**The Pennsylvania State University**

**The Graduate School**

**College of Health and Human Development**

# ASSESSING THE IMPACT OF PHYSICAL CONDITIONING, DIETARY INTAKE, BODY FAT, AND TOBACCO USE ON BLOOD PRESSURE PARAMETERS: A TWO-METHOD MEASUREMENT DESIGN APPROACH

A Thesis in

Biobehavioral Health

by

Allison E. Olchowski



Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

August 2007

The thesis of Allison E. Olchowski has been reviewed and approved* by the following:

John W. Graham, Ph.D.
Professor of Biobehavioral Health
Chair of Committee

Robert J. Turrisi, Ph.D.
Professor of Biobehavioral Health

Sheila G. West, Ph.D.
Associate Professor of Biobehavioral Health

Edward A. Smith, Ph.D.
Associate Professor of Human Development and Family Studies

Collins O. Airhihenbuwa, Ph.D.
Professor of Biobehavioral Health
Head of the Department of Biobehavioral Health

*Signatures are on file in the Graduate School.

## ABSTRACT

The two-method measurement design was applied to answer substantive questions pertaining to hypertension and several lifestyle-based risk factors; specifically, analyses involved estimating the cross-sectional association of three blood pressure parameters (systolic blood pressure (SBP); diastolic blood pressure (DBP); and pulse pressure (PP)) and four known hypertension risk factors (physical conditioning; dietary intake; body fat; and tobacco use). The two-method measurement design, a recent statistical advancement in the area of planned missingness approaches, measures constructs using several indicators of variable cost and validity. Cheaper, less valid measures of a construct and more expensive, valid measures of the same construct collectively serve as manifest indicators. All participants provide data for the cheap measures; a small proportion of participants also provide data for the expensive measures. When at least a subsample of participants provide complete data, bias correction models allow for the modeling of measurement bias (i.e., reduced construct validity) associated with the cheap measures; resulting parameter estimates are efficient and unbiased.

A simulation paradigm was used to apply the two-method measurement design to empirical NHANES data. The performance of the two-method measurement design was compared to that of financially-equivalent complete case designs. For two of the four predictors – body fat and tobacco use – application of the two-method measurement design produced statistical power advantages beyond those yielded by the financially-equivalent complete cases models. Under a hypothetical budget constraint of $20,000, complete case body fat data could be collected from N=333 participants; however, the two-method measurement design behaved as if the complete case sample sizes were N=1469, N=1625, and N=1565 for testing the effects between body fat and SBP, DBP, and PP, respectively. Under the same budget of $20,000, complete case tobacco use data could be collected from N=363 participants; the two-method measurement design behaved as if the sample sizes were N=1513, N=655, and N=872 for testing the effects between tobacco use and SBP, DBP, and PP, respectively. Application of the two-method measurement design was comparatively less effective for the physical conditioning and dietary intake variables. A potential explanatory factor involves the general lack of association between cheap and expensive measure indicators for these two variables.

In general, the strength of association between the independent and dependent variables was inversely correlated with the increase in statistical power produced by the two-method measurement design (results consistent with previous research). Results have implications for future application guidelines. To maximize the utility of the two-method measurement design, cheap and expensive measures for a given construct should be highly correlated. It is recommended that researchers collect small amounts of data from candidate cheap measures to determine, a priori, the set of cheap measures that best correlates with expensive measure data. It is also helpful if researchers are able to anticipate, to some degree, effect sizes between independent and dependent variables of interest; this offers researchers the opportunity to more accurately tailor data collection to achieve maximal cost-effectiveness.

Recent interest in the efficiency of health prevention programs, as well as limited external funding sources, has placed an increased emphasis on cost-effective research within many behavioral health disciplines. Whenever researchers have the opportunity to collect data for a particular construct using several measures of variable cost and construct validity, the two-method measurement design offers the potential for cost-effective data collection and unbiased and efficient parameter estimation.

## TABLE OF CONTENTS

# LIST OF TABLES

## LIST OF FIGURES

## ACKNOWLEDGEMENTS

I would first like to extend my thanks and gratitude to Dr. John Graham, my committee chair and mentor. Without your innovative ideas and assistance, this dissertation would not be a reality. I appreciate the care with which you reviewed drafts of my writing and for conversations that clarified my thinking. You have a gift for explaining difficult concepts with astounding simplicity. I am greatly indebted to you for all of your guidance and encouragement.

I would also like to thank Dr. Rob Turrisi, Dr. Sheila West, and Dr. Ed Smith for their professional collaboration and expertise. Throughout this process, your willingness to share your time and knowledge with me challenged me to meet your expectations. I am privileged to have had you as committee members.

Mom and Dad, I finished school! All it took was five years and all of your love, support, and patience. I think you should know that I'm where I am today because of the parents you are. I remember, one time, you mentioned to Adam, David, and I that the two of you wondered how we turned out, through your eyes, so well. I didn't answer you then, but I wanted to say this: it was you. You have always recognized our potential and made any and all sacrifices it took to help us reach whatever endpoint we were aiming for. I'm happy that, through your eyes, I have made you proud.

Liz, you are my Penn State family. Your top ten lists, listening skills, and superb taste in music have pulled me through some long days. I'm sad that you will not be down the hall from me in Delaware, but that just means the BWF legacy will enter its long-distance phase. I know that you will do amazing things next year and in the blink of an eye, you'll be writing your own acknowledgements page (don't forget our antics when you do). You're next! You will be great. Peace.

Rita and Julie, you are two of the greatest friends imaginable. Thank you for never making me feel as if I haven't seen you in months, even when it's true. It's been years and years, but I can pinpoint, with certainty, the very moments I met both of you. I think it's my brain's way of documenting exactly when my life changed for the much, much better.

To my mother- and father-in-law, Nancy and Dan, thank you for accepting me into your family. I deeply appreciate all of your efforts to make me feel comfortable in a new house, in a new state, in a new life. But, most importantly, I thank you for John.

John, I doubt I will ever be able to properly acknowledge your role in seeing this through to the end, but I'm going to give it a shot. Thank you for all of your sacrifices – the enormous ones and the small ones – over the past five years. Thank you for never making me seriously second guess my decision to finish school from Pennsylvania. This degree was my goal, but I thank you for your encouragement on the handful of occasions when you wanted this more than I did. Thank you for calling from family and friends' houses so I was included in gatherings; you probably won't ever realize how you make me feel at home, even from another state. Thank you for your humor, your patience, and Polar the Pirate Dog (!). Thank you for putting up with an unconventional first year of marriage. I promise that our second year will be filled with more Sunday night football and fewer Monday morning goodbyes. I love you very much!

# CHAPTER 1

## Introduction and Research Questions

The number of deaths attributable to cardiovascular disease (CVD) has substantially declined over the past several decades. The Centers for Disease Control and Prevention (CDC) report that, since 1950, age-adjusted CVD mortality rates have fallen approximately 60 percent (CDC, 1999). This striking decline prompted the CDC to recognize the successful reduction of CVD-related deaths as one of the most important public health achievements of the 20th Century (CDC, 1999). Declining CVD mortality rates are attributed to both primary and secondary prevention efforts. Primary prevention efforts, designed to prevent or delay the onset of CVD, are credited with reducing the prevalence of several modifiable CVD risk factors. Secondary prevention efforts, designed to stop or slow the progression of existing health conditions, have successfully increased risk factor screening within the population (Natarajan & Nietert, 2003).

However, despite the substantial decline in CVD mortality, heart disease and stroke remain the first and third most common causes of death, respectively, in the United States (CDC, 1999). Furthermore, research indicates that prevalence rates for several CVD risk factors are increasing at an alarming rate. Since the early 1960s, the prevalence of obesity among adults in the United States has more than doubled, rising from 13 percent to over 30 percent (Flegal, Carroll, Ogden, & Johnson, 2002). Similar trends are found among children and adolescents; since the late 1980s and early 1990s, the prevalence of overweight youth increased from approximately 12 to 16 percent (Munter, He, Cutler, Wildman, & Whelton, 2004). A corresponding increase in diabetes prevalence has occurred; in 2000, 8 percent of Americans were diabetic, up from 5 percent in the late 1970s (Gregg et al., 2005). Behavioral health data

offer insight into the causes driving the increase in risk factor prevalence. The proportion of Americans reporting no regular physical activity increased approximately 8 percent between 1990 and 1997 (Arnett et al., 2002). Dietary compensations for decreasing activity levels have not occurred; during the same time period, total energy intake (kcal) remained constant among men, but increased significantly among women (Arnett et al., 2002).

Hypertension, a well-substantiated CVD risk factor, benefited from public health prevention efforts in the latter half of the 20th Century; significant declines in hypertension occurred between the 1970s and early 1990s (Peterson, Czerwinski, & Siervogel, 2003; Qureshi, Suri, Kirmani, & Divani, 2005). Within the past decade, however, declining hypertension prevalence trends have reversed; since the early 1990s, the prevalence of pre-hypertension and hypertension has substantially increased, with minority populations bearing a disproportionate burden of illness (Fields et al., 2004; Qureshi et al., 2005). Especially alarming are the significant increases in pre-hypertension observed among adolescents and young adult populations (Munter et al., 2004; Qureshi et al., 2005). Currently, over 100 million Americans are pre-hypertensive or hypertensive, reflecting prevalence rates nearing 40 percent of the population (Qureshi et al., 2005).

Therefore, while CVD mortality rates have decreased substantially since the 1950s, the prevalence of several known CVD risk factors are on the rise. Experts cautiously note that current epidemiological trends may erase much of the progress made in reducing CVD mortality during the past several decades (Mensah & Brown, 2007). In the absence of successful, large-scale public health efforts, Mensah and Brown (2007) describe three current population-based trends with the potential to reverse declining CVD-related death rates: (1) the rapidly aging

population; (2) the dramatic increases in obesity and type 2 diabetes prevalence; and (3) the decline in prevalence of adults who display no CVD-related risk factors. In addition to the cascade of public health repercussions resulting from increased incidence of CVD, the enormous rise in CVD-related health care costs will have important economic implications for the United States. Future challenges regarding the reduction of CVD prevalence involve reducing racial and socioeconomic health disparities, establishing state- and local-level policies to promote healthy behaviors, and identifying emerging CVD risk factors (CDC, 1999). Re-organizing resources and launching efficient national-level initiatives to prevent morbidity and mortality requires knowledge of the complex associations among CVD predictors and outcomes (CDC, 1999). Additional research is necessary to identify the underlying mechanisms that account for current epidemiological patterns; namely, the simultaneous decline in CVD mortality in light of increasing CVD risk factor prevalence.

Research supports that hypertension partially mediates the relationship between behavioral risk factors and CVD (Cook et al., 2007; Wong et al., 2003). Thus, the recent dramatic increase in hypertension may be indicative of an impending reversal of CVD morbidity and mortality rates. Furthermore, since hypertension is an immediate CVD precursor, additional research is needed to determine how the prevalence rates of more distal CVD risk factors (e.g., health behaviors) affect more immediate CVD risk factors. This dissertation focuses on the cross-sectional association of hypertension and four known lifestyle-based CVD predictors: (1) physical conditioning; (2) dietary intake; (3) body fat; and (4) tobacco use. Analyses focus on estimating the degree of association between these predictors and three hypertension-related outcome measures: systolic blood pressure, diastolic blood pressure, and pulse pressure.

To assess these relationships, a two-method measurement framework is adopted. The two-method measurement design represents a recent statistical advancement in the area of planned missingness approaches. In a basic sense, planned missingness designs help researchers make the best use of limited resources by collecting maximum data for minimum costs. The two-method measurement design involves collecting complete data from a small proportion of participants and partial data from a much larger proportion of participants. Constructs are measured using several indicators of variable cost and construct validity. Cheaper, less valid measures of a construct and more expensive, valid measures of the same construct collectively serve as manifest indicators. All participants provide data for the cheap measures; a small proportion of participants also provide data for the expensive measures. Because the two-method measurement design capitalizes on the cost-effectiveness of collecting a substantial portion of partial data, data from more participants are able to be collected for the same cost as collecting complete data from a smaller sample. When at least a subsample of participants provide complete data, bias correction models allow for the modeling of measurement bias (i.e., reduced construct validity) associated with the cheap measures. The missing data generated from implementation of the two-method measurement design are handled appropriately using current missing data procedures; resulting parameter estimates are efficient and unbiased.

The two-method measurement design represents an important advancement in fields of research that rely on comparatively cheaper, less valid measures (e.g., self-report measures). Typically, the validity of self-report measures is assessed using a piecemeal approach in which researchers collect small amounts of data from self-report measures as well as from more valid, expensive measures; these two categories of data are used to assess convergent and/or

discriminate validity of the self-report measures. Provided an acceptable level of validity, researchers proceed with full-scale data collection using the self-report measures; often, the small amount of data collected from more valid indicators is discarded following preliminary assessment of cheap measure validity. The two-method measurement design is based on modified data collection principles; namely, cheaper, less valid measures, as well as more valid, expensive measures, are used to collect data. By using expensive measures to model the bias associated with the cheap measures, the two-method measurement design ensures that researchers do not discard any valuable data. In this respect, the two-method measurement design represents an innovative and comparatively efficient research design that concurrently capitalizes on the affordability of the cheaper measures and the validity of the expensive measures. Thus, it is important that the knowledge base pertaining to the two-method measurement design be extended by applying the design to diverse research scenarios.

The utility of the two-method measurement design is established for bivariate models involving simulated data (Graham, Taylor, Olchowski, & Cumsille, 2006). However, this design has not yet been applied to more complex models or models tested with empirical data. Thus, this dissertation involves the application of a new methodology to answer a substantive question pertaining to hypertension and related risk factors. The main model of interest includes seven variables; three dependent variables (systolic blood pressure (SBP), diastolic blood pressure (DBP), and pulse pressure (PP)) are predicted by four independent variables (physical conditioning, dietary intake, body fat, and tobacco use). Cheaper, less valid indicators of each independent variable are represented by self-report data. Estimated $VO_2$ max, total nutrient scores, body fat percentage and technician-obtained anthropometric measures, and serum

cotinine represent the expensive, more valid indicators of physical conditioning, dietary intake, body fat, and tobacco use, respectively.

A simulation paradigm (Graham, Cumsille, & Elek-Fisk, 2003; Graham et al., 2006) is used to apply the two-method measurement design to empirical data. The performance of the two-method measurement design is compared to that of financially-equivalent complete cases designs. Given the extension of the recently developed two-method measurement design to a new, diverse set of structural equation models, practical application guidelines are discussed. Research questions answered by this dissertation are formally stated below.

**1. What is the degree of association between established hypertension risk factors (physical conditioning, dietary intake, body fat, and tobacco use) and SBP, DBP, and PP in a large, nationally representative sample?**

**2. For each independent variable, what is the optimal ratio of partial to complete data that yields the most efficient and unbiased regression coefficients under the two-method measurement design?**

**3. What effect does the number of expensive measures for an independent variable, as well as the strength of correlation between cheap and expensive measures, have on the performance of the two-method measurement design?**

# CHAPTER 2

## Review of the Literature

### *The Two-Method Measurement Design*

The two-method measurement design belongs to a larger class of planned missingness designs in which researchers deliberately collect partial (incomplete) data for at least a subsample of participants (Graham et al., 2006). The resulting missing data are handled appropriately using full-information maximum likelihood (FIML) procedures found in most structural equation modeling packages[1]. With FIML procedures, missing data and parameter estimation are addressed simultaneously. Whereas multiple imputation is a two-step process in which missing data are handled prior to parameter estimation, FIML procedures achieve the same result in one analysis. Maximum likelihood refers to the fact that parameter estimates are calculated based on the probability that observed scores would occur given a particular population parameter estimate.

FIML procedures are well-suited to missing data that are missing at random (MAR); data are considered MAR if the cause of missingness is a variable that has been measured and included in analyses (Graham et al., 2003). A special case of MAR missingness occurs when the cause of missingness is uncorrelated with the variable containing missingness; missing values meeting this criteria are referred to as missing completely at random (MCAR) (Graham et al., 2003). For the two-method measurement design (and other planned missingness designs), missing values are under the researcher's control; the decision as to what participants provide partial versus complete data is random. Therefore, the MCAR assumption (specifically, that the

cause of missingness is uncorrelated with the variable containing missingness) is met under this design (Graham et al., 2006); accordingly, missing data generated generated by the two-method measurement design are appropriately analyzed using FIML procedures.

The two-method measurement design involves two categories of measures: relatively cheap, less valid indicators of a particular construct, and expensive, more valid indicators of the same construct (Graham et al., 2006). Thus, this design is applicable to situations in which two (or more) indicators of a given construct exist but differ substantially in their cost, ease of implementation, and/or construct validity. In such situations, and given a fixed budget scenario, researchers are seemingly faced with two options: (1) exhaust their resources (time and money) collecting a large amount of data using the cheaper, less valid measures or (2) exhaust their resources collecting a relatively small amount of data using the more expensive, valid measures. Neither option is ideal. Reduced construct validity resulting from cheap measure data has the potential to yield biased parameter estimates. Reduced statistical power resulting from a smaller amount of expensive measure data increases the likelihood of type II error.

Essentially, the two-method measurement design makes possible the combining of cheap and expensive measure data to simultaneously capitalize on the increased validity associated with the expensive measures and the affordability associated with the cheap measures. Under the two-method measurement design, the expensive, valid measures are used to model and control for the response bias associated with the cheaper measures (Graham et al., 2006). This research design directly addresses the primary drawback associated with collecting exclusively cheap measure data: reduced construct validity. If researchers can control for reduced construct validity

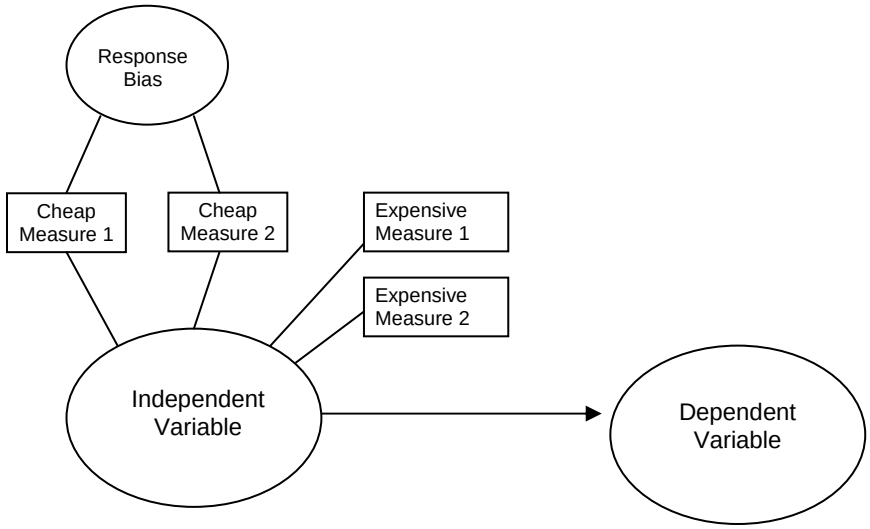
---

[1] Missing data are also commonly handled by multiple imputation; however, for the two-method measurement design, FIML procedures are preferred because missing data are addressed concurrently with parameter estimation.

using expensive measure data, they are no longer restricted to choosing between the two options previously described. Rather, they are able to expand their options to consider designs in which data are collected using both types of measures. Because of the opportunity for blended data collection, the two-method measurement design allows researchers to benefit from the desirable traits of each of measure (validity of the expensive measures, affordability of the cheap measures) without making sacrifices due to undesirable traits (cost of the expensive measures, reduced validity of the cheap measures).

The concept behind the two-method measurement design is illustrated by the bias factor model (Graham et al., 2006), shown below in Figure 1. The bias factor model is a structural equation model template in which the latent construct of substantive interest (the common factor comprised of cheap and expensive measures) predicts one or more dependent variables. Cheap and expensive measures load on the common factor (i.e., independent variable); additionally, cheap measures load on a second factor representing measurement bias. In this manner, the bias factor model specifies two sources of correlation between cheap measures and one source of correlation between expensive measures (Graham et al., 2006).

Figure 1
*Basic Bias Factor Model*




One important requirement for the bias factor model is that at least a subset of participants has both cheap and expensive measure data (i.e., complete data); this requirement is necessary for successful missing data analysis. When at least a subset of the sample has complete data, the remainder of the sample is permitted to have partial data in the form of exclusively cheap measure data. Given that only a subset of participants provide complete data, the two-method measurement design increases the cost-effectiveness of data collection; for the same cost as collecting complete case data from a certain number of participants, the two-method measurement design provides researchers with the equivalent of substantially more complete cases with which to test effects. Thus, the two-method measurement provides a statistical power advantage over the financially-equivalent complete case designs.

Accordingly, it is necessary to determine the optimal ratio of partial to complete data that yields the best parameter estimates when the two-method measurement design is applied. If the optimal ratio is very large, researchers will collect cheap measure data from all participants and expensive measure data from just a very small fraction of participants. If the optimal ratio is less

extreme, researchers will collect cheap measure data from all participants and expensive measure data from a comparatively larger group of participants. The optimal ratio is not static across data scenarios; rather, it varies as a function of the cost differential between the cheap and expensive measures and the strength of association between the independent and dependent variables of interest (Graham et al., 2006).

The statistical benefits derived from the two-method measurement design are quantified using a measure referred to as the Effective N Increase Factor; essentially, the Effective N refers to the number of complete cases necessary to produce the same standard error obtained under the most efficient two-method measurement design (Graham et al., 2006). The Effective N Increase Factor is computed by dividing the Effective N by the number of complete cases allowable under a specific data collection budget (Graham et al., 2006) (please refer to the following equation):

$$\text{Effective N Increase Factor} = \frac{\text{Effective N}}{\text{Nominal Complete Cases N}}$$

Using simulated data and bivariate models, Graham and colleagues (2006) obtained Effective N Increase Factors ranging from 1.09 to 3.47, depending on the data scenario, reflecting the advantage of the two-method measurement design over a complete cases design. Effective N Increase Factors greater than 1.0 indicate that, *for the same cost*, the two-method measurement design yields better (lower) standard errors than are possible using complete cases analysis. The higher the Effective N Increase Factor, the greater the benefits derived from the two-method measurement design.

### *Current Application of the Two-Method Measurement Design*

To evaluate the utility of the two-method measurement design across a variety of data

scenarios, Graham and colleagues (2006) altered the cost differential between the cheap and expensive measures, as well as the effect size between the independent and dependent variables. Regardless of cost differential and effect size, results indicated that the best design (the design producing the lowest standard errors for key parameter estimates) was always a version of the two-method measurement design (Graham et al., 2006). Standard errors obtained from financially-equivalent complete case models were comparably larger as a result of reduced sample size.

Findings from Graham et al. (2006) indicate that the two-method measurement design has the potential to provide researchers with substantial benefits in light of resource constraints. However, given its recent development, applications of the two-method measurement design have been limited to relatively simple models involving simulated data. It is important to examine how the two-method measurement design performs when applied to empirical data and within the context of complex models.

This dissertation represents the first application of the two-method measurement design involving empirical data. The proposed model assesses the impact of physical conditioning, dietary intake, body fat, and tobacco use on blood pressure parameters. The four predictors are comprised of a variable number of manifest indicators that differ in strength of intercorrelation; as a result, this dissertation also examines how the absolute number of cheap and expensive measures, as well as the intercorrelations among manifest indicators, impact the performance of the two-method measurement design.

*Hypertension: Definition, Prevalence, and Causal Factors*

Hypertension is clinically defined as chronically elevated blood pressure at or above 140 mm Hg (systolic) or 90 mm Hg (diastolic) (Silverthorn, 2001). Systolic blood pressure (SBP) represents the highest pressure in the circulatory system and reflects the pressure with which the ventricles contract (Silverthorn, 2001). Diastolic blood pressure (DBP) represents the lowest pressure in the circulatory system, occurring just prior to the next contraction of the heart muscle. Blood pressure is a direct function of three main cardiovascular determinants: arteriolar resistance, cardiac output, and blood volume. Internal and external factors that impact these three facets of the cardiovascular system have the ability to alter blood pressure (Silverthorn, 2001).

A third blood pressure parameter – pulse pressure (PP) – also offers insight into cardiovascular health. Mathematically, PP represents the difference between systolic and diastolic blood pressures (i.e., PP = SBP – DBP); physiologically, it serves as an indicator of arterial stiffness. In youth, arterial wall stress is supported predominately via elastin fibers; with advancing age and arterial pressure, arterial walls increasingly rely on stiffer collagen fibers for support (Nichols, 2005). Increased arterial stiffness, attributable to increased levels of collagen fibers, alters pressure wave velocity within the cardiovascular system; as a result, pressure waves are refracted much more quickly. Instead of wave refraction occurring during diastole, waves are refracted during periods of systole (Nichols, 2005). Rapid wave refraction results in increased SBP and duration of systole (Nichols, 2005), both of which are detrimental to cardiovascular health. Prolonged periods of systole delay relaxation of the heart, promote left ventricular hypertrophy, and place individuals at increased risk of heart failure (Cheitlin, 2003).

The utility of PP as a predictor of CVD is contentiously debated in the literature; multiple

studies report that PP is strongly predictive of cardiovascular outcomes (Haider, Larson, Franklin, & Levy, 2003); however, other studies suggest that PP is a comparatively less useful indicator of cardiovascular efficiency, as its predictive ability is inconsistent across cardiovascular risk profiles (Pastor-Barriuso, Banegas, Damian, Appel, & Guallar, 2003). The complexity surrounding PP as a predictor of CVD stems from its linear dependence on SBP and DBP; an increase in PP is achieved via an increase in SBP *or* a decrease in DBP *or* both (Pastor-Barriuso et al., 2003). The underlying SBP and DBP values dictating an individual's PP have important implications for PP as a CVD predictor. Increased PP attributable to elevated SBP is positively associated with CVD risk; on the other hand, increased PP attributable to reduced DBP is negatively associated with CVD risk (Pastor-Barriuso et al., 2003). Thus, it is suggested that PP be interpreted in the context of SBP and DBP, and not as a stand-alone indicator of CVD risk (Pastor-Barriuso et al., 2003).

Hypertension severity is categorized using a classification system developed by the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure (Chobanian et al., 2003). Stage 1 hypertension is defined as chronically elevated blood pressure at or between 140-159 mm Hg (SBP) and/or 90-99 mm Hg (DBP). Stage 2 hypertension, a more severe form of hypertension, refers to chronically elevated blood pressure at or above 160 mm Hg (SBP) and/or 100 mm Hg (DBP). Additionally, in light of the rapidly increasing prevalence of hypertension, a new "pre-hypertensive" category was recently created to include individuals with blood pressure at or between 120-139 mm Hg (SBP) and/or 80-89 mm Hg (DBP) (Chobanian et al., 2003). This new classification category reflects an increased emphasis on healthy lifestyle changes to prevent at-risk individuals from reaching the Stage 1

hypertension threshold. Normotensive individuals - those without hypertension - have blood pressure levels below 120 mm Hg (SBP) *and* 80 mm Hg (DBP) (Chobanian et al., 2003).

An estimated 100 million Americans (almost 40 percent of US adults) are pre-hypertensive or hypertensive (Qureshi et al., 2005). Furthermore, a substantial proportion of pre-hypertensive adults are likely to become hypertensive; one recent study found that over 93 percent of pre-hypertensive individuals also have at least one additional CVD risk factor (Liszka, Mainous, King, Everett, & Egan, 2005), dramatically increasing the likelihood of reaching the Stage 1 hypertension threshold. Hypertension prevalence estimates differ substantially as a function of age, race, and gender. Controlling for lifestyle factors that contribute to the development of high blood pressure, the onset of hypertension increases linearly with age; approximately 50 percent of adults between the ages of 60-69 are hypertensive while almost 75 percent of adults over the age of 70 are hypertensive (Burt, Whelton, Roccella, Brown, Cutler, Higgins, et al., 1995). Adults who reach the age of 50 hypertensive-free have a 90 percent chance of eventually developing the disorder in their lifetime (Vasan et al., 2002). Given recent epidemiological trends, hypertension prevalence is also likely to increase among young people; since 1988, mean SBP and DBP have increased 1.4 mm Hg and 3.3 mm Hg (both $p < 0.001$), respectively, among adolescents (Muntner et al., 2004). Hypertension prevalence is disproportionately higher among women and minority populations. African American and Hispanic adults are more likely to be affected than non-Hispanic whites (Fields et al., 2004; Qureshi et al., 2005). Higher prevalence estimates for certain minority populations (e.g., Mexican Americans) may appear in the future as the median age of these subpopulations increase (Fields et al., 2004).

Blood pressure is a potent predictor of CVD; this association is evident even within normotensive ranges of blood pressure (age- and sex-standardized), suggesting that there is no detectable risk threshold for blood pressure as a CVD precursor (Lewington et al., 2002). A meta-analysis of 61 prospective studies suggests that the causal association between hypertension and CVD mortality is stronger than previously thought (Lewington et al., 2002). After extensive adjustment for additional CVD risk factors, hypertensive adults are 2.37 times as likely to suffer a myocardial infarction, stroke, or congestive heart failure compared to normotensive adults (Liszka et al., 2005); even pre-hypertensive adults who have yet to reach the Stage I hypertension threshold are 1.32 times as likely to experience a CVD event than their normotensive peers (Liszka et al., 2005). For every 20 mm Hg increase in SBP, or 10 mm Hg increase in DBP, mortality rates attributed to stroke, ischemic heart disease, and other cardiovascular events increase two-fold among adults age 40-69 (Lewington et al., 2002).

For approximately 90 percent of hypertensive adults, no single etiological cause is determined (Beevers, Lip, & O'Brien, 2001). Rather, high blood pressure may reflect environmental factors (e.g., diet, activity level), genetic factors (e.g., kidney function), or interactions between factors (Appel et al., 2006). Past research has found a moderate heritability for hypertension. Specifically, the prevalence of hypertension is approximately two times higher among individuals with at least one hypertensive parent compared to peers without a hypertensive parent (Beevers et al., 2001). Twin concordance data also supports the notion of a genetic predisposition to hypertension (Beevers et al., 2001). However, it is important to consider the impact of shared lifestyle on familial hypertension. Shared diet, similar physical activity patterns, and pervasive attitudes toward general health are likely common within

families; thus, heritability may be confounded with shared lifestyle, obscuring the actual impact of genetic factors on the development of hypertension.

Several lifestyle-based risk factors for hypertension are well-substantiated. The following sections provide an overview of the literature pertaining to the four hypertension predictors considered here. Additionally, because the two-method measurement design emphasizes the use of multiple measures for assessing constructs, common methods of measurement are presented and critiqued for each risk factor.

*Physical Conditioning and Hypertension*

An individual's level of physical conditioning reflects several related, yet fundamentally different, constructs. Physical condition is not solely demonstrated by a person's level of physical activity, although physical activity certainly contributes to level of physical conditioning. Similarly, physical condition does not exclusively refer to an individual's fitness level, although physical fitness capabilities undoubtedly influence physical condition. These two elements – physical activity and physical fitness – are neither synonymous nor mutually exclusive. Individuals may be both physically active and physically fit. However, it is also plausible that highly physically fit individuals are not physically active; conversely, extremely physically active individuals may be physically unfit. Therefore, physical condition is a composite idea, reflecting several related dimensions of aerobic performance.

Physical activity refers specifically to bodily movements, produced by skeletal muscles, that result in energy expenditure (Casperson, Powell, & Christenson, 1985). A person's total energy expenditure is a continuous variable (Casperson et al., 1985), and is influenced by

function of mode, frequency, and intensity of that person's physical activity. Physical activity may be segmented into a series of mutually exclusive categories any number of ways (e.g., work-related physical activity versus leisure-time physical activity; light versus moderate versus vigorous physical activity), provided these categories sum to yield the total energy expenditure attributable to physical activity (Casperson et al., 1985).[2] On the other hand, physical fitness represents a set of natural attributes common to all people; these attributes relate to the ability to perform physical activity and are categorized into several fitness measures including cardiorespiratory fitness, muscular strength, and flexibility (Casperson et al., 1985).

