2012-05-04

A Latent Class Analysis of Metabolic Syndrome Among Hispanics/Latinos Living in the United States in Relation to Cardiovascular Disease Prevalence: Results from the Hispanic Community Health Study/Study of Latinos (HCHS/SOL)

William Arguelles
University of Miami, warguelles@psy.miami.edu

UNIVERSITY OF MIAMI

A LATENT CLASS ANALYSIS OF METABOLIC SYNDROME AMONG HISPANICS/LATINOS LIVING IN THE UNITED STATES IN RELATION TO CARDIOVASCULAR DISEASE PREVALENCE: RESULTS FROM THE HISPANIC COMMUNITY HEALTH STUDY/STUDY OF LATINOS (HCHS/SOL)

By

William Arguelles

A DISSERTATION

Submitted to the Faculty
of the University of Miami
in partial fulfillment of the requirements for
the degree of Doctor of Philosophy

Coral Gables, Florida

May 2012

UNIVERSITY OF MIAMI

A dissertation submitted in partial fulfillment of
the requirements for the degree of
Doctor of Philosophy

A LATENT CLASS ANALYSIS OF METABOLIC SYNDROME AMONG HISPANICS/LATINOS LIVING IN THE UNITED STATES IN RELATION TO CARDIOVASCULAR DISEASE PREVALENCE: RESULTS FROM THE HISPANIC COMMUNITY HEALTH STUDY/STUDY OF LATINOS (HCHS/SOL)

William Arguelles

Approved:

_____	_____
Maria M. Llabre, Ph.D. Professor of Psychology	Terri A. Scandura, Ph.D. Dean of the Graduate School
_____	_____
Neil Schneiderman, Ph.D. Professor of Psychology	Frank J. Penedo, Ph.D. Associate Professor of Psychology
_____	_____
Ralph L. Sacco, M.D., M.S., FAAN, FAHA Professor of Neurology	David J. Lee, Ph.D. Professor of Epidemiology and Public Health

ARGUELLES, WILLIAM (Ph.D., Psychology)
A Latent Class Analysis of Metabolic Syndrome Among Hispanics/Latinos Living in the United States in Relation to Cardiovascular Disease Prevalence: Results from the Hispanic Community Health Study/Study of Latinos (HCHS/SOL). (May 2012)

Abstract of a dissertation at the University of Miami.

Dissertation supervised by Professors Maria M. Llabre and Neil Schneiderman.
No. of pages in text. (80)

Using preliminary data from the Hispanic Community Health Study/Study of Latinos (HCHS/SOL), this cross-sectional study employed latent class analysis to investigate 1) whether distinct subtypes of metabolic syndrome (MetS) could be identified among a large and diverse sample of US Hispanic/Latinos (H/Ls), 2) how identified MetS subtypes differed in demographic, socioeconomic, clinical, and behavioral characteristics, and 3) the association between identified MetS subtypes and cardiovascular disease (CVD) prevalence. Incorporating continuous measures of MetS components (waist circumference, systolic and diastolic blood pressure, HDL cholesterol, triglycerides, and fasting glucose) and data on antihypertensive, lipid-lowering, and glucose-lowering medication use into analyses, two latent clusters were identified as best representing the data among the entire sample (n = 10970), as well as among men (n = 4429) and women (n = 6541) separately. One cluster was characterized by individuals exhibiting relatively healthy mean levels across most MetS components (Non-MetS cluster), while the other cluster was characterized by individuals exhibiting clinically elevated mean levels across most MetS components (MetS cluster). The presence of

additional, meaningful subtypes of MetS was not confirmed. This two-cluster model was associated with multiple covariates and prevalent CVD outcomes in a manner generally consistent with previous scientific knowledge, demonstrating adequate construct validity. For example, individuals who were older and had a positive family history of CHD exhibited greater odds of being classified into the MetS cluster, and, in turn, those classified into the MetS cluster demonstrated greater odds of having prevalent coronary heart disease. While study results largely converged with current conceptualizations of MetS as a distinct cardiometabolic state, valuable information pertaining to the presentation of MetS specifically among US H/Ls was obtained. For instance, compared to other MetS components, findings suggest that HDL cholesterol may poorly differentiate between individuals who have and do not have MetS among this ethnic group. Additionally, incipient evidence is provided suggesting that currently identified thresholds for some MetS components (i.e., the waist circumference cutoff proposed for US females by NCEP-ATP III criteria) might not be optimal for diagnosing MetS among US H/Ls. Given the exploratory nature of this methodology, and study design constraints (i.e., lack of a non-H/L comparison cohort), these results are tentative and warrant replication. The potential insights offered by adopting this analytic approach to the study of MetS are discussed, as are associated strengths, limitations, and directions for future research.

ACKNOWLEDGMENTS

The Hispanic Community Health Study/Study of Latinos (HCHS/SOL) was carried out as a collaborative study supported by contracts from the National Heart, Lung, and Blood Institute (NHLBI) to the University of North Carolina (N01-HC65233), University of Miami (N01-HC65234), Albert Einstein College of Medicine (N01-HC65235), Northwestern University (N01-HC65236), and San Diego State University (N01-HC65237). The following Institutes/Centers/Offices contribute to the HCHS/SOL through a transfer of funds to the NHLBI: National Institute on Minority Health and Health Disparities, National Institute on Deafness and Other Communication Disorders, National Institute of Dental and Craniofacial Research, National Institute of Diabetes and Digestive and Kidney Diseases, National Institute of Neurological Disorders and Stroke, NIH Institution-Office of Dietary Supplements.

The opinions, ideas, and interpretations included in this dissertation are mine alone, and not those of the HCHS/SOL investigators.

I thank the staff and participants of HCHS/SOL for their important contributions. Investigators website - http://www.cscc.unc.edu/hchs/

TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION

Cardiovascular Disease

Cardiovascular disease (CVD) is a general term for disorders affecting the heart and blood vessels, and encompasses conditions such as atherosclerosis, coronary heart disease (CHD), cerebrovascular disease, and hypertension (National Heart, Lung, and Blood Institute, 2009). Over recent years, death rates from CVD have declined in the United States (US) (National Center for Health Statistics, Centers for Disease Control and Prevention, 2009), likely attributable to advances in evidence-based medical therapies and modified risk factors in the population (Ford, Ajani, Croft, Critchley, Labarthe, et al., 2007). However, the burden of CVD in the US remains high.

Accounting for nearly 2,200 deaths each day, CVD continues to be the leading cause of mortality for both men and women in the US (Roger, Go, Lloyd-Jones, Adams, Berry, et al., 2011). Cardiovascular disorders are also among the leading causes of functional disabilities in the US population (Roger et al., 2011). It is estimated that over 1 in 3 Americans have at least one type of CVD, with associated total health costs – approximately $287 billion in 2007 – exceeding those of any other health condition (Roger et al., 2011). Estimates drawn from 2007 mortality data suggest that approximately 34% of all deaths in the US are attributable to CVD (Roger et al., 2011). The largest percentage of those deaths, approximately 50%, is purportedly due to CHD, with stroke accounting for an additional 17% (Roger et al., 2011).

CHD is a type of CVD resulting from the narrowing of the coronary arteries, which supply the heart muscle with the oxygenated blood it needs to function properly (National Heart, Lung, and Blood Institute, 2009). The underlying cause of CHD, as well

as most other clinical CVD events, is atherosclerosis. Atherosclerosis is a systematic disease process characterized by the build-up of fatty deposits, inflammatory factors, platelets, calcium, and scar tissue within the inside lining of artery walls, forming plaques which may harden and narrow arteries over time, consequently restricting blood flow to corresponding organs (Libby, 2003). Reduced or blocked blood flow to the heart can lead to angina (chest pain or discomfort) or myocardial infarction (sudden block in blood flow to a section of the heart muscle), which can cause serious damage and death, especially in the absence of rapid treatment (National Heart, Lung, and Blood Institute, 2009). Recent prevalence estimates suggest that approximately 7% of US adults have CHD, with a prevalence of 3.1% specific for myocardial infarction (Roger et al., 2011). About one in six US deaths are caused by CHD, and its associated total health costs – approximately $177.5 billion in 2007 – are the highest among all CVDs (Roger et al., 2011).

Along with lifestyle changes and medication, treatment for CHD may include undergoing medical procedures such as angioplasty (nonsurgical insertion of a balloon or other device into a narrowed artery which is inflated to compress plaque and restore blood flow), stenting (insertion of a small mesh tube into a narrowed artery during angioplasty to support the vascular wall for months to years after completion of the procedure), or coronary artery bypass grafting (surgical placement of arteries or veins from other body areas to bypass narrowed arteries for improved blood flow) (National Heart, Lung, and Blood Institute, 2009). Between 1997 and 2007, the number of these as well as other CVD procedures increased by approximately 27% (Roger et al., 2011).

Stroke, or cerebrovascular disease, is a type of CVD in which brain cells die or are damaged due to either a reduction in oxygenated blood flow to a portion of the brain

(ischemic stroke or, if short in duration, transient ischemic attack), or an increase in brain pressure as a result of arterial leaks or ruptures that lead to sudden bleeding (hemorrhagic stroke) (National Heart, Lung, and Blood Institute, 2009). Similar to CHD, the common underlying cause of ischemic stroke is atherosclerosis leading to plaque and blood clot formation (National Heart, Lung, and Blood Institute, 2009). Hemorrhagic strokes are commonly caused by high blood pressure, aneurysms, or arteriovenous malformations (National Heart, Lung, and Blood Institute, 2009). Recent prevalence estimates suggest that approximately 3% of US adults have had a stroke (Roger et al., 2011). Of all strokes, ischemic strokes are the most common (approximately 87%) followed by intracerebral hemorrhagic (10%) and subarachnoid hemorrhagic (3%) strokes (Roger et al., 2011). It is estimated that every 40 seconds, on average, an American suffers a stroke (Roger et al., 2011).

Treatments for ischemic stroke include medication (i.e., injection of tissue plasminogen activator to dissolve arterial blood clots, antiplatelet medication to reduce clot formation, and/or anticoagulant medication to prevent clot enlargement and the formation of new clots) or medical procedures (i.e., carotid endarterectomy or carotid artery angioplasty to open blocked arteries, or more investigative procedures such as intra-arterial thrombolysis or mechanical embolus removal in cerebral ischemia) (National Heart, Lung, and Blood Institute, 2009). Treatments for hemorrhagic stroke are concerned with controlling the cause of bleeding in the brain and may also include medication (i.e., antihypertensive medications to reduce blood pressure) or medical procedures (i.e., aneurysm clipping to prevent further blood leakage, coil embolization to block the flow of blood through the aneurysm, or arteriovenous malformation repair)

(National Heart, Lung, and Blood Institute, 2009). Lifestyle changes (i.e., quitting smoking, following a healthy diet, engaging in physical activity, etc.) are generally recommended following initial treatment of cerebrovascular disease (National Heart, Lung, and Blood Institute, 2009).

Metabolic Syndrome

Ever since the emergence and identification of CVD as a major cause of morbidity and mortality, significant efforts have been made to understand its underlying pathology and identify associated risk factors. Ensuing clinical and investigative observations noted that patients with CVD commonly presented with more than one risk factor, and often a cluster of them. A unified concept and formal hypothesis regarding this observation was first presented in 1988 by Reaven, who initially termed this phenomenon "syndrome X," and postulated insulin resistance as the central underlying abnormality (Reaven, 1988). Since then, many different terms (i.e., "insulin resistance syndrome," "deadly quartet," etc.) and definitions have emerged to characterize this phenomenon, which at present is most commonly referred to as metabolic syndrome (Johnson & Weinstock, 2006).

Metabolic syndrome (MetS) is currently conceptualized as the clustering of various interrelated risk factors – most commonly including obesity, hypertension, dyslipidemia, and hyperglycemia – that have been shown to increase the risk of developing CVD (Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults, 2001). While different criteria still abound (i.e., including obligatory components versus not, specifying varying threshold values for individual components, etc.), a unified definition has been recently proposed by various

organizations including the International Diabetes Foundation, the National Heart, Lung, and Blood Institute, and the American Heart Association (Alberti, Eckel, Grundy, Zimmet, Cleeman, et al., 2009). Based on this definition, MetS diagnosis requires that three or more of following five risk factors be present: 1) elevated waist circumference as suggested by population- and country-specific thresholds (i.e., ≥102 cm in US men and ≥88 cm in US women), 2) elevated blood pressure (systolic/diastolic blood pressure ≥130/≥85 mm Hg) or antihypertensive drug treatment in a patient with a history of hypertension, 3) low high-density lipoprotein (HDL) cholesterol (<40 mg/dL in men and <50 mg/dL in women) or drug treatment for low HDL cholesterol, 4) elevated triglycerides (≥150 mg/dL) or drug treatment for elevated triglycerides, and 5) elevated fasting glucose (≥100 mg/dL) or drug treatment for elevated glucose (Alberti et al., 2009). While abdominal obesity and insulin resistance are present in most individuals with MetS – and have been postulated to be major contributing factors for the development of MetS via their influence on atherosclerotic-promoting processes (i.e., increased vascular inflammation) – the exact pathophysiology and mechanisms underlying MetS are not completely understood, and are likely to be multifactorial and complex (Alberti et al., 2009).

Although not an indicator of absolute risk (given that it does not take factors such as age, sex, and smoking into account), MetS has consistently been shown to provide a greater risk of developing CVD (Alberti et al., 2009). For example, a meta-analysis of 37 prospective studies including 43 cohorts demonstrated that persons with MetS had a 1.78-fold increased risk (95% CI = 1.58 to 2.00) of cardiovascular events and death (Gami, Witt, Howard, Erwin, Gami, et al., 2007). This risk was about a third higher in women

(RR = 2.63) compared to men (RR = 1.98) (Gami et al., 2007). Furthermore, this association persisted after adjustment for individual MetS features (RR = 1.54; 95% CI = 1.32 to 1.79), suggesting MetS diagnosis confers risk for CVD beyond its specific components (Gami et al., 2007). Similar findings emerged in a more current meta-analysis that also incorporated data from other recently conducted prospective studies (87 studies total) and showed that MetS was associated with an approximately two-fold increased risk of CVD (RR = 2.35; 95% CI = 2.02 to 2.73), CVD mortality (RR = 2.40; 95% CI = 1.87 to 3.08), myocardial infarction (RR = 1.99; 95% CI = 1.61 to 2.46), and stroke (RR = 2.27; 95% CI = 1.80 to 2.85) (Mottillo, Filion, Genest, Joseph, Pilote, et al., 2010). Demonstrating a similarly increased risk in women compared to men across these estimates, this study also showed that MetS increased risk for CVD events even among patients without type 2 diabetes (i.e., RR = 1.62 and 1.86 for myocardial infarction and stroke, respectively) (Mottillo et al., 2010).

Additionally, CVD risk estimates have generally been shown to increase as the number of individual MetS components present increase. For example, in the Framingham Offspring Study, individuals with one to two components had a hazard ratio of 1.48 (men) and 3.39 (women) for CVD, while those with three or more components had a higher hazard ratio of 3.99 (men) and 5.95 (women), compared to men and women with no MetS components (Wilson, D'Agostino, Parise, Sullivan, & Meigs, 2005).[1] Moreover, different MetS component combinations may differentially increase risk for CVD. For instance, returning to published data from the Framingham Offspring Study, of all possible combinations of three MetS components, the clustering of central adiposity,

[1] While presented study results using both relative risk and hazard ratio estimates are consistent, it is important to note the subtle but important differences between them (i.e., hazard ratios take time-to-event into account whereas relative risk ratios do not, and are cumulative over an entire study period).

high blood pressure, and hyperglycemia was most strongly associated with CVD risk (HR = 2.36) in stepwise analyses (Franco, Massaro, Civil, Cobain, O'Malley, et al., 2009).

There is a high burden of MetS in the US. Recent prevalence estimates suggest that approximately one-third of US adults (35.1% of men and 32.6% of women) meet criteria for MetS, indicating this clinical problem is also of major public health concern (Roger et al., 2011; Alberti et al., 2009). Compared to individuals without MetS, individuals with MetS utilize healthcare services more often and, on average, have approximately 60% higher annual total costs (i.e., $5,732 versus $3,581 in 2005) (Boudreau, Malone, Raebel, Fishman, Nichols, et al., 2009). This translates to a 24% increase in overall healthcare costs per each MetS component present (Boudreau et al., 2009). Unfortunately, the prevalence of MetS is projected to rise along with continuing trends of increased obesity and sedentary lifestyles (Alberti et al., 2009). Other factors identified in either prospective or retrospective studies to increase risk of developing MetS included older age, male sex, low educational attainment and socioeconomic status, poor physical fitness, smoking, and no or heavy alcohol consumption (moderate alcohol intake protective), among others (Roger et al., 2011).

