Confirmatory factor analysis (CFA) of multivariate lipid distributions across time and race-ethnic groups: United States, 2003-2012

Ann Von Holle, Kari E. North, UNC, Chapel Hill, NC

UNC, Chapel Hill

March 1, 2017

Introduction

◇ High density lipoprotein cholesterol (HDL-C), low density lipoprotein cholesterol (LDL-C) and triglycerides (TG) are common measures of dyslipidemia.

  ◇ Common analytic practice includes separate evaluations of each lipoprotein when evaluating dyslipidemia.

  ◇ An alternative is to specify one latent factor, which predicts lipid values and incorporates their covariance structure.

◇ Before using the latent factor, important to determine if underlying structure of dyslipidemia construct is similar across groups (measurement invariance).

 ◇ For example, do racial/ethnic groups have the same construct?

Aims

Aim 1 Estimate dyslipidemia in a U.S. representative sample (NHANES) as one latent factor with three indicators (HDL-C, LDL-C, TG) via confirmatory factor analysis (CFA).

Aim 2 Test similarity of latent factor parameters (measurement invariance)[1] across age and racial/ethnic groups.

[1] R. van de Schoot et al. "A checklist for testing measurement invariance". In: *European Journal of Developmental Psychology* 9.4 (July 2012), pp. 486–492. DOI: 10.1080/17405629.2012.686740.

Confirmatory Factor Analysis (CFA) Model
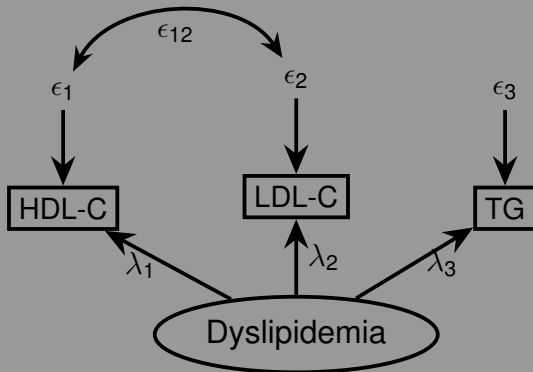
Equations for CFA[2]

$$\begin{aligned}
\text{HDL-C} &= \mu_1 + \lambda_1\eta_1 + \epsilon_1 \\
\text{LDL-C} &= \mu_2 + \lambda_2\eta_1 + \epsilon_2 \\
\text{TG-C} &= \mu_3 + \lambda_3\eta_1 + \epsilon_3
\end{aligned}$$

Parameters:

Factor loadings λ, represent the association between lipids and the latent factor for dyslipidemia, η_1.

Intercepts μ, represent the means of the indicators, lipids.

[2]T. A. Brown. *Confirmatory factor analysis for applied research*. Second edition. Methodology in the social sciences. New York ; London: The Guilford Press, 2015. 462 pp.

Four models[a]

$\Rightarrow$ Evaluate model fit across year and racial/ethnic groups in four models, holding different parameters constant.

Model 1 Test metric invariance: equal factor loadings but different intercepts across groups.

Model 2 Test intercept-only invariance: different factor loadings but equal intercepts across groups

Model 3 Test scalar invariance: equal factor loadings and equal intercepts across groups

Model 4 Full uniqueness invariance: equal loadings, intercept and variance across all groups

[a] All models include body mass index (BMI) and age covarying with factor (not shown in figure).

Model fit information by groups

Model	Chi-sq	df	LL	CFI	BIC	RMSEA
Groups: Race/Ethnicity						
1	118	49	-54014	0.946	108452	0.042
2	245	45	-54117	0.844	108690	0.075
3	1457	57	-55136	0.000	110628	0.176
4	276	57	-54143	0.829	108643	0.070
Groups: Time (year)						
1	110	49	-53436	0.931	107294	0.040
2	134	45	-53457	0.899	107370	0.050
3	1234	60	-54538	0.000	109408	0.157
4	147	57	-53470	0.897	107296	0.045

Best fit selections for racial/ethnic groups

◇ Model 1, allowing intercepts to vary, offers best fit.

    ◇ Acceptable levels include Root Mean Square Error of Approximation (RMSEA) < 0.05, Comparative Fit Index (CFI) = 0.95 and lowest Bayesian Information Criterion (BIC) of all models.

◇ Evidence for metric noninvariance (χ^2(df)=155 (8), p-value<0.001) $\Rightarrow$ Cannot make meaningful comparisons across racially and ethnically diverse populations.

Best fit selections for time groups

$\diamond$ Model 1, allowing intercepts to vary, offers best fit.

$\quad \diamond$ RMSEA < 0.05 and lowest Bayesian Information Criterion (BIC) of all models.

$\diamond$ Evidence for metric noninvariance (χ^2(df)=155 (12), p-value<0.001) $\Rightarrow$ Cannot make meaningful comparisons across time.