Physical activity and physical fitness are related to one another in a fairly complex manner. A large body of literature documents a moderate dose-response relationship between the two dimensions of physical conditioning. High-intensity physical activity is moderately correlated with cardiorespiratory fitness (measured by $VO_2$ max) (males: $r$=0.33; females: $r$ = 0.27) (Talbot, Metter, & Fleg, 2000). However, the degree of association is substantially weaker for moderate-intensity physical activity (males: $r$ = 0.12; females: $r$ = 0.17) and non-significant for low-intensity physical activity (males: $r$ = 0.08; females: $r$ = 0.06) (Talbot et al., 2000). Multiple regression reveals that high-intensity physical activity accounts for just over 10 percent and 7 percent of the variance in $VO_2$ max among males and females, respectively (Talbot et al., 2000). However, after controlling for age and BMI, only 1.6 percent and 1.8 percent of the variance in VO2 max among males and females, respectively, is explained by high-intensity physical activity (Talbot et al., 2000). These findings suggest that intensity of physical activity

[2] The terms physical activity and exercise are often used interchangeably; however, subtle differences in definitions highlight the distinction between the two concepts. Like physical activity, exercise also refers to bodily movement by skeletal muscles, resulting in energy expenditure (Casperson et al., 1985). Yet, exercise differs from physical activity in that exercise refers specifically to a category of "planned, structured, and repetitive bodily movement"

makes a small, yet non-trivial, contribution to cardiorespiratory fitness (Talbot et al., 2000).

Evidence that physical activity and physical fitness represent two related, yet different, concepts comes from studies in which the predictive value of the two constructs are assessed simultaneously for health outcomes. When controlling for one another, physical activity and physical fitness are both inversely associated with all-cause mortality (Myers et al., 2004). In fact, physical activity and physical fitness outperform smoking, hypertension, hyperlipidemia, and diabetes in their ability to predict mortality (Myers et al., 2004). However, physical fitness is a comparatively stronger predictor of cardiovascular risk and all-cause mortality compared to physical activity (Dvorak et al., 2000; Myers et al., 2004; Sternfeld et al., 1999).

Physical activity and physical fitness are both strongly associated with blood pressure. The association between physical activity and reduced blood pressure has been substantiated across age, sex, and racial strata; however, despite numerous national initiatives to promote physical activity, the proportion of individuals who report engaging in no physical activity approaches 50 percent for several subpopulations (Mensah, Hokdad, Ford, Greenlund, & Croft, 2005). While the U.S. Department of Health and Human Services (USDHHS) recommends a minimum of 30 minutes of physical activity most days of the week (USDHHS, 2002), much lower levels of physical activity confer cardiovascular benefits. For example, sedentary hypertensive participants who participated in as little as 30-60 minutes per week of moderate physical activity experienced significant reductions in both SBP and DBP (Ishikawa-Takata, Ohta, & Tanaka, 2003). Furthermore, research supports that cardioprotective effects may be obtained through intermittent bouts of physical activity; three short 10-minute periods of physical activity per day are associated with significant reductions in blood pressure (Staffileno,

---

performed in an attempt to maintain or increase level of physical fitness (Casperson et al., 1985, p.127)).

Braun, & Rosenson, 2001).

Physical fitness is also strongly protective against hypertension. After extensive adjustment for potential confounders, an inverse dose-response relationship exists for cardiorespiratory fitness and hypertension. Moderately fit and highly fit women have a 39 and 65 percent lower risk, respectively, of developing hypertension, compared to women with low cardiorespiratory fitness (Barlow et al., 2006). In fact, among women, cardiorespiratory fitness is as strong a predictor of hypertension as being pre-hypertensive, overweight, or over the age of 55 (Barlow et al., 2006). Similar patterns exist for males. Unfit men exhibit higher SBP and DBP compared to moderately fit or highly fit males (Carnethon, Gulati, & Greenland, 2005).

Unfortunately, low cardiorespiratory fitness is common among both adolescents and adults. In the United States, over 33 percent of adolescents and approximately 14 percent of adults are considered unfit (Carnethon et al., 2005). Despite the lack of gender differences in the prevalence of low cardiorespiratory fitness among adolescents (Carnethon et al., 2005), a disturbing trend exists for adolescent females. With increasing age, adolescent females display lower levels of cardiorespiratory fitness. This trend is reversed for adolescent males; namely, adolescent males display increased cardiorespiratory fitness in the later teenage years (Pate, Wang, Dowda, Farrell, & O'Neill, 2006). As adults, females are significantly less likely to be physically fit compared to adult males (Carnethon et al., 2005). Additionally, the prevalence of low cardiorespiratory fitness is higher among black and Hispanic populations compared to non-Hispanic white populations (Carnethon et al., 2005).

*Measuring physical conditioning.* Self-reports represent the most commonly used type of physical activity measure in related research (Sallis & Saelens, 2000). Major benefits include

their affordability, ease of data collection, ability to appropriately assess physical activity behavior within diverse populations (e.g., very young children, the elderly), and capacity to address physical activity across multiple contexts (Sallis & Saelens, 2000). Additionally, self-report recalls of physical activity do not alter physical activity behavior (Sallis & Saelens, 2000), reducing the likelihood of testing threat in experimental designs. Self-report physical activity data are more valid when collected by interviewers versus data obtained through self-administered surveys (Sallis & Saelens, 2000). Additionally, self-reports of vigorous physical activity are generally more valid than self-reported moderate physical activity (Sallis & Saelens, 2000).

Conversely, self-report physical activity measures have several noteworthy disadvantages. For example, these measures are often impacted by participants' attempts to provide socially desirable responses. One study found social desirability to be the strongest predictor of self-reported frequency of physical activity (Warnecke et al., 1997). Additional problems associated with these measures include the high probability of recall error and the potential for terminology ambiguity to affect participant responses (Sallis & Saelens, 2000). Self-report physical activity measures typically fail to collect data on low-level physical activities, increasing the likelihood of floor effects (i.e., a person's level of physical activity is not captured using a self-report measure because their activity level is lower than the lowest score available) (Tudor-Locke & Myers, 2001). Findings also indicate that self-reports of physical activity are inaccurate estimates of the total amount of physical activity in which an individual engages (Sallis & Saelens, 2000); thus, if researchers are interested in measuring participants' total amount of physical activity (as opposed to more isolated instances such the number of times per

week an individual lifts weights), it may be more appropriate to employ relatively more objective physical activity measures. Finally, due to the dearth of information regarding the internal consistency of self-report physical activity measures (Patterson, 2000), it is often difficult to judge reliability or compare findings across studies.

In contrast to self-report indicators, many objective measures of physical activity exist, including calorimetry, doubly labeled water, direct observation, accelerometers, and pedometers (Tudor-Locke & Myers, 2001). Within the class of objective physical activity measures, motion sensing technology (i.e., accelerometers and pedometers) is gaining popularity. Accelerometers and pedometers are commercially available; may be worn in several locations on the body; and do not require participants to manually record physical activity data. Furthermore, they are well-suited to individuals for which language barriers or illiteracy may render self-reports impractical (Tudor-Locke & Myers, 2001).

Accelerometers record individuals' activity levels in terms of activity counts (described as the product of movement frequency and movement intensity); from these data, researchers derive estimates of total energy expenditure (Tudor-Locke & Myers, 2001). However, accelerometers require specific software for calibration and data analysis; consequently, data collection is often tedious and requires technical expertise (Tudor-Locke & Myers, 2001). Additionally, the high cost of accelerometers serves as a major drawback in large-scale physical activity research; the price of an accelerometer is enormously expensive compared to that of pedometers ($50-$400 versus $10-$30 per unit, respectively) (Tudor-Locke & Myers, 2001). As a result, pedometers are becoming increasingly common in physical activity research. Pedometers are able to detect slight changes in physical activity patterns that self-report

measures may fail to capture. Additionally, pedometers demonstrate excellent step-counting accuracy at typical walking speeds (3.0 miles per hour or faster) (Melanson et al., 2004); this is important as walking represents one of the most popular forms of physical activity among adults (Wood, 2002).

However, experts caution that objective measures should not be considered gold standard methods for assessing physical activity (Patterson, 2000). This suggestion stems from the fact that multiple internal and external influences (e.g., chemical substances, emotional state, mode of activity) may compromise the validity of objective physical activity measures (Patterson, 2000; Tudor-Locke & Myers, 2001). In light of the drawbacks associated with both self-report and objective measures of physical activity, it is recommended that researchers employ a combination of methods (e.g., pedometers and self-report measures) to assess physical activity (Tudor-Locke & Myers, 2001). Similarly, Patterson (2000) suggests comparing data obtained from self-report and objective measures not to establish criterion validity for the self-report measures, but rather to evaluate convergent validity for both types of measures.

Unlike physical activity, physical fitness is generally assessed using relatively more complicated methodologies. $VO_2$ max, derived from maximal exercise testing, is considered a gold standard cardiorespiratory fitness measure (Noonan & Dean, 2001). Maximal exercise tests assess individuals' maximum oxygen consumption, which is reflective of the body's ability to transport and utilize oxygen for energy production (Hartung, Krock, Crandall, Bisson, & Myhre, 1993). However, several practical drawbacks exist regarding maximal exercise testing. Among individuals for which pain and exhaustion (rather than physical exertion) determine testing performance, $VO_2$ max measures may be invalid (Noonan & Dean, 2001). The time and expense

associated with collecting $VO_2$ max data present major difficulties in large-scale epidemiological investigations (Yamani et al., 1995). Additionally, collection of valid $VO_2$ max data requires high levels of participant motivation (Noonan & Dean, 2001).

In light of these obstacles, several other categories of exercise tests have been developed to measure cardiorespiratory fitness. One alternative category – predictive submaximal exercise testing – allows researchers to estimate participants' $VO_2$ max without requiring them to reach levels of maximum exertion. Participants' heart rate and oxygen consumption ($VO_2$) are measured at two more levels of physical work; using regression equations, estimated $VO_2$ max is determined by extrapolating the heart rate-oxygen consumption association to individuals' age-predicted maximum heart rate (Noonan & Dean, 2001). Multiple submaximal exercise tests exist (e.g., walking, cycle ergometer, and shuttle run tests); however, it is important to select an appropriate testing protocol to ensure participants' are not under- or over-stressed (Noonan & Dean, 2001). Many types of submaximal exercise tests produce data within a few minutes; several types do not require costly equipment such as treadmills (e.g., the 20-meter shuttle test) (Noonan & Dean, 2001). Furthermore, $VO_2$ max and estimated $VO_2$ max have been found to demonstrate excellent concurrent validity (Noonan & Dean, 2001). As such, submaximal exercise testing is highly applicable to large-scale investigations.

*Dietary Intake and Hypertension*

Dietary intake is strongly associated with blood pressure and is capable of evoking acute, as well as long-term, cardiovascular responses. Broad-level dietary patterns, as well as specific micronutrients, are important predictors of hypertension and other CVD risk factors. Research

demonstrates that nutrient intake affects biological factors such as kidney function and heart rate activity. Additionally, recent work suggests that nutrients affect cardiovascular health by impairing, maintaining, or improving endothelial function; more specifically, a healthy diet promotes endothelial-induced vasodilation and slows the atherosclerotic process (Brown & Hu, 2001). Maintaining endothelial function is crucial to general cardiovascular health, as abnormal endothelial function is an early risk factor for eventual CVD (Brown & Hu, 2001); evidence suggests that endothelial dysfunction is likely a mediating factor in the progression from hypertension to heart failure (Lapu-Bula & Ofili, 2007).

Regarding broad dietary patterns, the American Dietetic Association recommends a vegetarian diet to substantially lower the risk of developing hypertension, obesity, diabetes, and cancer (Mangels, Messina, & Melina, 2003). Additionally, the American Heart Association endorses reduced alcohol consumption and sodium intake as important points of intervention for reducing hypertension prevalence (Appel et al., 2006). Daily consumption of a multivitamin is associated with beneficial cardiovascular outcomes (McKay, Perrone, Rasmussen, Dallal, & Blumberg, 2000), as is consuming a diet rich in whole grains (Appel et al., 2005; Behall, Scholfield, & Hallfrisch, 2006). When saturated fat, salt, and cholesterol are concurrently reduced, protein-rich diets, as well as diets rich in unsaturated fats, are significantly associated with reductions in SBP and DBP (Appel et al., 2005). In contrast, certain broad-level dietary patterns negatively impact cardiovascular health. Disordered eating patterns (e.g., anorexia nervosa, binge-eating disorder) are associated with increased risk for hypertension and other negative cardiovascular outcomes (Johnson, Cohen, Kasen, & Brook, 2002). Diets high in energy density are directly related to BMI, suggesting an increased risk for hypertension (Kant &

Graubard, 2005).

Consumption of fruits and vegetables, characteristic of a vegetarian diet, increases the bioavailability of folate (Brouwer et al., 1999). Dietary folate is associated with reduced plasma concentrations of homocysteine, a naturally occurring substance known to reduce bioavailability of the vasodilator, nitric oxide (Brouwer et al., 1999; Symons, Rutledge, Simonsen, & Pattathu, 2006). Fruit and vegetable consumption also increases levels of antioxidant vitamins (especially vitamins C and E), shown to inversely correlate with coronary heart disease risk (Brown & Hu, 2001). Again, evidence supports improved endothelial function as the explanatory mechanism; vitamins C and E are associated with improved vasodilation and reduced arterial stiffness (Plantinga et al., 2007).

Significant reductions in SBP and DBP following long-term consumption of a diet rich in whole grains are attributed to the protective effects of fiber (Behall et al., 2006). A recent meta-analysis reports that fiber supplementation significantly reduces DBP; fiber supplementation is also associated with a reduction in SBP, though this association is generally non-significant (Streppel, Arends, van't Veer, Grobbee, & Geleijnse, 2005). While the exact mechanism through which fiber exerts its beneficial effects is unknown (Streppel et al., 2005), preliminary research suggests that fiber may attenuate insulin responses in the body (Wolever, Campbell, Geleva, & Anderson, 2004). Prolonged hyperinsulinemia (consistent with levels associated with type II diabetes) is associated with reduced nitric oxide-dependent endothelial dilation (Arcaro et al., 2002; Campia et al., 2004). Hyperinsulinemia likely results from the body's inability to respond to insulin; indeed, insulin resistance is a defining feature of type II diabetes and is predictive of eventual development of type II diabetes even among non-diabetics (Elder, Prigeon, Wadwa,

Dolan, & D'Alessio, 2006; Osei, Rhinesmith, Gaillard, & Schuster, 2004). Recent research indicates that consumption of a daily recommended level of insoluble fiber enhances whole-body insulin sensitivity (Weickert et al., 2006); thus, dietary fiber may promote vasodilation through improved insulin regulation and metabolism.

Consumption of foods rich in magnesium and potassium is also inversely associated with blood pressure. Magnesium supplementation produces small, but significant, reductions in SBP and DBP; blood pressure reductions are especially noteworthy among hypertensive adults (Kawano, Matsuoka, Takishita, & Omae, 1998). Research suggests that magnesium competes with calcium for membrane-binding sites along the vascular smooth muscle, thus promoting vasodilation (Kawano et al., 1998). Dietary potassium is also protective against hypertension (Appel et al., 2006). Research indicates that potassium supplementation is associated with increased production and release of the vasodilator, nitric oxide, within endothelial cells (Zhou, Kosaka, & Yoneyama, 2000).

In addition to long-term cardiovascular effects, dietary intake also induces acute physiological responses. For example, consumption of a single high-fat meal induces higher SBP and DBP reactivity to stress (Jakulj et al., 2007), indicating that dietary intake has the potential to alter the body's normal response to stressors. The acute physiological responses stemming from nutrient intake often involve endothelial function. Endothelial function is frequently assessed by measuring vascular reactivity (i.e., changes in cardiovascular function resulting from acute stressors (Beevers et al., 2001)). One common method for assessing cardiovascular reactivity, and thus, endothelial health, is flow-mediated dilation (FMD) (West, 2001). FMD of the brachial artery is assessed by briefly depriving the upper or lower arm of oxygen (by way of an inflated

blood pressure cuff). After removal of the cuff, increased blood flow through the brachial artery prompts the endothelium to increase arterial diameter; FMD is represented as the percent change between baseline arterial diameter and arterial diameter following removal of the cuff (West, 2001). Healthy arteries respond to increased blood flow by increasing in diameter; unhealthy or damaged blood vessels, on the other hand, demonstrate reduced dilation (i.e., lower FMD) (West, 2001).

FMD allows researchers to investigate the acute effects of dietary intake on blood pressure. Research shows that consumption of a single high-fat meal is associated with reduced endothelial functioning during the postprandial state (Ferreira et al., 2004; Vogel, Corretti, & Plotnick, 1997). Additionally, the specific type of fatty acid consumed has important implications for acute endothelial responses; consumption of a single meal high in saturated fat significantly reduces FMD three hours following consumption, whereas a single meal high in polyunsaturated fat does not (Nicholls et al., 2006). Research also indicates that certain dietary combinations acutely impact endothelial function in complex ways. For example, when small amounts of polyunsaturated fats are consumed at the same time, acute endothelial dysfunction associated with consumption of a single high-fat meal is effectively reversed (Cortes et al., 2006).

*Measuring dietary intake.* Dietary intake data are commonly collected via self-report measures (e.g., food records (3-, 4-, or 7-day), food frequencies, 24-hour dietary recalls) (Contento, Randell, & Basch, 2002). Self-report dietary intake measures differ drastically in their cost, ease of implementation, and respondent burden; food records and 24-hour dietary recalls are generally expensive and labor-intensive, potentially limiting their utility in large-scale studies

(Contento et al., 2002). Food frequency questionnaires (FFQs) collect data regarding average portion size and frequency of consumption for a given list of foods; FFQs have become increasingly popular in nutrition research as a result of their affordability and ease of data processing (Contento et al., 2002). However, several studies indicate that FFQs are not valid dietary intake measures at the individual level (Anderson, Tomten, Haggarty, Lovo, & Hustvedt, 2003). Anderson and colleagues (2003) suggest that the reduced validity associated with FFQs results from under- or over-reporting both portion size as well as food consumption frequency.

FFQs are often calibrated using comparatively more expensive 24-hour dietary recalls (Subar et al., 2003); with this measure, participants provide a detailed report of all food and beverages consumed in the 24 hours immediately preceding data collection. While 24-hour dietary recalls are often more valid than FFQs, they are nonetheless associated with biased estimates of dietary intake. On average, men and women underreport total energy intake by up to 14 and 20 percent, respectively, when measured using 24-hour dietary recalls. When assessed using FFQs, men and women underreport total energy intake by up to 36 and 38 percent, respectively (Subar et al., 2003).

Several individual-level factors influence the validity of dietary intake measures. Among women, fear of negative evaluation is significantly associated with underreporting of energy intake on FFQs as well as 24-hour dietary recalls (Tooze et al., 2004). Men who eat fewer than five times per day are also significantly more likely to underreport energy intake on both measures. For both men and women, social desirability is significantly associated with underreporting of energy intake for 24-hour dietary recalls, but not FFQs (Tooze et al., 2004). Typically, 24-hour dietary recalls are interviewer-administered whereas FFQs are commonly

self-administered. However, a sub-category of 24-hour dietary recalls – computer-assisted self-administered dietary interviews – eliminates the need for interviewers, potentially reducing the likelihood of social desirability bias. Self-administered dietary interviews are cheaper and less intrusive than traditional 24-hour dietary recalls and incorporate external cues such as prompting or portion size props to assist participants in providing valid dietary intake data (Kohlmeier, Mendez, McDuffie, & Miller, 1997).

Researchers argue that in order to detect causal associations involving dietary intake and health outcomes, valid and objective measures of exposure are necessary (Bingham, 2002). Compared to self-report measures, nutritional biomarkers may represent comparatively more objective indicators of dietary intake; however, few studies incorporate this class of measures (Subar et al., 2003). A major disadvantage associated with nutritional biomarkers is their relative lack of availability; currently, subcutaneous adipose tissue samples, doubly labeled water, 24-hour urine nitrogen, and 24-hour urine potassium provide valid measures of fatty acid intake, energy expenditure, protein intake, and potassium intake, respectively (Bingham, 2002). However, valid biomarkers are generally unavailable for measuring the intake of additional macro- and micronutrients. As nutritional biomarkers are developed, it is important to assess not only their reliability, but also their correspondence with actual exposure to dietary agents (Marshall, 2003). To be of maximal utility, nutritional biomarkers should be specific for exposure to a single dietary component (Marshall, 2003).

*Body Fat and Hypertension*

Body fat is directly associated with multiple negative health outcomes, including hypertension, cardiovascular mortality, cancer incidence, cancer mortality, and all-cause mortality (Gong, Agalliu, Lin, Stanford, & Kristal, 2007; Hu, Juneja, Maihle, & Cleary, 2002; Lee, Blair, & Jackson, 1999). Research shows that associations between excess body fat and negative CVD risk factors (e.g., elevated SBP, DBP, and triglycerides, and decreased HDL-cholesterol) are detectable as early as age 9 (Thompson et al., 2007). This is especially alarming given that adolescents identified as overweight before the age of 18 are 11 to 30 times more likely to be obese as young adults (Thompson et al., 2007). Therefore, the earlier the detrimental effects of excess body fat on the cardiovascular system begin, the more likely these effects will continue throughout a person's lifetime.

Among men and women, increase in body mass index (BMI; [weight (kg) / height $(m)^2$]) over an eight-year period positively predicted increases in SBP and DBP (Wilsgaard, Schirmer, & Arnesen, 2000). In addition to prospective BMI increases, baseline BMI also predicts future blood pressure change; given a uniform increase in BMI, women with higher initial BMIs show greater prospective increases in SBP and DBP than their peers with lower baseline BMIs (Wilsgaard et al., 2000). Associations between body fat and blood pressure are also observed among younger populations. Among youth ages 8-17, increases in mean BMI between 1988-1994 and 1999-2000 accounted for a substantial proportion of variance in the corresponding increases in SBP and DBP during the same time period (Munter et al., 2004).

Recent reviews highlight the metabolic, physiologically-active nature of adipose tissue (Chudek & Wiecek, 2006; Singhal, 2005), and provide evidence of the ability of fat tissue to

invoke atherosclerosis. Adipose tissue secretes a class of inflammatory interleukins and cytokines, including leptin, adiponectin, interleukin-6, and tumor necrosis factor (Fain, 2006). Excess body fat is associated with enhanced secretion of interleukins and cytokines, as well as elevated plasma concentrations of additional pro-inflammatory proteins (e.g., C-reactive protein (CRP), derived from the liver) (Mohamed-Ali et al., 1999; Tchernof, Nolan, Sites, Ades, & Poehlman, 2002). Chronic low-grade inflammation and persistent activation of immune response, characterized by elevated concentrations of cytokines and other pro-inflammatory proteins, is thought to mediate the relationship between excess adiposity and atherosclerosis (Chudek & Wiecek, 2006; Singhal, 2005).

Research also indicates that excess body fat contributes to endothelial dysfunction (Mizia-Stec, 2006); evidence exists for several explanatory mechanisms. Inflammatory responses stemming from excess adiposity interfere with the ability of the endothelium to dilate in response to increased blood flow (Arcaro et al., 1999). Additionally, recent evidence suggests that insulin resistance may mediate the observed association between obesity and endothelial dysfunction (Singhal, 2005). Insulin resistance is associated with reduced secretion of the vasodilator, nitric oxide (Cersosimo & DeFronzo, 2006), and thus contributes to endothelial dysfunction.

*Measuring body fat.* Health researchers have measured body fat for several decades, using assessment tools of variable reliability and validity. Perhaps the most easily obtained and often used body fat measure is BMI; the major benefit of BMI is the ease with which data are collected. While more state-of-the-art body fat measures are available (e.g., hydrostatic weighing and dual-energy x-ray absorptiometry), they are often impractical for large samples due to their higher cost and level of skill required for operation (Sampei, Novo, Juliano, & Sigulem, 2001).

As an alternative, BMI offers researchers an inexpensive and quick method for measuring adiposity.

However, existing literature is contentious regarding the validity of BMI as an indicator of body fat. It is important to note that BMI does not directly measure body fat (Piers, Soares, Frandsen, & O'Dea, 2000); rather, BMI approximates body fat based on participants' height and weight, and not based on the composition of participants' bodies (Heyward & Wagner, 2004). Several studies have found that BMI correlates strongly with gold standard body fatness measures, and thus represents a valid body fat indicator (Mei et al., 2002; Steinberger et al., 2005); however, body fat classifications (i.e., under-, normal-, or overweight) based exclusively on BMI may be invalid at the individual level (Piers et al., 2000). Among adults, BMI has poor sensitivity and poor predictive value of overweight and obese status compared to more invasive methods of body fat assessment (Piers et al., 2000). The poor sensitivity of BMI means that a percentage of actually overweight or obese individuals are not identified through the use of BMI alone. Conversely, a poor predictive value for BMI means that a non-trivial percentage of individuals identified as overweight or obese by BMI do not actually fall under these categories (Piers et al., 2000). Factors such as age and menarcheal status also affect BMI validity (Sampei et al., 2001).

Another relatively sophisticated body fat analysis method, bioelectrical impedance analysis (BIA), is gaining popularity in large-scale epidemiological studies. Advantages associated with BIA include its affordability, portability, safety, reliability, commercial availability, and ease of use (Guo, Chumlea, & Cockram, 1996; Houtkooper, Lohman, Going, & Howell, 1996). As a result, BIA is well-suited for epidemiological research involving large

sample sizes (Guo et al., 1996).

The collection of BIA data is non-invasive and typically lasts less than one minute (Houtkooper et al., 1996; NCHS, 2002b). A small alternating electrical current is introduced into a participant's body at a particular site (typically through electrodes placed on the palm); the current is mild enough that it is non-detectable to participants. Current flows through the body and the resulting voltage is measured at a second site (typically through electrodes placed on the ankle) (Heyward & Wagner, 2004). Impedance is measured as the opposition to electrical flow (Heyward & Wagner, 2004), and varies as a function of an individual's body geometry, size, and electrical properties (Foster & Lukaski, 1996). Impedance measures are broken down into two separate components: resistance and reactance. Resistance represents absolute opposition to current flow, while reactance is a measure of opposition to current flow due to the ability of cell membranes to hold electrical charge (Foster & Lukaski, 1996; Heyward & Wagner, 2004).

As current flows from entry to exit electrodes, it encounters tissues and fluids with variable levels of electrical conductivity. Electrical current is conducted most efficiently through tissues with high water and electrolyte contents. As a result, blood, muscle, and internal organs assist in carrying the charge from entry to exit electrodes; on the other hand, adipose tissue is anhydrous and contains very little water (Heyward & Wagner, 2004). Thus, the more body fat an individual has, the higher their impedance to current flow. Resistance and reactance data, as well as anthropometric data, are included as predictors in one of multiple existing regression equations to predict total body water (TBW), fat mass (FM), fat-free mass (FFM), or body fat percentage (BF)[3]. It is important that BIA equations are applied to samples representative of

[3] TBW represents all intracellular and extracellular fluid in the body. FM refers to all extractable lipids contained in adipose and other tissues. FFM includes all non-lipid fluids and tissues, including water, muscle, bone, organs, and

those from which the equations were derived. A large body of literature indicates that BIA data are much more valid predictors of body composition than are anthropometric measures, such as BMI (Kotler, Burastero, Wang, & Pierson Jr., 1996; Kyle, Genton, Karsegard, Slosman, & Pichard, 2001).

*Tobacco Use and Hypertension*

Tobacco use is capable of producing short- and long-term cardiovascular effects. Consistent findings support that nicotine produces an acute blood pressure reaction immediately following tobacco exposure. Long-term tobacco use is also strongly associated with cardiovascular morbidity and mortality (Mahmud & Feely, 2003; Streppel, Boshuizen, Ocke, Kok, & Kromhout, 2007; Teo et al., 2006). However, studies have cast doubt on nicotine-induced blood pressure responses as the cause of negative CVD outcomes following long-term tobacco use. Support for this theory comes from studies confirming that the negative cardiovascular effects stemming from long-term tobacco use vary as a function of intake method (e.g., cigarette smoke, smokeless tobacco, and transdermal nicotine patches).

Cigarette use is associated with comparatively more CVD risk factors than smokeless tobacco. Long-term cigarette use is significantly correlated with eight CVD risk factors (elevated BMI, waist circumference, waist-hip ratio, plasma triglycerides, plasma insulin, SBP, C-reactive protein, and reduced HDL cholesterol) whereas smokeless tobacco is only significantly associated with two risk factors (increased waist-hip ratio and plasma triglycerides) (Wallenfeldt, Hulthe, Bokemark, Wikstrand, & Fagerberg, 2001). Furthermore, cigarette use is significantly

---

connective tissue. BF represents FM as a percent of total body weight. (All definitions adapted from Heyward & Wagner, 2004).

associated with widespread atherosclerosis, whereas smokeless tobacco use is not (Wallenfeldt et al., 2001). Other studies have found similar results. For example, Westman (1995) found evidence of an acute increase in SBP and DBP following smokeless tobacco use, but substantially weaker evidence of a correlation between long-term smokeless tobacco use and chronic hypertension. Research into the effects of additional forms of tobacco use on CVD risk factors has produced comparable results. Compared to both nasal and transdermal nicotine use, cigarette smoking is more strongly associated with elevated heart rate, blood pressure, and concentration of plasma-based proteins that promote blood coagulation (including fibrinogen, and β-thromboglobulin) (Benowitz, Hansson, & Jacob III, 2002).

Because long-term cardiovascular effects vary with tobacco use intake method, it is suggested that the association between tobacco use and elevated CVD risk is not attributable to nicotine (Wallenfeldt et al., 2001); rather, components found in cigarette smoke other than nicotine may explain the negative effects of cigarette use on cardiovascular health. Several alternative hypotheses involving carbon monoxide have been put forth. While short-term exposure to carbon monoxide does not acutely impact cardiovascular parameters (Zevin, Saunders, Gourlay, Jacob III, & Benowitz, 2001), long-term carbon monoxide exposure may facilitate physiological processes that increase the likelihood of CVD morbidity or mortality.

Research supports that long-term carbon monoxide exposure directly promotes atherosclerosis. Level of cigarette smoking is directly associated with plasma concentrations of multiple pro-inflammatory proteins (Lind et al., 2004). Compared to heavy smokers with low concentrations of pro-inflammatory proteins, heavy smokers with high pro-inflammatory protein concentrations are 1.57 times more likely to experience a cardiac event and 1.50 times more

likely to die from a cardiac event (Lind et al., 2004). Research suggests that carboxyhemoglobin (COHb%), a carbon monoxide-hemoglobin complex formed in red blood cells following inhalation of cigarette smoke, is a mediating factor. COHb% is significantly correlated with both level of cigarette smoking and concentration of pro-inflammatory proteins (Lind et al., 2004). Thus, these findings suggest that carbon monoxide interacts with hemoglobin to increase COHb% levels; in turn, elevated COHb% levels invoke elevated concentrations of pro-inflammatory proteins, thus accelerating the atherosclerotic process.

*Measuring tobacco use.* Tobacco use is generally assessed using self-report measures or biochemical indicators. Self-report measures have the ability to assess a wide range of tobacco-related behaviors, including past and current use, and future use intentions. The flexibility and low cost associated with self-report tobacco use measures make them attractive option for researchers.

On the other hand, biochemical indicators (e.g., serum cotinine) quantify recent exposure to tobacco; as such, these measures assess a more specific facet of tobacco use. Nicotine is the predominant alkaloid in tobacco (Perez-Stable, Benowitz, & Marin, 1995). For any given dose of nicotine entering the body, approximately 70 percent is converted into cotinine, the major metabolite of nicotine; cotinine remains in the bloodstream for a substantially longer period of time than nicotine as a result of its 16-hour half-life (Vartiainen, Seppala, Lillsunde, & Puska, 2002). Smoker status is usually identified via cotinine concentrations greater than 14 ng/ml (Vartiainen et al., 2002; Wagenknecht, Burke, Perkins, Haley, & Friedman, 1992)); passive (second-hand) smoke typically produces serum cotinine levels ranging from 0.5 ng/ml to 10 ng/ml. Studies show that even non-smokers co-habitating with smokers have serum cotinine

levels below 7.5 ng/ml (Emmons et al., 1994). Therefore, determination of smoker status via the typical cotinine threshold of 14 ng/ml is not likely to misclassify actual non-smokers as current smokers based on environmental tobacco exposure.