The US Hispanic/Latino Population

Hispanics/Latinos (H/L) currently represent the largest and fastest growing minority population in the US. Between 1990 and 2000, the H/L population grew 22.4% to 35.3 million, accounting for 12.5% of the total US population (U.S. Census Bureau, The Hispanic Population Census, 2000 Brief; U.S. Census Bureau, Hispanic Population of the United States: 1970-2050). Between 2000 and 2006, the H/L growth rate was even

higher (24.3%) and over three times greater than that observed for the total US population during the same time period (U.S. Census Bureau, Hispanic Population of the United States: 1970-2050). As of July 2006, approximately 44.3 million H/Ls were residing in the US, accounting for 14.8% of the total US population (U.S. Census Bureau, Hispanic Population of the United States: 1970-2050). It is estimated that by 2050, the US H/L population will grow another three-fold to over 102 million, and account for 24% of the total US population (U.S. Census Bureau, Hispanic Population of the United States: 1970-2050).

Significant diversity exists within the US H/L population related to place of origin, cultural practices, lifestyles, and behaviors, as well as nativity, generation, length of time residing in the US, and level of acculturation (see Gallo, Penedo, Espinosa de los Monteros, & Arguelles, 2009, for a review). H/L designation by US statistics is generally inclusive of persons who trace their roots to Mexico, Puerto Rico, Cuba, the Dominican Republic, Spanish-speaking countries of Central and South America, Spain, or other Spanish cultures, regardless of race (Roger et al., 2011). The most recent census data on H/Ls living in the US indicated that 64% were of Mexican origin, 9% were of Puerto Rican origin, 3.4% were of Cuban origin, 2.8% were of Dominican origin, and 13.1% were of Central and South American origin, with approximately 7.7% labeled as "Other" (U.S. Census Bureau, Hispanic Population of the United States: 1970-2050). These H/L subgroups have different geographic distributions across the US (U.S. Census Bureau, Hispanic Population of the United States: 1970-2050).

As an aggregate group, H/Ls appear to have lower levels of educational attainment and English language proficiency, lower household income, and less

healthcare coverage compared to other minority groups and the general US population, albeit substantial variation among different H/L subgroups is evident (Centers for Disease Control and Prevention, 2011; Gallo et al., 2009). US H/Ls may also experience varying levels of psychosociocultural stress associated with immigration, unemployment, the acculturation process, and social marginalization and discrimination (Gallo et al., 2009). Thus, the H/L population as a whole may face disproportionate barriers to receiving adequate health knowledge, access, and care, which may contribute to health disparities (Centers for Disease Control and Prevention, 2011). On the other hand, H/L ethnicity may also be associated with certain health protective factors, ranging in scope from biological (i.e., genes) to environmental/behavioral (i.e., diet) to psychosociocultural (i.e., strong social networks and support) (Gallo et al., 2009).

Cardiovascular Disease Among US Hispanics/Latinos

CVD is also the leading cause of death among US H/L men and women (Roger et al., 2011). However, CVD appears to be lower among US H/Ls (prevalence rates of 30.7% and 30.9% in Mexican American adult men and women, respectively) compared to non-H/L whites (37.4% for men and 33.8%, for women) and non-H/L blacks (44.8% for men and 47.3%, for women) (Roger et al., 2011). US H/Ls also appear to have lower CVD mortality rates (165.0 for men and 118.8 for women) compared to non-H/L whites (294.0 for men and 205.7 for women) and non-H/L blacks (405.9 for men and 286.1 for women) (Roger et al., 2011).

The prevalence of CHD also appears to be lower among US H/Ls. The age-adjusted prevalence in US adult men and women is estimated to be 6.3% and 5.6%, respectively, for Mexican Americans compared to a prevalence of 8.5% and 6.1%,

respectively, for non-H/L whites and 7.9% and 7.6%, respectively, for non-H/L blacks (Roger et al., 2011). Among US H/Ls aged 18 years or older, CHD prevalence is 5.8% (Roger et al., 2011). In 2007, age-adjusted CHD death rates were also lower for US H/L men and women (122.3 and 77.8 per 100,000 population, respectively) compared to non-H/L whites (165.6 for men and 94.2 for women) and non-H/L blacks (191.6 for men and 121.5 for women) (Roger et al., 2011).

Similar trends specific to myocardial infarction were also observed, with Mexican American adult men and women showing a lower prevalence (3.0% and 1.1%, respectively) compared to non-H/L whites (4.3% for men and 2.1% for women) and non-H/L blacks (4.3% for men and 2.2% for women) (Roger et al., 2011). However, prevalence of angina does not appear lower among Mexican American adult men and women (3.6% and 3.7%, respectively) compared to non-H/L whites (3.8% for men and 3.7% for women) or non-H/L blacks (3.3% for men and 5.6% for women) (Roger et al., 2011).

The age-adjusted prevalence of stroke also appears to be lower among US H/Ls compared to non-H/L whites. The age-adjusted prevalence in US adult men and women is estimated to be 2.0% and 2.7%, respectively, for Mexican Americans compared to a prevalence of 2.4% and 3.3%, respectively, for non-H/L whites (Roger et al., 2011). Among US H/Ls aged 18 years or older, stroke prevalence is 2.0% (Roger et al., 2011). In 2007, age-adjusted stroke death rates were also lower for US H/L men and women (34.4 and 30.8 per 100,000 population, respectively) compared to non-H/L whites (40.2 for men and 39.9 for women). Compared to all other racial/ethnic groups, non-H/L blacks exhibit the highest age-adjusted prevalence of stroke (4.5% for men and 4.4% for

women) as well as the highest stroke death rates (67.1 and 55.0 per 100,000 for men and women, respectively).

Of important note, nationally representative estimates of CVD by specific H/L subgroups other than Mexican Americans are lacking. Compared to non-H/L whites, these generally lower or similar CVD estimates observed in the US H/L population that experiences more adversities (i.e., in terms of income, education, etc.) has been termed the Hispanic Paradox, and has received much attention and debate (Franzini, Ribble, & Keddie, 2001; Sorlie, Backlund, Johnson, & Rogot, 1993; Swenson, Trepka, Rewers, Scarbro, Hiatt, et al., 2002). Nonetheless, the projected growth and aging of the US H/L population – which exhibit a higher incidence and prevalence of CVD risk factors, as well as a greater lack of awareness regarding CVD risk (discussed below) – will substantially increase the national burden of CVD, with estimated healthcare costs for this group alone reaching $163 billion by 2050 (Davidson, Moreno, Badimon, Lopez-Candales, Giachello, et al., 2007).

Metabolic Syndrome Among US Hispanics/Latinos

The possible health advantage observed among US H/Ls in terms of CVD is not uniformly seen in terms of MetS and associated components. Compared to non-H/L whites and non-H/L blacks, Mexican Americans have been shown to have a higher prevalence of MetS, although this pattern appears to vary by gender (Razzouk & Muntner, 2009). For example, while Mexican American women exhibit a higher age-adjusted prevalence of MetS (40.6%) compared to non-H/L white (31.5%) and non-H/L black (38.8%) women, Mexican American men only exhibit a higher age-adjusted prevalence (33.2%) compared to non-H/L black (25.3%) but not non-H/L white (37.2%)

men (Roger et al., 2011). As with CVD, nationally representative estimates of MetS prevalence by specific H/L subgroups other than Mexican Americans are also lacking. One study of 204 women living in Northern Manhattan showed a higher prevalence of MetS among predominantly non-Mexican, Caribbean H/Ls (63.3%) compared to non-H/L whites (29.6%), suggesting that non-Mexican US H/L subgroups may also exhibit higher rates of MetS compared to non-H/L whites (Yala, Fleck, Sciacca, Castro, Joseph, et al., 2009).

Few studies have examined H/L subgroup differences in MetS prevalence. Examination of 1,437 H/L men and women in the Multi-Ethnic Study of Atherosclerosis (MESA) showed that Mexican Americans had the highest prevalence of MetS (49.1%) compared to Puerto Ricans (37.9%) and Other H/L Americans (37.6%), while Dominicans had the lowest prevalence (33.2%); Cuban Americans were not evaluated in this study due to small sample size (Allison, Budoff, Wong, Blumenthal, Schreiner, et al., 2008). A recent study of 419 middle-aged H/L women (not including Mexican Americans) living in New Jersey showed that Puerto Ricans had the highest prevalence (48.2%), followed by Central Americans (40.0%), South Americans (35.0%), and Cubans (29.3%), with Dominicans again exhibiting the lowest prevalence (13.9%) (Derby, Wildman, McGinn, Green, Polotsky, et al., 2010). Thus, compared to Puerto Rican women, Cuban women in this study were 60% less likely and Dominican women were 80% less likely to have MetS (Derby et al., 2010).

Prospective data specific to US H/Ls regarding the association between MetS and CVD events are also lacking. Still, consistent with the general literature, studies on US H/L populations have shown that MetS is associated with an increased 10-year risk for

CHD (Hoang, Ghandehari, Lopez, Barboza, & Wong, 2008; Meigs, Wilson, Nathan, D'Agostino, Williams, et al., 2003; Yala et al., 2009). In some studies, the strength of this predicted risk appeared slightly weaker in US H/Ls compared to non-H/L whites and non-H/L blacks (Hoang et al., 2008; Meigs et al., 2003). However, the opposite pattern was observed in a study that compared predominantly Caribbean H/L women to non-H/L white women and non-H/L black women, and showed that H/L women diagnosed with MetS exhibited the highest Framingham Risk Scores (Yala et al., 2009). In the Northern Manhattan Study, MetS was also associated with an increased risk of ischemic stroke among US H/Ls, and this effect was shown to be greater among H/Ls compared to non-H/L whites and blacks (Boden-Albala, Sacco, Lee, Grahame-Clarke, Rundek, et al., 2008).

US H/Ls have also been shown to have a higher incidence and prevalence of individual MetS components. For example, according to NHANES 2007 to 2008 data, Mexican American men had the highest prevalence of overweight or obesity (80%) compared to non-H/L white (73%) and non-H/L black (69%) men (Roger et al., 2011). Prevalence of overweight or obesity in Mexican American women (77%) was comparable to non-H/L black women (78%) and higher than non-H/L white women (61%) (Roger et al., 2011). Obesity-specific prevalence rates were also higher among Mexican American men (36%) and women (45%) compared to non-H/L white men (32%) and women (33%) (Roger et al., 2011). Thus, overall, US H/Ls are less likely to maintain a healthy weight compared to non-H/L whites (Roger et al., 2011). Among US H/Ls, Mexican American and Puerto Rican women appear to have higher rates of overweight and obesity compared to Cuban American and Dominican women, however

data are limited and these patterns appear to be different among men (Crespo, Loria, & Burt, 1996; Derby et al., 2010).

US H/Ls also bear a dramatically disproportionate burden of diabetes. The age-adjusted prevalence of diagnosed diabetes in US adults was 10.4% for H/Ls, compared with 6.6% for non-H/L whites and 11.8% for non-H/L Blacks (Roger et al., 2011). Prevalence of undiagnosed diabetes was highest among Mexican American men (6.3%) compared to non-H/L white (3.9%) and non-H/L black (4.8%) men (Roger et al., 2011). Mexican American women had a prevalence of undiagnosed diabetes (3.8%) comparable to non-H/L black women (4.0%) and higher than non-H/L white women (1.9%) (Roger et al., 2011). In MESA, the cumulative incidence of diabetes over a 5-year follow-up was highest in H/Ls (11.3%) compared to non-H/L blacks (9.5%) and non-H/L whites (6.3%) (Roger et al., 2011). Furthermore, estimates of projected increases in diabetes in the US are highest among H/Ls (127%) compared to non-H/L blacks (107%) and non-H/L whites (99%) (Roger et al., 2011). Additionally, among individuals with diabetes, US H/Ls have strikingly lower – and widening disparities in – glycemic control rates (37.8%) compared to non-H/L blacks (41.6%) and non-H/L whites (58.1%), and are also less likely to receive recommended comprehensive diabetes care (McWilliams, Meara, Zaslavsky, & Ayanian, 2009; Roger et al., 2011). Interestingly, prevalence of prediabetes does not appear to be substantially higher among US H/Ls compared to non-H/L whites, but may be due to H/Ls exhibiting faster progression to diabetes from prediabetic status (Roger et al., 2011). Limited data do not appear to show differences among US H/L subgroups in the prevalence of diabetes, hyperglycemia, or insulin resistance (Allison et al., 2008; Derby et al., 2010).

Compared to non-H/L whites and non-H/L blacks, Mexican American men also have a slightly higher prevalence of dyslipidemia as assessed by elevated total cholesterol (50.1% vs. 41.2% and 37.0%, respectively), elevated LDL cholesterol (41.9% vs. 30.5% and 34.4%), low HDL cholesterol (31.7% vs. 29.5% and 16.6%), as well as mean LDL cholesterol (121.2 vs. 114.5 and 114.6 mg/dL), HDL cholesterol (46.0 vs. 47.2 and 52.3 mg/dL), and triglyceride (169.4 vs. 150.2 and 120.1 mg/dL) levels (Roger et al., 2011). The lipid profiles of Mexican American women and non-H/L white women are more comparable, and in MESA, a comparable prevalence of dyslipidemia was observed among all H/Ls compared to non-H/L whites and non-H/L blacks (Goff, Bertoni, Kramer, Bonds, Blumenthal, et al., 2006; Roger et al., 2011). However, US H/Ls are less likely to be aware of having dyslipidemia compared to non-H/L whites (i.e., less than half of hypercholesterolemic Mexican Americans were aware of their condition), are less likely to be treated, and exhibit substantially lower rates of control (i.e., LDL cholesterol control rate of 16.5% compared to 26.9% in non-H/L whites) (Goff et al., 2006; McWilliams et al., 2010; Roger et al., 2011). Limited data suggest substantial variability between H/L subgroups in lipid profiles. In MESA, Mexican Americans had higher levels of triglycerides compared to other H/L subgroups (Allison et al., 200). In the Hispanic Health and Nutrition Examination Survey, Cuban Americans had the lowest rates (Crespo et al., 1996). In a smaller sample of H/L women living in New Jersey, Puerto Ricans had lower LDL cholesterol levels compared to Cubans, Dominicans, and South Americans, but also had lower HDL cholesterol compared to Cubans and Dominicans, and higher triglycerides compared to Dominicans (Derby et al., 2010).

Prevalence of hypertension is lower among US H/L adults (27.8% and 28.9% for Mexican American men and women, respectively) compared to non-H/L whites (33.9% for men and 31.3% for women) and non-H/L blacks (43.0% for men and 45.7% for women) (Roger et al., 2011). Limited data generally suggest a similar prevalence of hypertension and similar levels of systolic blood pressure among H/L subgroups (Allison et al., 2008; Crespo et al., 1996; Derby et al., 2010). However, data suggest that Puerto Rican Americans exhibit a greater hypertension-related death rate (154.0 per 100,000 population) compared to all other H/L subgroups and non-H/L whites (135.9), while Cuban Americans exhibit the lowest rate (82.5) (Roger et al., 2011). Across all H/L subgroups, these rates were higher in males compared to females (Roger et al., 2011). As with diabetes and dyslipidemia, compared to non-H/L whites, hypertensive US H/Ls are strikingly less likely to be aware of, receive treatment for, and have control over their condition (Cutler, Sorlie, Wolz, Thom, Fields, et al., 2008; Kramer, Han, Post, Goff, Diez-Roux, et al., 2004; McWilliams et al., 2010).

Reasons for these observed elevations in MetS and specific associated components among US H/Ls remain unclear, with various hypotheses positing the influential roles of acculturative processes (i.e., changes in traditional dietary and physical activity patterns to those of Western cultures), sociodemographic disparities (i.e., lower educational attainment and income), and barriers to quality healthcare (i.e., lower coverage rates compared to other US minority groups) (Derby et al., 2010; Gallo et al., 2009; Karlamangla et al., 2010). Besides exhibiting a generally lower prevalence of smoking (Roger et al., 2011), compared to non-H/L whites, US H/Ls report less optimal levels of other health behaviors shown to predict the development of MetS, such as

physical activity (i.e., 27.8% prevalence of regular leisure time physical activity vs. 38.1% in non-H/L whites) (King, Mainous, Carnemolla, & Everett, 2009; Roger et al., 2011).

The Metabolic Syndrome Debate

Recently, considerable debate has surrounded the validity and clinical utility of MetS diagnosis using currently proposed criteria (de Zeeuw & Bakker, 2008; Tenenbaum & Fisman, 2011). In 2005, a seminal statement by the American Diabetes Association highlighted several important limitations concerning the conceptualization and application of MetS, adamantly calling for rigorous research to address critical knowledge gaps (Kahn, Buse, Ferrannini, & Stern, 2005). At present, many of these gaps have yet to be adequately explored or delineated. Notwithstanding the consistent observation that particular CVD risk factors tend to cluster together more often than by chance alone, the following provides a brief description of cited criticisms regarding MetS that warrant continued attention.

