-91.

30. Smirnov NV. Tables for estimating the goodness of fit of empirical distributions. *Annals of Mathematical Statistics*. 1948;19:279-281.

31. SAS Institute. Proc univariate. http://support.sas.com/documentation/cdl/en/procstat/63104/HTML/default/viewer.htm#univariate_toc.htm. Accessed 7/27/2011.

32. Peng G, Lilly E. Testing normality of data using SAS, paper PO04. *SUGI.* 2004:1-6.

33. Bentler PM, Chou CP. Practical issues in structural modeling. *Sociological methods research*. 1987;16:78-117.

34. Little RJA. A test of missing completely at random for multivariate data with missing values. *JASA*. 1988;83:1198-1202.

35. Little RJA, Rubin DB. *Statistical Analysis with Missing Data.* 2nd ed. New York: John Wiley; 2002.

36. Geldhof GJ, Selig JP. Using SAS for multiple imputation: The MI and MIANALYZE procedures. *Fall Colloquium.* 2007.

37. Welch K. Introduction to proc MI and proc MIAnalyze for multiple imputation of missing data in SAS. *Michigan SAS Users Group Conference.*

38. Yung YF, Zhang W. Making use of incomplete observations in the analysis of structural equation models: The CALIS procedure's full information maximum likelihood method in SAS/STAT 9.3. *SAS Global Forum Proceedings.* 2011.

39. Enders CK, Bandalos DL. The relative performance of full information maximum likelihood estimation for missing data in structural equation models. *Structural Equation Modeling*. 2001;8:430-457.

40. Schafer JL, Olsen MK. Multiple imputation for multivariate missing data problems: A data analyst perspective. *Multivariate Behavioral Research*. 1998;33:545-571.