Studies typically find moderate levels of concordance between self-report and biochemical tobacco use measures. Perez-Stable and colleagues (1995) determined the correlation between self-reported tobacco use and serum cotinine to be $r = 0.695$ and $r = 0.576$ among men and women, respectively. However, research shows that misclassification of smoker status based on self-report data is common for a small percentage of participants. Vartiainen and colleagues (2002) reported that approximately 4 percent of respondents who reported not having smoked in the previous 24 hours had serum cotinine levels indicating recent tobacco exposure. Other studies have found similar results. In a study of young adults, 145 of 3445 self-reported non-smokers had serum cotinine levels above 14 ng/ml (a misclassification rate of 4.2 percent) (Wagenknecht et al., 1992). Additionally, misclassification rates were significantly greater among several subpopulations; African-American participants (compared to white participants), those with a high school education or less, and reported former smokers were respectively two-, three-, and four times more likely to have discordant self-report and serum cotinine data (Wagenknecht et al., 1992). The majority of misclassified smokers had cotinine levels below 50 ng/ml, suggesting that self-report tobacco use measures may be less valid among light smokers compared to heavy smokers (Vartiainen et al., 2002; Wagenknecht et al., 1992). Therefore, while a fairly high degree of association exists between self-report and biochemical tobacco use measures, self-report data may underestimate the prevalence of current tobacco use by up to 4 percent or more among several subpopulations.

# CHAPTER 3

## Method

### *Data*

*Overview of NHANES*

The National Health and Nutrition Examination Survey (NHANES) represents an ongoing surveillance effort by the National Center for Health Statistics (NCHS); the overall goals of NHANES are to monitor health-related incidence and prevalence rates, identify trends in risk behaviors, and explore the relationship between nutrition and health (NCHS, 2001a). Beginning in 1960, National Health Examination Survey (NHES) data were collected across a series of three separate national-level surveys (NHES I, II, and III); in 1970, the NHES was re-formatted with a new emphasis on the health effects of nutrition behaviors within the American population (NCHS, 2001a). Accordingly, NHES was re-named NHANES; four separate NHANES surveys (NHANES I, II, HHANES, and NHANES III) were conducted between 1971 and 1994. In 1999, NHANES was adapted to serve as a continuous, annual survey; during consecutive 12-month periods, cross-sectional health and nutrition data are collected from approximately 5,000 participants (NCHS, 2001a). NHANES uses a stratified, multi-stage probability cluster design to recruit subjects from approximately 15 locations across the United States each year. The sampling strategy ensures oversampling of minorities, low-income populations, adolescents, and individuals over the age of 60 so that data are representative of the non-institutionalized national population (NCHS, 2001a).

Following determination of eligibility and completion of informed consent, participants visited a Mobile Exam Center (MEC) that travels with NHANES staff to each data collection

site. MECs consisted of four linked trailers outfitted with state-of-the-art health examination equipment; each MEC was partitioned to ensure participant privacy. At the beginning of each MEC visit, participants were provided with bar-coded ID bracelets to facilitate their progression through all necessary data collection components (NCHS, 2001a). Data were collected from participants by way of three main MEC components. During an extensive interview portion, self-report data were collected for a wide range of demographic, socioeconomic, health, and nutrition variables. Following completion of the interview, more detailed health and dietary data were collected during the physical examination and laboratory portions of NHANES. Physical examinations were performed by trained and qualified health technicians and included dental, dermatological, vision, hearing, and mental health screenings; additional data were collected by way of a detailed 24-hour dietary recall, collection of anthropometric measures, and a cardiorespiratory fitness assessment. During the laboratory portion of NHANES, participants provided blood, urine, and hair samples for the purposes of quantifying health status biomarkers. MECs conducted two 4-hour data collection sessions per day; data from approximately 10 participants were collected per session (approximately 20 participants per day) (NCHS, 2001a). MEC examinations lasted approximately 3.5 hours, but varied depending on the participant's age. To accommodate participants, MEC appointments were scheduled on weekday mornings, afternoons, and evenings, as well as during weekends (NCHS, 2001a). Following completion of the examination and laboratory portions, participants were provided with a preliminary report of immediately available exam and lab results; a MEC physician reviewed the findings and discussed any abnormal results. Approximately 12-16 weeks following data collection,

participants received the remainder of test results via mail; seriously abnormal results were reported to participants via telephone by NCHS (NCHS, 2001a).

*Data Management*

Public use NHANES data were made available by the NCHS in two-year series; this dissertation incorporates NHANES data from the 1999-2000 and 2001-2002 series, reflecting four years of data. Individual NHANES questionnaire, examination, and laboratory data files, as well as related documentation (e.g., codebooks), were downloaded from main NHANES data dissemination website (http://www.cdc.gov/nchs/about/major/nhanes/datalink.htm). Data merging within and across NHANES series was performed using SAS 9.1. Preliminary analyses, including initial regressions, data management, and data reduction, were performed using STATA 8.0. Model simulations and FIML-missing data procedures were performed using LISREL 8.5.

*Sample Characteristics*

Pooled NHANES data from the 1999-2000 and 2001-2002 series yielded a total sample of N=21,004. Children under the age of 12 were excluded from analyses due to data collection procedures; NHANES protocol specified that proxy data be provided by an adult member of the household for children under the age of 6; children between the ages of 6-11 reported their own data with the help of an adult member of the household (NCHS, 2002c). The ability of young children to provide detailed physical activity and dietary intake data is questionable. Similarly, it is feasible that proxy reports of physical activity and dietary intake do not reflect children's

actual behaviors. Thus, the decision to exclude youth under the age of 12 was made in an attempt to limit analyses to participants who were most likely to reliably report behaviors.

Adults over the age of 65 were also excluded from current analyses. The decision to exclude older adults and elderly participants was based on the dependent variables of interest (SBP, DBP, and PP). The correlation between age and blood pressure varies across the lifespan. While SBP increases with age, research supports a threshold age range at which the increase of DBP with advancing age is no longer linear (Chobanian et al., 2003). Additionally, a recent meta-analysis found that among adults over the age of 60, DBP is inversely correlated with total mortality (Staessen et al., 2000). Including participants of advanced age in analyses would likely have attenuated parameter estimates involving the four lifestyle-based independent variables and DBP. Support for this decision was also found in hypertension literature in which analyses are commonly stratified by age group. Following exclusion of participants under the age of 12 or over the age of 65, the sample decreased to N=12,490.

Participants who reported taking medication to lower their blood pressure (N=1,017), or who indicated that a doctor had told them that they had experienced angina, heart attack, stroke, coronary heart disease, or ischemic heart disease (N=217) were excluded from further analyses. These exclusion criteria have been previously employed to decrease the likelihood of reverse causality (i.e., established disease causes hypertensive status) (Lewington et al., 2002). Finally, N=611 pregnant women were excluded from analyses. The final sample used for analyses included N=10,645 participants. Demographic characteristics of the final sample are displayed below in Table 1.

Table 1
*Descriptive Statistics: Combined Sample Data from 1999-2000 and 2001-2002 NHANES Series*

| | | N | % |
|---|---|---|---|
| **Total Sample** | | 10,645 | 100 |
| **Gender** | | | |
| | Male | 5,376 | 50.5 |
| | Female | 5,269 | 49.5 |
| | 12-20 | 4,985 | 46.8 |
| | 21-30 | 1,510 | 14.2 |
| | 31-40 | 1,485 | 13.9 |
| | 41-50 | 1,359 | 12.7 |
| | 51-65 | 1,306 | 12.3 |
| **Race / Ethnicity** | | | |
| | Non-Hispanic White | 3,807 | 35.7 |
| | Non-Hispanic Black | 2,467 | 23.2 |
| | Hispanic | 3,958 | 37.2 |
| | Other race (inc multiracial) | 413 | 3.9 |
| **Household Income** | | | |
| | $0 - $19,999 | 2,246 | 24.5 |
| | $20,000 - $44,999 | 2,934 | 32.0 |
| | $45,000 - $74,999 | 2,015 | 22.0 |
| | ≥ $75,000 | 1,963 | 21.4 |
| **Education*** | | | |
| | Less than high school diploma | 2,322 | 33.6 |
| | High school diploma | 1,683 | 22.4 |
| | More than high school diploma | 2,897 | 42.0 |
| **Marital Status*** | | | |
| | Married | 3,125 | 47.5 |
| | Widowed | 115 | 1.7 |
| | Divorced / Separated | 716 | 1.1 |
| | Never married | 2,212 | 33.6 |
| | Co-habitating | 416 | 6.3 |

*Note:* * indicates that variables were restricted participants 18 years or older

## *Data Reduction*

The complete model of interest included four independent variables (physical conditioning, dietary intake, body fat, and tobacco use) and three outcome variables (SBP, DBP, and PP). Each independent variable was comprised of cheap and expensive manifest indicators. Cheap measures reflected self-report data obtained during the interview portion of NHANES;

expensive measures reflected data obtained during the examination and laboratory portions of NHANES. Data reduction was performed in such a way that data from multiple manifest measures were combined to yield two cheap measures and either one or two expensive measures for each predictor. The following sections provide an overview of data reduction procedures and a summary of all variables comprising the full model.

*Physical Conditioning Measures*

Physical conditioning was modeled as a latent variable comprised of three manifest indicators (two cheap measures and one expensive measure). Cheap measures were derived from self-report physical activity data; estimated $VO_2$ max ($EVO_2$ max), obtained via submaximal treadmill testing, represented the expensive physical conditioning measure.

*Cheap measures*. Eight self-report items assessing physical activity habits were combined to form two cheap physical conditioning measures. Self-report physical activity items were presented in two interview sections: (a) the Physical Activity Questionnaire and (b) the Individual Activities Questionnaire. The Physical Activity Questionnaire assessed broad-level patterns of physical activity; example items included whether or not participants engaged in housework, yard work, moderate and vigorous physical activity, and muscle strengthening activities during the past 30 days; self-perceived level of physical activity; and average number of hours spent watching television per day.

If participants indicated that they had not engaged in at least 10 minutes of moderate or vigorous physical activity in the past 30 days, they were not presented with any additional physical activity items. If participants indicated that they had participated in at least 10 minutes

of moderate or vigorous physical activity in the past 30 days, they were provided with items from the Individual Activities Questionnaire. These additional items asked participants to provide detailed information regarding each session of physical activity that they had engaged in during the past 30 days. Thus, multiple records existed for individuals who had engaged in more than one session of physical activity during the past month; in such instances, multiple records were collapsed to form variables representing either the average response or the sum of responses pertaining to individual physical activity sessions.

The eight self-report items used to form the two cheap physical conditioning measures are located below in Table 2.

Table 2
*Self-Report Questionnaire Items Used to Generate Cheap Physical Conditioning Measures*

| Variable | Description | Response Range |
|---|---|---|
| palevel | (Collapsed average of) Reported intensity level of physical activity session during the previous 30 days | 1 – 2<br>(1 = moderate; 2 = vigorous) |
| pamet | (Collapsed average of) Estimated metabolic equivalent (MET) intensity level for each physical activity session during the previous 30 days | 0 – 10 METs |
| anyvig | Over the past 30 days, did you do any tasks in or around your house for at least 10 minutes that required moderate or greater physical activity? | 0 = No<br>1 = Yes |
| pasum | (Collapsed sum of) Number of minutes spent engaged in each physical activity session during the previous 30 days | 0 – 2070 Minutes |
| pamin | (Collapsed average of) Number of minutes spent engaged in each physical activity session during the previous 30 days | 0 – 600 Minutes |
| anylift | Over the past 30 days, did you do any physical activities specifically designed to strengthen your muscles such as lifting weights, push-ups, or sit-ups? | 0 = No<br>1 = Yes |
| anymod | Over the past 30 days, did you do any moderate physical activities for at least 10 minutes that caused only light sweating or a slight to moderate increase in breathing or heart rate? | 0 = No<br>1 = Yes |
| pafreq | (Collapsed average of) Number of occasions participant engaged in the specific physical activity during the previous 30 days | 0 – 210 occasions |

Data reduction involved subjecting the eight self-report variables to a principal components factor analysis in which items were forced to load on a single factor. Items were ordered according to strength of factor loading; ordered loadings appear below in Table 3.

Table 3
*Principal Components Factor Loadings for Self-Report Physical Activity Items*

| | |
|---|---|
| palevel | 0.9083 |
| pamet | 0.8964 |
| anyvig | 0.8153 |
| pasum | 0.7162 |
| pamin | 0.6553 |
| anylift | 0.5804 |
| anymod | 0.5412 |
| pafreq | 0.5063 |

Cheap physical conditioning measures were formed by creating two parcels from the eight items described above (parcel formation followed the procedure outlined in Rogers and Schmitt (2004)). Parcel 1, referred to **PA1**, represented the standardized average of palevel, pafreq, pasum, and pamin. Parcel 2, referred to as **PA2**, represented the standardized average of pamet, anymod, anyvig, and anylift. PA1 and PA2 served as the two cheap physical conditioning measures in the full model.

*Expensive measure.* **EVO$_2$ max**, obtained via submaximal treadmill testing, served as the expensive measure of physical conditioning. Treadmill testing was administered by trained, CPR-certified, radiologic technicians during MEC examinations (NCHS, 2004b). All participants between the ages of 12 and 49 were eligible. Exclusion criteria included the following: (1) pregnancy greater than 12 weeks; (2) physical functioning limitations; (3) cardiovascular conditions or symptoms; (4) lung conditions or symptoms; (5) asthma symptoms; (6) medication exclusions (including anti-arrhythmics, beta blockers, and calcium channel-blockers); (7) resting heart rate greater than 100 beats per minute; (8) resting SBP greater than 180 mmHg; (9) resting

DBP greater than 100 mmHg; (10) irregular heart beat; or (11) other MEC staff safety concerns (NCHS, 2004b). (A detailed list of specific cardiovascular fitness exclusion criteria may be found in the appendix of NCHS, 2004). If participants' footwear was not deemed appropriate for treadmill testing (i.e., footwear lacked arch support, heel cushioning, or traction), appropriate footwear was provided (NCHS, 2004b). Treadmill tests were performed on Quinton MedTrack ST65 Treadmills in conjunction with corresponding Quinton MedTrack ST Programmable Controllers. Heart rate and blood pressure were continuously monitored using Colin STBP-780 automated electronic monitors (NCHS, 2004b).

Prior to treadmill testing, technicians collected data regarding participants' age, height, weight, and self-reported level of regular physical activity. Information regarding level of physical activity was used to assign participants a physical activity readiness (PAR) code (0-7) (NCHS, 2004b). These data were used to calculate predicted $VO_2$ max by way of the following equation:

Predicted $VO_2$ Max = 56.63 + [1.921*(PAR)] – [0.38*(age)] – [0.754*(BMI)] + [10.987*(F=0, M=1)], where BMI equals body mass index, F equals female status, and M equals male status.

Predicted $VO_2$ max was used to assign participants to one of eight 10-minute treadmill test protocols (consisting of a 2-minute warm-up, two 3-minute exercise stages, and a 2-minute cool down) (NCHS, 2004b). At the end of each exercise stage, participants reported their rate of perceived exertion (RPE; 0-20, with 20 representing maximum exertion). The goal of the 2-minute warm-up was to increase heart rate to within 60 percent of particpants' age-predicted maximum. Following the warm-up, the first 3-minute exercise stage was designed to elicit 60-80 percent of the age-predicted maximum heart rate; the second 3-minute exercise stage was designed to increase heart rate to no greater than 85 percent the age-predicted maximum heart

rate (NCHS, 2004b). Protocol changes were permitted based on participants' heart rate during the warm-up stage; that is, if heart rate was greater than 60 percent of the age-predicted max at the conclusion of the warm-up, technicians decreased the intensity of the treadmill testing to the next lowest 10-minute protocol. Similarly, if heart rate did not reach 60 percent of the age-predicted max during the warm-up period, technicians increased treadmill testing intensity to the next highest 10-minute protocol (NCHS, 2004b). Treadmill testing was prematurely stopped if the participant exhibited discomfort or distress or safety concerns arose during the testing protocol.

Based on age- and gender-specific heart rate responses to known levels of exercise intensity, $EVO_2$ max was derived for each participant using participants' heart rate responses for both levels of submaximal work (NCHS, 2004b) (please refer to the appendix of NCHS (2004) for the specific equations used to derive $EVO_2$ max from submaximal treadmill testing data).

*Dietary Intake Measures*

Dietary intake was modeled as a latent construct comprised of four manifest indicators (two cheap and two expensive measures). Cheap measures represented self-reported broad-level dietary behaviors. Expensive measures were derived from data obtained during the detailed 24-hour dietary recall, performed as part of the MEC examination.

*Cheap measures*. Self-report items presented in the Dietary Behavior Questionnaire section of the NHANES interview addressed broad-level dietary behaviors such as the frequency with which particpants consumed vegetables, fish, beans, fruit, and dairy products; ate meals at restaurants; added table salt to food; and consumed meals at community centers or schools. Two

variables measured during the Dietary Behavior Questionnaire – drybeans and dgveg – served as the two cheap measure indicators of dietary intake for the full model. **Drybeans** represented a single item phrased as, "During the past 12 months, how often per month did you eat cooked dried beans or peas, such as the foods listed on this card?" (The prompt provided participants with a list of foods such as refried, baked, kidney, black, white, and navy beans, hummus, lentils, and chickpeas). **Dgveg** represented a single item phrased as, "During the past 12 months, how often per month did you eat dark green vegetables, such as the foods listed on this card?" (The prompt provided participants with a list of foods such as broccoli, kale, romaine lettuce, dark green lettuce, spinach, and mustard greens). Response coding was continuous; however, participants were permitted to respond that they consumed the indicated foods less frequently than once a month. These responses (N=177 for drybeans and N=83 for dgveg) were manually recoded to 0.5, representing bi-monthly consumption of the indicated food categories (i.e., a manually coded response of 0.5 corresponded to an average dry bean and dark green vegetable consumption rate of once every other month). It is important to note that these two items were presented only to participants taking part in the 2001-2002 cycles of NHANES. Participants taking part in the 1999-2000 cycles of NHANES were missing for drybeans and dgveg; as a result, correlations involving these two variables were computed using a reduced sample size.

*Expensive measures*. As part of the MEC examination, all participants were eligible to complete an intensive 24-hour dietary recall. Trained NHANES dietary interviewers administered the rigorous dietary recall sessions. Interviewers were required to have a Bachelor's degree in Food and Nutrition or Home Economics; prior to administering dietary recalls, interviewers also completed a one-week training course followed by supervised interview

sessions (NCHS, 2002c). The main objective of the dietary interview was to obtain a detailed list of all foods and beverages consumed by participants in the 24 hours immediately preceding the MEC-based dietary recall.

The 24-hour dietary recall was comprised of five main steps: (a) quick list; (b) forgotten foods list; (c) time and occasion recall; (d) detail and review cycle; and (e) final review probe (NCHS, 2002c). During the quick list portion of the recall, participants were asked to rapidly provide a verbal list of all easily remembered foods and beverages consumed in the past 24 hours. Next, dietary interviewers collected data regarding easily forgotten food categories (e.g., beverages, sweets, snacks, and breads); during this step, participants were presented with a handcard probe to prompt memory recall. For each recalled food item, interviewers collected detailed information regarding the time of day the food item was consumed, the amount of food consumed, and any relevant food preparation methods (e.g., no salt added, ingredient substitutions, type of liquid used to prepare food) (NCHS, 2002c). To assist in data collection, dietary interviewers referred to an extensive series of two-dimensional and three-dimensional measuring guides (e.g., glasses, bowls, measuring cups, rulers, empty food containers) designed to enhance participant recall.

Dietary interviewers entered data with the assistance of the Main Food List (MFL) – a computerized list of more than 2,600 foods; upon entering the first few letters of a food item, the MFL provided interviewers with an alphabetized list of all food items beginning with that particular letter sequence (NCHS, 2002c). Dietary intake data collected between 1999-2001 was initially coded using the University of Texas Food Intake Analysis System (FIAS, version 3.99); dietary intake data from 2002 was initially coded using Survey Net. These two software systems

assisted in food coding and data management by translating raw dietary interview intake data into an appropriate format for subsequent nutrient coding (NCHS, 2003, 2004a). The second stage of data coding involved processing the dietary intake data, determining portion size, and assigning specific nutrient values for each food item. For the 1999-2000 NHANES surveys, dietary data was processed using the USDA 1994-98 Survey Nutrient Database (NCHS, 2003). Dietary data obtained during 2001-2002 survey were processed using the USDA Food and Nutrient Database for Dietary Studies (FNDDS; version 1.0) (NCHS, 2004a); the FNDDS contained the most current information regarding food composition values available at the time of data processing. This second stage of data processing yielded total nutrient intake values for over 51 dietary intake measures.

Total intake values for four nutrients – magnesium (mg), fiber (gm), potassium (mg), and calcium (mg) – were used to form two expensive dietary intake measures. Data reduction involved subjecting these four nutrient variables to a principal components factor analysis in which items were forced to load on a single factor. Items were ordered according to strength of factor loading; ordered loadings appear below in Table 4.

Table 4
*Principal Components Factor Loadings for Total Nutrient Intake Items*

| | |
|---|---|
| magnesium | 0.9375 |
| potassium | 0.9353 |
| calcium | 0.8334 |
| fiber | 0.8008 |

Two expensive dietary intake measures were created by forming parcels (Rogers & Schmitt, 2004) from the four total nutrient intake scores. Parcel 1, referred to as **Min1**, represented the standardized average of magnesium and fiber. Parcel 2, referred to as **Min2**,

represented the standardized average of potassium and calcium. Min1 and Min2 represented the two expensive dietary intake measures in the full model.

*Body Fat Measures*

Body fat was modeled as a latent variable comprised of four manifest indicators (two cheap measures and two expensive measures). Cheap measures were derived from self-report data obtained during the interview portion of NHANES; expensive measures were obtained during the MEC examination portion of NHANES.

*Cheap measures.* Both cheap body fat measures were formed using several self-report items. As part of the MEC interview, participants provided self-reported height and weight values, which permitted the derivation of a self-reported BMI measure; this variable was referred to as **Srbmi** in the full model (Srbmi = [(self-reported weight (kg)) / (self-reported height $(m)^2$)]). Additionally, three individual items from the Weight History section of the interview were used to derive a self-reported measure of weight status; specifically, these three items were (a) "Has a doctor or health professional ever told you that you were overweight?"; (b) "Do you consider yourself to be overweight?"; and (c) "Would you ideally like to weigh less than you do now?" Responses were coded dichotomously for each item (0 = No; 1 = Yes) and a summary measure, referred to as **Srow**, was created by summing participants' responses to the three items. If participants were missing on any of the three items, Srow was set to missing. Srbmi and Srow represented the two cheap body fat measures in the full model.

*Expensive measures*. As part of the MEC examination, trained technicians obtained anthropometric measures via standardized assessment procedures. In addition to the self-reported

values provided during the interview, subjects' height and weight data were re-assessed electronically. Standing height was electronically measured using a fixed stadiometer with vertical backboard and moveable headboard; stadiometers were connected to computers to facilitate data collection (NCHS, 2002a). Subjects were positioned so that they faced forward, feet flat, with their heels, buttocks, shoulder blades, and head in contact with the backboard. Following correct positioning, participants were instructed to inhale and to hold their breath; the headboard was positioned on top of the head with enough pressure for hair compression (NCHS, 2002a). Recorders enabled screen captures for automated data collection. Weight was assessed using Toledo digital scales; scales were connected to computers to facilitate data collection. Participants were instructed to wear underwear, a paper gown, and foam slippers (NCHS, 2002a). NHANES health technicians positioned participants appropriately on the scale; NHANES recorders were responsible for enabling screen captures for automated data collection. For participants weighing greater than 440 pounds, weight was assessed using two Seca digital scales; weight was approximated by adding the values recorded by each scale (NCHS, 2002a). Based on electronically-obtained measures of height and weight, a second measure of BMI – **Ebmi** – was calculated (Ebmi = [(electronically-measured weight (kg)) / (electronically-measured height $(m)^2$)]).

A second expensive body fat measure, body fat percentage, was derived from multi-frequency bioelectrical impedance analysis (BIA) data. BIA was performed by trained NHANES health technicians as part of the MEC examination. Participants ages 8-49 were eligible for BIA testing; exclusion criteria included the following: (1) positive pregnancy status; (2) amputations (other than fingers or toes); (3) artificial joins, pins, plates, or other metal objects in the body; (4)

pacemakers or automatic defibrillators; (5) coronary stents; or (6) weight exceeding 300 pounds (NCHS, 2002b). BIA data were collected using HYDRA ECF/ICF Bio-Impedance Spectrum Analyzers (Model 4200) (Xitron Technologies, Inc., San Diego, California). Analyzers were equipped with serial ports and connected to computers for automated data collection (NCHS, 2002b).

Participants were instructed to lie in a supine position with arms and legs comfortably separated from the trunk of the body (NCHS, 2002b). After cleaning participants' skin with rubbing alcohol, four Xitron IS4000 disposable electrodes were affixed to participants. Current-injection electrodes were placed on the dorsal surface of the right hand and right foot; voltage-detector electrodes were placed on the dorsal surface of the right wrist and right ankle (NCHS, 2002b). Electrodes were connected to the analyzers by way of MC4200 measurement cables. A small alternating current was introduced into the body via the current-injection electrodes; impedance to the flow of current was assessed by way of the voltage-detector electrodes (NCHS, 2002b). The voltage drop between the two sets of electrodes was indicative of the level of opposition to flow of electric current (i.e., impedance). Impedance was measured at 50 frequencies spaced logarithmically from 5KHz to 1 MHZ; the BIA analysis procedure lasted approximately one minute (NCHS, 2002b).

Prior to generating body fat percentage, BIA data were first used to derive participants' fat-free mass (FFM) (kg). Participants' height, weight, and BIA resistance measured at 50 kHz, were incorporated into FFM gender-specific prediction equations developed by Sun and colleagues (2003). These equations are as follows:

Males: $FFM = -10.68 + 0.65\ (stature^2/resistance) + 0.26\ (weight) + 0.02\ (resistance)$
Females: $FFM = -9.53 + 0.69\ (stature^2/resistance) + 0.17\ (weight) + 0.02\ (resistance)$

The above FFM prediction equations were developed specifically for use with large-scale epidemiological studies involving diverse samples (Sun et al., 2003). NHANES data, by design, were representative of the population, and thus, included a diverse sample of participants. FFM equations were derived by Sun and colleagues using BIA data from a sample of approximately 2,000 black and white individuals between the ages of 12-94. Cross-validation and internal validation analyses verified that these prediction equations demonstrated good precision (Sun et al., 2003). Following estimation of FFM, body fat percentage (**BFat**) was calculated using the following equation (Houtkooper et al., 1996),

$$\text{BFat} = \frac{((W_{kg} - \text{FFM}))}{W_{kg}} * 100,$$

where $W_{kg}$ represents weight in kilograms.

*Tobacco Use Measures*

Tobacco use was modeled as a latent variable comprised of three manifest indicators (two cheap measures and one expensive measure). Cheap measures were derived from self-report data obtained during the interview portion of NHANES; the expensive tobacco use measure – serum cotinine (ng/ml) – was obtained during the MEC laboratory portion of NHANES.

*Cheap measures*. Self-reported tobacco use was assessed via questionnaire items focusing on past and current cigarette, pipe, cigar, chewing tobacco, snuff, and nicotine replacement product use. Eight individual self-report measures, primarily assessing cigarette use and secondarily assessing use of any additional tobacco products, were combined to yield two cheap tobacco use measures. Four of these eight items were taken directly from the Smoking

History Questionnaire; the remaining four items represented variables created specifically for analyses presented here. Created variables were formed by merging several items assessing individual product use. For example, items that separately assessed whether or not a participant had used cigarettes, cigars, pipes, chewing tobacco, snuff, or other nicotine products during the past 5 days were summed to create a measure reflecting the total number of tobacco products a participant used during the past 5 days (0 – 6 possible tobacco product categories). Table 5, located below, outlines the eight self-report tobacco use items used to derive the two cheap tobacco use measures.

Table 5
*Self-Report Questionnaire Items Used to Generate Cheap Tobacco Use Items*

| Variable | Description | Response Range |
|---|---|---|
| daycig5 | During the past 5 days, on how many days did you smoke cigarettes? | 0 – 5 days |
| nicex5* | (Represents the total sum of the number of days participant reports using any nicotine product over the past 5 days, divided by 5. A value of 1 reflects that, on average, the participant used 1 nicotine product per day during the past 5 days. Scores greater than or less than 1 reflect that, on average, a person used more or less than 1 nicotine product, respectively, during the past 5 days). | 0 – 2.6 |
| nowsmoke | Do you now smoke cigarettes? | 0 = No<br>1 = Yes |
| allday30* | During the past 30 days, on how many days did you smoke? | 0 – 30 days |
| nic5 | During the past 5 days, did you use any product containing nicotine, including cigarettes, pipes, cigars, chewing tobacco, snuff, nicotine patches, nicotine gum, or any other product containing nicotine? | 0 = No<br>1 = Yes |
| sumnic5* | (Total sum of the number of nicotine products participant reported using during the past 5 days). | 0 – 6 products |
| cperday5 | During the past 5 days, on how many days did you smoke cigarettes? | 0 – 5 days |
| allcig30 | During the past 30 days, on the days that you smoked, how many cigarettes did you smoke per day? | 0 – 95 cigarettes |

*Note*: Tobacco use measures specifically created for current analyses are indicated with an asterisk (*).

Data reduction involved subjecting the eight self-report variables to a principal components factor analysis in which items were forced to load on a single factor. Items were ordered according to strength of factor loading; ordered loadings appear below in Table 6.

Table 6
*Principal Components Factor Loadings for Self-Report Tobacco Use Items*

| | |
|---|---|
| daycig5 | 0.9646 |
| nicex5 | 0.9551 |
| nowsmoke | 0.9539 |
| allday30 | 0.9429 |
| nic5 | 0.9225 |
| sumnic5 | 0.9152 |
| cperday5 | 0.8518 |
| allcig30 | 0.8448 |

Two cheap tobacco use measures were formed by creating parcels from the eight items described above. Initially, parcels were created using the procedure created by Rogers and Schmitt (2004); however, resulting parcels were correlated $r = 0.984$. In an effort to represent a more realistic data scenario (i.e., to obtain two smoking measures that were slightly less correlated, yet still strongly associated), a second set of parcels was formed in the following manner. Parcel 1, referred to as **TUse1**, represented the standardized average of daycig5, nicex5, nowsmoke, and allday30. Parcel 2, referred to as **TUse2**, represented the standardized average of nic5, sumnic5, cigpday5, and allcig30. While TUse1 and TUse2 remained strongly correlated ($r = 0.918$), they did not represent virtually identical measures. TUse1 and TUse2 represented the two cheap tobacco use measures in the full model.

*Expensive measure*. Trained phlebotomists collected serum cotinine data during the laboratory portion of NHANES. All participants over the age of 3 were eligible for sample collection; participants were excluded if they met the following criteria: (a) positive hemophiliac status; (b) chemotherapy treatment within the previous four weeks; (c) physical limitations

preventing blood draw (e.g., no accessible veins, severe skin rash); (d) sickness or other emergency event precluding blood draw (e.g., fainting, vomiting); or (e) adamant participant refusal (NCHS, 2001a). Prior to blood draw, phlebotomists administered several items assessing participant fasting status[4]. If participants had not adhered to fasting guidelines, but fasting guidelines would be met if blood draws occurred at the end of the MEC examination, blood draw was postponed. If fasting status would not be met by the end of the MEC examination, participants were encouraged to return to complete their blood draw at a later point in time; if this was impossible, blood draw proceeded as scheduled and fasting guideline non-adherence was recorded (NCHS, 2001a).

For participants 12 years and older, 102 mL of blood (approximately 7 tablespoons) was drawn; the amount of blood drawn during the MEC examination was considerably less than that of typical blood donations (450 mL) (NCHS, 2001a). Phlebotomists applied a tourniquet several inches above the antecubital area of the participant's left arm (the right arm was selected only if blood draw was not possible using the left arm); the intended area was cleaned twice with alcohol wipes and dried with sterile gauze. The participant was asked to make a fist while phlebotomists inserted the appropriate size needle into the selected vein. All blood draw collection tubes were filled in priority order as outlined in the laboratory procedures manual (NCHS, 2001a).

Blood was permitted to clot for 30-45 minutes at room temperature; to separate serum from plasma, samples were centrifuged at 17-25 degrees Celsius at 2,900 rpm for 15 minutes by way of Beckman GS-6R refrigerated centrifuges (NCHS, 2001b). Serum separators were used to

[4] Participants with morning MEC appointments were asked to fast for 9 hours prior to their scheduled appointment; participants scheduled for afternoon and evening appointments were asked to fast for 6 hours (NCHS, 2001a).

transfer serum to appropriate storage tubes; serum samples allocated for serum cotinine analysis were subsequently frozen. Samples were shipped frozen (by way of FedEx), packed in dry ice, to the CDC CASPIR specimen repository in Lawrenceville, GA for cotinine analyis. Data processing involved infusing cotinine samples with a labeled cotinine isotope (methyl-D3 cotinine); following equilibration of the labeled and unlabeled cotinine, cotinine was extracted from serum samples. Participants' serum cotinine concentrations were derived by calculating the ratio of unlabeled to labeled cotinine with the use of a standard curve (NCHS, 2001b).