First, the etiologic basis of MetS remains unknown. While it was originally hypothesized that insulin resistance represented the underlying pathology influencing all other MetS components, research has shown that not all individuals meeting criteria for MetS present as insulin resistant, and that insulin resistance may in fact denote another abnormality related to a more fundamental causal process (Kahn et al., 2005). Current conceptualizations have now broadened their scope to include abdominal obesity as an additional driving force underlying MetS pathogenesis (Alberti et al., 2009). However, multiple factor analysis studies suggest that the development of MetS may occur via

several distinct pathophysiological pathways, some of which have not yet likely been identified (see Kahn et al., 2005, for a review).

Second, current MetS definitions differ in their proposed criteria, and there appears to be no clear basis or empirical rationale for the inclusion or exclusion of specific risk factors in such criteria (Kahn et al., 2005). For instance, while MetS is commonly cited as a proinflammatory and prothrombotic state, markers of such processes (i.e., C-reactive protein, interleukin-6, tumor necrosis factor alpha, plasminogen activator inhibitor 1, etc.) are not incorporated into existing definitions (Kahn et al., 2005). Further, markers that are common across definitions are not always assessed similarly or given the same weight. For example, while insulin resistance is obligatory for MetS diagnosis using the World Health Organization criteria and may be assessed in multiple ways (i.e., via euglycemic clamp, fasting insulin levels, etc.), the most recent "harmonized" criteria lists hyperglycemia (assessed solely via fasting glucose level or glucose-lowering medication use) as a possible but unrequired component (Alberti & Zimmet, 1998). Studies have additionally shown that a significant proportion of individuals are discrepantly classified when using different MetS criteria and that across all criteria, standardized methods or justification regarding their construction and/or potential for modification are not provided nor described (Kahn et al., 2005).

Third, there appears to be no rationale for the individual component cutoff values chosen and specified in existing MetS criteria (Kahn et al., 2005). While the threshold values listed in these definitions are generally comparable to those documented by well-established guidelines based on observed independent effects, the use of either higher or lower values to optimize the positive predictive power associated with this

multidimensional construct lacks investigation (Kahn et al., 2005). Moreover, any specified cut-point may be viewed as arbitrary given that biological levels of the individual MetS components are continuous and that the risk they confer is progressive, as opposed to purely absent or present (Kahn et al., 2005). Additionally, risk factor dichotomies associated with any MetS definition will capture both individuals with overt disease (i.e., hypertension, diabetes, etc.) as well as those with elevated but milder forms of such conditions, for which differential CVD risk gradients are likely (Kahn et al., 2005).

Fourth, proposed MetS criteria specify different cut-points for men and for women (i.e., for the individual components of waist circumference and HDL cholesterol), suggesting that the risk conferred by some MetS components differs as a function of gender (Kahn et al., 2005). However, solid empirical evidence warranting the use of these discrepant thresholds is again lacking (Kahn et al., 2005; Regitz-Zagrosek, Lehmkuhl, & Mahmoodzadeh, 2007). Furthermore, studies have shown that there are significant differences between men and women in the predictive power of MetS, as well as its individual components, on CVD outcomes (Hari, Nerusu, Veeranna, Sudhakar, Zalawadiya, et al., 2011). Similar differences have also been observed between different racial/ethnic groups, leading to the development of additional ethnic-specific cut-points (i.e., for waist circumference) (Zhu, Heymsfield, Toyoshima, Wang, Pietrobelli, et al., 2005).

Fifth, much debate has centered on whether individual MetS components act synergistically to increase risk, or whether MetS as a whole conveys information that is no greater than the sum of its individual parts. Studies examining this question have

yielded mixed findings, with some suggesting that MetS diagnosis adds little or no incremental value (see Kahn et al., 2005, for a review). Relatedly, it has been argued that different MetS component combinations may confer differential risk for disease. For example, while no study has systematically examined all possible combinations for meeting MetS criteria (i.e., the 16 profiles possible using the harmonized definition), several studies examining incident CVD and measures of subclinical disease have suggested that different individual MetS components and component combinations convey substantially variable degrees of risk (see Kahn et al., 2005, for a review). And although individual MetS components may have differential CVD predictive power (i.e., hyperglycemia or overt diabetes may confer a greater CVD risk compared to other MetS components), each is weighted equally by current criteria (Kahn et al., 2005). Presently, the hierarchy of risk related to each possible MetS combination has not been established and remains largely unknown (Kahn et al., 2005). Such knowledge, however, could be invaluable in helping identify individuals at most risk for disease.

Lastly, whether MetS diagnosis confers any added medical value or benefits has been critically questioned. Given the heterogeneity of MetS presentations across individuals, and the inherent complexity in identifying and studying such varying presentations, tailored treatments specific to different MetS component combinations are non-existent (Kahn et al., 2005). Thus treatment of MetS does not currently differ from treatment of each of its individual components (Kahn et al., 2005). In fact, some have argued that MetS diagnosis may impede the vigorous treatment of individual CVD risk factors, hindering optimal care (Kahn et al., 2005). In stark contrast, others have asserted that MetS diagnosis may enhance and facilitate treatment, by motivating physicians to

assess interrelated abnormalities and by motivating patients to engage in healthy lifestyle behaviors (i.e., weight management/reduction and physical activity) that can positively and simultaneously impact all components (Pratley, 2007).

In conclusion, it has been argued that fundamentally important information is lacking to merit the continued used of MetS as a clinical construct, and that further work must be done to critically and innovatively examine 1) current MetS characterizations and criteria, 2) the advantages and disadvantages of incorporating specific and varying component cutoff values, 3) the added benefits of including and/or replacing traditional and/or novel CVD risk factors in current definitions, and 4) whether an underlying cause(s) exist and can be identified (Kahn et al., 2005). While in agreement that much research is needed, others have been less dismissive of MetS as a clinical construct, stating that "variability in the presentation and course of any disease is the rule rather than the exception in medicine" and that such variability has been documented across many diseases both prior and subsequent to the identification of their etiology (i.e., AIDS) (Pratley, 2007). Such proponents argue that variability is expected due to patients' heterogeneous biological make-up in interaction with their diverse environmental experiences, and that this should not dampen interest in MetS as holding important clinical utility and the potential to substantively aid future CVD reduction efforts (Pratley, 2007; Tenenbaum & Fisman, 2011).

Present Study

In summary, although H/Ls represent the largest and fastest growing minority population in the US, research on this heterogeneous group has been largely constrained (i.e., to Mexican Americans, to groups of low socioeconomic status, etc.) and fraught

with limitations. While generally exhibiting a higher prevalence of MetS and associated components (i.e., obesity and diabetes), US H/Ls debatably appear to exhibit lower rates of CVD. Nonetheless, the burden of CVD and associated risk factors in this population is high and projected to increase, contributing substantially to national health costs and disparities. Amidst these trends, however, there has been a dire lack of research on representative and diverse US H/L subgroups examining CVD, MetS, and associated risk factors. Additionally, ensuing debate regarding the clinical utility of current MetS diagnostic criteria calls for the application of novel investigative methods to further elucidate and characterize this construct. Toward this aim, employing a latent class analysis (LCA) approach to the study of MetS may address some current criticisms, offer several advantages, and provide valuable clinical insights to better inform CVD prevention and treatment strategies.

LCA is a term often restricted to situations where latent classes or clusters are extracted on the basis of binary indicators and latent profile analysis (LPA) is used to describe when the classes are extracted on the basis of continuous indicators. When classes are extracted on the basis of a combination of both binary and continuous indicators, the term latent cluster analysis is sometimes used. In this dissertation, both continuous and categorical indicators will be used, but we will retain the simpler, more generic term LCA to describe the methodology, and will refer to the resulting groups as clusters.

LCA aims to classify similar individuals into groups (or latent clusters), in which each latent cluster is viewed as consisting of homogeneous individuals with regards to the observed variables being studied (in our case, components of MetS), and the different

latent clusters are viewed as representing the unobserved heterogeneity among individuals in these observed variables (Vermunt & Magidson, 2002). Considering the possibility that relationships among MetS components may meaningfully differ in subgroups of individuals, this person-centered analytic approach strives to identify quantitatively and qualitatively distinct profiles of individuals based on their presentation of MetS components, which may further be analyzed 1) with respect to other variables of interest (i.e., gender, age, H/L subgroup, smoking status, SES, etc.) and 2) in associating MetS with disease outcomes (i.e., prevalent CHD and cerebrovascular disease). Additionally, in contrast to current MetS diagnostic criteria which center solely on risk factor elevations for classification, LCA may allow for the identification of MetS subtypes that differ in which risk factors are elevated as well as the degree to which they are elevated, thus allowing for a more detailed examination of MetS by using continuous measures of its components (as opposed to cutoff scores or dichotomous classifications).

Thus, by employing a LCA approach and using data from the largest and most comprehensive study of diverse US H/Ls to date, the current study aimed to investigate 1) whether distinct subgroups of individuals can be identified based on their presenting levels of MetS components, 2) how the identified MetS subtypes are associated with various demographic, clinical, socioeconomic, and behavioral characteristics, and 3) whether the identified MetS subtypes are differentially associated with CHD and cerebrovascular disease prevalence in cross-sectional analysis. Results of such investigation (i.e., the identification of distinct MetS clusters, as well as their association with various risk factors and prevalent disease) may offer valuable insights into the cardiovascular health of diverse H/Ls living in the US, may help advance our current

conceptualization of MetS as a clinical construct, and may allow for improved identification of individuals at most risk for CVD potentially leading to refined targeted prevention efforts. Moreover, inspection of the resulting MetS subtypes may shed light on the underlying mechanisms involved in increasing CVD risk, and may generate hypotheses as to whether these are likely to be similar or different among distinct identified MetS presentations, as well as between different CVD outcomes (i.e., CHD versus cerebrovascular disease).

Since this study aims to investigate a relatively novel approach to the study of MetS, it is essentially exploratory. To our knowledge, this is the first study to employ a LCA approach to identifying subtypes of MetS using continuous measures of associated components. Thus, the hypotheses corresponding to each of the following aims are preliminary and tentative.

Specific Aim 1: To examine whether homogeneous subtypes of individuals exhibiting phenotypically different profiles of MetS components (waist circumference, systolic and diastolic blood pressure, high-density lipoprotein cholesterol, triglycerides, fasting glucose, and use of antihypertensive, lipid-lowering, and glucose-lowering medications) can be identified among a large and diverse sample of US H/Ls using a LCA approach.

Hypothesis 1: Several homogeneous subtypes of individuals will be identified as a function of exhibiting distinct profiles of MetS components. Specifically, one group may exhibit no clinically significant elevations across all components, one group may exhibit clinically significant elevations across all components, and other groups may exhibit clinically significant elevations in certain components and not others.

Specific Aim 2: To examine whether identified MetS subtypes of individuals differ in regards to various demographic (age, sex, H/L subgroup), clinical (family history of CHD, family history of stroke), socioeconomic (income, education), and behavioral (smoking status) characteristics.

Hypothesis 2: Identified MetS subtypes of individuals will differ in several characteristics. Specifically, subtypes exhibiting higher and/or certain MetS component elevations may be disproportionately represented by older individuals, men, individuals with a positive family history of CHD and stroke, current and former smokers, and individuals of low socioeconomic status compared to subtypes exhibiting lower and/or different patterns of MetS component elevations. In the face of limited data, it is not hypothesized that the composition of individuals within identified MetS subtypes will significantly differ as a function of H/L background.

Specific Aim 3: To cross-sectionally examine whether identified MetS subtypes of individuals differ in their prevalence of CHD and cerebrovascular disease.

Hypothesis 3: Identified MetS subtypes of individuals will differ in their prevalence of CHD and cerebrovascular disease. Specifically, subtypes exhibiting higher and/or certain MetS component elevations will have an increased prevalence of CVD compared to subtypes exhibiting lower and/or different patterns of MetS component elevations.

CHAPTER 2: METHODS

Participants

Participants were those of the Hispanic Community Health Study/Study of Latinos (HCHS/SOL), a multi-center epidemiologic study initiated in 2006 by the National Heart, Lung, and Blood Institute (NHLBI) and six other branches of the National Institutes of Health (NIH). The overall objectives of HCHS/SOL are to investigate the prevalence and development of disease in diverse US Hispanic/Latino (H/L) populations, as well as to identify associated risk and protective factors (Sorlie, Aviles-Santa, Wassertheil-Smoller, Kaplan, Daviglus, et al., 2010). Participants were recruited from four US communities: the Bronx, NY, Chicago, IL, Miami, FL, and San Diego, CA. These sites were selected to ensure adequate representation of persons from the following H/L backgrounds: Mexican, Puerto Rican and Dominican, Cuban, and Central and South American. At time of enrollment, participants had to self-identify as Hispanic or Latino and be between 18 and 74 years of age. A two-stage area probability sampling approach was employed with stratification and oversampling at each stage to efficiently provide a broadly diverse and representative sample meeting target age distribution objectives. Based on preliminary reports, the final sample included 16,479 participants (40.3% men; 59.2% aged 45 or older; 39.2% Mexican, 16.6% Puerto Rican, 14.3% Cuban, 10.5% Central American, 9% Dominican, 6.5% South American, and 4% more than one of these or Other). This study will analyze available preliminary data on a subset of these participants (N = 10970). This subset of participants excluded individuals who self-reported as “Other” or “More than one” in terms of H/L subgroup identification (N = 431).

Measures

All measures were administered by centrally trained and certified technicians following standardized protocols common to all study sites. Data collection was guided by a specialized Data Entry and Management System designed to enhance data accuracy and security. Measures collected via interview format were administered by bilingual personnel in either English or Spanish at the preference of the study participant. Translations of measures were reviewed by a Translation and Validation Subcommittee, which had representation from all study sites, the Coordinating Center, and the Project Office, as well as representation from Mexicans, Cubans, Puerto Ricans, and Central/South Americans. In addition, several questionnaires were tested on focus groups of community volunteers representing various H/L countries of origin at each study site. Final translations were certified by the Research Triangle Institute (RTI), with expertise in multilingual instrument development for large-scale surveys. Detailed information regarding all measures administered can be found in the HCHS/SOL protocols and manuals available at http://www.cscc.unc.edu/hchs/.

Metabolic Syndrome Components

Waist circumference (WC) was measured at the uppermost lateral border of the right ilium to the nearest 0.1 cm using a measuring tape. Resting blood pressure was measured three times in the right arm using an automatic sphygmomanometer (Omron model HEM-907 XL, Omron Healthcare Inc., Bannockburn, IL) and appropriate cuff sizes. Readings were taken after 5 minutes in the seated position. The three readings were averaged to obtain the systolic and diastolic blood pressure levels used in analyses.

Blood specimen samples were collected via venipuncture and processed at each of the four field centers, and then shipped daily to a Central Laboratory (located at the University of Minnesota Medical Center, Fairview in Minneapolis, MN) for assay. Triglycerides (TG) were measured in serum on a Roche Modular P Chemistry Analyzer (Roche Diagnostics Corporation) using a glycerol blanking enzymatic method (Roche Diagnostics, Indianapolis, IN 46250). High-density lipoprotein cholesterol (HDL-C) was measured in serum on a Roche Modular P Chemistry Analyzer (Roche Diagnostics Corporation) using a direct magnesium/dextran sulfate method (Roche Diagnostics, Indianapolis, IN 46250). Fasting glucose was measured in EDTA plasma on a Roche Modular P Chemistry Analyzer (Roche Diagnostics Corporation) using a hexokinase enzymatic method (Roche Diagnostics, Indianapolis, IN 46250). The Central Laboratory had quality control procedures in place to assess analytical performance.

A standard questionnaire and interview was used to collect information about participants' use of medications during the four weeks preceding their baseline examination. Participants were instructed and reminded to bring with them all prescription medications taken within the last month to their initial clinic visit. Ascertainment included scanning of twelve-digit Universal Product Code (UPC) bar code symbols when available. Medical Therapeutic Classification (coding) was automated where possible. Otherwise, manual coding was centralized (performed only in the Coordinating Center). Use of each medication class (antihypertensive, lipid-lowering, and glucose-lowering) was represented as a dichotomous variable (yes or no).

Cardiovascular Disease Prevalence

A standard digital 12-lead electrocardiogram (ECG; GEMSIT MAC 1200 portable electrocardiograph) was acquired for each participant. Procedures regarding the timing of ECG exams with respect to other study activities (i.e., glucose load and glucose tolerance test, snacks/meals, lung functioning test, etc.) were in place to assure that ECG results were not influenced by participant increases in blood sugar or bronchodilator use. ECG findings were electronically transmitted to a Central ECG Reading Center (The Epidemiological Cardiology Research Center (EPICARE) of Wake Forest University's School of Medicine). Clinical readings of the ECG pattern and ascertainment of possible old myocardial infarction (MI) was based on the Minnesota Code system of classification, following precise guidelines to determine wave duration and voltage (i.e., major Q wave abnormalities or minor Q,QS waves with ST,T abnormalities). The Central ECG Reading Center provided quality control feedback to ECG technicians throughout the study period and a Quality Control Committee periodically reviewed data quality.