Model 1 standardized parameters by race/ethnic groups

Parameter	HDL-C	LDL-C	TG
Mexican American			
Loading	-0.681 (0.027)	0.436 (0.038)	0.598 (0.024)
Intercept	-0.142 (0.035)	-0.874 (0.049)	-0.549 (0.036)
Residual variance	0.537 (0.037)	0.81 (0.033)	0.643 (0.028)
Non-Hispanic Black			
Loading	-0.596 (0.026)	0.374 (0.032)	0.598 (0.024)
Intercept	0.132 (0.031)	-0.692 (0.04)	-1.078 (0.04)
Residual variance	0.645 (0.031)	0.86 (0.024)	0.643 (0.028)
Non-Hispanic White			
Loading	-0.64 (0.026)	0.381 (0.035)	0.598 (0.024)
Intercept	-0.165 (0.036)	-0.697 (0.053)	-0.539 (0.035)
Residual variance	0.591 (0.033)	0.855 (0.027)	0.643 (0.028)
Other Hispanic			
Loading	-0.676 (0.035)	0.484 (0.045)	0.598 (0.024)
Intercept	-0.162 (0.078)	-1.07 (0.099)	-0.453 (0.083)
Residual variance	0.542 (0.047)	0.766 (0.044)	0.643 (0.028)

Latent factor characteristics by racial/ethnic groups

$\diamond$ Factor loadings: A one unit factor increase represents less favorable lipid value $\Rightarrow$ an increase in LDL-C (0.4) and TG (0.6) accompanied by a decrease in HDL ($\sim$ -0.6).

$\diamond$ Intercepts: Non-Hispanic Black group only one with positive HDL-C (0.132) and lowest TG (-1.078) relative to center of multivariate distribution.

Model 1 standardized parameters by year groups

Parameter	HDL-C	LDL-C	TG
2003-2004			
Loading	-0.588 (0.032)	0.414 (0.039)	0.606 (0.027)
Intercept	-0.162 (0.044)	-0.559 (0.057)	-0.435 (0.045)
Residual variance	0.655 (0.038)	0.829 (0.033)	0.633 (0.033)
2005-2006			
Loading	-0.603 (0.035)	0.439 (0.043)	0.606 (0.027)
Intercept	-0.142 (0.049)	-0.752 (0.059)	-0.57 (0.047)
Residual variance	0.637 (0.042)	0.808 (0.037)	0.633 (0.033)
2007-2008			
Loading	-0.664 (0.036)	0.438 (0.041)	0.606 (0.027)
Intercept	-0.047 (0.058)	-0.9 (0.09)	-0.734 (0.055)
Residual variance	0.559 (0.048)	0.808 (0.036)	0.633 (0.033)
2009-2010			
Loading	-0.656 (0.035)	0.389 (0.045)	0.606 (0.027)
Intercept	-0.024 (0.049)	-0.749 (0.103)	-0.738 (0.055)
Residual variance	0.569 (0.045)	0.848 (0.035)	0.633 (0.033)
2011-2012			
Loading	-0.691 (0.041)	0.41 (0.042)	0.606 (0.027)
Intercept	-0.145 (0.063)	-0.772 (0.073)	-0.607 (0.059)
Residual variance	0.523 (0.057)	0.832 (0.035)	0.633 (0.033)

Fit and parameters for groups by sex

Model fit

Model	Chi-sq	df	LL	CFI	BIC	RMSEA
Groups: Gender						
1	129	16	-53427	0.862	107052	0.060
2	192	15	-53483	0.783	107174	0.077
3	1267	21	-54661	0.000	109480	0.173
4	199	18	-53495	0.778	107172	0.071

Model 1 parameters

Parameter	HDL-C	LDL-C	TG
FEMALE			
Loading	-0.634 (0.032)	0.414 (0.047)	0.605 (0.03)
Intercept	0.078 (0.031)	-0.729 (0.043)	-0.627 (0.036)
Residual variance	0.598 (0.04)	0.829 (0.039)	0.635 (0.036)
MALE			
Loading	-0.68 (0.031)	0.378 (0.043)	0.605 (0.03)
Intercept	-0.287 (0.036)	-0.761 (0.057)	-0.602 (0.036)
Residual variance	0.538 (0.042)	0.857 (0.033)	0.635 (0.036)

Invariance testing indicates evidence for metric non-invariance when testing difference between a metric model (model 1) and a full invariance model (model 4) in which there are full equality constraints across groups (χ^2 = 68.5 (df=2); p-value<0.001).

Model fitting process

$\diamond$ Covariance between HDL-C and TG no longer significant when accounting for BMI covarying with latent lipid factor.

$\diamond$ Age positive covariance with lipid factor.

$\diamond$ BMI positive covariance with lipid factor.

Summary

◇ 3-indicator latent factor can function as a
  parsimonious and well-fitting measure for
  dyslipidemia in structural equation models.

◇ Important not to compare latent factor
  characteristics across groups by race/ethnic, time
  or sex status given metric measurement
  non-invariance.

Contact information: vonholle at unc dot edu
Further project information: https://avonholle.github.io/cfa-lipids/