*Outcome Measures*

Analyses involved three main outcome variables – SBP, DBP, and PP. Each outcome variable was modeled as a single item in the context of the complete model. Systolic and diastolic blood pressure data were obtained during the MEC examination and represent the average blood pressure reading over several blood pressure measurements. Pulse pressure was later calculated for the purposes of this dissertation using the following equation: PP = SBP – DBP.

MEC physicians oversaw collection of blood pressure data. Physicians were certified in blood pressure measurement in accordance with the guidelines developed by the American Heart Association (AHA) and were responsible for maintaining equipment integrity. All equipment was carefully inspected after each data collection session (NCHS, 2000). Blood pressure was measured for all participants over the age of 8. An inflation system – consisting of a latex inflation bag, Calibrated® V-Lok® cuff, inflation bulb, and Air-Flo® control valve – was attached to a wall-mounted calibrated Baumanometer® pressure gauge; Littman Classic II S.E.

Stethoscopes were used to detect Korotkoff sounds (NCHS, 2000). Physicians began the blood pressure examination by inquiring about participants' recent ingestion of food, coffee, cigarettes, and alcohol in the 30 minutes prior to testing; if participants indicated that they had consumed any of these items, physicians noted this information.

Physicians instructed participants to sit quietly in a chair for 5 minutes prior to blood pressure measurement; participants' legs were uncrossed with feet flat on the floor. The right arm was accessible, with the palm facing upward and elbow slightly bent (NCHS, 2000). Using a cosmetic pencil, physicians indicated the midpoint of the upper arm, measured between the shoulder and the elbow; at this midpoint, physicians measured the circumference of the upper arm. Using these two measurements, physicians consulted AHA guidelines to determine the appropriate inflation cuff size (child, adult, large adult, or adult thigh) (NCHS, 2000).

Three consecutive blood pressure measurements were obtained, with a minimum of 30 seconds elapsed time between readings. Between readings, manometer tubing was disconnected from the cuff to allow pressure in the cuff to return to zero (NCHS, 2000). SBP was recorded as the pressure on the manometer when the first repetitive sounds were detected via stethoscope; DBP was recorded as the pressure at which the last of these sounds disappeared. Because the manometer displayed readings in increments of 2 mm Hg, all blood pressure measurements were recorded as even numbers. If an SBP or DBP reading fell between two millimeter marks on the manometer, readings were rounded upward toward the nearest even digit (NCHS, 2000).

Average SBP and DBP were calculated using the following guidelines. If only one blood pressure reading was obtained, that reading was recorded as the average. If more than one blood pressure reading was obtained, the first reading was discarded from calculation of the average

reading. If only two blood pressure readings were obtained, the second blood pressure reading was recorded as the average (NCHS, 2000).

*Covariates*

Correlations among NHANES measures of primary interest were adjusted for several covariates. Specifically, partial correlations involving the independent and dependent measures of interest controlled for participant age, gender, ethnicity, education, household income, and martial status. Demographic data were collected during the NHANES interview portion. Age (in years) was reported as a continuous variable; gender was coded dichotomously (male or female). Participants indicated which of the following ethnic categories they most identified with: (a) non-Hispanic white; (b) non-Hispanic black; (c) Mexican American; (d) other Hispanic; and (e) other race, including multi-racial. For the purposes of this study, the Mexican American and other Hispanic categories were combined to yield an overall Hispanic category; the revised 4-category ethnicity variable was represented by three dichotomous dummy variables in later analyses. Education represented a discrete continuous variable in which participants indicated if they had completed less than a high school diploma, had earned a high school diploma (including GED), or had completed more than a high school diploma. Household income represented a discrete continuous variable; participants indicated if their annual household income fell within one of 11 monetary ranges (from $0 to over $75,000). For single-family households, household income was equivalent to total family income; for multi-family households, each family's total annual income was summed to derive total household income. Marital status was assessed via a single interview item; participants were considered married if they were married or living with a

partner. Participants were considered unmarried if they were widowed, divorced, separated, or had never married.

### *Preliminary Analyses: Generating Partial Correlation Matrices from Empirical NHANES Data*

Following data reduction and determination of the full model structure, preliminary analyses focused on quantifying the strength of association between cheap and expensive measures and the three outcome variables. Partial correlations (adjusted for age, gender, ethnicity, education, household income, and martial status) between cheap measures, expensive measures, and outcome variables were computed. In total, four partial correlation matrices (one per independent variable) were generated. All partial correlations were derived using complete case samples from NHANES data; that is, partial correlations were computed using a subsample of participants with no missing data on the cheap and expensive measure indicators for a given independent variable, the three outcome variables, or the six covariates. For this reason, the sample size used to compute correlations varied slightly across the four independent variables. Partial correlation matrices were later used as input data for all preliminary models and subsequent main analyses, including all two-method measurement design simulations.

### *Secondary Analyses: Baseline Model Performance using NHANES Data*

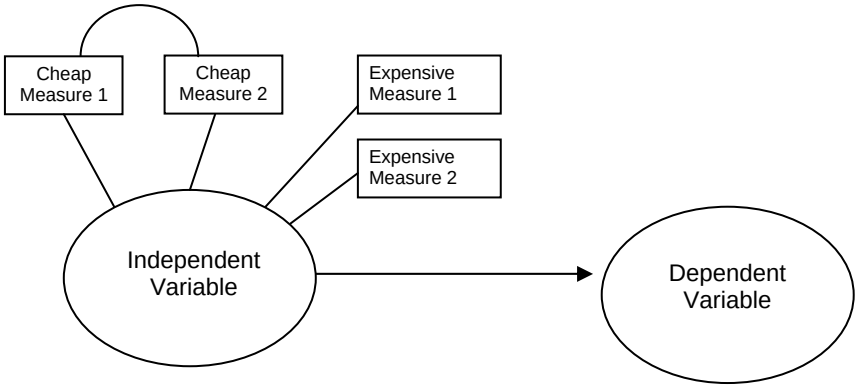
The efficiency of the two-method measurement design is assessed using the bias factor model framework; thus, secondary analyses focused on establishing plausibility of the bias factor model for each independent variable. Evaluating the viability of the bias factor model was a stepwise process in which three sequential sets of models were assessed. First, the **bias**

**parameter model** – a comparatively less rigorous form of the bias factor model – was evaluated. If models performed well at this preliminary stage, re-specification permitted the evaluation of the **Stage 1 bias factor model**. Feasibility of the Stage 1 bias factor model was of primary importance because this model template was equivalent to that used to assess performance of the two-method measurement design (i.e., the Stage 1 bias factor model template was later used for testing the two-method measurement design). Following the Stage 1 bias factor model, a final model – referred to as the **Stage 2 bias factor model** – was evaluated. The Stage 2 bias factor model served as a comparison model by which to judge the performance of the Stage 1 bias factor model. The following sections provide an overview of each category of models and discuss the relevance of comparisons among model categories.

*Bias Parameter Model*

Prior to testing the viability of the bias factor model, basic model structure was assessed using a relatively less stringent bias parameter model template. The bias parameter model differs from the bias factor model with respect to the method in which bias associated with the two cheap measures is modeled. Whereas the bias factor model specifies two sources of correlation between cheap measures (i.e., cheap measures are specified to load on a common factor as well as a bias factor), the bias parameter model controls for the bias associated with the cheap measures by estimating their residual correlation (i.e., the degree of association between cheap measures remaining after accounting for the common factor). The basic bias parameter model structure is displayed below in Figure 2.

Figure 2
*Basic Bias Parameter Model*




*Note*: The bias parameter model accounts for the bias associated with the cheap measures by estimating their item residual correlation.

Under the bias parameter model, the item residual correlation between cheap measures may be estimated as a negative value; this was helpful later when Stage 1 and Stage 2 bias factor models were estimated. As described more fully below, both versions of the bias factor model (Stage 1 and Stage 2) specified that cheap measures load on a second latent factor representing bias; for estimation purposes, equality constraints are placed on both bias factor loadings under these models. If, however, a negative item residual correlation between cheap measures is detected by the bias parameter model, bias factor loadings for the Stage 1 and Stage 2 bias factor models must be fixed to +1.0 and -1.0. Constraining bias factor loadings to be equal when the item residual correlation between cheap measures is a negative value will result in seriously poor model fit; commonly, the model will not even converge in this situation. For this reason, estimating the bias parameter model prior to the bias factor model was helpful for revealing scenarios in which sequential models may require specific modifications.

Eight bias parameter models (two models per independent variable) were estimated using LISREL 8.5. Because PP was calculated by subtracting DBP from SBP, PP was linearly dependent on the values of SBP and DBP. For this reason, it was not possible to estimate models

in which one independent variable predicted all three dependent variables. Instead, two bias parameter models per independent variable were estimated; the first model specified that the independent variable predict both SBP and DBP and the second model specified that the independent variable predict PP. The eight bias parameter models evaluated here are outlined below in Table 7.

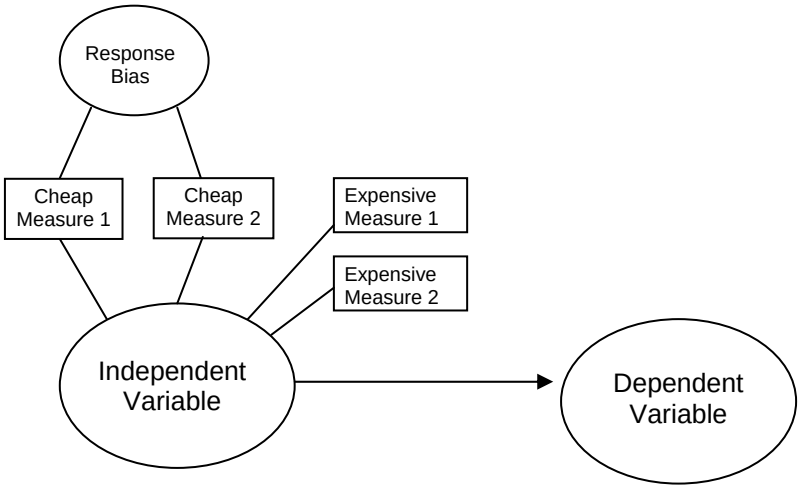
Table 7
*Eight Bias Parameter Models*

| | |
|---|---|
| (1) Physical Conditioning → SBP / DBP | (5) Body Fat → SBP / DBP |
| (2) Physical Conditioning → PP | (6) Body Fat → PP |
| (3) Dietary Intake → SBP / DBP | (7) Tobacco Use → SBP / DBP |
| (4) Dietary Intake → PP | (8) Tobacco Use → PP |

*Stage 1 Bias Factor Model*

After the eight bias parameter models were tested for adequate model performance, subsequent analyses evaluated performance of the Stage 1 bias factor model. Stage 1 bias factor models were technically equivalent to bias parameter models; both sets of models estimated the bias associated with cheap measures, albeit different bias modeling approaches were employed. The Stage 1 bias factor model specified that cheap measure indicators load on two factors: (a) a common factor (i.e., the independent variable of interest – physical conditioning, dietary intake, body fat, or smoking) and (b) a latent bias factor. Thus, instead of estimating the item residual correlation between cheap measures (like the bias parameter model), the Stage 1 bias factor model estimated the bias associated with cheap measures through the use of a second latent factor. Importantly, the bias factor was not specified to predict the dependent variables (i.e., b-weights for the bias factor predicting SBP, DBP, and PP were not estimated). The basic structure of the Stage 1 bias factor model is displayed below in Figure 3.

Figure 3
*Basic Stage 1 Bias Factor Model*




Assessing the performance of Stage 1 bias factor models was necessary prior to the main simulation analyses because the utility of the two-method measurement design depended on the ability of data to conform to the Stage 1 bias factor model. Simulation analyses were performed using the Stage 1 bias factor model template[5]. Because they were technically equivalent models, b-weights and standard errors involving the independent and dependent variables were expected to be identical for each of the eight bias parameter model and its corresponding Stage 1 bias factor model.

Because Stage 1 bias factor models did not control for the effect of bias on the outcome variables, (i.e., the bias factor was not specified to predict the dependent variables), resulting b-weights involving the independent and dependent variables were not adjusted for the effect of response bias associated with the cheap measures. To determine if the Stage 1 bias factor model produced biased parameter estimates (as a result of not estimating the pathway from the bias

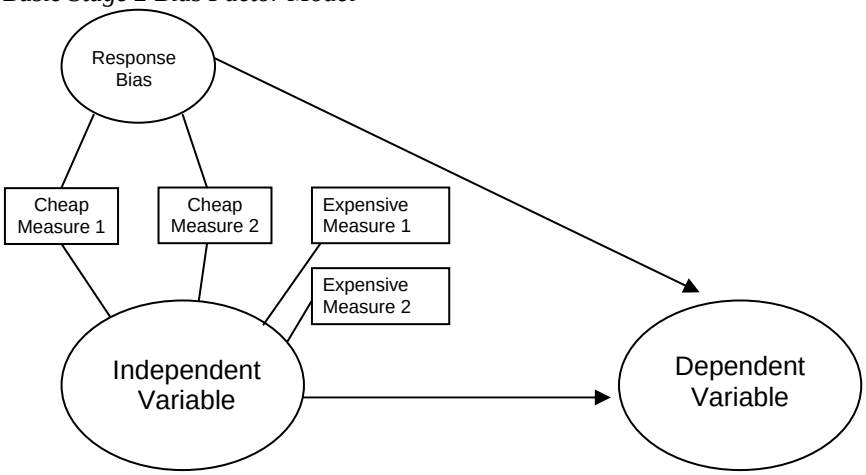
[5] Please note that the Stage 1 bias factor model displayed above in Figure 3 is equivalent to the basic bias factor model previously displayed in Figure 1.

factor to the outcomes), a second set of bias factors (referred to as Stage 2 bias factor models) was estimated.

*Stage 2 Bias Factor Model*

Stage 2 bias factor models differed from Stage 1 bias factor models in that b-weights for the bias factor predicting the outcome variables were estimated. Consequently, resulting b-weights for the common factors predicting the three blood pressure parameters were adjusted for the effect of bias on the outcome variables. The basic structure of the Stage 2 bias factor model is displayed below in Figure 4.

Figure 4
*Basic Stage 2 Bias Factor Model*




As alluded to above, estimation of Stage 2 bias factor models was important for one key reason: for the two-method measurement design to be maximally beneficial to researchers, the Stage 1 bias factor models must be plausible. Plausibility of the Stage 1 bias factor models was assessed by comparing b-weights obtained from the Stage 1 and Stage 2 bias factor models. If the b-weights involving the independent and dependent variables were approximately equivalent between the Stage 1 and Stage 2 bias factor models, it was implied that the parameter estimates

derived from Stage 1 bias factor models were unbiased. Thus, the two-method measurement design approach would be of maximal value in such instances. On the other hand, if the b-weights obtained from Stage 2 bias factor models differed appreciably from those derived from Stage 1 bias factor models, the two-method measurement design approach would be of less value because resulting parameter estimates would be considered unreliable.

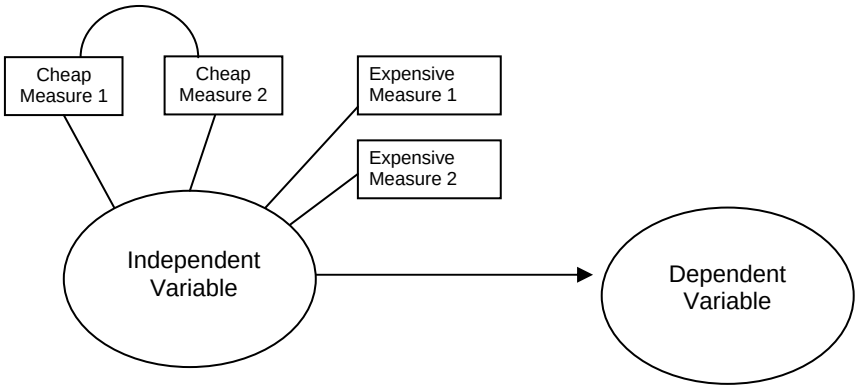
*Basic Specification of Sequential Models*

Basic specification of the bias parameter models, Stage 1 bias factor models, and Stage 2 bias factor models was relatively straightforward. For all eight models, at each of the three sequential model stages (24 models in total), the partial correlation matrix involving that particular independent variable (calculated from empirical NHANES data) was used as input data. Sample size was set to N=10,000; this sample size was selected based on the approximate size of the combined 1999-2002 NHANES sample[6].

*Bias parameter model specification.* For the eight bias parameter models (see Figure 5, below), the number of factors specified in each model represented the total number of independent and dependent variables in the model; as a result, the total number of factors depended on the particular bias parameter model being assessed. For example, in the model in which physical conditioning was specified to predict both SBP and DBP (Physical Conditioning → SBP / DBP), three factors were estimated. In the model in which dietary intake was specified to predict PP only (Dietary Intake → PP), two factors were estimated.

[6] Another N could have been employed just as well, as secondary analyses did not focus on drawing conclusions from the data; rather, secondary analyses focused on establishing model performance, which is unrelated to sample size.

Figure 5
*Basic Bias Parameter Model*

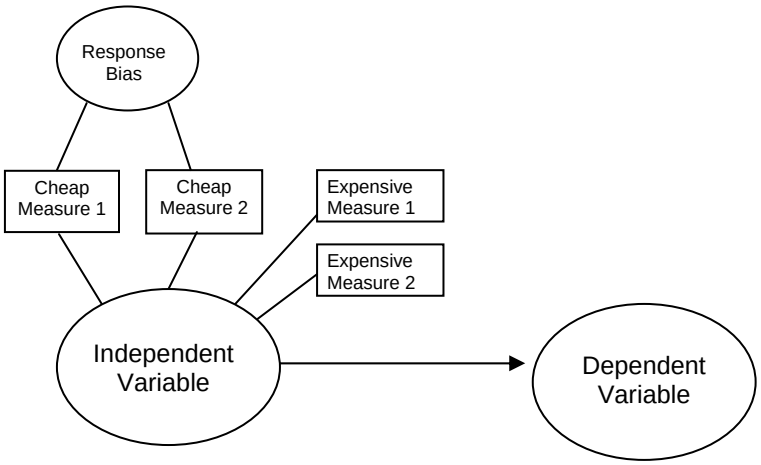



All cheap and expensive measure factor loadings for the common factor were estimated. Because SBP, DBP, and PP were single-item factors, factor loadings for these three variables were fixed at 1.0 (i.e., not estimated) in accordance with LISREL model specification guidelines. All b-weights involving the independent and dependent variables were estimated. All residual factor-level and item-level variances (except as noted above) were estimated. To model the bias associated with the two cheap measures, their residual covariance was estimated[7]. Please refer to Table 8 at the end of Chapter 3 for sample LISREL code for estimating the bias parameter models.

*Stage 1 bias factor model specification.* Basic LISREL specification of the eight Stage 1 bias factor models (see Figure 6, below) differed from that of the bias parameter models.

[7] As described above, estimating this parameter represented the defining feature of the bias parameter model.

Figure 6
*Basic Stage 1 Bias Factor Model*



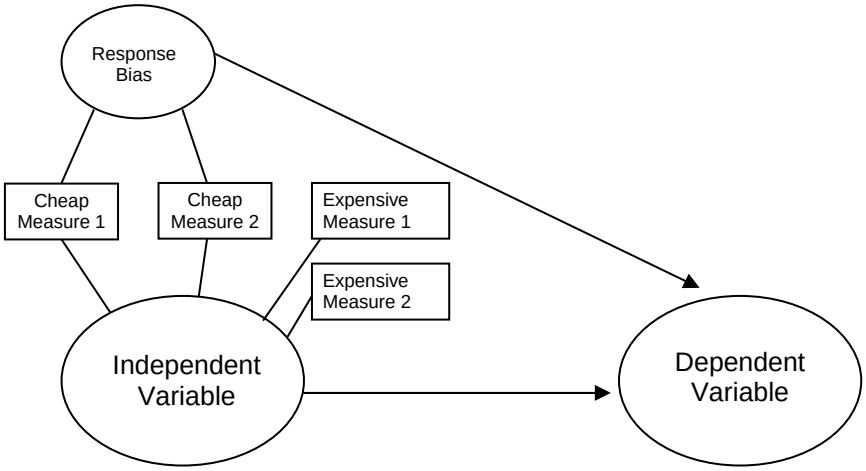

As before, the partial correlation matrices calculated from NHANES data were used as input and sample size was set to N=10,000. However, for each Stage 1 bias factor model, the total number of factors specified was one more than that specified in the corresponding bias parameter model; this additional factor represented the bias factor – the second source of covariation between cheap measures specified in the Stage 1 bias factor models. As an example, for the Stage 1 bias factor model in which physical conditioning predicted SBP and DBP, four factors were specified: physical conditioning, the bias factor, SBP, and DBP. All b-weights involving the independent and dependent variables were estimated; importantly, b-weights quantifying the association between the bias factor and outcome variables were *not* estimated in the Stage 1 bias factor models. Residual factor-level and item-level variances were estimated (except that item residuals for the DVs were fixed at 0); unlike the bias parameter model, the residual correlation between the cheap measures was not estimated.

Two sets of factor loadings were estimated for Stage 1 bias factor models. Cheap and expensive measure factor loadings were estimated for the common factor (physical conditioning,

dietary intake, body fat, or tobacco use). Additionally, cheap measure factor loadings were estimated for the bias factor; for model identification purposes, equality constraints were placed on this second set of loadings. Again, because SBP, DBP, and PP were single-item factors, factor loadings for these three variables were set to 1.0. Please refer to Table 9 at the end of Chapter 3 for sample LISREL code for estimating the Stage 1 bias factor models.

*Stage 2 bias factor model specification.* Basic Stage 2 bias factor model specification in LISREL was virtually identical to that of Stage 1 bias factor models, with one important difference. In Stage 2 bias factor models, b-weights for the bias factor predicting the outcome variables (i.e., Bias → SBP / DBP / PP) were estimated (see Figure 7, below). Please refer to Table 10 at the end of Chapter 3 for sample LISREL code for estimating the Stage 2 bias factor models.

Figure 7
*Basic Stage 2 Bias Factor Model*



*Hierarchical Process for Assessing Model Performance*

As mentioned previously, preliminary analyses followed a sequential format in which bias parameter models were tested, followed by Stage 1, and then Stage 2 bias factor models. The ability to test a model at particular stage in the sequence was conditional on acceptable model performance for all previous stages; for example, Stage 2 bias factor models were tested only after they had previously demonstrated an admissible solution and acceptable model fit at both the bias parameter model and Stage 1 bias factor model stages.

This process of assessing model admissibility and fit began by testing basic factor structure with no additional assumptions placed on the model. Model performance was judged primarily by evaluating item- and factor-level residual variances and, secondarily, by the general pattern of several fit indices. Admissibility of the models was first established by confirming that none of the resulting item- or factor-level variances were negative. After establishing model admissibility, further performance assessment involved four indices of practical fit. The chi-square ($\chi^2$) statistic – a commonly used model fit index – is sensitive to sample size; with very large Ns (such as those used in the preliminary analyses), even minor deviations from a perfect model will result in statistically significant $\chi^2$ values. Because of this, model fit was also assessed using three additional indices of practical fit: (a) NNFI (Bentler & Bonett, 1980) (also known as RHO (Tucker & Lewis, 1973)); (b) CFI (Bentler, 1990), and (c) RMSEA (Browne & Cudeck, 1993; Steiger & Lind, 1980).

If results indicated negative item-level or factor-level residual variances, assumptions were placed on the model, one at a time, and model admissibility and fit were re-assessed. Estimating a model with any number of assumptions reduces the number of parameters

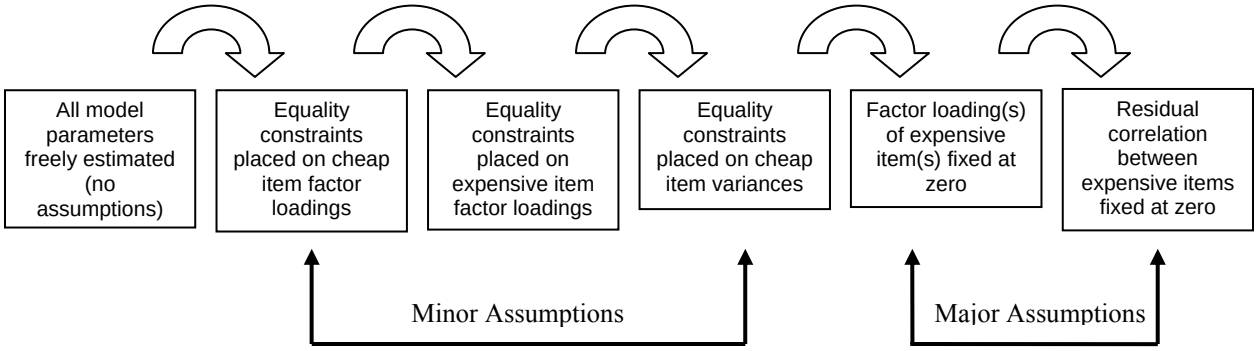
estimated; in effect, less information is asked of the model, increasing the likelihood of acceptable model performance. Assumptions varied in their degree of severity; one category included relatively minor assumptions. These assumptions did not place strong restrictions on model estimation. On the other hand, another class of assumptions was considered more severe; placing major assumptions on a model was equivalent to making very strong a priori statements about model parameters. Major model assumptions had the ability to greatly influence resulting parameter estimates.

The hierarchical process of performance analysis begins by testing models in which all appropriate parameters were estimated[8]; thus, this first model contained no additional assumptions. If necessary (i.e., if results indicated negative item-level or factor-level residual variances), a second model was estimated in which equality constraints were placed on both cheap measure factor loadings. If this second model performed unacceptably, a third model was estimated, in which equality constraints were placed on expensive measure factor loadings in addition to the equality constraints on cheap measure factor loadings. It is important to note that for this third assumption to be viable (i.e, equality constraints on expensive measure factor loadings), more than one expensive measure item must be specified for an independent variable. If necessary, a fourth model may be estimated in which equality constraints are placed on cheap measure item variances; thus, this fourth model specifies three sets of equality constraints: equality constraints on cheap measure factor loadings (assumption 1); equality constraints on expensive measure factor loadings (assumption 2); and equality constraints on cheap measure

[8] The number of appropriate parameters varied across sequential model stages (bias parameter models, Stage 1 bias factor models, and Stage 2 bias factor models) and is reviewed in the preceding section.

residual item variances (assumption 3). The hierarchical process of increasing the number of assumptions placed on a model is outlined below in Figure 8.

Figure 8
*Hierarchical Process of Assessing Bias Parameter Model Performance*




The three assumptions described above represented minor modifications (and thus, minor assumptions) to the model. However, if additional versions of the model were necessary, with increasingly more assumptions, the nature of model assumptions increased in severity. For example, if the three assumptions described above were placed on the model, and negative residual variances persisted, the next (fifth) model tested in the hierarchical sequence was one in which all factor loadings for expensive measure items were fixed at a particular value. This model assumption was relatively more severe than placing equality constraints on expensive measure factor loadings. Equality constraints permitted parameters to be estimated, albeit parameters with equality constraints were constrained to be equal. On the other hand, fixing parameters at specific values bypassed estimation altogether.

If the model continued to perform poorly (indicated by negative item- and factor-level residual variances), a sixth model was estimated in which the residual correlation between expensive measure items was fixed at zero. As with fixing expensive measure factor loadings at

specific values, assuming that the residual correlation between expensive measures was zero was a comparatively serious model modification. This assumption stated that the independent variable (the common factor) accounted for 100 percent of the variance between expensive measures, and thus had the potential to greatly overstate the construct validity of expensive measures. If model performance remained poor following this final assumption, the basic model structure was assumed unviable and alternative models were considered.

### *Main Analyses: Two-Method Measurement Design Simulations*

The first step in evaluating the performance of the two-method measurement design involved generating estimated costs of data collection. From these estimates, multiple ratios of partial to complete data were generated using an artificial budget scenario; all data ratios were computed using the same hypothetical budget of $20,000. Because data collection costs differed across independent variables, the nature of the data ratios varied as a function of the main construct of interest (physical conditioning, dietary intake, body fat, or tobacco use).

After determining a series of viable data ratios for each independent variable, main analyses used simulations to demonstrate performance of the two-method measurement design. Two-method measurement simulations followed the basic steps outlined in Graham et al. (2006). Simulations were performed for each independent variable (i.e., the two-method measurement design was applied to each individual predictor to evaluate design performance given that predictor's unique characteristics). Similar to Graham et al. (2006), parameters of primary interest included the b-weights and standard errors for each independent variable predicting SBP,

DBP, and PP. The optimal ratio (i.e., the ratio of partial to complete data that yielded the lowest standard errors) was identified for each predictor.

For comparison purposes, b-weights and standard errors were derived under financially-equivalent complete cases scenarios. The number of complete cases allowable given the $20,000 budget was determined for each predictor. This sample size was used to estimate one-group models in which the independent variables predicted SBP, DBP, and PP. For a fixed cost, the allowable complete case sample size was substantially lower than the total N allowable under a two-method measurement design. Resulting parameter estimates from complete case simulations were compared to those obtained under the optimal two-method measurement design. Lower standard errors obtained under the two-method measurement design indicated that the two-method measurement design was more efficient than the corresponding complete case model of equivalent cost.

*Estimated Data Collection Costs*

Data collection costs were estimated for all cheap and expensive measures associated with the four predictors; these estimates were then used to create a series of partial to complete data ratios that yielded a total cost equal to or slightly greater or less than $20,000. The following sections briefly review the method in which costs were estimated for cheap and expensive measures.

*Physical conditioning.* PA1 and PA2 – the two cheap physical conditioning measures – were assigned a combined total data collection cost of $10 per participant. This value was chosen based on the following logic: for incentive purposes, participants are commonly paid to complete

self-report surveys. It was reasonable to assume that $40 per participant was an appropriate monetary compensation for an hour's worth of providing self-report data. However, most participants would be able to provide data for the eight items used to generate PA1 and PA2 in 15 minutes or less. Based on the hourly compensation rate of $40, 15 minutes worth of self-report data was valued at $10.

The data collection cost for the expensive physical conditioning measure – $EVO_2$ max – was derived using estimates provided by the Penn State General Clinical Research Center (GCRC). The GCRC reported a treadmill stress test cost of approximately $250 per participant; however, clinical staff indicated that this measure may cost between $500 and $700 when performed by private physicians. To derive data ratios, $250 was used as the estimated per-participant submaximal treadmill testing fee; as a result, this cost estimate may be considered fairly conservative. Based on these estimates, the expensive measure (E) to cheap measure (C) cost differential for physical conditioning indicators was 25:1 ($250:$10), or E=25C.

Twenty-one data ratios were generated for the physical conditioning variable; ratios are displayed below in Table 11. For a total cost of $20,020 (slightly larger than the hypothetical $20,000 budget), complete case data could be collected from 77 participants. Additional ratios were generated by increasing the number of cheap measures by increments of 100, and reducing the number of expensive measures to keep costs at $20,000.