A standard questionnaire and interview was used to collect self-reported information on angina ("Has a doctor ever said that you have angina?"), heart attack ("Has a doctor ever said that you had a heart attack?"), and coronary procedures ("Have you had a balloon angioplasty, a stent, or bypass surgery to the arteries in your heart to improve the blood flow to your heart?"). A standard questionnaire and interview was also used to collect self-reported information on stroke ("Has a doctor ever said that you had a stroke?"), mini-stroke or transient ischemic attack ("Has a doctor ever said that you had a mini-stroke or TIA (transient ischemic attack)?"), and cerebrovascular procedures ("Have

you had a balloon angioplasty or surgery to the arteries of your neck to prevent or correct a stroke?").

Prevalent CHD was represented as a dichotomous variable (yes or no) that combined ECG reports of possible old MI as well as self-report of heart attack and coronary procedures (angioplasty, stent, or bypass). Prevalent CHD further including self-report of angina was also examined. Prevalent cerebrovascular disease was represented as a dichotomous variable that combined self-reported information on stroke, mini-stroke or transient ischemic attack, and cerebrovascular procedures.

Covariates

Standard questionnaires and interviews were used to collect information about demographic characteristics (age, sex, and H/L subgroup), smoking status, family history of CHD and stroke, and socioeconomic status (income and education). Age was examined as a continuous variable. Sex was examined as a dichotomous variable (male or female). H/L subgroup was represented by four dummy coded variables: Puerto Rican, Cuban, Dominican, and Central/South American, with Mexican serving as the reference group. Smoking status was represented by two dummy coded variables: former smoking and current smoking, with never smoking serving as the reference group. Family history of CHD (history of myocardial infarction in either parents or siblings) and family history of stroke (history of stroke in either parents or siblings) were examined as dichotomous variables (positive or negative). Total gross family income was examined as a 5-level categorical variable (<$10,000, $10,000 to $20,000, >$20,000 to $40,000, >$40,000 to $75,000, or >$75,000). Education was examined as a dichotomous variable (no high school diploma/GED or at least high school diploma/GED).

Procedure

The sampling and recruitment plan for HCHS/SOL was designed to support two analytical objectives: 1) that the study sample support estimates of prevalence of baseline risk factors, both overall and by H/L background and other demographic subgroups, and 2) that the study sample support evaluation of the relationship between the various risk factors and disease outcomes measured during follow-up. The study thus aimed to recruit a cohort of 16,000 H/Ls aged 18-74 years (with 62.5% over 44 years of age), with adequate representation of H/L subgroups to support inferences by H/L background. A detailed description of the HCHS sample design and cohort selection has been previously published (LaVange, Kalsbeek, Sorlie, Aviles-Santa, Kaplan, et al., 2010).

Sampling

Target areas in each study site (the Bronx, NY, Chicago, IL, Miami, FL, and San Diego, CA) were defined by groups of neighboring census tracts to provide geographical balance and diversity with respect to H/L background. These target areas were purposefully selected based on their proximity to clinic sites, tract-level demographic distributions available from the 2000 Decennial US Census, and local information about neighborhoods. To ensure broad representation of the target population and minimize sources of bias, a hybrid approach to cohort identification and selection was used that combined deliberate selection of community areas and random selection of households within those areas.

In each of the four study sites, a stratified two-stage area probability sample of household addresses was selected. At the first stage, a stratified simple random sample of census block groups was selected and served as primary sampling units (PSU). PSU

sampling strata were defined by the cross-classification of 1) high and low H/L concentration, and 2) high and low socioeconomic status. Block groups in the high H/L concentration stratum were oversampled to maximize efficiencies in field center operations by increasing the probability of selecting H/L households. Additional special strata were created as needed to target specific neighborhoods in the Bronx and Miami field centers. At the second stage, stratified samples of household addresses were selected within each sample PSU from non-overlapping lists of postal addresses and H/L surnames. Household addresses using H/L surname lists were oversampled to further maximize field center operations. To minimize bias resulting from temporal trends and ensure that each yearly sample over the three-year recruitment period was representative of the target areas, the sample of households in each target area was randomly subsampled to form three waves that corresponded to the three years of recruitment. To meet age distribution objectives for the final cohort (62.5% aged 45-74 years), eligible households or persons within households were further subsampled during the screening stage according to the household's age distribution. Once a household was selected, all eligible members of the household were invited to participate.

Recruitment

Each field center aimed to enroll 4,000 participants with predominant representation of one or more H/L subgroup: Puerto Ricans and Dominicans in the Bronx field center; Mexicans in the San Diego field center; Cubans and Central and South Americans in the Miami field center; and Mexicans, Puerto Ricans, and Central and South Americans in the Chicago field center. Extensive community engagement efforts and recruitment procedures were selected to optimize the ability to establish contact with,

determine eligibility of, and actively engage households at every sample address, regardless of neighborhood characteristics or living conditions. The recruitment protocol consisted of three steps: 1) advanced mailings to sample addresses describing the study, 2) optional telephone contacts for households with telephone numbers available, and 3) in-person contacts. Recruitment teams informed potential participants of the study objectives and associated benefits of their participation, emphasizing the nature of study, the information it is designed to provide, as well as the impact the study results may have on policy making and healthcare for future US H/L generations.

Screening

The target population in HCHS/SOL corresponded to all noninstitutionalized H/Ls aged 18-74 years residing in the four sampled areas. Once contact with sample households was established, a brief household screener was administered to determine eligibility. Persons were considered ineligible if they were 1) on active military duty, 2) not currently living at home, 3) planning to move from the area in the next 6 months, or 4) physically unable to attend the clinic examination. Eligible women who were pregnant were rescheduled for a clinic visit approximately three months postpartum. Otherwise, there was no exclusion of persons based on existing health status.

Assessments

Eligible individuals attended a baseline clinic examination to assess cardiovascular and other disease risk factors, both known and potential. This examination consisted of a series of fixed and flexible components organized to accommodate appropriate collection of measures, and lasted approximately seven hours. A detailed description of the standardized examination content and its typical flow and duration has

been previously published (Sorlie, Aviles-Santa, Wassertheil-Smoller, Kaplan, Daviglus, et al., 2010). The institutional review board at each study site and the Coordinating Center approved the study protocol. All sites provided a van or taxi service to assist participant attendance, and participants were reimbursed for expenses involved in attending the clinic examination. Informed consent was obtained for all participants.

CHAPTER 3: DATA ANALYSIS PLAN

Preliminary Analyses

Data Screening

Univariate distributions for all observed measures were examined for normality (i.e., skew < 3 and kurtosis < 10). Additionally, the relative variances between variables were assessed to examine the potential for convergence problems when estimating the proposed models due to ill-scaled variance-covariance matrices.

Missing Data

Missing data were handled using the full information maximum likelihood (FIML) approach. FIML estimation uses all available data to estimate group parameters by obtaining a likelihood function for each participant based on the data that is present for that participant (Arbuckle, 1996). The likelihoods are then summed across participants. Thus, all participants with any available data relevant to a given parameter can contribute to that parameter's estimation. The use of FIML assumes that missing data are either missing completely at random (MCAR; missing data in a variable are unrelated to other observed variables and to the values of that variable itself) or missing at random (MAR; missing data in a variable are related to other observed variables available for analysis).

Primary Analyses

All analyses were cross-sectional using participant data from the baseline clinic examination. The stratification, clustering, and sampling weights determined by the Coordinating Center and described in their Analysis Methods Document were applied. All analyses were conducted using Mplus version 6 (Muthen & Muthen, 1998-2010).

Latent Class Analysis

A latent class analysis (LCA) approach was used to identify subtypes of individuals with similar patterns of MetS components. The aim of LCA is to classify similar individuals into groups (or latent clusters), in which each latent cluster is viewed as consisting of homogeneous individuals with regards to observed variables being studied (in our case, MetS components), and the different latent clusters are viewed as representing the unobserved heterogeneity among individuals with regards to these observed variables (Vermunt & Magidson, 2002). LCA is a person-centered analytical approach (as opposed to a variable-centered technique), which allows for the consideration that relationships among MetS components may differ in meaningful ways for distinct subgroups of individuals. In other words, individuals may come from distinct subpopulations in which the observed associations between components of MetS may differ both quantitatively as well as qualitatively (Morin, Morizot, Boudrias, & Madore, 2011).

The parameters of interest in LCA are the cluster probabilities (cluster sizes) and the conditional probabilities of a specific level on the observed variables (i.e., MetS components) given the cluster membership (cluster structure). The means, variances, and covariances of the observed variables are estimated for each latent cluster using a maximum-likelihood or maximum-posterior method (Vermunt & Magidson, 2002). The number of clusters, their sizes, and their characteristics are examined and do not need to be known a priori (Vermunt & Magidson, 2002).

LCA has many advantages compared to standard cluster analysis techniques, including 1) flexibility in parameter restrictions on the covariance structure that may

allow for more accurate and realistic conceptualizations of the data, 2) the ability to simultaneously include varying scaled and scale-type data (i.e., continuous, ordinal, and categorical measurement scales) in the same model, and 3) more formal criteria for selecting best-fitting models, among others (Vermunt & Magidson, 2002).

Additionally, LCA can be used to examine covariates associated with observed latent clusters (i.e., age, gender, H/L subgroup, socioeconomic status, smoking status, etc.). This is achieved by directly including covariates (or predictors) in the models, thus limiting the potential for type 1 errors and biased observations by estimating all relationships in a single analytical step (i.e., taking into account the model-estimated posterior probabilities versus dichotomously assigning individuals into latent clusters using their most likely probabilities) (Morin et al., 2011). Although they may be assumed to alter the cluster probabilities (cluster sizes), it has been posited that the inclusion of covariates should not define or change the nature or shape of identified profiles, thus allowing for the stability of models to be examined (Morin et al., 2011). The inclusion of covariates in LCA has also been used to assess the construct validity of classifications by assessing whether identified profiles relate to such variables in theoretically meaningful ways (Morin et al., 2011). It has also been suggested that the inclusion of covariates can lead to improvements in parameter coverage and the accuracy of classification (Lubke & Muthen, 2007).

Furthermore, LCA can also be used to examine how latent clusters predict concurrent and distal outcomes (i.e., CHD and cerebrovascular disease prevalence) (McCutcheon, 2002). This is also achieved by directly and simultaneously including outcomes in the models, again reducing error and bias. Similar to the inclusion of

covariates, examination of identified profiles in relation to various outcomes has also been used to verify the construct validity of classification (Morin et al., 2011).

Analysis of Aim 1

LCA was used to investigate the number and types of latent clusters underlying the distribution and associations among the following MetS components: waist circumference, systolic blood pressure, diastolic blood pressure, high-density lipoprotein cholesterol, triglycerides, fasting glucose, antihypertensive medication use, lipid-lowering medication use, and glucose-lowering medication use (see Figure 1). These variables were used as continuous indicators with the exception of medication use data, which were used as binary indicators. Models were first tested on the entire sample, and were then tested on men and women separately. This allowed for the examination of whether MetS subtypes differed quantitatively and/or qualitatively between genders, and importantly allowed for MetS component estimates to be compared with currently employed gender-specific cutoff values.

While traditional latent class analysis assumes 1) local independence, or that all covariances within classes are equal to zero (i.e., the latent profiles suffice to account for the observed correlations between indicators), and 2) homogeneity of variance, or that the error structure across classes has the same form while exhibiting different locations of the indicators, these assumptions are often unrealistic with real-life data and research questions (Vermunt & Magidson, 2002). Moreover, imposing such restrictions may lead to the emergence of spurious classes, especially when indicators (i.e., some components of MetS) are known to be interrelated (Morin et al., 2011). LCA allows for the examination of different model constraints, and the importance of assessing various

model specifications pertaining to these assumptions has been documented (Vermunt & Magidson, 2002). Therefore, 4 different model specifications with regards to restrictions imposed on the variance-covariance matrix were tested. In Model 1, the traditional assumptions described above were retained, and thus covariance terms were constrained to zero within clusters and variances were assumed to be equal across clusters. In Model 2, selected covariances between the blood pressure variables (diastolic and systolic blood pressure) and the lipid variables (HDL cholesterol and triglycerides) were freely estimated while variances were assumed to be equal across clusters. In Model 3, cluster-dependent variances were freely estimated while all covariance terms were constrained to zero. In Model 4, selected covariances between the blood pressure variables (diastolic and systolic blood pressure) and the lipid variables (HDL cholesterol and triglycerides) were freely estimated along with cluster-dependent variances.

In the current study, models from 1 to 20 latent clusters were specified. This decision was informed by current MetS criteria requiring elevations in at least 3 of 5 components for diagnosis, allowing for 20 possible combinations of MetS profiles. The following indices were used to guide, evaluate, and select best-fitting models: Akaike Information Criteria (AIC), Bayesian Information Criteria (BIC), sample size adjusted BIC (ABIC), entropy, and the Adjusted Lo-Mendell-Rubin Likelihood-Ratio Test (Adjusted LRT). Comparatively lower values on parsimony criteria such as AIC, BIC, and ABIC indicate better fitting models. A measure of classification uncertainty, entropy values range from 0 to 1, with values near 1 indicating high certainty in classification. The Adjusted LRT was used to compare models of differing cluster sizes, with a significant p value ($p < .05$) indicating that a larger number of clusters fit the data better

whereas a p value > .05 indicates a smaller number of clusters are a better fit for the data. Models were also carefully examined and evaluated with respect to cluster sizes and, importantly, theoretical and clinical meaningfulness in relation to previous research.

An important concern in LCA regards the convergence of models on a local solution, or a false maximum likelihood, often stemming from the use of inadequate starting values (Vermunt & Magidson, 2002). To avoid the likelihood of this occurring, and following recommendations to use multiple random sets of starting values, this study requested a priori that 100 random sets of starting values to be used in the initial generation stage (also allowing for a maximum of 20 iterations), retaining the best 10 for final optimization. Moreover, the final retained model was estimated a second time using 2,000 random sets of starting values to increase confidence in having achieved a true global maximum likelihood.

Analysis of Aim 2

After the final model from the analysis of Aim 1 was chosen, the covariates of interest (age, gender, H/L subgroup, smoking status, family history of CHD, family history of stroke, income, and education) were directly included as predictors of MetS latent cluster assignment (see Figure 2). LCA uses logistic regression to relate the latent clusters to the covariates, in which the covariates are not directly included in the classification procedure (Morin et al., 2011). The paths from each covariate to the latent clustering variable were examined for significance ($p < .05$).

Analysis of Aim 3

To examine the cross-sectional association between identified MetS subtypes and prevalent CVD, each outcome of interest (prevalent CHD without self-reported angina,

prevalent CHD including self-reported angina, and prevalent cerebrovascular disease) was directly included in the final retained model, separately (see Figure 3). LCA uses logistic regression to relate the latent clusters to the outcome. In these analyses, each outcome was also regressed on the covariates to control for their effects. For each outcome model, the path from the latent clustering variable to the outcome variable was examined for significance ($p < .05$).

CHAPTER 4: RESULTS

Preliminary Analyses

Data Screening and Transformations

Non-normal distributions were observed for triglycerides and fasting glucose values, and these variables were log-transformed. These log-transformed variables were further multiplied by a constant of 100 to make their variances less discordant with the variances of the other variables in the model.

Missing Data and Sample Description

Two participants had missing data on at least one of the complex design variables (the primary sampling unit cluster variable or the 2-stage sample design stratification variable), and were excluded from analyses. Furthermore, five participants were also excluded in the covariate-free analyses due to missing data on all MetS components. The baseline demographic, risk factor, and prevalent disease information for the entire sample, as well as for men and women separately, are presented in Table 1.

Primary Analyses

Number of Latent Clusters

Fit indices for the 4 differentially specified LCA models conducted on the entire sample and extracting increasing numbers of clusters are presented in Table 2. These results were analogous to those obtained for the gender-specific models (data not presented). Across all models, loglikelihood (LL), AIC, BIC, and sample size adjusted BIC (ABIC) indices were observed to decrease as the number of clusters extracted increased for the full range of number of clusters considered, suggesting that a greater number of clusters fit the data progressively better. This was also suggested by significant

adjusted LRT (ALRT) *p* values, which similarly indicated that a larger number of clusters represented a better fit to the data. However, this was not surprising given the large size of our study sample and the dependency of these indices on sample size. It is important to note that across all models, the largest decrease in these fit indices occurred when comparing the two-cluster to the one-cluster model, with indices further decreasing at much smaller magnitudes when examining models extracting three or more clusters. Additionally, across all models, the quality of classification as indicated by the entropy values was best for the two-cluster solution. These relative commonalities suggested the potential superiority of a two-cluster model. Moreover, the two-cluster model estimates were the most stable across the different model specifications. Detailed examination of models extracting an increasing number of clusters revealed further – albeit less systemtatic and interpretable – re-classification of individuals within each of the two profiles initially identified in the two-cluster solution. Thus, the two-cluster model appeared to be consistently and substantially more interpretable according to current theoretical, empirical, and clinical conceptualizations of MetS.