Table 11
*Physical Conditioning Data Ratios*

| | N Cheap | N Expensive | Cost Cheap | Cost Expensive | Total Cost |
|---|---|---|---|---|---|
| 1 | 77* | 77 | $ 10.00 | $ 250.00 | $ 20,020 |
| 2 | 100 | 76 | $ 10.00 | $ 250.00 | $ 20,000 |
| 3 | 200 | 72 | $ 10.00 | $ 250.00 | $ 20,000 |
| 4 | 300 | 68 | $ 10.00 | $ 250.00 | $ 20,000 |
| 5 | 400 | 64 | $ 10.00 | $ 250.00 | $ 20,000 |
| 6 | 500 | 60 | $ 10.00 | $ 250.00 | $ 20,000 |
| 7 | 600 | 56 | $ 10.00 | $ 250.00 | $ 20,000 |
| 8 | 700 | 52 | $ 10.00 | $ 250.00 | $ 20,000 |
| 9 | 800 | 48 | $ 10.00 | $ 250.00 | $ 20,000 |
| 10 | 900 | 44 | $ 10.00 | $ 250.00 | $ 20,000 |
| 11 | 1000 | 40 | $ 10.00 | $ 250.00 | $ 20,000 |
| 12 | 1100 | 36 | $ 10.00 | $ 250.00 | $ 20,000 |
| 13 | 1200 | 32 | $ 10.00 | $ 250.00 | $ 20,000 |
| 14 | 1300 | 28 | $ 10.00 | $ 250.00 | $ 20,000 |
| 15 | 1400 | 24 | $ 10.00 | $ 250.00 | $ 20,000 |
| 16 | 1500 | 20 | $ 10.00 | $ 250.00 | $ 20,000 |
| 17 | 1600 | 16 | $ 10.00 | $ 250.00 | $ 20,000 |
| 18 | 1700 | 12 | $ 10.00 | $ 250.00 | $ 20,000 |
| 19 | 1800 | 8 | $ 10.00 | $ 250.00 | $ 20,000 |
| 20 | 1900 | 4 | $ 10.00 | $ 250.00 | $ 20,000 |
| 21 | 2000 | 0 | $ 10.00 | $ 250.00 | $ 20,000 |

*Note:* * indicates the number of complete cases allowable by budget.

*Dietary intake.* Using the same logic described above, the two cheap measures of dietary intake – drybeans and dgveg – were assigned a combined total data collection cost of $10 per participant. Similarly, this cost estimate may be considered conservative as it is unlikely that participants would require 15 minutes to provide data for two self-report measures.

Data collection costs associated with the 24-hour dietary recalls were derived from estimates provided by the Penn State Dietary Assessment Center (DAC). Current figures provided by the DAC Coordinator indicated that the per-participant cost for a 24-hour dietary recall typically ranged from $80-100. However, per NHANES protocol, dietary interviewers were required to complete intensive training prior to data collection; additionally, a percentage of

data collection sessions were supervised by trained dieticians. During dietary recalls, interviewers employed preparation props, visual handcards, and measuring tools. Based on these additional fees, estimates provided by the DAC were slightly inflated to yield a 24-hour dietary recall cost of $150 per participant. Thus, the cost differential for expensive measure (E) to cheap measure (C) dietary intake indicators was 15:1 ($150:$10), or E=15C.

Twenty data ratios were derived for the dietary intake variable; ratios are displayed below in Table 12. For $20,000, complete case data may be collected from 125 participants. Additional ratios were generated by increasing the number of complete cases by increments of 100, and reducing the number of expensive measures to keep costs at $20,000.

Table 12
*Dietary Intake Data Ratios*

| | **N Cheap** | **N Expensive** | **Cost Cheap** | **Cost Expensive** | **Total Costs** |
|---|---|---|---|---|---|
| 1 | 125* | 125 | $ 10.00 | $ 150.00 | $ 20,000 |
| 2 | 200 | 120 | $ 10.00 | $ 150.00 | $ 20,000 |
| 3 | 300 | 113 | $ 10.00 | $ 150.00 | $ 19,950 |
| 4 | 400 | 106 | $ 10.00 | $ 150.00 | $ 19,900 |
| 5 | 500 | 100 | $ 10.00 | $ 150.00 | $ 20,000 |
| 6 | 600 | 93 | $ 10.00 | $ 150.00 | $ 19,950 |
| 7 | 700 | 86 | $ 10.00 | $ 150.00 | $ 19,900 |
| 8 | 800 | 80 | $ 10.00 | $ 150.00 | $ 20,000 |
| 9 | 900 | 73 | $ 10.00 | $ 150.00 | $ 19,950 |
| 10 | 1000 | 66 | $ 10.00 | $ 150.00 | $ 19,900 |
| 11 | 1100 | 60 | $ 10.00 | $ 150.00 | $ 20,000 |
| 12 | 1200 | 53 | $ 10.00 | $ 150.00 | $ 19,950 |
| 13 | 1300 | 46 | $ 10.00 | $ 150.00 | $ 19,900 |
| 14 | 1400 | 40 | $ 10.00 | $ 150.00 | $ 20,000 |
| 15 | 1500 | 33 | $ 10.00 | $ 150.00 | $ 19,950 |
| 16 | 1600 | 26 | $ 10.00 | $ 150.00 | $ 19,900 |
| 17 | 1700 | 20 | $ 10.00 | $ 150.00 | $ 20,000 |
| 18 | 1800 | 13 | $ 10.00 | $ 150.00 | $ 19,950 |
| 19 | 1900 | 6 | $ 10.00 | $ 150.00 | $ 19,900 |
| 20 | 2000 | 0 | $ 10.00 | $ 150.00 | $ 20,000 |

*Note:* * indicates the number of complete cases allowable by budget.

*Body fat.* Srbmi and Srow – the two cheap body fat measures – were assigned a total data collection cost of $10 per participant. As with the cheap dietary intake measures, this cost estimate may be considered conservative as it is based on 15 minutes worth of providing self-report data.

Local- and national-level estimates were used to derive data collections costs for Ebmi and BFat, the two expensive body fat measures. The combined cost of technician-obtained height and weight measures (used to derive Ebmi) was estimated at $20 per participant. Per NHANES protocol, two trained health technicians were responsible for pre-measurement explanations, precise positioning of participants on the scale and stadiometer, and data recording. Thus, it was reasonable to assume that participants' height and weight could be collected in approximately 15 minutes. To approximate an appropriate hourly wage for trained health technicians, recent estimates from by the U.S. Department of Labor, Bureau of Labor Statistics (BLS) were considered. In 2004, the national median income for a physician assistant was $69,410 (BLS, 2004). Assuming 40 hours of work per week, and 50 weeks of work per year, this estimate yielded a gross hourly wage of $34.71. Thus, the combined total cost for 15 minutes worth of work by two trained technicians (at the physician assistant level) was estimated at $17.36. For simplicity, this figure was rounded upward to $20. The second expensive body fat measure – BFat – was derived from BIA data. The Penn State GCRC estimated the cost of BIA at approximately $30 per participant. Therefore, the combined total cost of collecting data consistent with the two expensive body fat measures – Ebmi and BFat – was estimated at $50. Based on these estimates, the cost differential for expensive measure (E) to cheap measure (C) body fat indicators was 5:1 ($50:$10), or E=5C.

Eighteen data ratios were derived for the body fat variable; ratios are displayed below in Table 13. For $19,980, complete case data could be collected from 333 participants. Additional data ratios were generated by increasing the number of cheap measures by increments of 100, and reducing the number of expensive measures to keep costs at $20,000.

Table 13
*Body Fat Data Ratios*

| | **N Cheap** | **N Expensive** | **Cost Cheap** | **Cost Expensive** | **Total Costs** |
|---|---|---|---|---|---|
| 1 | 333* | 333 | $ 10.00 | $ 50.00 | $ 19,980 |
| 2 | 400 | 320 | $ 10.00 | $ 50.00 | $ 20,000 |
| 3 | 500 | 300 | $ 10.00 | $ 50.00 | $ 20,000 |
| 4 | 600 | 280 | $ 10.00 | $ 50.00 | $ 20,000 |
| 5 | 700 | 260 | $ 10.00 | $ 50.00 | $ 20,000 |
| 6 | 800 | 240 | $ 10.00 | $ 50.00 | $ 20,000 |
| 7 | 900 | 220 | $ 10.00 | $ 50.00 | $ 20,000 |
| 8 | 1000 | 200 | $ 10.00 | $ 50.00 | $ 20,000 |
| 9 | 1100 | 180 | $ 10.00 | $ 50.00 | $ 20,000 |
| 10 | 1200 | 160 | $ 10.00 | $ 50.00 | $ 20,000 |
| 11 | 1300 | 140 | $ 10.00 | $ 50.00 | $ 20,000 |
| 12 | 1400 | 120 | $ 10.00 | $ 50.00 | $ 20,000 |
| 13 | 1500 | 100 | $ 10.00 | $ 50.00 | $ 20,000 |
| 14 | 1600 | 80 | $ 10.00 | $ 50.00 | $ 20,000 |
| 15 | 1700 | 60 | $ 10.00 | $ 50.00 | $ 20,000 |
| 16 | 1800 | 40 | $ 10.00 | $ 50.00 | $ 20,000 |
| 17 | 1900 | 20 | $ 10.00 | $ 50.00 | $ 20,000 |
| 18 | 2000 | 0 | $ 10.00 | $ 50.00 | $ 20,000 |

*Note:* * indicates the number of complete cases allowable by budget.

*Tobacco use.* The two cheap measure tobacco use indicators – TUse1 and TUse2 – were assigned a combined total data collection cost of $10 per participant. TUse1 and TUse2 represented parcels formed from eight self-report items; it was assumed that participants would be able to provide data for all eight items in 15 minutes. Thus, this cost estimate followed the same participant hourly compensation logic described above.

Serum cotinine – the expensive tobacco use measure – was assigned a per-participant cost of $45. This figure was derived using current cost estimates provided by Salimetrics, LLC, a

locally-based company providing salivary immunoassay products and analytical services[9]. Necessary collection equipment, including oral swabs and swab storage receptacles, cost $50 per 50 participants, yielding a data collection equipment fee of $1 per participant. Following data collection, samples must be shipped frozen to laboratory facilities for analysis. Effective July 1, 2007, Salimetrics, LLC charges $20.21 per sample for duplicate salivary assay analysis. In addition to collection, storage, shipping, and analysis costs, fees associated with technicians' time (including administration of the pre-test participant questionnaire and post-test sample preparation) were summed to yield the $45 serum cotinine estimate. Thus, the cost differential for expensive measure (E) to cheap measure (C) tobacco use indicators was 4.5:1 ($45:$10), or E=4.5C.

Eighteen data ratios were derived for the tobacco use variable; ratios are displayed below in Table 14. For $19,965, complete case data could be collected from 363 participants. Additional data ratios were generated by increasing the number of cheap measures by increments of 100, and reducing the number of expensive measures to keep costs at $20,000.

[9] While cost estimates provided by Salimetrics, LLC pertained to salivary cotinine rather than serum cotinine, serum cotinine collection costs were nonetheless derived using salivary cotinine figures in the interest of realistic financial estimates. The cost of serum cotinine data collection was likely more expensive than that of salivary cotinine because trained phlebotomists, venipuncture equipment, and biohazard protection measures were required for blood draws; therefore, the per-participant serum cotinine estimate of $45 was most likely slightly undervalued.

Table 14
*Tobacco Use Data Ratios*

| | N Cheap | N Expensive | Cost Cheap | Cost Expensive | Total Costs |
|---|---|---|---|---|---|
| 1 | 363* | 363 | $ 10.00 | $ 45.00 | $ 19,965 |
| 2 | 400 | 355 | $ 10.00 | $ 45.00 | $ 19,975 |
| 3 | 500 | 333 | $ 10.00 | $ 45.00 | $ 19,985 |
| 4 | 600 | 311 | $ 10.00 | $ 45.00 | $ 19,995 |
| 5 | 700 | 288 | $ 10.00 | $ 45.00 | $ 19,960 |
| 6 | 800 | 266 | $ 10.00 | $ 45.00 | $ 19,970 |
| 7 | 900 | 244 | $ 10.00 | $ 45.00 | $ 19,980 |
| 8 | 1000 | 220 | $ 10.00 | $ 45.00 | $ 19,900 |
| 9 | 1100 | 200 | $ 10.00 | $ 45.00 | $ 20,000 |
| 10 | 1200 | 177 | $ 10.00 | $ 45.00 | $ 19,965 |
| 11 | 1300 | 155 | $ 10.00 | $ 45.00 | $ 19,975 |
| 12 | 1400 | 133 | $ 10.00 | $ 45.00 | $ 19,985 |
| 13 | 1500 | 111 | $ 10.00 | $ 45.00 | $ 19,995 |
| 14 | 1600 | 88 | $ 10.00 | $ 45.00 | $ 19,960 |
| 15 | 1700 | 66 | $ 10.00 | $ 45.00 | $ 19,970 |
| 16 | 1800 | 44 | $ 10.00 | $ 45.00 | $ 19,980 |
| 17 | 1900 | 22 | $ 10.00 | $ 45.00 | $ 19,990 |
| 18 | 2000 | 0 | $ 10.00 | $ 45.00 | $ 20,000 |

*Note:* * indicates the number of complete cases allowable by budget.

*Two-Group Analyses: Evaluating the Performance of the Two-Method Measurement Design*

Efficiency of the two-method measurement design was evaluated using the bias factor model template and a simulation framework. The multiple group procedure outlined in Allison (1987) was employed to handle missing data produced via differential data collection procedures; as a result, the multiple-group analysis strategy was appropriate for missingness consistent with the two-method measurement design. The multiple-group procedure involved estimating key parameters by simultaneously running one model in two groups; for the purposes of this dissertation, the two groups reflected the subset of cases with complete data (i.e., cheap and expensive measure data) and the subset of cases with partial data (i.e., cheap measures only).

The two-method measurement design was applied to each predictor individually. As a result, it was possible to evaluate performance of the two-method measurement design across diverse data scenarios. Diversity among independent variables stemmed from factors such as the number of expensive measures (either one or two), the strength of association between cheap and expensive measures, and the predictive value of the manifest measures on the outcome variables. However, the driving factor behind estimating the efficiency of the two-method measurement design involved the cost differential between cheap and expensive measure items. The magnitude of cost differential varied substantially across independent variables (from 4.5:1 to 25:1); this had important implications for the number of partial and complete case data allowable under a fixed budget scenario. If data collection costs for the expensive measure(s) are high, researchers are limited in the amount of complete case data they may collect. On the other hand, if the expensive measure(s) are only slightly more expensive than for the cheap measure indicators, researchers are able to collect substantially more complete case data.

Per the two-group modeling procedure outlined in Allison (1987), Group 1 represented the subset of complete data cases. Group 2 represented the subset of partial data cases (those with data for the cheap measures only). The specific sample sizes for Groups 1 and 2 were derived using the data ratio tables displayed above (Tables 11-14). Because data ratios were generated by increasing the number of cheap measures by increments of 100, for each ratio, the absolute number of cheap measures was greater than that of expensive measures. Expensive measures served as a limiting factor; the number of complete cases for a given data ratio was equivalent to the number of expensive measures. In turn, the number of partial data cases was calculated by subtracting the number of expensive measures from the number of cheap measures

(i.e., $N_{partial} = N_{cheap} - N_{expensive}$). For example, as shown above in Table 11, if cheap physical conditioning measures were provided to N=500 participants, to remain within the $20,000 budget, expensive physical conditioning data were collected from N=60 participants. Thus, according to this particular data ratio, researchers would be able to collect complete case data from 60 participants and partial data from 440 (i.e., 500 – 60) participants.

Group 1 (representing complete data cases) was modeled using empirical NHANES matrices as input; model specification was analogous to that of the Stage 1 bias factor models. Cheap and expensive measures were specified to load on the common factor; cheap measures were also specified to load on the bias factor. Model assumptions were retained from the Stage 1 bias factor model analyses; any assumptions necessary for acceptable performance of the Stage 1 bias factor models (e.g., factor loading equality constraints) were also specified for the two-group simulation models. All b-weights (and standard errors) for the common factor predicting the outcome variables were estimated; b-weights for the bias factor predicting the outcome variables were *not* estimated. With the exception of the residual item variances for the three dependent variables (modeled as one-item factors), all factor- and item-level residual variances were estimated.

Group 2 (representing partial data cases) was also modeled using empirical NHANES correlation matrices as input; however, these matrices differed from those used as input for Group 1. Because Group 1 represented complete cases, input matrices contained estimated covariances for each pair of items (i.e., a full and complete covariance matrix). However, Group 2 represented partial data cases with (theoretically) missing values for the expensive measures; as a result, all covariances involving expensive measures would not be able to be estimated. The

two-group modeling procedure addressed this by employing a modified covariance matrix for Group 2. This modified matrix had the same number of elements as the input matrix used for Group 1, with several important differences. Specifically, all covariances involving the missing expensive measure(s) were set to 0.0; additionally, the variance(s) for the missing expensive measure(s) was fixed at 1.0 (Allison, 1987). This second matrix was used as input for Group 2. For illustration purposes, examples of the Group 1 and Group 2 physical conditioning input covariance matrices are displayed below in Table 15.

Table 15
*Group 1 and Group 2 Physical Conditioning Input Correlation Matrices*

Group 1 (Complete Cases) Input Matrix

| | PA1 | PA2 | $EVO_2$ Max | SBP | DBP | PP |
|---|---|---|---|---|---|---|
| PA1 | 1.00000 | 0.69357 | 0.08859 | -0.00727 | -0.02283 | 0.01326 |
| PA2 | 0.69357 | 1.00000 | 0.07487 | -0.01322 | -0.02398 | 0.00945 |
| $EVO_2$ Max | 0.08859 | 0.07487 | 1.00000 | -0.07940 | -0.04957 | -0.02234 |
| SBP | -0.00727 | -0.01322 | -0.07940 | 1.00000 | 0.25735 | 0.58837 |
| DBP | -0.02283 | -0.02398 | -0.04957 | 0.25735 | 1.00000 | -0.62994 |
| PP | 0.01326 | 0.00945 | -0.02234 | 0.58837 | -0.62994 | 1.00000 |

Group 2 (Partial Cases) Input Matrix

| | PA1 | PA2 | $EVO_2$ Max | SBP | DBP | PP |
|---|---|---|---|---|---|---|
| PA1 | 1.00000 | 0.69357 | *0.00000* | -0.00727 | -0.02283 | 0.01326 |
| PA2 | 0.69357 | 1.00000 | *0.00000* | -0.01322 | -0.02398 | 0.00945 |
| $EVO_2$ Max | *0.00000* | *0.00000* | *1.00000* | *0.00000* | *0.00000* | *0.00000* |
| SBP | -0.00727 | -0.01322 | *0.00000* | 1.00000 | 0.25735 | 0.58837 |
| DBP | -0.02283 | -0.02398 | *0.00000* | 0.25735 | 1.00000 | -0.62994 |
| PP | 0.01326 | 0.00945 | *0.00000* | 0.58837 | -0.62994 | 1.00000 |

*Note:* For Group 2, all correlations involving the (missing) expensive measure(s) were set to zero; all variances involving the (missing) expensive measure(s) were set to 1.0.

Group 2 sample size equaled the number of partial data cases specified by a particular data ratio. Many key parameters were constrained to be invariant between Groups 1 and 2. In Group 2, all factor-level parameters (including factor variances, factor covariances, and factor

regressions) were constrained to equal those in Group 1. Also, in Group 2, where data were available, all item-level parameters (including factor loadings and item-level residual variances) were constrained to equal those in Group 1. Because the expensive measures were missing for Group 2, expensive measure factor loadings were fixed at 0.0; residual variances for the (missing) expensive measures were fixed at 1.0. When the two-group bias factor model was specified in the manner described above, the procedure produced maximum-likelihood estimates of the regression coefficients of interest (Allison, 1987; Graham, Hofer, & Piccinin, 1994; Muthen, Kaplan, & Hollis, 1987). Please refer to Table 16 at the end of Chapter 3 for sample LISREL code for the two-group bias factor model.

*Comparison Analyses: Complete Case Models*

Complete case models were utilized for two main purposes. First, complete case models estimated parameters using the sample size allowable under the $20,000 budget; these complete case Ns are represented by Ratio 1 in Tables 11-14. Estimating these models provided a backdrop against which to compare performance of the two-method measurement design; resulting parameter estimates obtained using the two-method measurement design were compared to those produced using the maximum number of complete cases allowable under $20,000.

Second, after the two-method measurement design was applied to all partial to complete data ratios, one-group models were employed to judge performance of the two-method measurement design in another context. To quantify the benefits produced by the two-method measurement design, it was necessary to determine the number of complete cases that *would* be

required to produce standard errors as efficient as those obtained using the two-method measurement design. Because the two-method measurement design capitalizes on the cost-effectiveness of collecting a substantial portion of partial data, it produces efficiency in standard errors using a smaller N than would be required to produce the same degree of efficiency under less cost-effective designs. The complete case design represents a comparatively less cost-effective design; thus, comparison analyses determined the sample size necessary to produce standard errors on par with those obtained using the two-method measurement design. The necessary complete case sample sizes represented the Effective Ns for that particular independent variable; in other words, Effective Ns reflected the sample size "effectively" produced through use of the (more cost-effective) two-method measurement design.

Complete case models were tested using one-group structural equation models and a simulation framework. One-group models were estimated because these analyses were limited to complete case data; thus, only one pattern of missingness (namely, no missingness) was represented. One-group models were equivalent to Stage 1 bias factor models; as before, empirical NHANES correlation matrices were used as input. However, whereas the Stage 1 bias factor models were previously estimated with sample sizes of 10,000, the sample sizes for complete case models were dictated by additional factors. For testing the number of complete cases allowable under the $20,000 budget, sample size was set to the number of complete cases displayed in Ratios 1 (Tables 11 – 14). Because of the variation in cost differential between cheap and expensive measures, complete case N varied across independent variables (from N=77 for physical conditioning to N=363 for tobacco use). For generating Effective Ns, sample size was increased monotonically until resulting standard errors were just below those produced by

the most efficient two-method measurement design. Please refer to Table 17 at the end of Chapter 3 for sample LISREL code for the one-group complete case bias factor model.

Table 8
*Sample LISREL Code for Bias Parameter Model (Physical Conditioning → SBP / DBP)*

```
da ni=6 no=10000 ma=cm
Labels
c1 c2 e1 o1 o2 o3
cm fu
  1.000000000000  0.693573564348  0.088592057615 -0.007269302775 -0.022826628798  0.013258230924
  0.693573564348  1.000000000000  0.074874516978 -0.013217608109 -0.023984808833  0.009446623062
  0.088592057615  0.074874516978  1.000000000000 -0.079403242125 -0.049574414437 -0.022335083371
 -0.007269302775 -0.013217608109 -0.079403242125  1.000000000000  0.257346347780  0.588374511633
 -0.022826628798 -0.023984808833 -0.049574414437  0.257346347780  1.000000000000 -0.629938629134
  0.013258230924  0.009446623062 -0.022335083371  0.588374511633 -0.629938629134  1.000000000000
Selection
c1 c2 e1 o1 o2/
mo ny=5 ne=3 ly=fu,fr ps=sy,fr te=sy,fr be=fu,fr
Le
pcon SBP DBP
pa ly
1 0 0
1 0 0
1 0 0
0 0 0
0 0 0
start 1.0 ly 4 2 ly 5 3
pa ps
0
0 1
0 1 1
ma ps
1
0 1
0 0 1
pa be
0 0 0
1 0 0
1 0 0
pa te
1
1 1
0 0 1
0 0 0 0
0 0 0 0 0
ou sc mi ad=off it=750 nd=5
```

Table 9
*Sample LISREL Code for Stage 1 Bias Factor Model (Physical Conditioning → SBP / DBP)*

```
da ni=6 no=10000 ma=cm
Labels
c1 c2 e1 o1 o2 o3
cm fu
  1.000000000000  0.693573564348  0.088592057615 -0.007269302775 -0.022826628798  0.013258230924
  0.693573564348  1.000000000000  0.074874516978 -0.013217608109 -0.023984808833  0.009446623062
  0.088592057615  0.074874516978  1.000000000000 -0.079403242125 -0.049574414437 -0.022335083371
 -0.007269302775 -0.013217608109 -0.079403242125  1.000000000000  0.257346347780  0.588374511633
 -0.022826628798 -0.023984808833 -0.049574414437  0.257346347780  1.000000000000 -0.629938629134
  0.013258230924  0.009446623062 -0.022335083371  0.588374511633 -0.629938629134  1.000000000000
Selection
c1 c2 e1 o1 o2/
mo ny=5 ne=4 ly=fu,fr ps=sy,fr te=sy,fr be=fu,fr
Le
pcon Bias SBP DBP
pa ly
1 2 0 0
1 2 0 0
1 0 0 0
0 0 0 0
0 0 0 0
start 1.0 ly 4 3 ly 5 4
pa ps
0
0 0
0 0 1
0 0 1 1
ma ps
1
0 1
0 0 1
0 0 0 1
pa be
0 0 0 0
0 0 0 0
1 0 0 0
1 0 0 0
pa te
1
0 1
0 0 1
0 0 0 0
0 0 0 0 0
ou sc mi ad=off it=750 nd=5
```

Table 10
*Sample LISREL Code for Stage 2 Bias Factor Model (Physical Conditioning → SBP / DBP)*

```
da ni=6 no=10000 ma=cm
Labels
c1 c2 e1 o1 o2 o3
cm fu
  1.000000000000  0.693573564348  0.088592057615 -0.007269302775 -0.022826628798  0.013258230924
  0.693573564348  1.000000000000  0.074874516978 -0.013217608109 -0.023984808833  0.009446623062
  0.088592057615  0.074874516978  1.000000000000 -0.079403242125 -0.049574414437 -0.022335083371
 -0.007269302775 -0.013217608109 -0.079403242125  1.000000000000  0.257346347780  0.588374511633
 -0.022826628798 -0.023984808833 -0.049574414437  0.257346347780  1.000000000000 -0.629938629134
  0.013258230924  0.009446623062 -0.022335083371  0.588374511633 -0.629938629134  1.000000000000
Selection
c1 c2 e1 o1 o2/
mo ny=5 ne=4 ly=fu,fr ps=sy,fr te=sy,fr be=fu,fr
Le
pcon Bias SBP DBP
pa ly
1 2 0 0
1 2 0 0
0 0 0 0
0 0 0 0
0 0 0 0
start 1.0 ly 4 3 ly 5 4
start 0.66127 ly 3 1
pa ps
0
0 0
0 0 1
0 0 1 1
ma ps
1
0 1
0 0 1
0 0 0 1
pa be
0 0 0 0
0 0 0 0
1 1 0 0
1 1 0 0
pa te
1
0 1
0 0 1
0 0 0 0
0 0 0 0 0
ou sc mi ad=off it=5000 nd=5
```

Table 16
*Sample LISREL Code for Two-Group Bias Factor Model (Physical Conditioning → SBP / DBP)*

```
GROUP 1: Complete Cases
da ni=6 no=56 ma=cm ng=2
Labels
c1 c2 e1 o1 o2 o3
cm fu
  1.000000000000  0.693573564348  0.088592057615 -0.007269302775 -0.022826628798  0.013258230924
  0.693573564348  1.000000000000  0.074874516978 -0.013217608109 -0.023984808833  0.009446623062
  0.088592057615  0.074874516978  1.000000000000 -0.079403242125 -0.049574414437 -0.022335083371
 -0.007269302775 -0.013217608109 -0.079403242125  1.000000000000  0.257346347780  0.588374511633
 -0.022826628798 -0.023984808833 -0.049574414437  0.257346347780  1.000000000000 -0.629938629134
  0.013258230924  0.009446623062 -0.022335083371  0.588374511633 -0.629938629134  1.000000000000
Selection
c1 c2 e1 o1 o2/
mo ny=5 ne=4 ly=fu,fr ps=sy,fr te=sy,fr be=fu,fr
Le
pcon Bias SBP DBP
pa ly
1 2 0 0
1 2 0 0
1 0 0 0
0 0 0 0
0 0 0 0
start 1.0 ly 4 3 ly 5 4
pa ps
0
0 0
0 0 1
0 0 1 1
ma ps
1
0 1
0 0 1
0 0 0 1
pa be
0 0 0 0
0 0 0 0
1 0 0 0
1 0 0 0
pa te
1
0 1
0 0 1
0 0 0 0
0 0 0 0 0
ou nd=5
```

(Table 16 Cont'd)

**GROUP 2: Partial Data Cases**

```
da ni=6 no=544 ma=cm ng=2
Labels
c1 c2 e1 o1 o2 o3
cm fu
1.000000000000  0.693573564348  0.000000000000 -0.007269302775 -0.022826628798  0.013258230924
0.693573564348  1.000000000000  0.000000000000 -0.013217608109 -0.023984808833  0.009446623062
0.000000000000  0.000000000000  1.000000000000  0.000000000000  0.000000000000  0.000000000000
-0.007269302775 -0.013217608109  0.000000000000  1.000000000000  0.257346347780  0.588374511633
-0.022826628798 -0.023984808833  0.000000000000  0.257346347780  1.000000000000 -0.629938629134
0.013258230924  0.009446623062   0.000000000000  0.588374511633 -0.629938629134  1.000000000000
Selection
c1 c2 e1 o1 o2/
mo ny=5 ne=4 ly=fu,fr ps=in te=di,fr be=in
Le
pcon Bias SBP DBP
pa ly
1 2 0 0
1 2 0 0
0 0 0 0
0 0 0 0
0 0 0 0
start 1.0 ly 4 3 ly 5 4
eq ly 1 1 1 ly 1 1
eq ly 1 2 1 ly 2 1
eq ly 1 3 1 ly 3 1
eq ly 1 1 2 ly 1 2
eq ly 1 2 2 ly 2 2
pa te
1
0 1
0 0 0
0 0 0 0
0 0 0 0 0
start 1.0 te 3 3
eq te 1 1 1 te 1 1
eq te 1 2 2 te 2 2
ou nd=5
```

Table 17

*Sample LISREL Code for One-Group Complete Case Bias Factor Model ( Physical Conditioning → SBP / DBP)*

```
da ni=6 no=77 ma=cm
Labels
c1 c2 e1 o1 o2 o3
cm fu
1.000000000000  0.693573564348  0.088592057615 -0.007269302775 -0.022826628798  0.013258230924
0.693573564348  1.000000000000  0.074874516978 -0.013217608109 -0.023984808833  0.009446623062
0.088592057615  0.074874516978  1.000000000000 -0.079403242125 -0.049574414437 -0.022335083371
 -0.007269302775 -0.013217608109 -0.079403242125  1.000000000000  0.257346347780  0.588374511633
 -0.022826628798 -0.023984808833 -0.049574414437  0.257346347780  1.000000000000 -0.629938629134
  0.013258230924  0.009446623062 -0.022335083371  0.588374511633 -0.629938629134  1.000000000000
Selection
c1 c2 e1 o1 o2/
mo ny=5 ne=4 ly=fu,fr ps=sy,fr te=sy,fr be=fu,fr
Le
pcon Bias SBP DBP
pa ly
1 2 0 0
1 2 0 0
1 0 0 0
0 0 0 0
0 0 0 0
start 1.0 ly 4 3 ly 5 4
pa ps
0
0 0
0 0 1
0 0 1 1
ma ps
1
0 1
0 0 1
0 0 0 1
pa be
0 0 0 0
0 0 0 0
1 0 0 0
1 0 0 0
pa te
1
0 1
0 0 1
0 0 0 0
0 0 0 0 0
ou sc mi ad=off it=750 nd=5
```

# CHAPTER 4

## Results

### *Preliminary Analyses: Empirical NHANES Correlation Matrices*

Partial correlations (adjusted for age, gender, race, education, household income, and marital status) are summarized below by independent variable. The sample size used to derive partial correlations is indicated for each predictor.

### *Physical Conditioning Measures*

Physical conditioning partial correlations were calculated using a complete cases sample of N=3,861; results are displayed below in Table 18. PA1 and PA2 – the two cheap physical conditioning measures – were strongly correlated ($r$ = 0.69357). PA1 and PA2 were also significantly correlated with the expensive physical conditioning measure, estimated $EVO_2$ max; however, these associations were substantially weaker than the correlation between PA1 and PA2. SBP, DBP, and PP were all significantly correlated; whereas SBP was directly associated with DBP and PP ($r$ = 0.25735 and $r$ = 0.58837, respectively), DBP was inversely associated with PP ($r$ = -0.62994). The partial correlations involving PA1, PA2, and $EVO_2$ max with the outcome variables were generally weak. PA1 and PA2 were negatively associated with SBP and DBP, though these correlations did not reach significance. PA1 and PA2 were positively associated with PP, though the correlations were non-significant. While $EVO_2$ max was significantly and inversely associated with SBP and DBP, its association with PP was non-significant.

Overall, the partial correlation matrix indicated that physical activity (i.e., PA1 and PA2) was positively associated with cardiorespiratory fitness (i.e., $EVO_2$ max). Furthermore, physical activity and cardiorespiratory fitness were associated with reduced SBP and DBP (although the association reached statistical significance only for cardiorespiratory fitness). Physical conditioning was virtually uncorrelated with PP.

Table 18
*Physical Conditioning Measures: Inter-Item Partial Correlation Matrix*

| | PA1 | PA2 | $EVO_2$ Max | SBP | DBP | PP |
|---|---|---|---|---|---|---|
| PA1 | 1.0000 | | | | | |
| PA2 | 0.69357*** | 1.0000 | | | | |
| $EVO_2$ Max | 0.08859*** | 0.07487*** | 1.0000 | | | |
| SBP | -0.00727 | -0.01322 | -0.07940*** | 1.0000 | | |
| DBP | -0.02283 | -0.02398 | -0.04957** | 0.25735*** | 1.0000 | |
| PP | 0.01326 | 0.00945 | -0.02234 | 0.58837*** | -0.62994*** | 1.0000 |

*Note:* N=3,861. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

*Dietary Intake Measures*

Dietary intake partial correlations were derived from a complete cases sample of N=3,922; results are displayed below in Table 19. The two cheap dietary indicators – drybeans and dgveg – were significantly correlated, although the strength of association was modest ($r$ = 0.16911). Conversely, the two expensive dietary indicators – Min1 and Min2 – were strongly associated ($r$ = 0.68345). Min1 was more strongly correlated with drybeans and dgveg ($r$ = 0.12157 and $r$ = 0.10511, respectively) than was Min2 ($r$ = 0.04455 and $r$ = 0.03552, respectively). SBP, DBP, and PP were significantly correlated. SBP was positively associated with DBP and PP, while DBP was inversely associated with PP. None of the four dietary intake measures was significantly associated with SBP. Furthermore, only drybeans was significantly associated with DBP and PP ($r$ = -0.05236 and $r$ = 0.04090, respectively).

Taken together, the results from the partial correlation matrix indicated that specific total nutrient intake scores (i.e., Min1 and Min2) were moderately indicative of self-reported broad-level dietary behaviors (i.e., drybeans and dgveg). In general, while the dietary intake measures were non-significantly correlated with blood pressure parameters, the trends observed in Table 19 indicated that increased consumption of dry beans, dark green vegetables, magnesium and fiber (Min1), and potassium and calcium (Min2) was associated with reduced SBP and DBP, and increased PP.

Table 19
*Dietary Intake Measures: Inter-Item Partial Correlation Matrix*

| | Drybeans | Dgveg | Min1 | Min2 | SBP | DBP | PP |
|---|---|---|---|---|---|---|---|
| Drybeans | 1.0000 | | | | | | |
| Dgveg | 0.16911*** | 1.0000 | | | | | |
| Min1 | 0.12157*** | 0.10511*** | 1.0000 | | | | |
| Min2 | 0.04455** | 0.03552* | 0.68345*** | 1.0000 | | | |
| SBP | -0.00038 | -0.02129 | -0.02481 | -0.03084 | 1.0000 | | |
| DBP | -0.05236** | -0.02733 | -0.02780 | -0.02355 | 0.28060*** | 1.0000 | |
| PP | 0.04090* | 0.00287 | 0.00016 | -0.00847 | 0.65490*** | -0.54159*** | 1.0000 |

*Note:* N=3,922. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

*Body Fat Measures*

Body fat partial correlations were calculated using a complete cases sample of N=4,024; results are displayed below in Table 20. These data reflected an interesting scenario. While the two cheap body fat measures – Srbmi and Srow – were strongly correlated ($r = 0.64628$), Srbmi was comparatively more strongly associated with both expensive body fat indicators ($r_{\text{Srbmi , Ebmi}} = 0.92228$ and $r_{\text{Srbmi , BFat}} = 0.79246$). Srow was also highly correlated with Ebmi and BFat ($r = 0.67374$ and $r = 0.64455$, respectively), though not to the same degree as Srbmi. Regarding the expensive measures, Ebmi and BFat were strongly correlated ($r = 0.86383$). SBP, DBP, and PP

were significantly correlated; SBP was positively associated with both DBP and PP while DBP was inversely associated with PP. All four body fat measures were significantly associated with SBP, DBP, and PP.

Overall, body fat was strongly associated with increased SBP, DBP, and PP; this association was most robust for SBP. All four body fat measures were highly correlated, regardless of their differential costs of data collection; BMI calculated from self-report measures (Srbmi), as well as a summary score of self-reported overweight status (Srow), were strongly correlated with BMI calculated from technician-obtained measures (Ebmi) and body fat percentage derived from BIA-prediction equations (BFat).

Table 20
*Body Fat Measures: Inter-Item Partial Correlation Matrix*

| | Srbmi | Srow | Ebmi | BFat | SBP | DBP | PP |
|---|---|---|---|---|---|---|---|
| Srbmi | 1.0000 | | | | | | |
| Srow | 0.64628*** | 1.0000 | | | | | |
| Ebmi | 0.92228*** | 0.67374*** | 1.0000 | | | | |
| BFat | 0.79246*** | 0.64455*** | 0.86383*** | 1.0000 | | | |
| SBP | 0.25923*** | 0.19061*** | 0.28502*** | 0.23552*** | 1.0000 | | |
| DBP | 0.09152*** | 0.07256*** | 0.09697*** | 0.08939*** | 0.35431*** | 1.0000 | |
| PP | 0.15305*** | 0.10799*** | 0.17144*** | 0.13367*** | 0.58932*** | -0.54669*** | 1.0000 |

*Note:* N=4,024. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

*Tobacco Use Measures*

Tobacco use partial correlations were derived using a complete cases sample of N=3,833; results are displayed below in Table 21. The two cheap tobacco use measures – TUse1 and TUse2 – were strongly correlated ($r = 0.91780$). Additionally, the expensive tobacco use measure – serum cotinine – was strongly associated with both cheap smoking measures (average $r = 0.73366$). SBP, DBP, and PP were significantly correlated; SBP was positively associated with DBP and PP while DBP was negatively associated with PP. Results indicated that all three

smoking measures were uncorrelated with SBP; however, all three smoking measures were significantly and inversely associated with DBP. TUse1 and TUse2 were significantly associated with PP; however, the association between serum cotinine and PP did not reach significance.

Taken together, the data indicated that self-reported tobacco use behaviors (TUse1 and TUse2) were consistent with a biochemical indicator of recent tobacco exposure. Furthermore, tobacco use was significantly associated with reduced DBP and increased PP. Tobacco use, as assessed by the three measures in this study, was uncorrelated with SBP.

Table 21
*Tobacco Use Measures: Inter-Item Partial Correlation Matrix*

| | TUse1 | TUse1 | Cotinine | SBP | DBP | PP |
|---|---|---|---|---|---|---|
| TUse1 | 1.0000 | | | | | |
| TUse2 | 0.91780*** | 1.0000 | | | | |
| Cotinine | 0.73185*** | 0.73548*** | 1.0000 | | | |
| SBP | -0.01542 | -0.01211 | -0.02858 | 1.0000 | | |
| DBP | -0.08079*** | -0.08076*** | -0.06784*** | 0.41267*** | 1.0000 | |
| PP | 0.04977** | 0.05309** | 0.02606 | 0.67622*** | -0.39199*** | 1.0000 |

*Note:* N=3,833. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

Table 22, located below, summarizes the basic factor structure and pattern of partial correlations for each independent variable.

Table 22
*Summary of Basic Factor Structure and Pattern of Partial Correlations, by Independent Variable*

| | N Cheap Measures | N Expensive Measures | $r_{cheap,\ cheap}$ | $r_{ex,\ ex}$ | $r_{cheap,\ ex}$ | $r_{IV,\ DVs}$ |
|---|---|---|---|---|---|---|
| Physical Conditioning | 2 | 1 | Moderate ($r$ = 0.694) | N / A | Low ($r$ = 0.082) | SBP: $r$ = -0.033<br>DBP: $r$ = -0.032<br>PP: $r$ = 0.015[a] |
| Dietary Intake | 2 | 2 | Low ($r$ = 0.169) | Moderate ($r$ = 0.683) | Low ($r$ = 0.077) | SBP: $r$ = -0.019<br>DBP: $r$ = -0.033<br>PP: $r$ = 0.013[a] |
| Body Fat | 2 | 2 | Moderate ($r$ = 0.646) | High ($r$ = 0.864) | High ($r$ = 0.758) | SBP: $r$ = 0.243<br>DBP: $r$ = 0.088<br>PP: $r$ = 0.142 |
| Tobacco Use | 2 | 1 | High ($r$ = 0.918) | N / A | High ($r$ = 0.734) | SBP: $r$ = -0.019<br>DBP: $r$ = -0.076<br>PP: $r$ = 0.043 |

*Note:* $r_{cheap,\ cheap}$ refers to the strength of correlation between the two cheap measure indicators; $r_{ex,\ ex}$ refers to the strength of correlation between the expensive measure indicators (if two expensive measure indicators were present in the model); $r_{cheap,\ ex}$ represents the average correlation between cheap and expensive measures for a given independent variable; and $r_{IV,DVs}$ represents the average correlation between manifest measures for a given independent variable and the three outcome variables. [a] reflects the absolute value of the average correlation; in such instances, manifest indicators were correlated with the outcome variables in opposite directions (i.e., a combination of negative and positive item-outcome correlations).

### *Secondary Analyses: Bias Parameter Models, Stage 1 and Stage 2 Bias Factor Models*

Secondary analyses focused on establishing plausibility of the bias factor model for each independent variable as the two-method measurement design is assessed using a bias factor model framework. Admissibility of the Stage 1 bias factor model was of primary importance because the Stage 1 bias factor model template was later used for testing the two-method measurement design. Analyses began by estimating the series of eight bias parameter models; assuming models performed well at this preliminary stage, models were re-specified to evaluate

performance of the Stage 1 bias factor models. Stage 2 bias factor models served as comparison models by which to judge the performance of the Stage 1 bias factor model.

Preliminary results from two bias parameter models prompted model re-specification. For six of the eight bias parameter models ((a) Physical Conditioning → SBP / DBP; (b) Dietary Intake → SBP / DBP; (c) Dietary Intake → PP; (d) Body Fat → SBP / DBP; (e) Body Fat → PP; and (f) Tobacco Use → SBP / DBP), no negative item- or factor-level residual variances were detected and the overall pattern of fit indices indicated acceptable performance. However, initial analyses indicated that the Physical Conditioning → PP bias parameter model failed to converge after 5,000 iterations. Additionally, preliminary results from the Tobacco Use → PP bias parameter model indicated a negative item-level residual variance for TUse1. These findings were attributed to the comparatively weak correlations between physical conditioning and PP and tobacco use and PP; whereas physical conditioning and tobacco use were more strongly associated with SBP and DBP, they were virtually uncorrelated with PP. To add stability, these two bias parameter models were re-estimated with physical conditioning and tobacco use specified to predict both SBP and PP (i.e., Physical Conditioning → SBP / PP; Tobacco Use → SBP / PP). This modification was retained for the Stage 1 and Stage 2 bias factor models, as well. The following sections summarize the performance of the bias parameter models, Stage 1 bias factor models, and Stage 2 bias factor models by independent variable. Tables allow for the comparison of parameter values across sequential models.

*Physical Conditioning*

Two physical conditioning models (Physical Conditioning → SBP / DBP and Physical Conditioning → SBP / PP) were evaluated at each stage of the sequential model series. Results are displayed below in Table 23.

Table 23
*Physical Conditioning: Sequential Model Results*

| | | BPM | Stage 1 BFM | Stage 2 BFM |
|---|---|---|---|---|
| PC Factor Loadings | | | | |
| | PA1 | 0.133 | 0.133 | 0.133 |
| | PA2 | 0.114 | 0.114 | 0.114 |
| | $EVO_2$ Max | 0.661 | 0.661 | 0.661± |
| Bias Factor Loadings | | | | |
| | PA1 | ----- | 0.824 | 0.824 |
| | PA2 | ----- | 0.824 | 0.824 |
| B-weights | | | | |
| | PC → SBP | -0.119 | -0.119 | -0.120 |
| | PC → DBP | -0.077 | -0.077 | -0.075 |
| | PC → PP | -0.031 | -0.031 | -0.034 |
| Model Fit Indices | | | | |
| | Chi-Square (df) | 4.364 (3) | 4.364 (3) | 1.162 (2) |
| | RMSEA | 0.007 | 0.007 | 0.000 |
| | CFI | 0.999 | 0.999 | 1.000 |
| | NNFI | 0.999 | 0.999 | 1.000 |

*Note:* PC Factor Loadings=physical conditioning common factor loadings; BPM=bias parameter model; BFM=bias factor model; ± refers to a fixed parameter value

*Bias parameter models*. The two physical conditioning bias parameter models required no additional model assumptions. $EVO_2$ max loaded comparatively more highly on the physical conditioning common factor than did PA1 and PA2; thus, $EVO_2$ max dominated the physical conditioning factor relative to the two self-report physical activity measures. Resulting b-weights indicated that the association between physical conditioning and SBP fell under the category of "small effects" (i.e., $\rho = .10$) in Cohen's (1977) terms. By Cohen's (1977) standards, the

associations between physical conditioning and DBP and PP failed to reach the "small effect" threshold. The general pattern of fit indices indicated excellent model fit.

*Stage 1 bias factor models*. Initial assessment of the two physical conditioning Stage 1 bias factor models produced no negative item- or factor-level residual variances; therefore, no additional model modifications were necessary. Common factor loadings for PA1, PA2, and $EVO_2$ max were identical to those obtained from the bias parameter models. Bias factor loadings for PA1 and PA2 (constrained to be equal) were estimated at 0.824, indicating that a substantial correlation remained between PA1 and PA2 after taking into account the common physical conditioning factor. Resulting b-weights were identical to those obtained from the bias parameter models. Fit indices indicated excellent model fit for both physical conditioning Stage 1 bias factor models.

*Stage 2 bias factor models*. Preliminary results indicated that when the bias factor was specified to predict the three outcome variables, neither of the two physical conditioning Stage 2 bias factor models converged. However, after fixing the $EVO_2$ max common factor loading at 0.661 (the factor loading estimated by the bias parameter and Stage 1 bias factor models), both Stage 2 bias factor models performed well. By fixing the $EVO_2$ factor loading at 0.661, stability was added to the model by way of a reduced model estimation load.

Common factor loadings for PA1 and PA2 were identical to those obtained from the bias parameter models and Stage 1 bias factor models. Bias factor loadings for the cheap measures were identical to those obtained from the Stage 1 bias factor models. Again, the bias factor loadings were estimated at 0.824, indicating that a large degree of residual correlation between PA1 and PA2 was unaccounted for by the physical conditioning common factor.

After controlling for the effect of the bias factor on the outcome variables (i.e., estimating the Bias → SBP / DBP / PP b-weights), the b-weights for physical conditioning predicting SBP and DBP remained virtually unchanged. Again, the overall pattern of fit indices indicated excellent fit of the Stage 2 bias factor models.

Parameter estimates from the Stage 2 bias factor models completely controlled for the bias associated with the cheap physical conditioning measures; thus, resulting b-weights from the Stage 2 bias factor models were completely unbiased. Analyses revealed that parameter estimates from Stage 1 and Stage 2 bias factor model were virtually identical, indicating that the Stage 1 bias factor model also produced unbiased parameter estimates. Based on these results, it was determined that the physical conditioning Stage 1 bias factor model was a viable template for subsequent evaluation of the two-method measurement design.

*Dietary Intake*

Sequential model results pertaining to the two dietary intake models (Dietary Intake → SBP / DBP and Dietary Intake → PP) are found below in Table 24.

Table 24
*Dietary Intake: Sequential Model Results*

| | | BPM | Stage 1 BFM | Stage 2 BFM |
|---|---|---|---|---|
| DI Factor Loadings | | | | |
| | drybeans | 0.094 | 0.094 | 0.094 |
| | dgveg | 0.094 | 0.094 | 0.094 |
| | Min2 | 0.827 | 0.827 | 0.827 |
| | Min2 | 0.827 | 0.827 | 0.827 |
| Bias Factor Loadings | | | | |
| | drybeans | ----- | 0.400 | 0.400 |
| | dgveg | ----- | 0.400 | 0.400 |
| B-weights | | | | |
| | DI → SBP | -0.034 | -0.034 | -0.034 |
| | DI → DBP | -0.032 | -0.032 | -0.031 |
| | DI → PP | -0.004 | -0.004 | -0.005 |
| Model Fit Indices | | | | |
| | Chi-Square (df) | 181.530 (9) | 165.907 (6) | 157.233 (5) |
| | RMSEA | 0.044 | 0.052 | 0.055 |
| | CFI | 0.978 | 0.977 | 0.978 |
| | NNFI | 0.963 | 0.963 | 0.958 |

*Note:* DI=dietary intake; BPM=bias parameter model; BFM=bias factor model

*Bias parameter models*. Initial estimation of the dietary intake bias parameter models resulted in a negative residual variance for Min1. Equality constraints were placed on cheap measure factor loadings and the models were re-assessed; however, the negative residual variance persisted. When a second set of equality constraints was placed on the expensive measure factor loadings, no negative residual variances were detected. Thus, the dietary intake bias parameter models required two additional minor assumptions to improve model performance to an acceptable level.

Results indicated that the two expensive measures (Min1 and Min2) loaded more highly on the common factor than did the cheap measures (drybeans and dgveg). Thus, the expensive

measures dominated the dietary intake common factor relative to the cheap measures. Resulting b-weights revealed that the associations between dietary intake and SBP, DBP, and PP, in the context of Cohen (1977), failed to reach the "small effect" threshold (i.e., $\rho = .10$). Fit indices indicated overall good model fit.

*Stage 1 bias factor models.* The two sets of equality constraints specified in the bias parameter models were retained for the Stage 1 bias factor models. Additionally, as with all bias factor models (both Stage 1 and 2), equality constraints were placed on cheap measure bias factor loadings for identification purposes. Thus, the Stage 1 bias factor models incorporated three sets of factor loading equality constraints ((a) cheap measure common factor loadings; (b) expensive measure common factor loadings; and (c) cheap measure bias factor loadings).

Common factor loadings for all four dietary intake measures were identical to those obtained from the bias parameter models. Bias factor loadings were estimated at 0.40, indicating that a rather substantial portion of variance between drybeans and dgveg was unaccounted for by the dietary intake factor. Resulting b-weights were identical to those estimated by the bias parameter models. The overall pattern of fit indices indicated good model fit.

*Stage 2 bias factor models.* All four common factor loadings were estimated to be identical to those obtained from the Stage 1 bias factor models; similarly, estimated bias factor loadings were identical to those from the Stage 1 bias factor models. After controlling for the effect of the bias factor on the outcome measures, the pattern of b-weights involving dietary intake and SBP, DBP, and PP was virtually unchanged from Stage 1 bias factor model results. Model fit remained acceptable, as judged by the pattern of fit indices.

Because the Stage 1 and 2 bias factor models yielded virtually identical results, the Stage 1 bias factor models were found to be unbiased. Thus, the dietary intake Stage 1 bias factor model was a viable template for subsequent evaluation of the two-method measurement design.

*Body Fat*

Two body fat models (Body Fat → SBP / DBP and Body Fat → PP) were evaluated at each stage of the sequential model series. Results are displayed below in Table 25.

Table 25
*Body Fat: Sequential Model Results*

| | | BPM | Stage 1 BFM | Stage 2 BFM |
|---|---|---|---|---|
| BF Factor Loadings | | | | |
| | Srbmi | 0.925 | 0.925 | 0.925 |
| | Srow | 0.677 | 0.677 | 0.677 |
| | Ebmi | 0.997 | 0.997 | 0.997 |
| | BFat | 0.867 | 0.867 | 0.867 |
| Bias Factor Loadings | | | | |
| | Srbmi | ----- | -0.144 | -0.142 |
| | Srow | ----- | -0.144 | -0.142 |
| B-weights | | | | |
| | BF → SBP | 0.285 | 0.285 | 0.285 |
| | BF → DBP | 0.097 | 0.097 | 0.097 |
| | BF → PP | 0.171 | 0.171 | 0.172 |
| Model Fit Indices | | | | |
| | Chi-Square (df) | 306.285 (7) | 306.285 (7) | 301.687 (3) |
| | RMSEA | 0.065 | 0.065 | 0.099 |
| | CFI | 0.993 | 0.993 | 0.992 |
| | NNFI | 0.984 | 0.984 | 0.978 |

*Note:* BF=body fat; BPM=bias parameter model; BFM=bias factor model

*Bias parameter models.* Model specification for the body fat bias parameter models required no additional assumptions. Both cheap body fat measures (Srbmi and Srow), as well as both expensive body fat measures (Ebmi and BFat), loaded highly on the common factor (factor

loading range: 0.677 – 0.997). Interestingly, self-reported BMI (Srbmi; a cheap measure) loaded more highly on the body fat factor than did BIA-derived body fat percentage (BFat; an expensive measure). The b-weights indicated that the associations between body fat and SBP approached what Cohen (1977) termed a "medium effect" (i.e., ρ = .30). On the other hand, the associations between body fat and DBP and PP represented "small effects" (Cohen, 1977). The overall pattern of fit indices indicated excellent model fit.

*Stage 1 bias factor models*. Baseline assessment of the Stage 1 bias factor models resulted in no negative item- or factor-level residual variances; therefore, no additional model modifications were necessary. The four common factor loadings were estimated to be identical to those obtained from the bias parameter models. Cheap measures loaded on the bias factor with loadings of -0.144, indicating that only a relatively small proportion of variance between Srbmi and Srow was unaccounted for by the common body fat factor. Resulting b-weights were identical to those obtained from the bias parameter models. The overall pattern of fit indices reflected good model fit.

*Stage 2 bias factor models*. All common factor and bias factor loadings were estimated to be identical to those obtained from previous bias parameter and Stage 1 bias factor models. Resulting b-weights and standard errors quantifying the effect of body fat on the three outcomes were identical across both Stage 1 and Stage 2 bias factor models. Again, the overall pattern of fit indices indicated good model fit.

Stage 1 and Stage 2 bias factor models produced virtually identical b-weights and standard errors for the pathways involving the independent and dependent variables; thus, the Stage 1 bias factor model demonstrated the ability to yield unbiased parameter estimates. The

Stage 1 body fat bias factor model was found to be a viable template for subsequent application of the two-method measurement design.

*Tobacco Use*

Sequential model results pertaining to the two tobacco use models (Tobacco Use → SBP / DBP and Tobacco Use → SBP / PP) are found below in Table 26.

Table 26
*Tobacco Use: Sequential Model Results*

| | | BPM | Stage 1 BFM | Stage 2 BFM |
|---|---|---|---|---|
| TU Factor Loadings | | | | |
| | TUse1 | 0.967 | 0.967 | 0.967 |
| | TUse2 | 0.967 | 0.967 | 0.967 |
| | Serum Cotinine | 0.759 | 0.759 | 0.759 |
| Bias Factor Loadings | | | | |
| | TUse1 | ----- | 1.000± | 1.000± |
| | TUse2 | ----- | -1.000± | -1.000± |
| B-weights | | | | |
| | TU → SBP | -0.015 | -0.015 | -0.015 |
| | TU → DBP | -0.084 | -0.084 | -0.084 |
| | TU → PP | 0.053 | 0.053 | 0.053 |
| Model Fit Indices | | | | |
| | Chi-Square (df) | 9.021 (4) | 9.021 (4) | 8.208 (2) |
| | RMSEA | 0.011 | 0.011 | 0.018 |
| | CFI | 0.999 | 0.999 | 1.000 |
| | NNFI | 0.999 | 0.999 | 0.999 |

*Note:* TU=Tobacco Use; BPM=bias parameter model; BFM=bias factor model; ± refers to a fixed parameter value

*Bias parameter models*. Initial results from the tobacco use bias parameter models indicated acceptable model performance. However, as described below, Stage 1 and Stage 2 bias factor models required equality constraints for cheap measure common factor loadings. While the bias parameter models did not require this set of equality constraints, they were nevertheless placed on the models to establish consistency of model specification across the sequential model

series. Thus, the tobacco use bias parameter models were re-specified to constrain cheap measure factor loadings on the common factor to be equal.[10]

All three tobacco use indicators (TUse1, TUse2, and serum cotinine) loaded highly on the common factor; however, the two self-report measures loaded more highly than did serum cotinine. Resulting b-weights revealed that the association between tobacco use and DBP represented a "small effect"; however, the associations between tobacco use and SBP and PP did not reach the "small effect" threshold (Cohen, 1977). Fit indices revealed excellent model fit.

*Stage 1 bias factor models*. Initial assessments indicated that the Stage 1 bias factor models failed to converge. Further examination of previous results (those from the bias parameter models) revealed an interesting finding: the residual covariance between the two cheap measures (TUse1 and TUse2) was negative. Stage 1 and 2 bias factor models estimate the residual covariation between cheap measures by specifying that they load on a latent bias factor. Furthermore, for model identification purposes, the bias factor loadings for cheap measures are constrained to be equal. However, because the residual covariation between TUse1 and TUse2 (i.e., the residual association unaccounted for by the tobacco use common factor) was negative, bias factor loadings for TUse1 and TUse2 could not be constrained to be equal. To resolve this issue, Stage 1 bias factor models were re-specified; specifically, the TUse1 bias factor loading was fixed at 1.00, the TUse2 bias factor loading was fixed at -1.00, and the bias factor variance was estimated. This re-specification permitted the estimation of the negative residual correlation between the two cheap measures. Following this modification, the Tobacco Use → SBP / PP

[10] This model modification may be considered especially minor in the case of the two tobacco use bias parameter models; even without equality constraints, cheap measure factor loadings were very similar to one another. Specifically, without equality constraints on the bias parameter models, common factor loadings for TUse1 and TUse2 were estimated at 0.965 and 0.970, respectively.

Stage 1 bias factor model performed well; however, the Tobacco Use → SBP / DBP Stage 1 bias factor model still failed to converge. At this point, equality constraints were placed on the cheap measure common factor loadings. Following this modification, the Tobacco Use → SBP / DBP Stage 1 bias factor model converged. For model equivalency purposes, cheap measure common factor equality constraints were retained for all tobacco use models in the sequential model series.

All common factor loadings were estimated to be identical to those obtained from the bias parameter models. Furthermore, b-weights resulting from the Stage 1 bias factor models were identical to those obtained from the bias parameter models. Fit indices suggested excellent model fit.

*Stage 2 bias factor models*. All model assumptions from the Stage 1 bias factor models were retained; that is, equality constraints were placed on cheap measure common factor loadings; the bias factor loadings for TUse1 and TUse2 were fixed at values of 1.00 and -1.00, respectively; and the bias factor variance was estimated. Estimated tobacco use factor loadings were identical to those obtained from the bias parameter models and Stage 1 bias factor models. Additionally, b-weights and standard errors pertaining to tobacco use, SBP, DBP, and PP were identical to those from previous models. Fit indices revealed excellent model fit.

Taken together, findings from the Stage 2 bias factor models indicated that the Stage 1 bias factor models produced unbiased b-weights involving tobacco use, SBP, DBP, and PP. The tobacco use Stage 1 bias factor model represented a viable model template for simulations evaluating the performance of the two-method measurement design.

## *Main Analyses: Evaluating the Performance of the Two-Method Measurement Design Using Two-Group Structural Equation Modeling Approach*

Using a simulation framework and a Stage 1 bias factor model template, the two-method measurement design was applied to each predictor. The two-group modeling procedure (Allison, 1987; Graham et al., 1994; Muthen et al., 1987) was employed to permit parameter estimation among data with missingness patterns consistent with the two-method measurement design (i.e., a substantial portion of partial data cases and a smaller portion of complete data cases). For each predictor, the two-method measurement design was applied to all data ratios generated under the hypothetical $20,000 budget. Results are discussed below.

### *Presentation of Results*

Please note that the presentation of results deviates from that of previous sections. For two of the four predictors – body fat and tobacco use – application of the two-method measurement design produced statistical advantages beyond those yielded by the financially-equivalent complete cases models; results pertaining to these predictors are presented first. However, application of the two-method measurement design was comparatively less effective for the physical conditioning and dietary intake variables; results pertaining to these two predictors – as well as a discussion of explanatory factors – follow after.

### *Body fat*

Eighteen ratios of partial to complete data cases were generated for the body fat predictor; all ratios remained within the hypothetical budget of $20,000. Ratio 1, corresponding

to the number of complete cases allowable under the budget (N=333), was tested using a one-group design. The two-method measurement design was applied to ratios 2-17 using the two-group modeling procedure; resulting b-weights and standard errors are displayed below in Table 27.

Table 27
*Body Fat Parameter Estimates Obtained using the Two-Method Measurement Design*

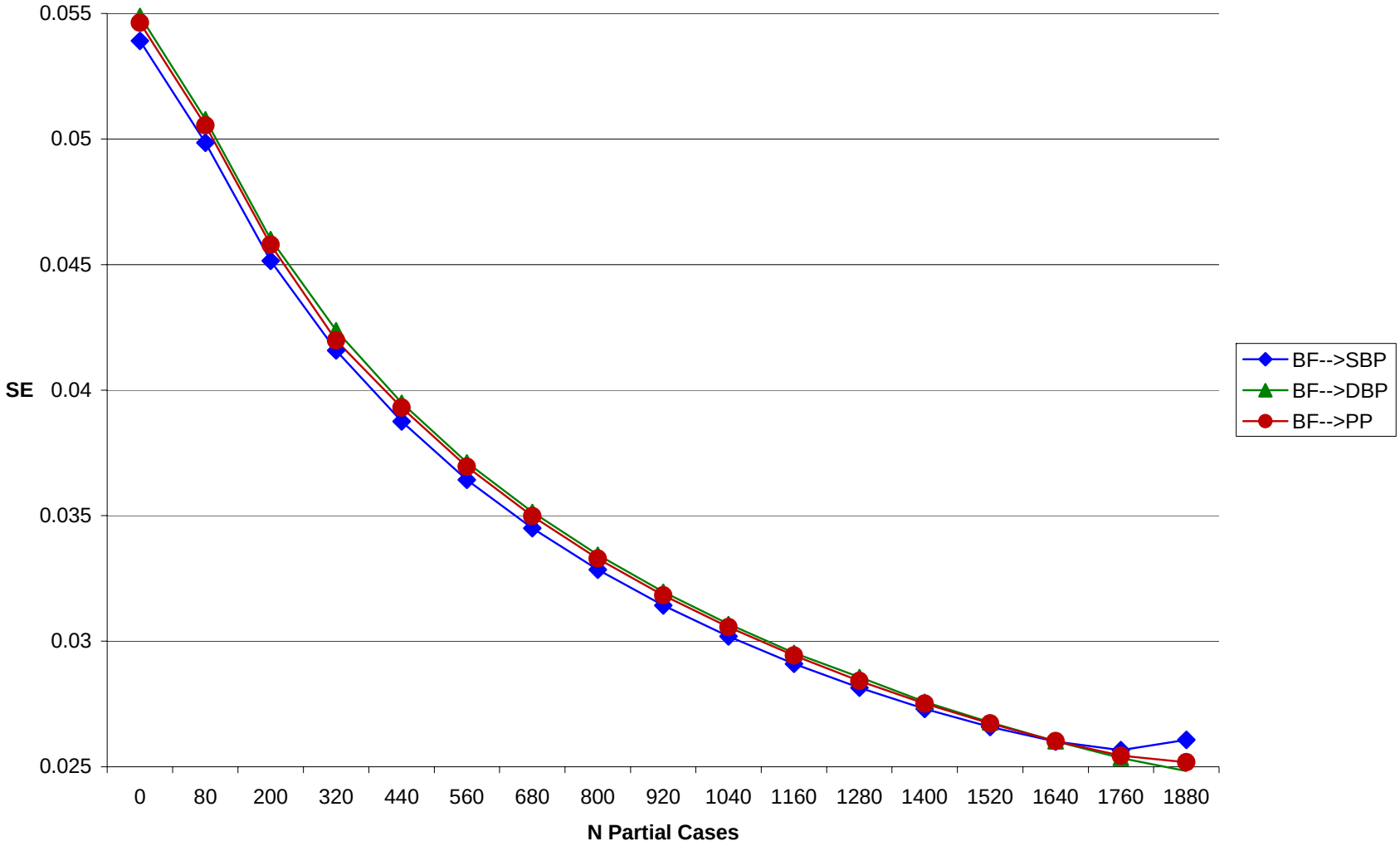
| | N Cheap | N Expensive | Cost Cheap | Cost Expensive | Total Cost | N Partial | N Complete | BF→ SBP | | BF→ DBP | | BF→ PP | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | b | SE | b | SE | b | SE |
| 1 | 333 | 333 | $ 10.00 | $ 50.00 | $ 19,980 | 0 | 333 | 0.28540 | 0.05391 | 0.09741 | 0.05490 | 0.17139 | 0.05463 |
| 2 | 400 | 320 | $ 10.00 | $ 50.00 | $ 20,000 | 80 | 320 | 0.28455 | 0.04985 | 0.09784 | 0.05078 | 0.17024 | 0.05054 |
| 3 | 500 | 300 | $ 10.00 | $ 50.00 | $ 20,000 | 200 | 300 | 0.28363 | 0.04515 | 0.09831 | 0.04601 | 0.16901 | 0.04579 |
| 4 | 600 | 280 | $ 10.00 | $ 50.00 | $ 20,000 | 320 | 280 | 0.28299 | 0.04158 | 0.09863 | 0.04238 | 0.16859 | 0.04198 |
| 5 | 700 | 260 | $ 10.00 | $ 50.00 | $ 20,000 | 440 | 260 | 0.28251 | 0.03875 | 0.09887 | 0.03949 | 0.16752 | 0.03930 |
| 6 | 800 | 240 | $ 10.00 | $ 50.00 | $ 20,000 | 560 | 240 | 0.28215 | 0.03643 | 0.09905 | 0.03712 | 0.16704 | 0.03695 |
| 7 | 900 | 220 | $ 10.00 | $ 50.00 | $ 20,000 | 680 | 220 | 0.28186 | 0.03450 | 0.09920 | 0.03514 | 0.16665 | 0.03498 |
| 8 | 1000 | 200 | $ 10.00 | $ 50.00 | $ 20,000 | 800 | 200 | 0.28162 | 0.03285 | 0.09931 | 0.03344 | 0.16634 | 0.03329 |
| 9 | 1100 | 180 | $ 10.00 | $ 50.00 | $ 20,000 | 920 | 180 | 0.28142 | 0.03143 | 0.09941 | 0.03197 | 0.16608 | 0.03183 |
| 10 | 1200 | 160 | $ 10.00 | $ 50.00 | $ 20,000 | 1040 | 160 | 0.28126 | 0.03019 | 0.09949 | 0.03068 | 0.16587 | 0.03056 |
| 11 | 1300 | 140 | $ 10.00 | $ 50.00 | $ 20,000 | 1160 | 140 | 0.28112 | 0.02910 | 0.09956 | 0.02953 | 0.16568 | 0.02943 |
| 12 | 1400 | 120 | $ 10.00 | $ 50.00 | $ 20,000 | 1280 | 120 | 0.28099 | 0.02815 | 0.09963 | 0.02857 | 0.16553 | 0.02842 |
| 13 | 1500 | 100 | $ 10.00 | $ 50.00 | $ 20,000 | 1400 | 100 | 0.28088 | 0.02731 | 0.09968 | 0.02759 | 0.16539 | 0.02752 |
| 14 | 1600 | 80 | $ 10.00 | $ 50.00 | $ 20,000 | 1520 | 80 | 0.28079 | 0.02658 | 0.09973 | 0.02676 | 0.16528 | 0.02672 |
| 15 | 1700 | 60 | $ 10.00 | $ 50.00 | $ 20,000 | 1640 | 60 | 0.28070 | 0.02600 | 0.09977 | 0.02601 | 0.16518 | 0.02602 |
| 16 | 1800 | 40 | $ 10.00 | $ 50.00 | $ 20,000 | 1760 | 40 | 0.28062 | **0.02566** | 0.09981 | 0.02535 | 0.16511 | 0.02544 |
| 17 | 1900 | 20 | $ 10.00 | $ 50.00 | $ 20,000 | 1880 | 20 | 0.28054 | 0.02607 | 0.09984 | 0.02483 | 0.16509 | 0.02518 |
| 18 | 2000 | 0 | $ 10.00 | $ 50.00 | $ 20,000 | 2000 | 0 | ---- | ---- | ---- | ---- | ---- | ---- |

*Note:* Bolded parameters indicate the lowest standard error obtained for a particular independent-dependent variable pathway. For Body Fat → SBP and Body Fat → PP, standard errors monotonically decreased as the ratio of partial to complete increased; as a result, no minimum standard error was reached.