Furthermore, according to most indices, the least restrictive model allowing for class-dependent variances as well as covariances between the 2 blood pressure variables and the 2 lipid variables (Model 4) consistently exhibited a better fit to the data compared to the more restrictive models (Models 1-3). This was observed in lower LL, AIC, BIC, and ABIC values across all class number specifications. In fact, only the entropy values were slightly lower for Model 4 compared to the other models, suggesting a small reduction in the precision with which individuals were classified into the various extracted latent clusters. Nonetheless, while solutions from the different models (in terms

of estimated cluster means and overall profiles) were substantively alike, the results for Model 4 appeared to be the clearest and made the most sense in relation to previous research and the nature of the groups.

On these bases, the least restrictive two-cluster model (Model 4) solution was selected as best representing the data. This model yielded adequate classification of individuals, with an entropy value of 0.836 for the entire sample (0.850 for men and 0.818 for women), indicating the extraction of distinct clusters. As presented in Table 3, this was further evidenced by elevated average posterior probabilities of cluster membership ranging from 0.927 to 0.964 for the entire sample (0.936 to 0.967 for men and 0.937 to 0.956 for women), and very low cross-probabilities ranging from 0.036 to 0.073 for the entire sample (0.033 to 0.064 for men and 0.044 to 0.063 for women). This two-cluster model was also replicated using a greater number of starting values, increasing confidence that this solution did not converge on a local maximum.

Description of Latent Clusters

The means of the MetS components in the two latent clusters of this retained model conducted on the entire sample, as well as on men and women separately, are reported in Table 4. The gender-specific model results and interpretations paralleled those observed for the entire sample. Given this observation, as well as the fact that commonly used cutoff values for some MetS components (i.e., waist circumference and HDL cholesterol) differ for men and women, the gender-specific results will be described to aid readers in the interpretation of the findings. The commonly used NCEP-ATP III criteria for diagnosing MetS are reproduced in Table 5 to further assist readers in making

comparisons (Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults, 2001).

The first latent cluster was named "Non-MetS" based on the mean profile of MetS components. In both the male- and female-specific models, individuals exhibiting relatively healthy mean levels across all MetS components characterized this cluster. The single exception was the average waist circumference observed for women in this cluster, which was slightly elevated compared to the NCEP-ATP III threshold of 88cm (Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults, 2001). This cluster described 81.3% of the participants among men, and 72.5% of the participants among women. The means of MetS components observed among men and women, respectively, were 96.0 and 93.7 cm for waist circumference (SD = 12.8 and 13.3), 119.9 and 109.9 mm Hg for systolic blood pressure (SD = 11.4 and 12.2), 72.0 and 68.2 mm Hg for diastolic blood pressure (SD = 9.8 and 9.4), 44.7 and 52.3 mg/dL for HDL cholesterol (SD = 10.9 and 13.6), 116.1 and 92.4 mg/dL for triglycerides (transformed means: 206.5 and 196.6, SD = 24.9 and 21.7), and 95.1 and 90.5 mg/dL for fasting glucose (transformed means: 197.8 and 195.7, SD = 3.6 and 3.4). With the exception of waist circumference among women, each of the mean values is below the diagnostic threshold proposed by all major MetS definitions. Additionally, the estimated proportion of individuals that were taking examined medications was low. Among men, 3.8% were on antihypertensive medication, 3.0% were on lipid-lowering medication, and 0.2% were on glucose-lowering medication. The corresponding proportions among the women in this cluster were 1.8%, 1.5%, and 0.1%.

The second latent cluster was named "MetS" based on their mean profile of MetS components. In both gender-specific models, and in contrast to the Non-MetS latent cluster described above, this cluster was characterized by individuals exhibiting clinically elevated mean levels across most MetS components. This cluster described 18.7% of the male participants and 27.5% of the female participants. The means of the MetS components observed among men and women, respectively, were 105.8 and 103.8 cm for waist circumference (SD = 16.0 and 14.7), 135.8 and 132.4 mm Hg for systolic blood pressure (SD = 21.1 and 20.6), 79.1 and 76.1 mm Hg for diastolic blood pressure (SD = 13.4 and 11.7), 44.9 and 50.0 mg/dL for HDL cholesterol (SD = 14.1 and 12.8), 156.1 and 142.3 mg/dL for triglycerides (transformed means: 219.3 and 215.3, SD = 26.3 and 21.5), and 132.2 and 116.0 mg/dL for fasting glucose (transformed means: 212.1 and 206.4, SD = 16.6 and 13.7). With the exception of diastolic blood pressure and HDL cholesterol among men and women, as well as triglycerides among women, each of the other mean MetS component values were above proposed MetS diagnostic thresholds. Additionally, the estimated proportion of individuals taking examined medications was high. Among men, 45.7% were on antihypertensive medication, 34.1% were on lipid-lowering medication, and 36.6% were on glucose-lowering medication. The corresponding proportions among the women in this cluster were 40.6%, 30.5%, and 29.0%.

Effects of Covariates on Latent Cluster Membership

The direct inclusion of covariates (age, H/L subgroup, education, income, and smoking status) in both of the gender-specific retained models did not change the characteristics of the clusters, further confirming the stability of the two-cluster model.

The parameter estimates for the models including covariates are also presented in Table 4, alongside those of the covariate-fee models previously described. As can be observed, the estimated means and standard deviations were essentially the same between the models that included and excluded covariates, with the models including covariates showing only small decreases in mean triglyceride and glucose levels in the MetS clusters for both men and women, as well as slight changes in medication use proportions in both clusters. The direct inclusion of covariates did change the proportion of individuals classified into each cluster in both men and women, decreasing the proportion of individuals classified into the Non-MetS cluster (77.3% vs. 81.3% of men and 66.8% vs. 72.5% of women in the covariate vs. covariate-free models, respectively) and consequently increasing the proportion of individuals classified as belonging to the MetS cluster (22.7% vs. 18.7% of men and 33.2% vs. 27.5% of women in the covariate vs. covariate-free models, respectively).

The relationships between the covariates and the clusters are presented in Table 6. Among both men and women, being older (OR = 1.119, $p < .001$ for men; OR = 1.162, $p < .001$ for women) and having a positive family history of CHD (OR = 1.478, $p < .01$ for men; OR = 1.371, $p < .05$ for women) was associated with significantly greater odds of belonging to the MetS latent cluster compared to the Non-MetS latent cluster. Among women but not men, having a lower education level (OR = 0.604, $p < .01$), a lower family income (OR = 0.779, $p < .001$), never smoking relative to current smoking (OR = 0.668, $p < .05$), and being of Puerto Rican descent relative to Mexican descent (OR = 2.070, $p < .001$) was also associated with significantly greater odds of being classified into the MetS latent cluster compared to the Non-MetS latent cluster. With the exception

of the observed smoking effect among women, these results are consistent with previous research on predictors of MetS (Roger et al., 2011), suggesting adequate construct validity for the extracted latent clusters. Among both men and women, latent cluster classification was not significantly related to family history of stroke, former smoking relative to never smoking, or being of Cuban, Dominican, or Central/South American descent relative to Mexican descent.

Latent Cluster Membership and Prevalent Cardiovascular Disease

The direct inclusion of all outcomes (prevalent CHD without and with inclusion of self-reported angina, and self-reported cerebrovascular disease) in both of the gender-specific retained models did not change the characteristics of the clusters, again confirming the stability of the two-cluster model (data not shown). The covariate-adjusted relationships between the prevalent CVD outcomes and the two latent clusters are presented in Table 7. Among men, the odds of having any type of prevalent CVD were significantly higher among individuals classified as belonging to the MetS group compared to those classified as belonging to the Non-MetS group (CHD OR = 1.092, $p < .05$; CHD including self-reported angina OR = 1.115, $p < .01$; cerebrovascular disease OR = 1.051, $p < .05$). Among women, the odds of having prevalent CHD including self-reported angina was also significantly higher among individuals classified as belonging to the MetS group compared to those classified as belonging to the Non-MetS group (OR = 1.068, $p < .05$). However, among women, the associations between the extracted latent clusters and both the CHD without self-reported angina (OR = 1.048, $p = 0.097$) and the self-reported cerebrovascular disease (OR = 1.028, $p = 0.129$) prevalence outcomes were non-significant. For interested readers, the Mplus input statement of the retained two-

cluster model predicting CHD prevalence adjusting for covariates among men is provided in the Appendix.

CHAPTER 5: DISCUSSION

Using preliminary data from HCHS/SOL, this study employed a LCA approach to investigate whether distinct subtypes of MetS could be identified among a large and diverse sample of US H/Ls aged 18 to 74 years. Incorporating continuous measures of MetS components, along with corresponding medication use data, two distinct clusters of individuals were identified among the entire sample, as well as among men and women separately. This two-cluster model was selected as best representing the data on the basis of various statistical fit indices, parsimony, as well as correspondence to substantive theory and previous scientific knowledge. Consistent with hypotheses, one cluster of individuals was characterized by exhibiting relatively healthy mean levels across all MetS components, whereas a second and comparatively smaller cluster of individuals was characterized by exhibiting mean clinical elevations across most MetS components. These two latent clusters proved to be stable across different model specifications, and were associated with multiple covariates and concurrent CVD prevalence outcomes in a coherent manner, adding further support to their adequacy and construct validity. Contrary to hypotheses, the presence of additional MetS subtypes of individuals that were deemed by the investigators to be well-defined, clearly interpretable, and to possess additional and substantial empirical or clinical utility was not verified via this analytic methodology.

When conducting and evaluating LCA models, it has been well-documented that different model specifications may lead to different results and subsequent conclusions regarding the number and types of latent clusters underlying relationships among a given set of observed variables (Bauer & Curran, 2004; Vermunt & Magidson, 2002). To

address such concerns and reduce the potential of retaining a faulty model, four differentially specified models were tested that imposed varying, systematic restrictions on the variance-covariance matrix. Several other model specifications were examined in addition to these but were not reported. Additionally, to decrease the possibility of having retained a model that converged on a local or "false" maximum likelihood, multiple random starts were used to estimate all models. As anticipated, the model that allowed for the free estimation of cluster-dependent variances among all MetS components as well as covariances between the blood pressure variables and the lipid variables consistently showed a better fit to the data – albeit slightly reduced entropy values – compared to the more restrictive models tested, and was thus retained.

The decision to select a two-cluster solution as best representing the data was aided by detailed examination of various statistical fit and classification indices, and, importantly, by subjective evaluation of the nature of the extracted clusters in relation to current theoretical, empirical, and clinical conceptualizations of MetS. Although certain fit indices suggested that models extracting an increasingly greater numbers of clusters might better explain the data, the application and utility of such sample size dependent criteria was rendered suspect given the large size of our sample (for a more comprehensive discussion of this issue, readers are referred to pp. 214 – 216 in Marsh, Ludtke, Trautwein, & Morin, 2009). Relatedly, the practice of relying solely on statistical fit indices to guide and evaluate LCA model selection has been criticized, and investigators have urged that strong precedence must also be given to the subjective examination of associated model parameter estimates in relation to theory and a priori expectations (Marsh, Hau, & Wen, 2004; Marsh, Hau, & Grayson, 2005; Marsh et al.,

2009). The two-cluster model was retained after careful consideration of all these recommendations, and was strongly guided by its stability and convergence with current scientific knowledge regarding MetS, as well as observed support for its construct validity (described below).

The first identified cluster, labeled "Non-MetS," comprised approximately 80% of the entire sample (81% of men and 73% of women). With the exception of waist circumference among women, individuals classified as belonging to this cluster exhibited mean levels of MetS components that were below currently identified diagnostic threshold values and were significantly less likely to be on antihypertensive, lipid-lowering, and glucose-lowering medications. However, significant variability across these mean levels was observed in both men and women. Thus, while MetS component levels mostly fell within clinically healthy ranges on average among individuals belonging to this cluster, it is likely that certain participants within this cluster may have had clinical elevations across some of these variables. Nonetheless, the aggregate MetS profile of such individuals was not characteristic of an unhealthy cardiometabolic group. In fact, the observed elevation in mean waist circumference among H/L women classified into the Non-MetS cluster may likely reflect the increased prevalence of overweight and obesity previously documented in national statistics for this group, which might not necessarily cluster with similar elevations across other MetS components (Roger et al., 2011).

The second identified cluster, labeled "MetS," comprised approximately 20% of the entire sample (19% of men and 28% of women). Individuals classified as belonging to this cluster were more likely to be taking medication and exhibited mean levels of

MetS components that were noticeably more elevated than those observed for the Non-MetS cluster (i.e., for the entire sample, a waist circumference of 105.7 vs. 94.9 cm, a systolic blood pressure of 134.1 vs. 115.7 mm Hg, a diastolic blood pressure of 77.4 vs. 70.4 mm Hg, a triglyceride level of 152.8 vs. 104.2 mg/dL, and a fasting glucose level of 126.2 vs. 92.9 mg/dL). However, such discrepancy was less apparent for HDL-cholesterol among women in the MetS vs. the Non-MetS cluster (50.0 vs. 52.3 mg/dL, respectively), and among men, respective mean HDL-cholesterol levels were nearly identical (44.9 vs. 44.7 mg/dL). This suggests that among H/L populations, HDL cholesterol may poorly differentiate between individuals at high and low cardiometabolic risk compared to other MetS features, including triglyceride levels. This observation is consistent with unpublished analyses conducted in our lab on the same sample using a latent variable approach (a variable- as opposed to person-centered technique) to investigate MetS, which suggested that HDL cholesterol conferred the least amount of information toward defining a latent MetS construct compared to other components. Given national data suggesting that US H/Ls exhibit lower HDL cholesterol levels compared to other racial/ethnic groups, it would be interesting to further investigate how this component of dyslipidemia relates to other CVD risk factors and outcomes among this ethnic group (Roger et al., 2011).

It should be noted that in both men and women, the MetS cluster exhibited more heterogeneity across most MetS components (particularly fasting glucose and systolic blood pressure) compared to the Non-MetS cluster, suggesting that substantial variability exists in MetS presentation among affected individuals. This supports the use of current diagnostic criteria that require elevations in only a subset of MetS components, as

opposed to all. Interestingly, relative to other MetS components, individuals classified into the Non-MetS cluster were most homogeneous with respect to non-elevated fasting glucose values, lending additional support to the postulated role of insulin resistance in the development and pathophysiology of MetS.

Convergence of the current study's results with previous empirical findings was further witnessed across other domains. First, many consistencies were observed between the mean MetS component estimates for the MetS cluster identified in this study and the diagnostic threshold values currently proposed by existing MetS definitions. For example, comparing observed mean estimates to NCEP-ATP III cutoff scores, similarities were seen among men in terms of waist circumference (104.2 compared to 102 cm), systolic blood pressure (135.6 compared to 130 mm Hg), HDL cholesterol (45.4 compared to 40 mg/dL), and triglycerides (152.1 compared to 150 mg/dL), as well as among women in terms of systolic blood pressure (131.6 vs. 130 mm Hg), HDL-cholesterol (51.0 vs. 50 mg/dL), and fasting glucose (112.6 vs. 110 mg/dL). In fact, it has been suggested that LCAs and similar methods can provide useful information regarding cutoff values for diagnostic categories when, as in the case of the current study, the indicators appear to measure primarily one underlying construct (Marsh et al., 2009). However, compared to their wide use across psychological and other social sciences research, these statistical techniques have been seldom applied to medical investigations (Llabre & Fitzpatrick, 2012). Their potential usefulness is highlighted in the present study, which not only showed several similarities between obtained parameter estimates and previously established guidelines, but also revealed some important differences. For instance, mean diastolic blood pressure values for both men and women within the MetS

cluster, as well as mean triglyceride levels among women, were lower than currently proposed NCEP-ATP III thresholds. Additionally, the mean fasting glucose level in men was higher than currently proposed criteria, as was the mean waist circumference in women.

Thus, a valuable contribution of the current study is the provision of incipient evidence suggesting that currently identified threshold values for some components may not optimize the diagnosis of MetS among US H/L populations. For example, the mean waist circumference value of 102.6 cm observed among H/L women within the MetS cluster is markedly higher than the cutoff value of 88 cm currently proposed for evaluating US females by the NCEP-ATP III definition. Although there is recent consensus among several leading scientific organizations advocating the use of ethnic-specific waist circumferences thresholds for MetS diagnosis, specific cutoffs for H/Ls have not been formally proposed, likely due to a lack of informative data on representative US H/L samples (Alberti et al., 2009). The complicated nature of defining optimum threshold values for abdominal obesity among different ethnic groups has been documented, and in order to generate an appropriate evidence base to aid this process, there has been a call for both cross-sectional and prospective data relating waist circumference to CVD and diabetes across diverse populations (Alberti et al., 2009). Our cross-sectional findings tentatively support utilizing a higher waist circumference threshold for diagnosing MetS in US H/L women. Whether such recommendations extend to women of all US H/L subgroups remains unclear, as differences between these subgroups were not directly assessed (i.e., separate models were not conducted among different H/L subgroups).