*Regression coefficients*. As the partial to complete data ratio became more extreme (i.e., from Ratio 1 to Ratio 17), the b-weights for each of the three regression paths of interest (i.e., (a) Body Fat → SBP; (b) Body Fat → DBP; and (c) Body Fat → PP) changed in a monotonic fashion; the regression coefficients for body fat predicting SBP and PP increased monotonically, while the b-weight for body fat predicting DBP decreased monotonically. This pattern of b-weights indicated that as the ratio of partial to complete data cases became more extreme, parameter estimates became increasingly biased. However, between the complete case design and the most extreme two-method measurement design (i.e., 1880 partial data cases and 20 complete data cases), the absolute difference in b-weights was 0.00486, 0.00243, and 0.02129 for body fat predicting SBP, DBP, and PP, respectively (reflecting a 2.0, 2.4, and 12.4 percent change in b-weights for SBP, DBP, and PP, respectively). Thus, the degree of bias associated with parameter estimates derived under the two-method measurement design was small.

*Standard errors.* The two-method measurement design produced a substantial increase in power by way of reduced standard errors. As the ratio of partial to complete data became more extreme, resulting standard errors decreased monotonically for all three regression paths. The standard errors pertaining to the Body Fat → SBP path reached a minimum (SE = 0.02566) when the two-method measurement design was applied using 1760 partial data cases and 40 complete data cases. For the Body Fat → DBP and Body Fat → PP regression coefficients, no minimum standard error was detected; rather, as the ratio of partial to complete data increased to the most extreme data ratio, standard errors steadily decreased. Please refer to Figure 9, located below, for a graphical representation of how standard errors decreased as the two-method measurement design was applied to increasingly extreme partial to complete data ratios.

**Figure 9: Simulation Standard Errors for the Regression Coefficients of Body Fat Predicting SBP, DBP, and PP**



*Statistical power benefits*. To quantify the benefits derived from the two-method measurement design, the Effective N was calculated for each of the three body fat regression pathways. Effective N was ascertained by determining the number of complete cases that, when tested using the one-group complete cases model, yielded a standard error lower than that produced by the most efficient two-method measurement design (Graham et al., 2006). The lowest standard error associated with the Body Fat → SBP pathway was 0.02566; thus, the Effective N for the Body Fat → SBP pathway represented the number of complete cases that yielded a standard error just below 0.02566. Unlike the standard error for the Body Fat → SBP pathway, the Body Fat → DBP and Body Fat → PP pathways did not demonstrate a minimum standard error. Effective Ns for these two pathways were calculated using the standard errors obtained when the two-method measurement design was tested with the most extreme partial to complete data ratio (i.e., 1880 partial cases, 20 complete cases); these standard errors were 0.02438 and 0.02518 for body fat predicting DBP and PP, respectively. Table 28, located below, displays resulting standard errors as the one-group model was tested with an increasingly large number of complete cases; for each of the three pathways, the number of cases representing the Effective N is indicated in bold.

Table 28
Standard Errors and Effective Ns for Body Fat Regression Pathways Obtained under a One-Group Complete Cases Design

| Complete Cases N | BF → SBP | BF → DBP | BF → PP |
|---|---|---|---|
| 333 | 0.05391 | 0.05490 | 0.05463 |
| 500 | 0.04397 | 0.04478 | 0.04456 |
| 650 | 0.03856 | 0.03926 | 0.03908 |
| 800 | 0.03475 | 0.03539 | 0.03522 |
| 950 | 0.03189 | 0.03247 | 0.03231 |
| 1100 | 0.02963 | 0.03017 | 0.03003 |
| 1250 | 0.02779 | 0.0283 | 0.02817 |
| 1400 | 0.02626 | 0.02674 | 0.02661 |
| 1469 | **0.02565** | 0.02611 | 0.02598 |
| 1500 | | 0.02583 | 0.02571 |
| 1550 | | 0.02541 | 0.02529 |
| 1565 | | 0.02529 | **0.02517** |
| 1600 | | 0.02501 | |
| 1625 | | **0.02482** | |

*Note:* Bolded values represent complete case standard errors lower than those obtained using the most efficient two-method measurement design.

To obtain statistical power equivalent to that produced by the two-method measurement design, researchers would require 1469, 1625, and 1565 complete cases to test the associations between body fat and SBP, DBP, and PP, respectively. Another way to conceptualize this is that the two-method measurement design behaved as if the sample sizes were N=1469, N=1625, and N=1565 for testing the body fat regression weights involving SBP, DBP, and PP, respectively.

To compute Effective N Increase Factors, the Effective N for each regression path was divided by 333 (the nominal complete case N permitted by the hypothetical $20,000 budget). Effective N Increase Factors were 4.4, 4.9, and 4.7, respectively, for the SBP, DBP, and PP regression pathways. Taken together, results indicated that the two-method measurement design produced a substantial increase in statistical power beyond that produced by the financially-equivalent complete cases models. The utility of the two-method measurement design was demonstrated by the sizeable reduction in standard errors and stability of estimated b-weights.

The benefits of the two-method measurement design were greatest for the Body Fat → DBP pathway (Effective N Increase Factor: 4.9).

*Tobacco Use*

Eighteen partial to complete data ratios, all remaining within the $20,000 hypothetical budget, were generated for the tobacco use predictor. Ratio 1 reflected the number of complete cases permitted by the budget (N=363); this data ratio was tested using the one-group complete cases design. The two-method measurement design was applied to ratios 2-17 using the two-group modeling procedure; resulting b-weights and standard errors are displayed below in Table 29.

Table 29
*Tobacco Use Parameter Estimates Obtained using the Two-Method Measurement Design*

| | N Cheap | N Expensive | Cost Cheap | Cost Expensive | Total Costs | N Partial | N Complete | TU → SBP | | TU → DBP | | TU → PP | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | b | SE | b | SE | b | SE |
| 1 | 363 | 363 | $ 10.00 | $ 45.00 | $ 19,965 | 0 | 363 | -0.01494 | 0.05332 | -0.08371 | 0.05699 | 0.05261 | 0.05472 |
| 2 | 400 | 355 | $ 10.00 | $ 45.00 | $ 19,975 | 45 | 355 | -0.01485 | 0.05086 | -0.08367 | 0.05485 | 0.05267 | 0.05240 |
| 3 | 500 | 333 | $ 10.00 | $ 45.00 | $ 19,985 | 167 | 333 | -0.01467 | 0.04550 | -0.08361 | 0.05038 | 0.05280 | 0.04741 |
| 4 | 600 | 311 | $ 10.00 | $ 45.00 | $ 19,995 | 289 | 311 | -0.01456 | 0.04156 | -0.08356 | 0.04734 | 0.05288 | 0.04386 |
| 5 | 700 | 288 | $ 10.00 | $ 45.00 | $ 19,960 | 412 | 288 | -0.01448 | 0.03852 | -0.08353 | 0.04525 | 0.05293 | 0.04122 |
| 6 | 800 | 266 | $ 10.00 | $ 45.00 | $ 19,970 | 534 | 266 | -0.01443 | 0.03607 | -0.08351 | 0.04381 | 0.05297 | 0.03922 |
| 7 | 900 | 244 | $ 10.00 | $ 45.00 | $ 19,980 | 656 | 244 | -0.01439 | 0.03406 | -0.08349 | 0.04290 | 0.05300 | 0.03771 |
| 8 | 1000 | 220 | $ 10.00 | $ 45.00 | $ 19,900 | 780 | 220 | -0.01435 | 0.03238 | -0.08348 | 0.04253 | 0.05302 | 0.03662 |
| 9 | 1100 | 200 | $ 10.00 | $ 45.00 | $ 20,000 | 900 | 200 | -0.01433 | 0.03094 | -0.08347 | **0.04242** | 0.05304 | 0.03580 |
| 10 | 1200 | 177 | $ 10.00 | $ 45.00 | $ 19,965 | 1023 | 177 | -0.01431 | 0.02971 | -0.08346 | 0.04289 | 0.05305 | 0.03539 |
| 11 | 1300 | 155 | $ 10.00 | $ 45.00 | $ 19,975 | 1145 | 155 | -0.01429 | 0.02866 | -0.08346 | 0.04381 | 0.05307 | **0.03529** |
| 12 | 1400 | 133 | $ 10.00 | $ 45.00 | $ 19,985 | 1267 | 133 | -0.01427 | 0.02776 | -0.08345 | 0.04534 | 0.05308 | 0.03560 |
| 13 | 1500 | 111 | $ 10.00 | $ 45.00 | $ 19,995 | 1389 | 111 | -0.01426 | 0.02701 | -0.08344 | 0.04771 | 0.05309 | 0.03644 |
| 14 | 1600 | 88 | $ 10.00 | $ 45.00 | $ 19,960 | 1512 | 88 | -0.01425 | 0.02644 | -0.08344 | 0.05158 | 0.05309 | 0.03818 |
| 15 | 1700 | 66 | $ 10.00 | $ 45.00 | $ 19,970 | 1634 | 66 | -0.01424 | **0.02610** | -0.08344 | 0.05755 | 0.05310 | 0.04123 |
| 16 | 1800 | 44 | $ 10.00 | $ 45.00 | $ 19,980 | 1756 | 44 | -0.01423 | 0.02624 | -0.08343 | 0.06836 | 0.05311 | 0.04724 |
| 17 | 1900 | 22 | $ 10.00 | $ 45.00 | $ 19,990 | 1878 | 22 | -0.01422 | 0.02799 | -0.08343 | 0.09465 | 0.05311 | 0.06286 |
| 18 | 2000 | 0 | $ 10.00 | $ 45.00 | $ 20,000 | 2000 | 0 | ---- | ---- | ---- | ---- | ---- | ---- |

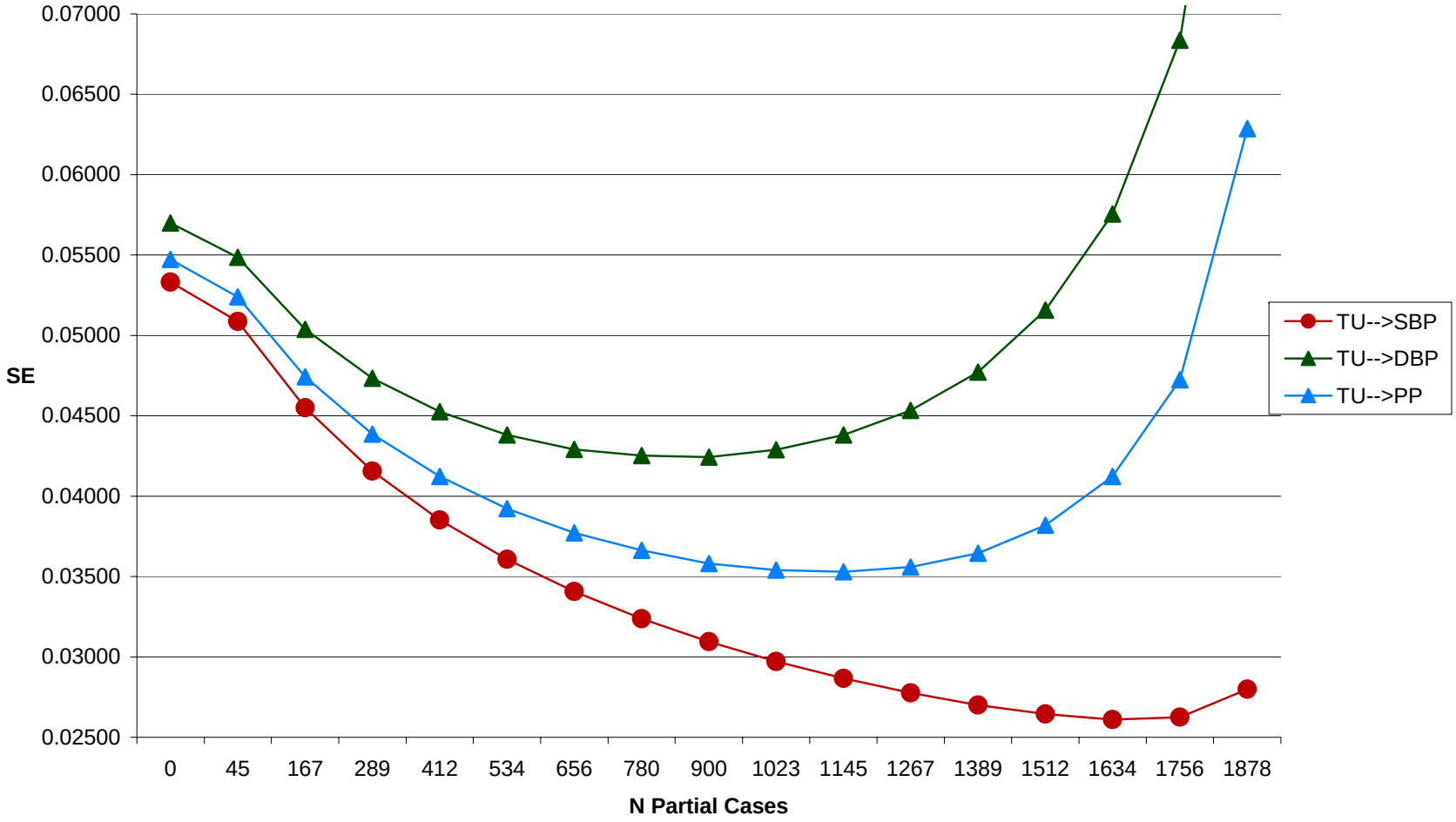
*Note:* Bolded parameters indicate the lowest standard error obtained for a particular independent-dependent variable pathway.

*Regression coefficients.* As the partial to complete data ratio became more extreme (i.e., from Ratio 1 to Ratio 17), the b-weights for all three regression paths of interest (i.e, (a) Tobacco Use → SBP; (b) Tobacco Use → DBP; and (c) Tobacco Use → PP) increased monotonically. Thus, the two-method measurement design produced slightly biased parameter estimates. However, the absolute difference in b-weights between the complete cases ratio and the most extreme partial to complete data ratio (i.e., 1878 partial data cases and 22 complete data cases) was 0.00072, 0.00028, and 0.0005 for tobacco use predicting SBP, DBP, and PP, respectively (reflecting a 4.8, 0.33, and 0.95 percent change in b-weights for SBP, DBP, and PP, respectively). Thus, the degree of bias associated with parameter estimates derived under the two-method measurement design was small.

*Standard errors*. As the partial to complete data ratio became more extreme, the standard errors for all three regression paths initially decreased monotonically to some minimum value before they began to increase; therefore, all three sets of standard errors displayed an inflection point at which the two-method measurement design no longer produced an increase in statistical power. For the Tobacco Use → SBP regression weight, the lowest standard error (0.02610) was obtained when the two-method measurement design was applied using 1634 partial data cases and 66 complete data cases. For tobacco use predicting DBP, the lowest standard error (0.04242) was obtained by applying the two-method measurement design to 900 partial and 200 complete data cases. The lowest standard error for tobacco use predicting PP (0.03529) was obtained when the two-method measurement design was applied using 1145 partial and 155 complete data cases. Please refer to Figure 10, located below, for a graphical representation of how standard

errors decreased as the two-method measurement design was applied to increasingly extreme partial to complete data ratios.

**Figure 10: Simulation Standard Errors for the Regression Coefficients of Tobacco Use Predicting SBP, DBP, and PP**



*Statistical power benefits*. Using the minimum standard errors as reference, Effective Ns were computed for all three regression paths. The one-group design was used to determine the number of complete cases that yielded lower standard errors than those obtained using the most efficient two-method measurement designs. Table 30, located below, displays resulting standard errors as the one-group model was estimated with an increasingly large number of complete cases; for each of the three regression paths, the number of cases representing the Effective N is indicated in bold.

Table 30
*Standard Errors and Effective Ns for Tobacco Use Regression Pathways Obtained under a One-Group Complete Cases Design*

| Complete Cases N | TU → SBP | TU → DBP | TU → PP |
|---|---|---|---|
| 363 | 0.05332 | 0.05699 | 0.05472 |
| 400 | 0.05078 | 0.05428 | 0.05212 |
| 500 | 0.04541 | 0.04854 | 0.04661 |
| 600 | 0.04145 | 0.04430 | 0.04254 |
| 655 | 0.03967 | **0.04240** | 0.04071 |
| 700 | 0.03837 | | 0.03938 |
| 800 | 0.03589 | | 0.03683 |
| 872 | 0.03437 | | **0.03528** |
| 900 | 0.03383 | | |
| 1000 | 0.03209 | | |
| 1100 | 0.03060 | | |
| 1200 | 0.02930 | | |
| 1300 | 0.02815 | | |
| 1400 | 0.02712 | | |
| 1500 | 0.02620 | | |
| 1513 | **0.02609** | | |

To obtain statistical power equivalent to that produced by the two-method measurement design, 1513, 655, and 872 complete cases would be necessary to test the associations between tobacco use and SBP, DBP, and PP, respectively. Thus, the two-method measurement design behaved as if the sample sizes were N=1513, N=655, and N=872 for testing the tobacco use regression weights involving SBP, DBP, and PP, respectively.

Effective N Increase Factors were computed by dividing the Effective Ns by 363 (the nominal complete cases N permitted by the hypothetical $20,000 budget); for the SBP, DBP, and PP regression paths, Effective N Increase Factors were 4.2, 1.8, and 2.4, respectively. Overall, results indicated that the two-method measurement design produced a substantial increase in statistical power by way of lower standard errors for regression paths of interest. The benefits of the two-method measurement design were greatest for the Tobacco Use → SBP pathway (Effective N Increase Factor: 4.2).

*Physical Conditioning*

Twenty-one partial to complete data ratios were computed for the physical conditioning variable; with the exception of Ratio 1 (total cost: $20,020), all ratios remained within the $20,000 hypothetical budget. Ratio 1, corresponding to the complete cases scenario, was tested using a one-group design. The two-method measurement design was applied to ratios 2-20 using the two-group modeling procedure; resulting b-weights and standard errors are displayed below in Table 31.

Table 31
*Physical Conditioning Parameter Estimates Obtained using the Two-Method Measurement Design*

| | N Cheap | N Expensive | Cost Cheap | Cost Expensive | Total Costs | N Partial | N Complete | PC → SBP | | PC → DBP | | PC → PP | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | b | SE | b | SE | b | SE |
| 1 | 77 | 77 | $ 10.00 | $ 250.00 | $ 20,020 | 0 | 77 | -0.11887 | 0.41175 | -0.07751 | 0.29894 | -0.03069 | 0.19843 |
| 2 | 100 | 76 | $ 10.00 | $ 250.00 | $ 20,000 | 24 | 76 | -0.15906 | 0.31028 | -0.12787 | 0.30291 | -0.02084 | 0.28935 |
| 3 | 200 | 72 | $ 10.00 | $ 250.00 | $ 20,000 | 128 | 72 | -0.11086 | 0.33353 | -0.13036 | 0.34485 | 0.01999 | 0.30305 |
| 4 | 300 | 68 | $ 10.00 | $ 250.00 | $ 20,000 | 232 | 68 | -0.08798 | 0.32921 | -0.13257 | 0.38412 | 0.04022 | 0.28981 |
| 5 | 400 | 64 | $ 10.00 | $ 250.00 | $ 20,000 | 336 | 64 | -0.07688 | 0.32742 | -0.13374 | 0.44147 | 0.05012 | 0.28691 |
| 6 | 500 | 60 | $ 10.00 | $ 250.00 | $ 20,000 | 440 | 60 | -0.07128 | 0.33494 | -0.13429 | 0.51196 | 0.05508 | 0.29746 |
| 7 | 600 | 56 | $ 10.00 | $ 250.00 | $ 20,000 | 544 | 56 | -0.06825 | 0.35254 | -0.13456 | 0.59318 | 0.05774 | 0.31920 |
| 8 | 700 | 52 | $ 10.00 | $ 250.00 | $ 20,000 | 648 | 52 | -0.06650 | 0.38001 | -0.13472 | 0.68575 | 0.05929 | 0.35056 |
| 9 | 800 | 48 | $ 10.00 | $ 250.00 | $ 20,000 | 752 | 48 | -0.06543 | 0.41748 | -0.13483 | 0.79196 | 0.06023 | 0.39121 |
| 10 | 900 | 44 | $ 10.00 | $ 250.00 | $ 20,000 | 856 | 44 | -0.06474 | 0.46602 | -0.13488 | 0.91586 | 0.06082 | 0.44195 |
| 11 | 1000 | 40 | $ 10.00 | $ 250.00 | $ 20,000 | 960 | 40 | -0.06430 | 0.52737 | -0.13493 | 1.06264 | 0.06124 | 0.50475 |
| 12 | 1100 | 36 | $ 10.00 | $ 250.00 | $ 20,000 | 1064 | 36 | -0.06399 | 0.60492 | -0.13495 | 1.24068 | 0.06150 | 0.58286 |
| 13 | 1200 | 32 | $ 10.00 | $ 250.00 | $ 20,000 | 1168 | 32 | -0.06380 | 0.70405 | -0.13498 | 1.46192 | 0.06167 | 0.68165 |
| 14 | 1300 | 28 | $ 10.00 | $ 250.00 | $ 20,000 | 1272 | 28 | -0.06365 | 0.83351 | -0.13497 | 1.74585 | 0.06181 | 0.80996 |
| 15 | 1400 | 24 | $ 10.00 | $ 250.00 | $ 20,000 | 1376 | 24 | -0.06358 | 1.00914 | -0.13502 | 2.12650 | * | * |
| 16 | 1500 | 20 | $ 10.00 | $ 250.00 | $ 20,000 | 1480 | 20 | -0.07214 | 1.37715 | -0.15331 | 2.91215 | * | * |
| 17 | 1600 | 16 | $ 10.00 | $ 250.00 | $ 20,000 | 1584 | 16 | -0.07087 | 1.76508 | -0.15075 | 3.74436 | * | * |
| 18 | 1700 | 12 | $ 10.00 | $ 250.00 | $ 20,000 | 1688 | 12 | -0.07230 | 2.52847 | -0.15388 | 5.37416 | * | * |
| 19 | 1800 | 8 | $ 10.00 | $ 250.00 | $ 20,000 | 1792 | 8 | -0.07325 | 4.14108 | -0.15594 | 8.81186 | * | * |
| 20 | 1900 | 4 | $ 10.00 | $ 250.00 | $ 20,000 | 1896 | 4 | -0.08041 | 11.28585 | -0.17122 | 24.02815 | * | * |
| 21 | 2000 | 0 | $ 10.00 | $ 250.00 | $ 20,000 | 2000 | 0 | ---- | ---- | ---- | ---- | ---- | ---- |

*Regression coefficients*. As the partial to complete data ratio became more extreme (i.e., from Ratio 1 to Ratio 20), the b-weights for all three regression paths of interest (i.e., (a) Physical Conditioning → SBP; (b) Physical Conditioning → DBP; and (c) Physical Conditioning → PP) dramatically changed. The b-weights for physical conditioning predicting DBP and PP decreased and increased, respectively, in a monotonic manner. However, the b-weight for physical conditioning predicting SBP initially increased monotonically to a maximum value (i.e., -0.06358, obtained using 1376 partial data cases and 24 complete data cases) before reversing direction and monotonically decreasing across the remainder of data ratios. The instability of b-weights for all three regression paths indicated that the two-method measurement design produced biased parameter estimates. Additionally, the large increases and decreases in b-weights across the data ratios indicated that the degree of bias was fairly substantial.

*Standard errors*. Resulting standard errors also indicated that the application of the two-method measurement design did not produce any statistical advantage over the financially-equivalent complete case models. Standard errors pertaining to the Physical Conditioning → DBP and Physical Conditioning → PP pathways increased monotonically as the partial to complete data ratio became more extreme. For the most extreme data ratio (i.e., 1896 partial data cases and 4 complete data cases), the standard error for the DBP regression path was 24.02815, nearly 80 times greater than the standard error obtained using the complete cases model with a sample size of N=77. The two-group model estimating the effect of physical conditioning on PP failed to converge when data ratios more extreme than 1272 partial data cases and 28 complete data cases were tested; thus, for 6 of the 20 data ratios, parameter estimates involving physical conditioning and PP were not able to be estimated.

*Dietary Intake*

Twenty-one partial to complete data ratios were computed for the dietary intake variable; all ratios remained within the $20,000 hypothetical budget. Ratio 1, corresponding to the complete cases scenario, was tested using a one-group design. The two-method measurement design was applied to ratios 2-19 using the two-group modeling procedure; resulting b-weights and standard errors are displayed below in Table 32.

Table 32
*Dietary Intake Parameter Estimates Obtained using the Two-Method Measurement Design*

| | N Cheap | N Expensive | Cost Cheap | Cost Expensive | Total Costs | N Partial | N Complete | DI → SBP | | DI → DBP | | DI → PP | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | b | SE | b | SE | b | SE |
| 1 | 125 | 125 | $ 10.00 | $ 150.00 | $ 20,000 | 0 | 125 | -0.03382 | 0.09958 | -0.03252 | 0.09959 | -0.00402 | 0.09961 |
| 2 | 200 | 120 | $ 10.00 | $ 150.00 | $ 20,000 | 80 | 120 | -0.03482 | 0.10096 | -0.03737 | 0.10095 | -0.00113 | 0.10106 |
| 3 | 300 | 113 | $ 10.00 | $ 150.00 | $ 19,950 | 187 | 113 | -0.03629 | 0.10298 | -0.04463 | 0.10294 | 0.00310 | 0.10322 |
| 4 | 400 | 106 | $ 10.00 | $ 150.00 | $ 19,900 | 294 | 106 | -0.03793 | 0.10500 | -0.05292 | 0.10493 | 0.00776 | 0.10550 |
| 5 | 500 | 100 | $ 10.00 | $ 150.00 | $ 20,000 | 400 | 100 | -0.03968 | 0.10644 | -0.06207 | 0.10635 | 0.01274 | 0.10739 |
| 6 | 600 | 93 | $ 10.00 | $ 150.00 | $ 19,950 | 507 | 93 | -0.04167 | 0.10808 | -0.07292 | 0.10801 | 0.01851 | 0.10983 |
| 7 | 700 | 86 | $ 10.00 | $ 150.00 | $ 19,900 | 614 | 86 | -0.04378 | 0.10929 | -0.08525 | 0.10935 | 0.02506 | 0.11229 |
| 8 | 800 | 80 | $ 10.00 | $ 150.00 | $ 20,000 | 720 | 80 | -0.04580 | 0.10945 | -0.09797 | 0.10987 | 0.03207 | 0.11414 |
| 9 | 900 | 73 | $ 10.00 | $ 150.00 | $ 19,950 | 827 | 73 | -0.04781 | 0.10935 | -0.11197 | 0.11063 | 0.04051 | 0.11641 |
| 10 | 1000 | 66 | $ 10.00 | $ 150.00 | $ 19,900 | 934 | 66 | -0.04958 | 0.10860 | -0.12577 | 0.11146 | 0.05004 | 0.11846 |
| 11 | 1100 | 60 | $ 10.00 | $ 150.00 | $ 20,000 | 1040 | 60 | -0.05093 | 0.10707 | -0.13773 | 0.11223 | 0.05963 | 0.11972 |
| 12 | 1200 | 53 | $ 10.00 | $ 150.00 | $ 19,950 | 1147 | 53 | -0.05209 | 0.10559 | -0.14950 | 0.11436 | 0.07055 | 0.12151 |
| 13 | 1300 | 46 | $ 10.00 | $ 150.00 | $ 19,900 | 1254 | 46 | -0.05299 | 0.10401 | -0.15999 | 0.11781 | 0.08156 | 0.12369 |
| 14 | 1400 | 40 | $ 10.00 | $ 150.00 | $ 20,000 | 1360 | 40 | -0.05360 | 0.10230 | -0.16835 | 0.12207 | 0.09114 | 0.12609 |
| 15 | 1500 | 33 | $ 10.00 | $ 150.00 | $ 19,950 | 1467 | 33 | -0.05413 | 0.10108 | -0.17654 | 0.12983 | 0.10106 | 0.13116 |
| 16 | 1600 | 26 | $ 10.00 | $ 150.00 | $ 19,900 | 1574 | 26 | -0.05452 | 0.10040 | -0.18377 | 0.14189 | 0.11013 | 0.13981 |
| 17 | 1700 | 20 | $ 10.00 | $ 150.00 | $ 20,000 | 1680 | 20 | -0.05478 | 0.10031 | -0.18941 | 0.15822 | 0.11739 | 0.15238 |
| 18 | 1800 | 13 | $ 10.00 | $ 150.00 | $ 19,950 | 1787 | 13 | -0.05500 | 0.10310 | -0.19514 | 0.19325 | 0.12483 | 0.18123 |
| 19 | 1900 | 6 | $ 10.00 | $ 150.00 | $ 19,900 | 1894 | 6 | -0.05517 | 0.11750 | -0.20024 | 0.29122 | 0.13152 | 0.26591 |
| 20 | 2000 | 0 | $ 10.00 | $ 150.00 | $ 20,000 | 2000 | 0 | ---- | ---- | ---- | ---- | ---- | ---- |

*Regression coefficients*. As the partial to complete data ratio became more extreme (i.e., from Ratio 1 to Ratio 19), the b-weights for all three regression paths of interest (i.e., (a) Dietary Intake → SBP; (b) Dietary Intake → DBP; and (c) Dietary Intake → PP) changed monotonically; the b-weights for dietary intake predicting SBP and DBP monotonically decreased while the b-weight for dietary intake predicting PP monotonically increased. The magnitude of change was relatively large, indicating that the two-method measurement design produced biased parameter estimates. Between the complete case ratio and the most extreme partial to complete data ratio (i.e., 1894 partial data cases and 6 complete data cases), the b-weights for dietary intake predicting SBP, DBP, and PP changed 0.02135, 0.16772, and 0.1275 respectively (reflecting a 63 percent decrease, a 515 percent decrease, and a 3171 percent increase, respectively, in b-weights). Thus, the large increases and decreases in b-weights across the data ratios indicated that the degree of bias was fairly substantial.