Given a lack of published data on similarly representative US H/L samples, it is uncertain how other observed discrepancies between currently endorsed guidelines and the estimates derived from our study should inform the need for different MetS cutoff scores specific to US H/L populations. Nonetheless, results from the present study can guide future research aimed at optimizing the identification of MetS among this particular ethnic group. For instance, the use of different diagnostic thresholds informed by our results can be tested against other MetS criteria to examine associated changes in sensitivity and specificity. Such investigations can be carried out within a latent class framework, which does not require a "gold" standard referent (which for MetS has not been agreed upon) for comparisons (readers are referred to Llabre & Fitzpatrick, 2012, for an informative discussion regarding the use of latent class analysis to examine sensitivity and specificity for various cutoff scores, as well as its associated advantages).

Second, consistent with recent age-adjusted national prevalence rates of MetS among US H/Ls, this study demonstrated that a higher proportion of women (33.2%) were classified into the MetS cluster compared to men (22.7%). It is difficult to directly compare these observed estimates to those previously reported using current diagnostic definitions because, among other factors, 1) national prevalence estimates for US H/Ls other than Mexican Americans are lacking, and 2) there are differences in the age range and composition of the US H/Ls studied in HCHS/SOL compared to the US H/L samples used to generate previously published estimates. Nonetheless, the estimates observed in our study represent a plausible approximation of the burden of MetS among US H/Ls of diverse backgrounds. However, it is important to note that this approximation was noticeably lower than estimates drawn from the same study sample using NCEP-ATP III

criteria, which classified 41.7% of women and 33.8% of men as having MetS. Moreover, while the proportion of men and women identified as belonging to the MetS cluster in the current study who also met NCEP-ATP III criteria was high (62.1% and 70.6%, respectively) compared to those belonging to the Non-MetS cluster (21.8% and 17.3%, respectively), discordance between these two operationalizations of MetS are noted. Such observed differences in classification merit further investigation, and may in part reflect the previously discussed adequacy or inadequacy of using currently identified thresholds for MetS diagnosis that are non-specific to this ethnic group.

Third, the latent clusters identified in the present study generally related to both covariates and CVD outcomes in a manner consistent with previous research. For instance, in a multivariate model, both men and women of older age and who had a positive family history of CHD were at greater odds of being classified into the MetS cluster, as were women who had lower education and income levels. The observation that Puerto Rican women, relative to Mexican women, exhibited greater odds of belonging to the MetS cluster is consistent with at least one previous investigation of US H/L women reporting that Puerto Ricans exhibited the highest prevalence of MetS compared to other H/L subgroups, except that Mexicans were not represented in that particular study sample (Derby et al., 2010). In MESA, which had representation from both Mexicans and Puerto Ricans, Mexicans were shown to have the highest prevalence of MetS, 49.1%, followed by Puerto Ricans, 37.9% (Allison et al., 2008). However, these estimates were not gender-specific, and it is likely that H/L subgroup differences in MetS prevalence may also vary as a function of sex. Such observed differences may also be driven by generational status or length of residence in the US, which has been positively associated

with obesity and other cardiometabolic risk factors (Abraido-Lanza, Chao, & Florez, 2005). The Puerto Rican sample in HCHS/SOL, recruited mainly from New York communities, may have consisted of less immigrants and/or more acculturated individuals than the study's Mexican sample. This warrants further investigation, as does the unanticipated observation among women that current smokers, relative to never smokers, exhibited lower odds of being classified into the MetS cluster.

Additionally, while both men and women within the MetS cluster showed significantly greater odds of having prevalent CHD that included self-reported angina, statistically significant observations between latent cluster membership and the other prevalent CVD outcomes (CHD not including self-reported angina and cerebrovascular disease) were observed only among men and not women. Such findings, however, may reflect well-documented gender differences in clinical CVD onset as a function of age (Mosca, Barrett-Connor, & Wenger, 2011).

Although not considered an a priori aim of the current study, the examination of how extracted latent clusters relate to multiple predictors and outcomes is also the best-proposed method to rigorously verify the adequacy of selected latent class models (Bauer & Curran, 2003; Muthen, 2003; Morin et al., 2011). The convergence of such relationships with previous empirical work on MetS, as just described, in addition to the observed stability and unaltered qualitative nature of the MetS clusters across models that both excluded and included multiple covariates and outcomes, further supports the acceptability of the clusters extracted in our analyses. These observations are important and reassuring given the exploratory nature of the current investigation, which employed a new statistical approach to the study of MetS. It should be noted that other covariates

(i.e., physical activity, alcohol use, etc.) and specific outcomes (i.e., stroke versus transient ischemic attack) of initial interest were not examined in the current study due to the unavailability of such data at the time of analysis. Thus, while hypothesized to be unlikely, it is unknown whether the inclusion of such variables may have considerably weakened conclusions regarding the stability and validity of the extracted latent clusters.

To our knowledge, this is the first study to incorporate a latent class analytical approach to the study of MetS using continuous measures of its associated components. Boyko, Doheny, McNeely, Kahn, Leonetti, and colleagues (2010) recently published findings from a latent class analysis of MetS using dichotomous measures of MetS components collected on a sample of non-diabetic Japanese Americans. They concluded that a three-class model represented the best fit to their data, and also reported support for this model when diabetics were included in analyses. Similar to the current study, they identified a class exhibiting low probabilities for all MetS components as well as a class exhibiting high probabilities for all MetS components. They also identified a third and much smaller class of individuals exhibiting a high probability for hypertension and hyperglycemia but a low probability for increased abdominal adiposity and insulin resistance, suggesting the possibility of an additional MetS subtype potentially associated with a distinct pathophysiologic pathway. These conclusions were somewhat different than those arrived at in the current study. However, aside from the ethnicities of the samples and the measures of MetS components analyzed, there were many other differences between these two latent class studies that should be noted. For example, in contrast to the current study, Boyko and colleagues 1) did not examine different model specifications and adhered to the traditional latent class analysis assumptions of local

independence and homogeneity of variance, 2) could not examine more than three classes due to insufficient degrees of freedom, 3) did not examine medication use separately from their corresponding MetS indicators, 4) did not have representation of individuals younger than 34 years of age in their sample, 5) relied solely on statistical fit indices to determine the adequacy of tested models, and 6) did not assess relationships between extracted latent classes and relevant covariates or outcomes to assess model stability and construct validity, to name but a few.

One of the most debated issues regarding MetS diagnosis has centered on the appropriateness of including or excluding individuals with overt disease, such as diabetes, and the associated investigative and clinical implications of either approach (Kahn et al., 2005). In line with current MetS diagnostic definitions, the current study did not exclude from analyses individuals that had diabetes (n = 2255, taking into account information on serum glucose levels adjusted for fasting time, glucose-lowering medication use, and, if available, post-OGTT glucose levels and A1C percentages). To assess whether the exclusion of such individuals may have led to different results, post-hoc analyses were conducted on the entire sample of non-diabetic participants. Models ranging from 1 to 5 latent clusters were examined. Results of these analyses showed that the exclusion of diabetic patients greatly altered the ability to clearly identify any distinct and meaningful subgroups of individuals on the basis of MetS components. Given that insulin resistance and obesity have been posited as major contributing factors for the development of both MetS and diabetes, and that individuals with diabetes may represent an important subsample of individuals with overt MetS at an advanced pathophysiological state, such post-hoc observations should not be surprising. In fact,

nearly all diabetics were classified as belonging to the MetS cluster (accounting for approximately 50.4% of individuals in this cluster compared to only 3.0% of individuals in the Non-MetS cluster) in initial analyses conducted on the entire sample. However, given the high variability in fasting glucose scores and glucose-lowering medication use observed among individuals within the identified MetS cluster, it is likely that diabetes status alone did not entirely drive the formation of this group, but rather that diabetic participants also exhibited elevations across other MetS features that importantly influenced our ability to capture a valid picture of the syndrome as a whole. However, whether observed associations between latent cluster membership and CVD prevalence would have changed upon the exclusion of diabetic patients from the originally identified clusters was not assessed, as this would have required outputting data to conduct additional tests – a process that has been criticized for introducing substantial bias and error into analyses (Morin et al., 2011).

In conclusion, results obtained from this study appear to converge with previous research and current conceptualizations of MetS as a distinct cardiometabolic state. For example, the significant variability observed across MetS component levels among individuals within the MetS cluster support current criteria requiring elevations in some, but not all, MetS indicators. Additionally, the homogeneity in non-elevated glucose levels observed among individuals within the Non-MetS cluster, as well as the inability to adequately differentiate among MetS groups in post-hoc analyses that excluded diabetic participants, offer some support for the etiologic role of insulin resistance in MetS development. Importantly, this study also provides specific and valuable information regarding the presentation of MetS among US H/Ls. For instance, findings suggest that

among this ethnic group, HDL cholesterol may confer little utility in helping differentiate among individuals with and without MetS relative to other MetS components. Furthermore, observed results provide an incipient evidence base that could help inform the adequacy of developing H/L-specific threshold values to optimize MetS diagnosis among this group. For example, among US H/L women, utilizing a higher waist circumference cutoff than is currently endorsed by existing criteria may help better identify those with MetS and who may thus be at greater risk for developing CVD.

In comparison to previous research on MetS, several notable strengths as well as limitations of the current study exist. First, this investigation utilized community-based data from the most comprehensive study to date on US H/Ls, allowing for improved inference regarding the cardiometabolic health of this generally understudied group. However, such inferences cannot be extended to the US H/L community at-large, as the study sample was not nationally representative (Sorlie et al., 2010). Second, a novel approach to studying MetS was employed which allowed for corresponding components be analyzed as continuous rather than dichotomous variables, addressing a major criticism of previous and ongoing MetS investigations (Kahn et al., 2005). This permitted detailed evaluation of MetS component levels within identified clusters, as opposed to just examining proportions of individuals meeting pre-specified elevations defined by potentially arbitrary cutoff values. However, our results were dependent on the MetS components under study, which while consistent with those currently cited in the majority of existing MetS definitions, were not exhaustive of all previously proposed indicators (i.e., measures of insulin resistance/sensitivity, inflammation, coagulation, etc.). Such indicators may be incorporated in future analyses to further evaluate the

adequacy and construct validity of identified clusters, as well as to garner evidence regarding their added utility in classifying individuals with MetS. Additionally, while many systematic precautions were taken to circumvent commonly cited problems associated with LCA and analogous methods (i.e., testing variously specified models, increasing starting values, incorporating covariates and outcomes, examining relationships in a single analytic step reducing the potential for bias and error, etc.), these results remain preliminary and warrant replication. Third, the current study relied on cross-sectional data, and thus conclusions cannot be made regarding the directionality of observed relationships. The prospective design of HCHS/SOL will present investigators with future opportunities to address this and similar issues in subsequent analyses. Lastly, given that HCHS/SOL did not include a non-H/L cohort, the current study could not directly compare obtained results to non-H/L populations (Sorlie et al., 2010). Future research may be carried out to examine this indirectly using data collected from other epidemiologic cohorts (i.e., NHANES, MESA, etc.) (Sorlie et al., 2010).

As debate ensues regarding the validity and utility of MetS as a diagnostic construct, continued research is needed to fill current knowledge gaps (Kahn et al., 2005). Employing novel statistical approaches to the study of MetS can offer valuable insights into the pathophysiology and clinical significance of this seemingly elusive entity. Such investigation may be greatly enhanced by exploring further the similarities and differences in MetS presentation among various racial/ethnic groups. Research on US H/Ls is particularly warranted given the size, aging, and projected growth of this minority population, factors which are likely to greatly impact our nation's future burden of CVD and associated healthcare expenditures (Davidson et al., 2007).

REFERENCES

Abraido-Lanza, A. F., Chao, M. T., & Florez, K. R. (2005). Do healthy behaviors decline with greater acculturation? Implications for the Latino mortality paradox. *Social Science & Medicine, 61,* 1243-1255.

Alberti, K. G. M. M., Eckel, R. H., Grundy, S. M., Zimmet, P. Z., Cleeman, J. I., Donato, K. A., et al. (2009). Harmonizing the metabolic syndrome: a joint statement of the International Diabetes Federation Task Force on Epidemiology and Prevention; National Heart, Lung, and Blood Institute; American Heart Association; World Heart Federation; International Atherosclerosis Society; and International Association for the Study of Obesity. *Circulation, 120,* 1640-1645.

Alberti, K. G. M. M., & Zimmet, P. Z. (1998). Definition, diagnosis and classification of diabetes mellitus and its complications, part 1: diagnosis and classification of diabetes mellitus: provisional report of a WHO consultation. *Diabetic Medicine, 15,* 539-553.

Allison, M. A., Budoff, M. J., Wong, N. D., Blumenthal, R. S., Schreiner, P. J., & Criqui, M. H. (2008). Prevalence of and risk factors for subclinical cardiovascular disease in selected US Hispanic ethnic groups. *American Journal of Epidemiology, 167,* 962-969.

Arbuckle, J. L. (1996). Full information estimation in the presence of incomplete data. In G. A. Marcoulides & R. E. Schumacker (Eds), *Advanced structural equation modeling* (pp. 243-277). Mahwah, NJ: Erlbaum.

Bauer, D. J., & Curran, P. J. (2003). Distributional assumptions of growth mixture models over-extraction of latent trajectory classes. *Psychological Methods, 8,* 338-363.

Bauer, D. J., & Curran, P. J. (2004). The integration of continuous and discrete latent variable models: Potential problems and promising opportunities. *Psychological Methods, 9,* 3-29.

Boden-Albala, B., Sacco, R. L., Lee, H., Grahame-Clarke, C., Rundek, T., Elkind, M. V., et al. (2008). Metabolic syndrome and ischemic stroke risk: Northern Manhattan Study. *Stroke, 39,* 30-35.

Boudreau, D. M., Malone, D. C., Raebel, M. A., Fishman, P. A., Nichols, G. A., Feldstein, A. C., et al. (2009). Health care utilization and costs by metabolic syndrome risk factors. *Metabolic Syndrome and Related Disorders, 7,* 305-314.

Boyko, E. J., Doheny, R. A., McNeely, M. J., Kahn, S. E., Leonetti, D. L., & Fujimoto, W. Y. (2010). Latent class analysis of the metabolic syndrome. *Diabetes Research and Clinical Practice, 89,* 88-93.

Centers for Disease Control and Prevention. (2011). Surveillance of health status in minority communities – racial and ethnic approaches to community health across the U.S. (REACH U.S.) Risk Factor Survey, United States, 2009. *Morbidity and Mortality Weekly Report. Surveillance Summaries, 60,* 1-12.

Clark, S. L., & Muthen, B. (2009). Relating latent class analysis results to variables not included in the analysis. Submitted for publication. Article available at http://www.statmodel.com/papers.shtml

Crespo, C. J., Loria, C. M., & Burt, V. L. (1996). Hypertension and other cardiovascular disease risk factors among Mexican Americans, Cuban Americans, and Puerto Ricans from the Hispanic Health and Nutrition Examination Survey. *Public Health Reports, 111,* 7-10.

Cutler, J. A., Sorlie, P. D., Wolz, M., Thom, T., Fields, L. E., & Roccella, E. J. (2008). Trends in hypertension prevalence, awareness, treatment, and control rates in United States adults between 1988-1994 and 1999-2004. *Hypertension, 52,* 818-827.

Davidson, J. A., Moreno, P. R., Badimon, J. J., Lopez-Candales, A., Giachello, A. L. M., Ovalle, F., et al. (2007). Cardiovascular disease prevention and care in Latino and Hispanic subjects. *Endocrine Practice, 13,* 77-85.

de Zeeuw, D., & Bakker, S. J. L. (2008). Does the metabolic syndrome add to the diagnosis and treatment of cardiovascular disease? *Nature Clinical Practice Cardiovascular Medicine, 5,* S10-S14.

Derby, C. A., Wildman, R. P., McGinn, A. P., Green, R. R., Polotsky, A. J., Ram, K. T., et al. (2010). Cardiovascular risk factor variation within a Hispanic cohort: SWAN, the Study of Women's Health Across the Nation. *Ethnicity & Disease, 20,* 396-402.

Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults. (2001). Executive Summary of the Third Report of the National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III). *Journal of the American Medical Association, 285,* 2486-2497.

Ford, E. S., Ajani, U. A., Croft, J. B., Critchley, J. A., Labarthe, D. R., Kottke, T. E., et al. (2007). Explaining the decrease in U.S. deaths from coronary disease, 1980-2000. *New England Journal of Medicine, 356,* 2388-2398.

Franco, O. H., Massaro, J. M., Civil, J., Cobain, M. R., O'Malley, B., & D'Agostino, R. B. (2009). Trajectories of entering the metabolic syndrome: the Framingham Heart Study. *Circulation, 120,* 1943-1950.

Franzini, L., Ribble, J. C., & Keddie, A. M. (2001). Understanding the Hispanic paradox. *Ethnicity & Disease, 11,* 496-518.

Gallo, L. C., Penedo, F. J., Espinosa de los Monteros, K., & Arguelles, W. (2009). Resiliency in the face of disadvantage: do Hispanic cultural characteristics protect health outcomes? *Journal of Personality, 77,* 1707-1746.

Gami, A. S., Witt, B. J., Howard, D. E., Erwin, P. J., Gami, L. A., Somers, V. K., et al. (2007). Metabolic syndrome and risk of incident cardiovascular events and death. A systematic review and meta-analysis of longitudinal studies. *Journal of the American College of Cardiology, 49,* 403-414.

Goff, D. C., Bertoni, A. G., Kramer, H., Bonds, D., Blumenthal, R. S., Tsai, M. Y., et al. (2006). Dyslipidemia prevalence, treatment, and control in the Multi-Ethnic Study of Atherosclerosis (MESA): gender, ethnicity, and coronary artery calcium. *Circulation, 113,* 647-656.

Gutt, M., Davis, C. L., Spitzer, S. B., Llabre, M. M., Kumar, M., Czarnecki, E. M., et al. (2000). Validation of the insulin sensitivity index ($ISI_{0,\ 120}$): comparison with other measures. *Diabetes Research and Clinical Practice, 47,* 177-184.

Hari, P., Nerusu, K., Veeranna, V., Sudhakar, R., Zalawadiya, S., Ramesh, K., et al. (2011). A gender-stratified comparative analysis of various definitions of metabolic syndrome and cardiovascular risk in a multi-ethnic U.S. population. *Metabolic Syndrome and Related Disorders, epub ahead of print,* 1-10.

Hoang, K. C., Ghandehari, H., Lopez, V. A., Barboza, M. G., & Wong, N. D. (2008). Global coronary heart disease risk assessment of individuals with the metabolic syndrome in the U.S. *Diabetes Care, 31,* 1405-1409.

Johnson, L. W., and Weinstock, R. S. (2006). The metabolic syndrome: concepts and controversy. *Mayo Clinic Proceedings, 81,* 1615-1620.

Kahn, R., Buse, J., Ferrannini, E., & Stern, M. (2005). The metabolic syndrome: time for a critical appraisal. *Diabetes Care, 28,* 2289-2304.

King, D. E., Mainous, A. G., Carnemolla, M., & Everett, C. J. (2009). Adherence to healthy lifestyle habits in US adults, 1988-2006. *The American Journal of Medicine, 122,* 528-534.

Kramer, H., Han, C., Post, W., Goff, D., Diez-Roux, A., Cooper, R., et al. (2004). Racial/ethnic differences in hypertension and hypertension treatment and control in the Multi-Ethnic Study of Atherosclerosis (MESA). *American Journal of Hypertension, 17,* 963-970.

LaVange, L. M., Kalsbeek, W. D., Sorlie, P. D., Aviles-Santa, L. M., Kaplan, R. C., Barnhart, J., et al. (2010). Sample design and cohort selection in the Hispanic Community Health Study/Study of Latinos. *Annals of Epidemiology, 20,* 642-649.

Libby, P. (2003). Vascular biology of atherosclerosis: overview and state of the art. *The American Journal of Cardiology, 91,* 3A-6A.

Llabre, M. M., & Fitzpatrick, S. L. (2012). Revisiting measurement models in psychosomatic medicine research: A latent variable approach. *Psychosomatic Medicine, 74,* 169-177.

Lubke, G. H., & Muthen, B. (2007). Performance of factor mixture models as a function of model size, criterion measure effects, and class-specific parameters. *Structural Equation Modeling, 14,* 26-47.

Marsh, H. W., Hau, K. T., & Grayson, D. (2005). Goodness of fit evaluation in structural equation modeling. In A. Maydeu-Olivares & J. McArdle (Eds.), *Contemporary Psychometrics. A festschrift for Roderick P. McDonald* (pp. 275-340). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Marsh, H. W., Hau, K. T., & Wen, Z. (2004). In search of golden rules: Comment on hypothesis testing approaches to setting cutoff values for fit indexes and dangers of overgeneralising Hu & Bentler's (1999) findings. *Structural Equation Modeling, 11,* 320-341.

Marsh, H. W., Ludtke, O., Trautwein, U., & Morin, A. J. S. (2009). Classical latent profile analysis of academic self-concept dimensions: Synergy of person- and variable-centered approaches to theoretical models of self-concept. *Structural Equation Modeling, 16,* 191-225.

McCutcheon, A. L. (2002). Basic concepts and procedures in single- and multiple-group latent class analysis. In J. A. Hagenaars & A. L. McCutcheon (Eds.), *Applied latent class analysis* (pp. 56-85). Cambridge, UK: Cambridge University Press.

McWilliams, J. M., Meara, E., Zaslavsky, A. M., & Ayanian, J. Z. (2009). Differences in control of cardiovascular disease and diabetes by race, ethnicity, and education: U.S. trends from 1999 to 2006 and effects of Medicare coverage. *Annals of Internal Medicine, 150,* 505-515.

Meigs, J. B., Wilson, P. W. F., Nathan, D. M., D'Agostino, R. B., Williams, K., & Haffner, S. M. (2003). Prevalence and characteristics of the metabolic syndrome in the San Antonio Heart and Framingham Offspring Studies. *Diabetes, 52,* 2160-2167).

Morin, A. J. S., Morizot, J., Boudrias, J., & Madore, I. (2011). A multifoci person-centered perspective on workplace affective commitment: A latent profile/factor mixture analysis. *Organizational Research Methods, 14,* 58-90.

Mosca, L., Barrett-Connor, E., & Wenger, N. K. (2011). Sex/gender differences in cardiovascular disease prevention: What a difference a decade makes. *Circulation, 124,* 2145-2154.

Mottillo, S., Filion, K. B., Genest, J., Joseph, L., Pilote, L., Poirier, P., et al. (2010). The metabolic syndrome and cardiovascular risk. A systematic review and meta-analysis. *Journal of the American College of Cardiology, 56,* 1113-1132.

Mozumdar, A., & Liguori, G. (2011). Persistent increase of prevalence of metabolic syndrome among U.S. adults: NHANES III to NHANES 1999-2006. *Diabetes Care, 34,* 216-219.

Muthen, B. O. (2003). Statistical and substantive checking in growth mixture modeling: Comment on Bauer and Curran (2003). *Psychological Methods, 8,* 369-377.

Muthen, L. K., & Muthen, B. O. (1998-2010). Mplus User's Guide. Sixth Edition. Los Angeles, CA: Muthen & Muthen.

National Center for Health Statistics, Centers for Disease Control and Prevention. *Compressed Mortality File: Underlying Cause of Death, 1979 to 2005*. Retrieved June 6, 2009 from http://wonder.cdc.gov/mortSQL.html.

National Heart, Lung, and Blood Institute. *Clinical Guidelines on the Identification, Evaluation, and Treatment of Overweight and Obesity in Adults: The Evidence Report*. Retrieved June 6, 2009, from http://www.nhlbi.nih.gov/guidelines/obesity/e_txtbk/appndx/apndx8.htm.

Pratley, R. E. (2007). Metabolic syndrome: why the controversy? *Current Diabetes Reports, 7,* 56-59.

Razzouk, L., & Muntner, P. (2009). Ethnic, gender, and age-related differences in patients with the metabolic syndrome. *Current Hypertension Reports, 11,* 127-132.

Regitz-Zagrosek, V., Lehmkuhl, E., & Mahmoodzadeh, S. (2007). Gender aspects of the role of the metabolic syndrome as a risk factor for cardiovascular disease. *Gender Medicine, 4,* S162-S177.

Roger, V. L., Go, A. S., Lloyd-Jones, D. M., Adams, R. J., Berry, J. D., Brown, T. M., et al. (2011). Heart disease and stroke statistics – 2011 update: a report from the American Heart Association. *Circulation, 123,* e18-e209.

Sorlie, P. D., Aviles-Santa, L., Wassertheil-Smoller, S., Kaplan, R. C., Daviglus, M. L., Giachello, A. L., et al. (2010). Design and Implementation of the Hispanic Community Health Study / Study of Latinos. *Annals of Epidemiology, 20,* 629-641.

Sorlie, P. D., Backlund, E., Johnson, N. J., & Rogot, E. (1993). Mortality by Hispanic status in the United States. *Journal of the American Medical Association, 270,* 2464-2468.

Swenson, C. J., Trepka, M. J., Rewers, M. J., Scarbro, S., Hiatt, W. R., & Hamman, R. F. (2002). Cardiovascular disease mortality in Hispanics and non-Hispanic whites. *American Journal of Epidemiology, 156,* 919-928.

Tenenbaum, A., & Fisman, E. Z. (2011). "The metabolic syndrome… is dead": these reports are an exaggeration. *Cardiovascular Diabetology, 10,* 11.

U.S. Census Bureau. Hispanic Population of the United States: 1970-2050. Available at: http://www.census.gov/population/www/socdemo/hispanic/hispanic.html. Accessed November 18, 2011.

U.S. Census Bureau. The Hispanic Population Census. 2000 Brief. Available at: http://www.census.gov/prod/2001pubs/c2kbr01-3.pdf. Accessed November 18, 2011.

Vermunt, J. K., & Magidson, J. (2002). Latent class cluster analysis. In J. A. Hagenaars & A. L. McCutcheon (Eds.), *Applied latent class analysis* (pp. 89-106). Cambridge, UK: Cambridge University Press.

Wilson, P. W., D'Agostino, R. B., Parise, H., Sullivan, L., & Meigs, J. B. (2005). Metabolic syndrome as a precursor of cardiovascular disease and type 2 diabetes mellitus. *Circulation, 112,* 3066-3072.

Yala, S. M., Fleck, E. M., Sciacca, R., Castro, D., Joseph, Z., & Giardina, E. V. (2009). Metabolic syndrome and the burden of cardiovascular disease in Caribbean Hispanic women living in Northern Manhattan: a red flag for education. *Metabolic Syndrome and Related Disorders, 7,* 315-322.

Zhu, S., Heymsfield, S. B., Toyoshima, H., Wang, Z., Pietrobelli, A., & Heshka, S. (2005). Race-ethnicity-specific waist circumference cutoffs for identifying cardiovascular disease risk factors. *American Journal of Clinical Nutrition, 81,* 409-415.

APPENDIX: FIGURES

Figure 1. Schematic representation of the latent class analysis model of metabolic syndrome.

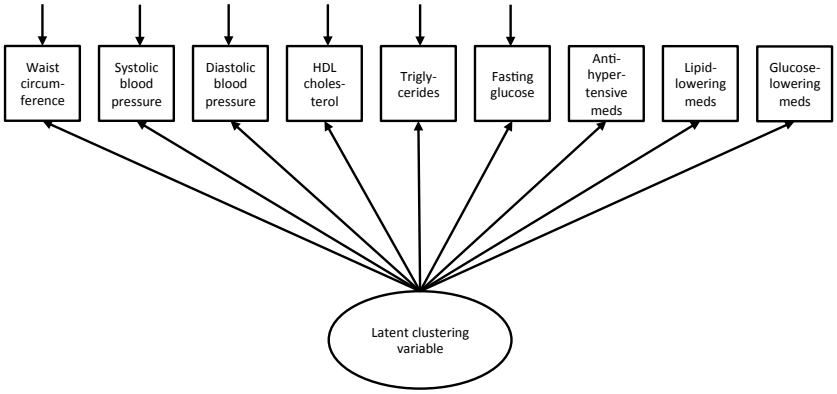


Figure 2. Schematic representation of the latent class analysis model of metabolic syndrome including covariates.

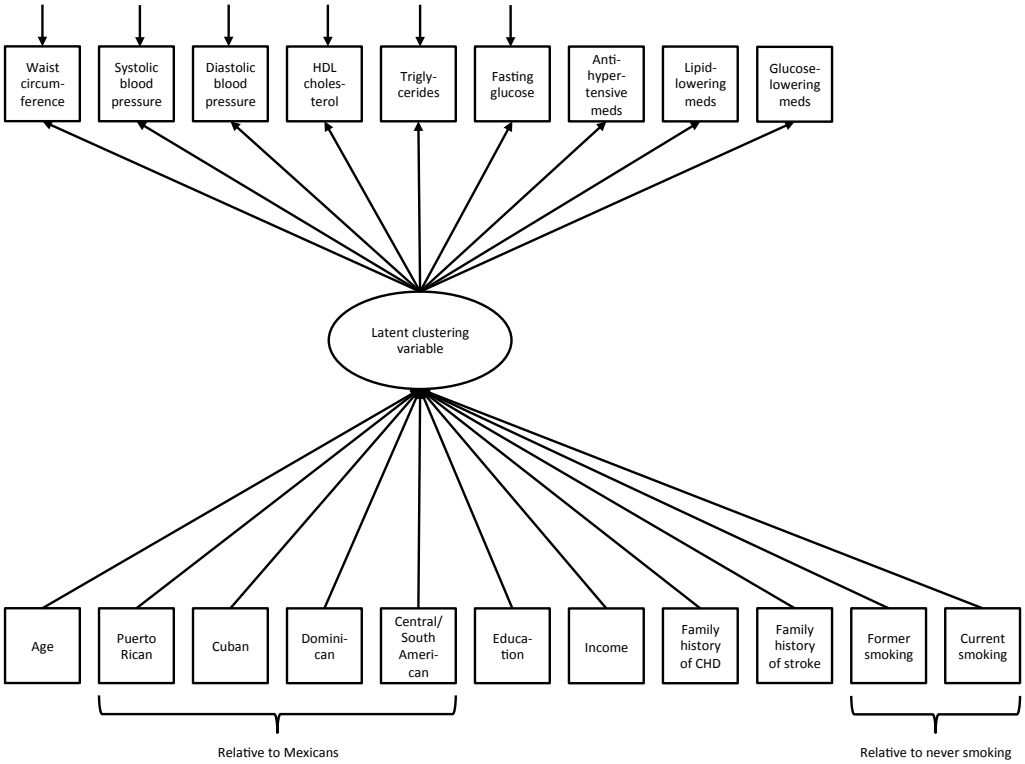
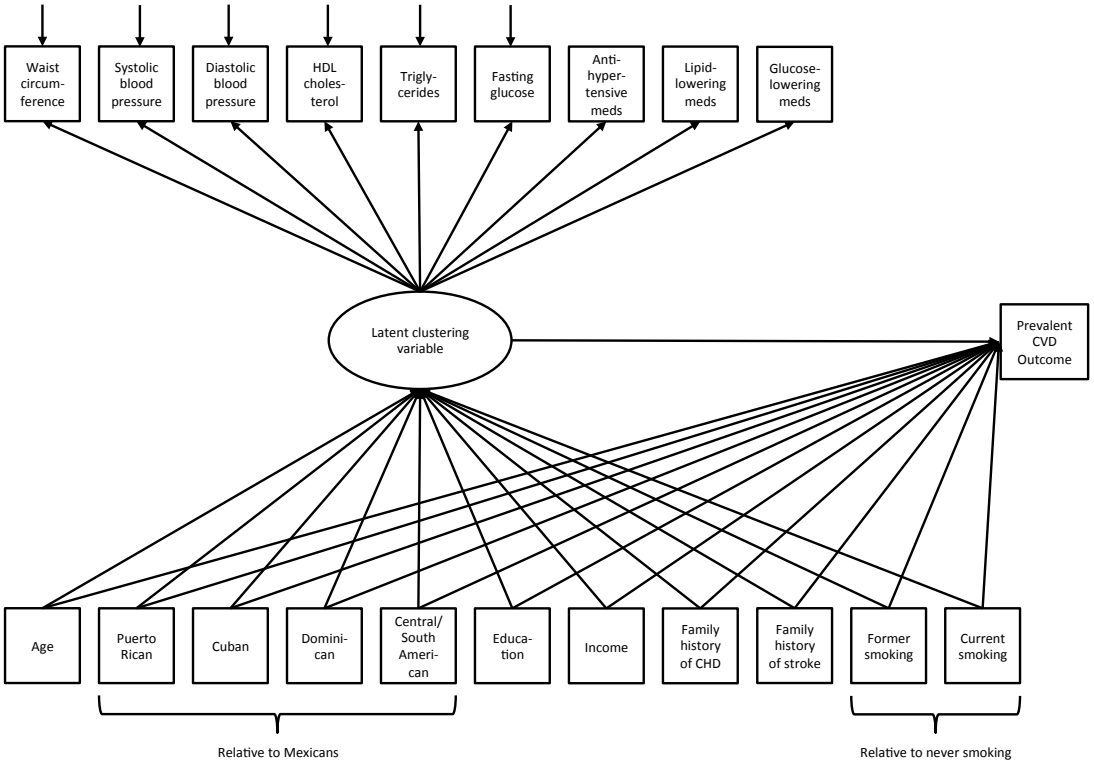


Figure 3. Schematic representation of the latent class analysis model of metabolic syndrome including covariates and a prevalent cardiovascular disease outcome.



APPENDIX: TABLES

Table 1. Baseline demographic, risk factor, and prevalent disease information.

Variable	Entire sample (n = 10970) Mean (SD) or proportion	Men (n = 4429) Mean (SD) or proportion	Women (n = 6541) Mean (SD) or proportion
Gender			
Male	40.4%	…..	…..
Female	59.6%	…..	…..
Age (years)	45.9 (14.0)	44.9 (14.3)	46.6 (13.7)
Hispanic/Latino subgroup			
Mexican	41.3%	38.6%	43.1%
Puerto Rican	17.3%	18.2%	16.8%
Cuban	14.5%	17.1%	12.7%
Dominican	8.7%	7.6%	9.5%
Central/South American	18.2%	18.5%	18.0%
Education			
< HS/GED	38.1%	37.2%	38.7%
≥ HS/GED	61.9%	62.8%	61.3%
Income			
< $10K	15.7%	12.7%	17.9%
$10K - $20K	32.3%	29.8%	34.1%
> $20K - $40K	33.3%	35.4%	31.8%
> $40K - $75K	14.0%	16.1%	12.4%
> $75K	4.7%	6.0%	3.8%
Family history of CHD			
Yes	32.6%	28.4%	35.5%
No	67.4%	71.6%	64.5%
Family history of stroke			
Yes	15.1%	12.9%	16.6%
No	84.9%	87.1%	83.4%
Smoking status			
Never	60.6%	48.7%	68.6%
Former	19.9%	25.8%	15.9%
Current	19.5%	25.5%	15.5%
Waist circumference (cm)	98.0 (14.0)	98.5 (13.7)	97.6 (14.2)
Systolic blood pressure (mm Hg)	121.6 (18.0)	124.7 (16.0)	119.4 (19.0)
Diastolic blood pressure (mm Hg)	72.8 (11.0)	74.3 (11.0)	71.8 (10.9)
HDL cholesterol (mg/dL)	49.0 (13.2)	44.8 (12.0)	51.8 (13.2)
Triglycerides (mg/dL)	139.6 (104.2)	154.3 (128.5)	129.7 (82.4)
Glucose (mg/dL)	104.2 (36.2)	107.2 (38.5)	102.2 (34.4)
Antihypertensive medication use			
Yes	16.9%	15.4%	17.9%
No	83.1%	84.6%	82.1%
Lipid-lowering medication use			
Yes	12.7%	11.9%	13.3%
No	87.3%	88.1%	86.7%
Glucose-lowering medication use			
Yes	11.2%	9.9%	12.1%
No	88.8%	90.1%	87.9%
CHD			
Yes	5.6%	7.3%	4.5%
No	94.4%	92.7%	95.5%
CHD, including self-reported angina			
Yes	7.5%	8.7%	6.6%
No	92.5%	91.3%	93.4%
Cerebrovascular disease			
Yes	2.7%	2.8%	2.6%
No	97.3%	97.2%	97.4%

Table 2. Fit indices for the four differentially specified latent class models conducted on the entire sample.

Model	LL	# parameters	AIC	BIC	ABIC	Entropy	ALRT (*p*)
Model 1							
One class	-276494.6	15	553019.2	553128.7	553081.1	Na	Na
Two class	-271093.2	25	542236.4	542419.0	542339.5	0.985	< .001
Three class	-266930.8	35	533931.7	534187.3	534076.0	0.808	< .001
Four class	-265349.3	45	530788.7	531117.3	530974.3	0.799	< .001
Five class	-263895.6	55	527901.2	528302.8	528128.0	0.814	< .001
Six class	-263235.8	65	526601.6	527076.2	526869.7	0.821	< .001
Seven class	-262620.9	75	525391.9	525939.6	525701.2	0.796	< .001
Eight class	-262054.3	85	524278.6	524899.3	524629.2	0.785	< .001
Nine class	-261692.7	95	523575.5	524269.2	523967.3	0.779	< .001
…							
Twenty class	-259416.9	205	519243.7	520740.8	520089.3	0.769	< .001
Model 2							
One class	-271300.0	17	542633.9	542758.1	542704.0	Na	Na
Two class	-265909.9	29	531877.9	532089.7	531997.5	0.985	< .001
Three class	-262811.7	41	525705.3	526004.7	525874.5	0.833	< .001
Four class	-261573.4	53	523252.7	523639.8	523471.3	0.741	< .001
Five class	-260500.3	65	521130.7	521605.3	521398.8	0.764	< .001
Six class	-260045.2	77	520244.3	520806.6	520561.9	0.765	< .001
Seven class	-259605.6	89	519389.3	520039.2	519756.4	0.778	< .001
Eight class	-259212.6	101	518627.3	519364.9	519043.9	0.781	< .001
Nine class	-258836.2	113	517898.3	518723.5	518364.4	0.779	< .001
…							
Twenty class	-257402.3	245	515294.7	517083.8	516305.2	0.741	< .001
Model 3							
One class	-276494.6	15	553019.2	553128.7	553081.1	Na	Na
Two class	-266565.9	31	533193.8	533420.2	533321.7	0.790	< .001
Three class	-263078.3	47	526250.6	526593.8	526444.4	0.777	< .001
Four class	-261489.0	63	523104.1	523564.1	523363.9	0.765	< .001
Five class	-260591.6	79	521341.2	521918.1	521667.1	0.748	< .001
Six class	-259807.3	95	519804.6	520498.3	520196.4	0.741	< .001
Seven class	-259355.7	111	518933.4	519744.0	519391.2	0.748	< .001
Eight class	-258951.5	127	518157.0	519084.5	518680.9	0.747	< .001
Nine class	-258650.3	143	517586.7	518630.9	518176.5	0.734	< .001
…							
Twenty class	-256708.9	319	514055.7	516385.2	515371.5	0.736	0.620
Model 4							
One class	-271300.0	17	542633.9	542758.1	542704.0	Na	Na
Two class	-261904.3	35	523878.7	524134.2	524023.0	0.836	< .001
Three class	-259463.6	53	519033.3	519420.3	519251.9	0.738	< .001
Four class	-258275.1	71	516692.2	517210.7	516985.1	0.725	< .001
Five class	-257755.3	89	515688.5	516338.5	516055.6	0.68	< .001
Six class	-257332.0	107	514878.0	515659.4	515319.3	0.705	< .001
Seven class	-257055.6	125	514361.1	515274.0	514876.7	0.67	< .001
Eight class	-256819.7	143	513925.4	514969.7	514515.3	0.672	< .001
Nine class	-256622.8	161	513567.6	514743.3	514231.6	0.677	< .001
…							
Twenty class	-255566.8	359	511851.5	514473.2	513332.3	0.676	0.240

Table 3. Average latent cluster probabilities for most likely latent cluster membership (row) by latent cluster (column).

	Latent cluster 1	Latent cluster 2
Entire sample		
Latent cluster 1	0.964	0.036
Latent cluster 2	0.073	0.927
Men		
Latent cluster 1	0.967	0.033
Latent cluster 2	0.064	0.936
Women		
Latent cluster 1	0.956	0.044
Latent cluster 2	0.063	0.937

Table 4. Parameter estimates of the latent clusters on the metabolic syndrome component indicators.

	Model without covariates		Model with covariates	
	Latent cluster 1: Non-MetS *Entire sample: 79.7%* *Men: 81.3%* *Women: 72.5%*	Latent cluster 2: MetS *Entire sample: 20.3%* *Men: 18.7%* *Women: 27.5%*	Latent cluster 1: Non-MetS *Entire sample: 73.9%* *Men: 77.3%* *Women: 66.8%*	Latent cluster 2: MetS *Entire sample: 26.1%* *Men: 22.7%* *Women: 33.2%*
MetS component	Mean (SD) or proportion	Mean (SD) or proportion	Mean (SD) or proportion	Mean (SD) or proportion
Entire sample				
WC (cm)	94.9 (13.0)	105.7 (15.3)	94.8 (13.5)	103.9 (14.3)
SBP (mm Hg)	115.7 (13.4)	134.1 (21.4)	114.4 (12.4)	133.8 (20.0)
DBP (mm Hg)	70.4 (9.9)	77.4 (12.7)	69.9 (9.7)	77.3 (12.2)
HDL-C (mg/dL)	48.8 (13.1)	46.3 (12.7)	48.6 (12.9)	47.6 (13.3)
Triglycerides (transformed)	201.8 (23.7)	218.4 (24.0)	201.1 (23.8)	217.0 (23.5)
Triglycerides (mg/dL)	104.2 (----)	152.8 (----)	102.6 (----)	147.8 (----)
Glucose (transformed)	196.8 (3.7)	210.1 (15.6)	196.7 (3.6)	207.9 (14.7)
Glucose (mg/dL)	92.9 (----)	126.2 (----)	92.6 (----)	120.0 (----)
Antihypertensive med use	3.1%	47.3%	1.2%	43.1%
Lipid-lowering med use	2.6%	34.9%	1.2%	32.0%
Glucose-lowering med use	0.2%	36.2%	0.2%	28.9%
Men				
WC (cm)	96.0 (12.8)	105.8 (16.0)	96.0 (13.3)	104.2 (14.4)
SBP (mm Hg)	119.9 (11.4)	135.8 (21.1)	119.2 (10.9)	135.6 (20.0)
DBP (mm Hg)	72.0 (9.8)	79.1 (13.4)	71.7 (9.7)	79.1 (12.9)
HDL-C (mg/dL)	44.7 (10.9)	44.9 (14.1)	44.5 (10.8)	45.4 (14.0)
Triglycerides (transformed)	206.5 (24.9)	219.3 (26.3)	206.2 (25.1)	218.2 (25.5)
Triglycerides (mg/dL)	116.1 (----)	156.1 (----)	115.3 (----)	152.1 (----)
Glucose (transformed)	197.8 (3.6)	212.1 (16.6)	197.7 (3.6)	210.2 (15.9)
Glucose (mg/dL)	95.1 (----)	132.2 (----)	94.9 (----)	126.4 (----)
Antihypertensive med use	3.8%	45.7%	1.6%	46.0%
Lipid-lowering med use	3.0%	34.1%	1.7%	33.3%
Glucose-lowering med use	0.2%	36.6%	0.1%	30.9%
Women				
WC (cm)	93.7 (13.3)	103.8 (14.7)	93.5 (13.6)	102.6 (14.2)
SBP (mm Hg)	109.9 (12.2)	132.4 (20.6)	108.6 (11.3)	131.6 (19.5)
DBP (mm Hg)	68.2 (9.4)	76.1 (11.7)	67.8 (9.3)	75.7 (11.4)
HDL-C (mg/dL)	52.3 (13.6)	50.0 (12.8)	52.0 (13.5)	51.0 (13.1)
Triglycerides (transformed)	196.6 (21.7)	215.3 (21.5)	195.7 (21.7)	214.1 (21.3)
Triglycerides (mg/dL)	92.4 (----)	142.3 (----)	90.6 (----)	138.5 (----)
Glucose (transformed)	195.7 (3.4)	206.4 (13.7)	195.5 (3.3)	205.2 (13.1)
Glucose (mg/dL)	90.5 (----)	116.0 (----)	90.1 (----)	112.6 (----)
Antihypertensive med use	1.8%	40.6%	0.7%	36.8%
Lipid-lowering med use	1.5%	30.5%	0.6%	27.9%
Glucose-lowering med use	0.1%	29.0%	0.2%	24.4%

Table 5. Threshold values for metabolic syndrome components adapted from the Third Report of the National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III).

Risk factor	Defining level
Abdominal obesity	Waist circumference
Men	> 102 cm
Women	> 88 cm
Blood pressure	≥ 130/85 mmHg
HDL cholesterol	
Men	< 40 mg/dL
Women	< 50 mg/dL
Triglycerides	≥ 150 mg/dL
Fasting glucose	≥ 110 mg/dL

Table 6. Results from the logistic regression evaluating the effect of covariates on latent cluster membership.

Covariate	Coefficient (SE)	Odds ratio	*p* value
Entire sample			
Gender	0.356 (0.095)**	1.427**	< 0.001
Age	0.121 (0.005)**	1.128**	< 0.001
H/L subgroup (relative to Mexicans)			
Puerto Rican	0.446 (0.134)**	1.562**	0.001
Cuban	0.005 (0.142)	1.005	0.974
Dominican	0.202 (0.162)	1.224	0.212
Central/South American	-0.211 (0.135)	0.810	0.118
Education	-0.334 (0.100)**	0.716**	0.001
Income	-0.142 (0.042)**	0.867**	0.001
Family history of CHD	0.406 (0.094)**	1.500**	< 0.001
Family history of stroke	0.083 (0.112)	1.087	0.459
Smoking status (relative to never smoking)			
Former	0.057 (0.103)	1.059	0.580
Current	-0.208 (0.125)	0.812	0.096
Men			
Age	0.113 (0.007)**	1.119**	< 0.001
H/L subgroup (relative to Mexicans)			
Puerto Rican	0.195 (0.181)	1.215	0.281
Cuban	-0.063 (0.188)	0.939	0.739
Dominican	0.415 (0.224)	1.513	0.064
Central/South American	-0.111 (0.184)	0.895	0.548
Education	-0.071 (0.141)	0.931	0.612
Income	-0.070 (0.064)	0.932	0.269
Family history of CHD	0.390 (0.135)**	1.478**	0.004
Family history of stroke	0.177 (0.190)	1.193	0.352
Smoking status (relative to never smoking)			
Former	0.046 (0.149)	1.047	0.757
Current	-0.076 (0.176)	0.927	0.664
Women			
Age	0.150 (0.007)**	1.162**	< 0.001
H/L subgroup (relative to Mexicans)			
Puerto Rican	0.728 (0.189)**	2.070**	< 0.001
Cuban	-0.096 (0.180)	0.908	0.593
Dominican	-0.002 (0.238)	0.998	0.993
Central/South American	-0.204 (0.175)	0.815	0.242
Education	-0.504 (0.145)**	0.604**	0.001
Income	-0.250 (0.059)**	0.779**	< 0.001
Family history of CHD	0.315 (0.129)*	1.371*	0.014
Family history of stroke	-0.058 (0.139)	0.944	0.679
Smoking status (relative to never smoking)			
Former	0.139 (0.159)	1.150	0.382
Current	-0.403 (0.170)*	0.668*	0.018

* $p < .05$

** $p < .01$

Table 7. Results from the logistic regression evaluating the effect of the latent clustering variable on prevalent cardiovascular disease outcomes, adjusting for covariates.

Prevalent CVD outcome	Coefficient (SE)	Odds ratio	*p* value
Entire sample			
CHD	0.073 (0.024)**	1.076**	0.002
CHD, including self-reported angina	0.101 (0.026)**	1.106**	< 0.001
Cerebrovascular disease	0.045 (0.014)**	1.046**	0.002
Men			
CHD	0.088 (0.038)*	1.092*	0.020
CHD, including self-reported angina	0.109 (0.040)**	1.115**	0.006
Cerebrovascular disease	0.050 (0.025)*	1.051*	0.043
Women			
CHD	0.047 (0.028)	1.048	0.097
CHD, including self-reported angina	0.066 (0.031)*	1.068*	0.032
Cerebrovascular disease	0.028 (0.018)	1.028	0.129

* $p < .05$

** $p < .01$

APPENDIX: SAMPLE MPLUS INPUT STATEMENT

TITLE: Retained Two-Cluster Model with Covariates and Prevalent CHD Outcome Among Men

DATA: FILE IS "HCHS_Dataset.dat";

VARIABLE: NAMES ARE id psu strat weight wave center age gender bkgrd income educ fhchd fhstrok ciguse bmi wc sbp dbp hdl trig ldl fglu fins pglu pins htnmed lipmed glumed chd chdang cdcr diab3 dumDom dumCSA dumCub dumPR dumOth dumForm dumCurr dumPDiab dumDiab triglog fglulog pglulog finslog pinslog bmix ldlx triglgx fglulgx pglulgx finslgx pinslgx;
USEVARIABLES ARE wc sbp dbp hdl triglgx fglulgx htnmed lipmed glumed age dumDom dumCSA dumCub dumPR dumForm dumCurr income educ fhchd fhstrok chd;
MISSING ARE ALL (-999);
STRATIFICATION IS strat;
CLUSTER IS psu;
WEIGHT IS weight;
SUBPOPULATION = bkgrd LE 4 AND gender EQ 1;
CLASSES = c(2);
CATEGORICAL = htnmed lipmed glumed;

ANALYSIS:
TYPE=MIXTURE COMPLEX;
STARTS = 100 10;
STITERATIONS = 20;
ALGORITHM=INTEGRATION;
INTEGRATION=MONTECARLO;

MODEL:
%overall%
sbp WITH dbp;
hdl WITH triglgx;
c ON age dumDom dumCSA dumCub dumPR dumForm dumCurr income educ fhchd fhstrok;
age dumDom dumCSA dumCub dumPR dumForm dumCurr income educ fhchd fhstrok;
chd ON age dumDom dumCSA dumCub dumPR dumForm dumCurr income educ fhchd fhstrok;
%c#1%
sbp WITH dbp;
hdl WITH triglgx;
wc sbp dbp hdl triglgx fglulgx;
%c#2%
sbp WITH dbp;
hdl WITH triglgx;
wc sbp dbp hdl triglgx fglulgx;

OUTPUT: SAMPSTAT RESIDUAL TECH1 TECH8 TECH11;