*Standard errors*. Resulting standard errors also indicated that the application of the two-method measurement design did not produce any statistical advantage over the financially-equivalent complete cases model. Standard errors pertaining to the Dietary Intake → DBP and Dietary Intake → PP pathways increased monotonically as the partial to complete data ratio became more extreme. For the most extreme data ratio (i.e., 1894 partial data cases and 6 complete data cases), the standard errors for the DBP and PP regression paths were 0.29122 and 0.26591, respectively (reflecting standard errors 2.9 and 2.6 times greater than the standard error obtained using the complete cases model with a sample size of N=125). The standard error pertaining to the Dietary Intake → SBP pathway did not change monotonically; rather, the standard error increased monotonically between data ratios 1-9, and then decreased

monotonically between ratios 10-17; between ratios 18 and 19, the standard error increased. Importantly, among data ratios in which the standard error pertaining to the Dietary Intake → SBP pathway decreased, the standard error still remained greater than that produced by the financially-equivalent complete case design. Taken together, the instability of b-weights and increases in standard errors demonstrated that the two-method measurement design did not produce any appreciable statistical power benefit when applied to the dietary intake variable.

# CHAPTER 5

## Discussion and Conclusions

The following sections address formally stated research questions, provide an overview of key findings, and discuss opportunities for future research related to the two-method measurement design.

*Research Questions*

**1. What is the degree of association between established hypertension risk factors (physical conditioning, dietary intake, body fat, and tobacco use) and SBP, DBP, and PP in a large, nationally representative sample?**

Partial NHANES correlations offered insight into the strength of association between the four hypertension risk factors and blood pressure parameters. Physical conditioning – as assessed by PA1, PA2, and $EVO_2$ max – was generally uncorrelated with SBP, DBP, and PP. Partial correlations involving the three physical conditioning measures indicated that their associations with SBP, DBP, and PP did not reach the "small effect" threshold in the context of Cohen (1977). However, despite their small magnitude, the general pattern of partial correlations indicated that increased levels of physical conditioning were associated with reduced SBP and DBP and elevated PP. $EVO_2$ max was most strongly associated with SBP and DBP, indicating that the strength of association between physical fitness and blood pressure was more robust than that between physical activity and blood pressure; this finding is consistent with previous research (Dvorak et al., 2000; Myers et al., 2004; Sternfeld et al., 1999).

Dietary intake – as assessed by drybeans, dgveg, Min1, and Min2 – was generally uncorrelated with SBP, DBP, and PP. Partial correlations indicated that none of the associations involving the dietary intake measures and the blood pressure parameters reached Cohen's (1977)

"small effect" threshold. Despite their small magnitude, however, results indicated that increased consumption of dry beans, dark green vegetables, magnesium, and fiber was associated with reduced SBP and DBP and elevated PP. These findings are consistent with previous literature stating that total nutrient intakes confer small reductions in SBP and DBP (Kawano et al., 1998; Zhou et al., 2000).

Body fat was strongly associated with blood pressure. All four body fat indicators – Srbmi, Srow, Ebmi, and BFat – were significantly correlated with SBP, DBP, and PP; effect sizes generally reached Cohen's (1977) "medium effect" threshold. These findings are consistent with the extensive body of literature implicating excess body fat as a strong contributing factor to the onset of hypertension (Wilsgaard et al., 2000).

Tobacco use – as measured by TUse1, TUse2, and serum cotinine – was negatively associated with DBP and positively associated with PP; on the other hand, tobacco use was generally uncorrelated with SBP. Partial correlations indicated that the associations between the three tobacco use measures and SBP, DBP, and PP were considered "small effects" in the context of Cohen (1977). These findings were generally inconsistent with previous literature indicating the detrimental effects of tobacco use on blood pressure and general cardiovascular risk profile (Mahmud & Feely, 2003).

**2. For each independent variable, what is the optimal ratio of partial to complete data that yields the most efficient and unbiased regression coefficients under the two-method measurement design?**

When applied to the body fat and tobacco use variables, the two-method measurement design produced smaller standard errors and larger Effective Ns for testing regression coefficients of substantive interest compared to financially-equivalent complete case models. For

these two independent variables, it was possible to determine the optimal ratios of partial to complete data that produced the most unbiased and efficient regression coefficients.

For body fat predicting SBP, the lowest standard error (0.02566) was obtained when the two-method measurement design was applied using 1760 partial data cases and 40 complete data cases; thus, the optimal partial to complete data ratio was 1760:40, or 40:1. When body fat predicted DBP and PP, standard errors monotonically decreased through the most extreme partial to complete case data ratio; thus, the lowest standard errors (0.02483 and 0.02518 for DBP and PP, respectively) were obtained when the two-method measurement design was applied using 1880 partial data cases and 20 complete data cases, reflecting an optimal ratio of 1880:20, or 94:1.

For tobacco use, standard errors for each of the three regression paths of substantive interest decreased monotonically to some minimum value before beginning to increase. When tobacco use predicted SBP, the lowest standard error (0.02610) was obtained when the two-method measurement design was applied using 1634 partial data cases and 66 complete data cases, reflecting an optimal partial to complete data ratio of 1634:66, or 24.8:1. The lowest standard error for the b-weight involving tobacco use predicting DBP (0.04242) was obtained when the two-method measurement design was applied to a comparatively less extreme data ratio (i.e., 900 partial data cases and 200 complete data cases); thus, the optimal ratio for estimating the Tobacco Use → DBP b-weight was 900:200, or 4.5:1. When tobacco use predicted PP, the two-method measurement design produced the lowest standard error (0.03529) when the two-group model was estimated using 1145 partial data cases and 155 complete data cases, reflecting an optimal data ratio of 1145:155, or 7.4:1.

Application of the two-method measurement design was less effective for the physical conditioning and dietary intake variables; as a result, no optimal ratios of partial data to complete data cases were observed. As the two-method measurement design was applied using increasingly extreme partial to complete data ratios, the pattern of physical conditioning b-weights changed substantially, indicating that the two-method measurement design produced biased parameter estimates. Additionally, increasingly large standard errors reflected that the two-method measurement design failed to produce an increase in statistical power when applied to the physical conditioning factor.

Similar patterns were observed for the dietary intake factor; substantial changes in estimated b-weights occurred with increasingly extreme data ratios, indicating that the two-method measurement design produced biased estimates of dietary intake b-weights. Additionally, standard errors for all three dietary intake regression paths increased substantially as the partial to complete case data ratios became more extreme; consequently, no statistical power advantage was provided by the two-method measurement design when applied to the dietary intake factor.

**3. What effect does the number of expensive measures for an independent variable, as well as the strength of correlation between cheap and expensive measures, have on the performance of the two-method measurement design?**

Body fat and tobacco use – the variables for which the two-method measurement design produced increases in statistical power – differed from one another in several important ways, including their number of expensive measures. Body fat was represented by two expensive measures (i.e., Ebmi and BFat), whereas tobacco use was represented by one expensive measure (i.e., serum cotinine). Body fat and tobacco use also differed in the strength of correlation between cheap measures. Srow and Srbmi – the two cheap body fat measures – were strongly

correlated ($r$ = 0.646); however, the degree of association between TUse1 and TUse2 – the two cheap tobacco use indicators – was comparatively stronger ($r$ = 0.917). Additionally, whereas all four body fat indicators were strongly associated with SBP, DBP, and PP, partial correlations indicated that tobacco use measures were less strongly correlated with blood pressure parameters. The following sections summarize key results related to the performance of the two-method measurement design when applied to the body fat and tobacco use factors.

*Body fat.* As shown in Table 33, below, Effective N Increase Factors for the three body fat coefficients of substantive interest were 4.4, 4.9, and 4.7. Essentially, these values reflect that the best two-method measurement designs provided the equivalent of 4.4 - 4.9 times the number of complete cases than were allowable by the financially-equivalent complete cases model (nominal complete case N=333).

Another way to conceptualize the benefits provided by the two-method measurement design is to consider the Effective Ns; as shown in Table 33, Effective Ns were 1469, 1635, and 1565 for testing the effects involving body fat and SBP, DBP, and PP, respectively. These results illustrate the enormous statistical advantage provided by the two-method measurement design. As an example, researchers could test the Body Fat → SBP effect using a complete cases design and statistical power based on a sample size of N=333; however, for the same cost, researchers could implement the two-method measurement design and test the same effect with power equivalent to N=1469 complete cases.

Table 33
*Summary of Key Results: Application of the Two-Method Measurement Design to the Body Fat Construct*

| | **BF → SBP** | **BF → DBP** | **BF → PP** |
|---|---|---|---|
| Average correlation between manifest measures and outcome | $r$ = 0.243 | $r$ = 0.088 | $r$ = 0.142 |
| Optimal ratio (partial:complete data cases) | 1760:40 (40:1) | 1880:20 (94:1) | 1880:20 (94:1) |
| SE (Nominal Complete Cases Design) | 0.05391 | 0.05490 | 0.05463 |
| SE (Most Efficient Two-Method Measurement Design) | 0.02566 | 0.02483 | 0.02518 |
| % Decrease in SE | 52 | 55 | 54 |
| Effective N | **1469** | **1625** | **1565** |
| Effective N Increase Factor | 4.4 | 4.9 | 4.7 |

The increase in statistical power provided by the two-method measurement design varied across the three body fat regression paths. Effective N Increase Factors correlated inversely with the average effect size between body fat and the three blood pressure parameters. That is, the largest Effective N Increase Factor (4.9) was obtained for the Body Fat → DBP pathway; correspondingly, of the three blood pressure parameters considered, body fat was least strongly associated DBP (average $r$ = 0.088). In the same manner, body fat measures were most strongly correlated with SBP ($r$ = 0.243); subsequently, the two-method measurement design produced the smallest Effective N Increase Factor (4.4) for the coefficient pertaining to Body Fat → SBP.

*Tobacco use.* As shown below in Table 34, Effective N Increase Factors for the three tobacco use coefficients of substantive interest were 4.2, 1.8, and 2.4; thus, for the same cost, the two-method measurement design behaved as if a substantially larger number of complete cases were available for parameter estimation than the complete case N allowable under the hypothetical budget constraint (nominal complete case N=363). Using the same logic above, the considerable statistical advantages provided by the two-method measurement design are

demonstrated by the Effective Ns (1513, 655, and 872 for tobacco use predicting SBP, DBP, and PP, respectively). As an example, for $20,000, the Tobacco Use → SBP effect could be tested with statistical power based on a sample size of N=363; or, for the same cost, under the two-method measurement design, the same effect could be tested with power equivalent to N=1513 complete cases.

Table 34
*Summary of Key Results: Application of the Two-Method Measurement Design to the Tobacco Use Construct*

| | **TU → SBP** | **TU → DBP** | **TU → PP** |
|---|---|---|---|
| Average correlation between manifest measures and outcome | $r$ = -0.019 | $r$ = -0.076 | $r$ = 0.043 |
| Optimal ratio (partial:complete data cases) | 1634:66 (24.7:1) | 900:200 (4.5:1) | 1145:155 (7.4:1) |
| SE (Nominal Complete Cases Design) | 0.05332 | 0.05699 | 0.05472 |
| SE (Most Efficient Two-Method Measurement Design) | 0.02610 | 0.04242 | 0.03529 |
| % Decrease in SE | 51 | 26 | 36 |
| Effective N | **1513** | **655** | **872** |
| Effective N Increase Factor | 4.2 | 1.8 | 2.4 |

As with body fat, Effective N Increase Factors correlated inversely with the average effect size between tobacco use and the three blood pressure parameters. The largest Effective N Increase Factor (4.2) was obtained for estimating the association between tobacco use and SBP; correspondingly, of the three blood pressure parameters considered, tobacco use was least strongly associated SBP (average $r$ = -0.019). Tobacco use measures were most strongly correlated with DBP (average $r$ = -0.076); subsequently, the two-method measurement design produced the smallest Effective N Increase Factor (1.8) for estimating the association between tobacco use and DBP.

The present findings indicate that for bivariate models, the strength of association

between the independent and dependent variables is inversely correlated with the increase in statistical power produced by the two-method measurement design. These results are consistent with previous research; Graham and colleagues (2006) demonstrated that the benefit of the two-method measurement design was comparatively greater when the effect size between independent and dependent variables was $r = 0.10$ versus $r = 0.40$.

Effective N Increase Factors provide researchers with the ability to compare the performance of the two-method measurement design across diverse data scenarios and variable sample sizes (Graham et al., 2006); therefore, it was possible to compare the performance of the two-method measurement design for the body fat and tobacco use factors. As indicated by the magnitude of Effective N Increase Factors, in general, the two-method measurement design produced a comparatively greater statistical power advantage for body fat relative to tobacco use. However, because body fat and tobacco use factors differed in several respects (e.g., number of expensive measures; strength of correlation between cheap measures; strength of correlation between cheap and expensive measures; and strength of correlation between manifest measures and outcome variables), it is difficult to determine the impact that these factors make on performance of the two-method measurement design. An area for future research involves using a simulation framework to manipulate these factors, one at a time, to gain a better understanding of how common factor characteristics influence performance of the two-method measurement design.

*Cost-Effectiveness Implications*

Effective Ns for the three body fat regression coefficients highlight the degree of cost-effectiveness produced the two-method measurement design. Under the hypothetical budget of $20,000, it is possible to collect complete body fat data from N=333 participants. However, the two-method measurement design provided the equivalent of N=1513, N=1625, and N=1565 complete cases for testing the effect of body fat on SBP, DBP, and PP, respectively. In order to collect complete data from sample sizes this large (without the use of the two-method measurement design), data collection costs would equal $88,140, $97,500, and $93,900 for testing the effects involving body fat and SBP, DBP, and PP, respectively.

Similarly, for $20,000, complete tobacco use data may be collected from N=363 participants. However, the two-method measurement design provided the equivalent of N=1513, N=655, and N=872 complete cases for testing the effect of tobacco use on SBP, DBP, and PP, respectively. Without the use of the two-method measurement design, to collect complete data from sample sizes this large, data collection costs would equal $83,215, $36,025, and $47,960 for testing the effects involving tobacco use and SBP, DBP, and PP, respectively.

Body fat and tobacco use simulations were performed using expensive to cheap measure cost ratios of 5:1 and 4.5:1, respectively. Thus, these findings suggest that the cost-effectiveness of the two-method measurement design is enhanced as the cost differential between cheap and expensive measures increases. Graham and colleagues (2006) found similar results; as the expensive measure to cheap measure cost ratio increased from 1.6:1 to 10:1, Effective N Increase Factors increased proportionately (Graham et al., 2006). Because the degree of cost-effectiveness produced by the two-method measurement design is correlated with the cost differential between

cheap and expensive measures, the utility of this design is likely enhanced for data scenarios in which researchers would normally be severely limited in the amount of complete case data able to be collected.

*Improving Performance of the Two-Method Measurement Design*

The two-method measurement design produced biased and inefficient parameter estimates when applied to the physical conditioning and dietary intake variables. Several factors likely influenced the comparatively poor performance of the two-method measurement design for these two predictors. Related to physical conditioning, cheap and expensive measures were only modestly correlated (average $r = 0.082$). Despite a strong correlation between cheap measures (PA1 and PA2) ($r = 0.694$), the physical conditioning factor was dominated by $EVO_2$ max; consequently, a large portion of the correlation between the two cheap measures was attributed to response bias (as indicated by bias factor loadings of 0.840 for PA1 and PA2). This created a scenario in which physical conditioning was represented by three manifest measures, only one of which ($EVO_2$ max) loaded highly on the common factor.

The relative lack of association between cheap and expensive physical conditioning indicators may reflect that the measures were indicative of separate, yet related, underlying constructs. PA1 and PA2 – the two cheap physical conditioning measures – focused on participants' physical activity behaviors; $EVO_2$ max, on the other hand, represented participants' level of physical fitness. The lack of power benefits produced by the two-method measurement design may indicate that physical conditioning represented too broad of a construct; perhaps the two-method measurement design would have performed comparatively more efficiently had the

independent variable represented physical activity or physical fitness, rather than a more complex physical conditioning construct. Modifying the physical conditioning factor to represent a narrower physical activity or fitness construct would have required an expensive measure of physical activity (e.g., accelerometers, pedometers) or cheap measures of participants' fitness level (e.g., self-report items assessing cardiorespiratory fitness status), respectively.

Similar to the physical conditioning indicators, cheap and expensive dietary intake measures were weakly correlated (average $r = 0.077$); additionally, unlike the cheap physical conditioning measures, the cheap dietary intake measures (drybeans and dgveg) were only modestly correlated (0.169). Because of the strong association between the expensive dietary intake measures ($r = 0.683$), the common factor was dominated by Min1 and Min2 (both common factor loadings estimated to be 0.827). On the other hand, common factor loadings for drybeans and dgveg were estimated to be substantially lower (both 0.094). The two-method measurement design is configured in such a way that expensive measures are used to model response bias associated with cheap measures; however, for the dietary intake factor, expensive measures were weakly correlated with cheap measures. This created a situation in which the dietary intake was comparatively less stable; resulting biased b-weights reflected the instability of the factor.

The relatively poor performance of the two-method measurement design for the physical conditioning and dietary intake factors has implications for future applications. Findings highlight several opportunities for researchers to maximize the utility of the two-method measurement design. To begin with, manifest measures should be selected carefully and with a specific purpose in mind; specifically, cheap and expensive measures for a given construct

should be highly correlated. It is likely that for a large percentage of two-method measurement design applications, cheap measures will be represented by self-report items; thus, researchers should select self-report measures that are most likely to yield the strongest correlations with expensive measures of the same construct. It is recommended that researchers collect small amounts of data from candidate cheap measures to determine, a priori, the set of cheap measures that best correlates with expensive measure data.

It is also helpful if researchers are able to anticipate, to some degree, effect sizes between independent and dependent variables of interest. As the two-method measurement design has been shown to be especially useful for testing smaller effects, researchers may perform small-scale simulations prior to data collection to more precisely estimate the optimal ratio of partial to complete data for a given effect. This offers researchers the opportunity to more accurately tailor data collection to achieve maximal cost-effectiveness; researchers are able to consider, prior to data collection, for which effects the two-method measurement design will be of maximum utility.

*Strengths*

This dissertation represents the first application of the two-method measurement design involving empirical data. Bivariate models assessed the impact of physical conditioning, dietary intake, body fat, and tobacco use on SBP, DBP, and PP. The four predictors were comprised of a variable number of manifest indicators that differed in strength of intercorrelation; as a result, this dissertation also examined how the absolute number of cheap and expensive measures, as

well as the intercorrelations among manifest indicators, impacted the performance of the two-method measurement design.

Additionally, this dissertation explored several challenges related to real-world implementation of the two-method measurement design. Whereas Graham and colleagues (2006) explored the utility of the two-method measurement design using simulated data, results from this dissertation revealed scenarios for which the two-method measurement design was both more and less efficient than financially-equivalent complete cases designs. Empirical data were obtained from an ongoing large-scale epidemiological survey; thus, the cheap and expensive measures used in this dissertation represented actual, real-world indicators of physical conditioning, dietary intake, body fat, and tobacco use. As a result, findings from this dissertation are highly applicable to several areas of current research. Findings extend the previous state of the current literature regarding the two-method measurement design and provide several guiding principles for researchers considering using a version of the design for future research.

Recent interest in the efficiency of health prevention programs, as well as limited external funding sources, has placed an increased emphasis on cost-effective research within many behavioral health disciplines. When researchers are faced with several measures for assessing particular constructs, they must determine a data collection strategy that balances data quantity and quality with budget constraints. The two-method measurement design helps researchers to spend their resources in the most efficient manner, thus producing the greatest power for testing substantive effects. Across disciplines, it is extremely common that multiple indicators exist for assessing constructs of interest; the methodology presented in this dissertation is relevant for

scenarios in which researchers are limited in the number of data cases that may be collected, but desire greater statistical power for testing effects. Results presented here may yield new efforts to re-organize financial resources for tailored data collection procedures consistent with the two-method measurement design.

*Limitations*

As this dissertation employed empirical data, key measures (including cheap and expensive measures, as well as outcome variables) were limited to those collected as part of the ongoing NHANES surveillance effort. In several instances, cheap and expensive measures were weakly correlated because a wide selection of variables from which to choose ideal manifest measures was lacking. For example, while NHANES data provided a substantial number of total nutrient intake scores from which to select expensive dietary intake measures, data were particularly scarce regarding cheap (self-report) dietary intake measures. Even after selecting cheap measures based on their magnitude of association with expensive measures, intercorrelations remained generally weak between cheap measures (drybeans, dgveg) and expensive measures (Min1, Min2).

Physical conditioning was comprised of manifest measures reflecting two separate, yet related, constructs (i.e., physical activity and physical fitness). NHANES data contained cheap (but not expensive) physical activity measures; likewise, NHANES data included expensive (but not cheap) physical fitness measures. Therefore, the physical conditioning factor was comprised of modestly correlated manifest measures indicative of fundamentally different constructs,

limiting the ability of the two-method measurement design to reliably estimate regression coefficients of substantive interest.

Another potential limitation involves the estimated costs of data collection for cheap and expensive measures. The most efficient two-method measurement design was determined by repeated applications of the design using series of partial to complete data ratios; therefore, the cost differential between cheap and expensive measures had important implications for the number of partial and complete cases permitted under the hypothetical $20,000 budget. It is possible that the estimated costs of data collection used to generate ratios of partial to complete data over- or under-estimated actual data collection costs. Future research will be able to extend the current state of the literature by performing sensitivity analyses to assess how performance of the two-method measurement design varies as a function of the cost differential between cheap and expensive measures.

### *Areas for Future Research*

One important finding from the present study is that, consistent with Graham et al. (2006), the strength of association between the independent and dependent variables is inversely correlated with the statistical power benefit provided by the two-method measurement design. Determining the precise underlying cause for this finding represents an essential area for future research. The two-method measurement design represents a recent statistical advancement in the field of planned missingness; this design has the potential to transform and streamline default data collection methods that have been used for decades. Accordingly, as more information is gained regarding the performance of the two-method measurement design across diverse data

scenarios, researchers will be able to better tailor application of the design to their unique data collection needs.

Future research will involve examining the impact of multiple factors (e.g., cost differential between cheap and expensive measures; number of expensive measures) on the efficiency of the two-method measurement design. Disentangling such factors, and examining their impact individually, will assist in determining data scenarios most appropriately suited for implementation of the two-method measurement design. Additionally, as this dissertation and previous research (Graham et al., 2006) have assessed the performance of the two-method measurement design when applied to bivariate models, future work will assess the utility of the design for multivariate models. This extension also invites multiple other research questions related to design application. For example, in a two-method measurement design involving several independent variables, is there any statistical benefit to having complete case data from the same sample of participants? How does the degree of association between multiple independent variables affect efficiency of estimated parameters under the two-method measurement design?

Future research will also explore the utility of other potential two-method measurement design configurations. The translation of the two-method measurement design to mediation and moderation models offers the opportunity to achieve cost-effectiveness in longitudinal research. The two-method measurement design may also be combined with other planned missingness designs (e.g., the three-form design; (Graham et al., 2006)) to create innovative approaches to data collection that capitalize on current missing data analysis procedures.

The two-method measurement design is highly applicable to multiple research disciplines. Whenever researchers have the opportunity to collect data for a particular construct using several measures of variable cost and construct validity, the two-method measurement design offers the potential for cost-effective data collection and unbiased and efficient parameter estimation. Additionally, as state-of-the-art measures continue to advance across research fields, the cost differential between cheap and expensive measures will likely widen, further emphasizing the need for cost-effective designs. With the ongoing development and refinement of valid measures across research disciplines, the utility of the two-method measurement design will continue to increase.

## References

Allison, P. (1987). Estimation of linear models with incomplete data. *Sociological Methodology, 17*, 71-103.

Anderson, L. F., Tomten, H., Haggarty, P., Lovo, A., & Hustvedt, B. (2003). Validation of energy intake estimated from a food frequency questionnaire: a doubly-labelled water study. *European Journal of Clinical Nutrition, 57*, 279-284.

Appel, L. J., Brands, M. W., Daniels, S. R., Karanja, N., Elmer, P. J., & Sacks, F. M. (2006). Dietary approaches to prevent and treat hypertension: A scientific statement from the American Heart Association. *Hypertension, 47*, 296-308.

Appel, L. J., Sacks, F. M., Carey, V. J., Obarzanek, E., Swain, J. F., Miller, E. R., et al. (2005). Effects of protein, monounsaturated fat, and carbohydrate intake on blood pressure and serum lipids. *Journal of the American Medical Association, 294*(12), 2455-2464.

Arcaro, G., Cretti, A., Balzano, S., Lechi, A., Muggeo, M., Bonora, E., et al. (2002). Insulin causes endothelial dysfunction in humans: Sites and mechanisms. *Circulation, 105*, 576-582.

Arcaro, G., Zamboni, M., Rossi, L., Turcato, E., Covi, G., Armellini, F., et al. (1999). Body fat distribution predicts the degree of endothelial dysfunction in uncomplicated obesity. *International Journal of Obesity and Related Metabolic Disorders, 23*, 936-924.

Arnett, D. K., McGovern, P. G., Jacobs, D. R., Shahar, E., Duval, S., Blackburn, H., et al. (2002). Fifteen-year trends in cardiovascular risk factors (1980-1982 through 1995-1997): The Minnesota Heart Study. *American Journal of Epidemiology, 156*, 929-935.

Barlow, C. E., LaMonte, M. J., Fitzgerald, S. J., Kampert, J. B., Perrin, J. L., & Blair, S. N. (2006). Cardiorespiratory fitness is an independent predictor of hypertension incidence among initially normotensive healthy women. *American Journal of Epidemiology, 163*, 142-150.

Beevers, G., Lip, G., & O'Brien, E. (2001). ACB of hypertension: The pathophysiology of hypertension. *British Medical Journal, 322*, 912-916.

Behall, K. M., Scholfield, D. J., & Hallfrisch, J. (2006). Whole-grain diets reduce blood pressure in mildly hypercholesterolemic men and women. *Journal of the American Dietetic Association, 106*, 1445-1449.

Benowitz, N. L., Hansson, A., & Jacob III, P. (2002). Cardiovascular effects of nasal and transdermal nicotine and cigarette smoking. *Hypertension, 39*, 1107-1112.

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin, 107*, 238-246.

Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin, 88*, 588-606.

Bingham, S. A. (2002). Biomarkers in nutritional epidemiology. *Public Health Nutrition, 5*, 821-827.

BLS. (2004). Occupation report: 2004 median annual earnings.

Brouwer, I., Dusseldorp, M., West, C., Meyboom, S., Thomas, C., Duran, M., et al. (1999). Dietary folate from vegetables and citrus fruit decreases plasma homocysteine concentration in humans in a dietary controlled trial. *The Journal of Nutrition, 129*, 1135-1139.

Brown, A. A., & Hu, F. B. (2001). Dietary modulation of endothelial function: Implications for cardiovascular disease. *American Journal of Clinical Nutrition, 73*, 673-686.

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing stuctural equation models* (pp. 136-162). Newbury Park, CA: Sage.

Campia, U., Sullivan, G., Bryant, M. B., Waclawiw, M. A., Quon, M. J., & Panza, J. A. (2004). Insulin impairs endothelium-dependent vasodilation independent of insulin sensitivity or lipid profile. *American Journal of Physiology: Heart and Circulatory Physiology, 286*, H76-H82.

Carnethon, M. R., Gulati, M., & Greenland, P. (2005). Prevalence and cardiovascular disease correlates of low cardiorespiratory fitness in adolescents and adults. *Journal of the American Medical Association, 294*, 2981-2988.

Casperson, C., Powell, K., & Christenson, G. (1985). Physical activity, exercise, and physical fitness: Definitions and distinctions for health-related research. *Public Health Reports, 100*, 126-131.

CDC. (1999). Decline in deaths from heart disease and stroke, United States, 1900-1999. *Morbidity and Mortality Weekly Report (MMWR), 48*, 649-656.

Cersosimo, E., & DeFronzo, R. A. (2006). Insulin resistance and endothelial dysfunction: The road map to cardiovascular disease *Diabetes/Metabolism Research and Reviews, 22*, 423-436.

Cheitlin, M. D. (2003). Cardiovascular physiology -- changes with aging. *The American Journal of Cardiology, 12*, 9-13.

Chobanian, A. V., Bakris, G. L., Black, H. R., Cushman, W. C., Green, L. A., Izzo, J. L., et al. (2003). Seventh report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of blood pressure. *Hypertension, 42*, 1206-1252.

Chudek, J., & Wiecek, A. (2006). Adipose tissue, inflammation, and endothelial dysfunction. *Pharmacological Reports, 58, suppl.*, 81-88.

Cohen, J. (1977). *Statistical power analysis for the behavioral sciences*. New York: Academic Press.

Contento, I. R., Randell, J. S., & Basch, C. E. (2002). Review and analysis of evaluation measures used in nutrition education intervention research. *Journal of Nutrition Education and Behaviora, 34*, 2-25.

Cook, N. R., Cutler, J. A., Obarzanek, E., Buring, J. E., Rexrode, K. M., Kumanyika, S. K., et al. (2007). Long term effects of dietary sodium reduction on cardiovascular disease outcomes: Observational follow-up of the trials of hypertension prevention (TOHP). *British Medical Journal, 334*, 885-893.

Cortes, B., Nunez, I., Cofan, M., Gilabert, R., Perez-Heras, A., Casals, E., et al. (2006). Acute effects of high-fat meals enriched with walnuts or olive oil on postprandial endothelial function. *Journal of the American College of Cardiology, 48*, 1666-1671.

Dvorak, R. V., Tchernof, A., Starling, R. D., Ades, P. A., DiPietro, L., & Poehlman, E. T. (2000). Respiratory fitness, free living physical activity, and cardiovascular disease risk in older individuals: A doubly labeled water study. *The Journal of Clinical Endocrinology and Metabolism, 85*, 957-963.

Elder, D. A., Prigeon, R. L., Wadwa, R. P., Dolan, L. M., & D'Alessio, D. A. (2006). B-cell function, insulin sensitivity, and glucose tolerance in obese diabetic and nondiabetic adolescents and young adults. *Journal of Clinical Endocrinology and Metabolism, 91*, 185-191.

Emmons, K. M., Abrams, D. B., Marshall, R., Marcus, B. H., Kane, M., Novotny, T. E., et al. (1994). An evaluation of the relationship between self-report and biochemical measures of environmental tobacco smoke exposure. *Preventive Medicine, 23*, 35-39.

Fain, J. N. (2006). Release of interleukins and other inflammatory cytokines by human adipose tissue is enhanced in obesity and primarily due to the nonfat cells. *Vitamins and Hormones, 74*, 443-477.

Ferreira, A. C., Peter, A. A., Mendez, A. J., Jimenez, J. J., Mauro, L. M., Chirinos, J. A., et al. (2004). Postprandial hypertriglyceridemia increases circulating levels of endothelial cell microparticles. *Circulation, 110*, 3599-3606.

Fields, L. E., Burt, V. L., Cutler, J. A., Hughes, J., Roccella, E. J., & Sorlie, P. (2004). The burden of adult hypertension in the United States 1999 to 2000: A rising tide. *Hypertension, 44*, 398-404.

Flegal, K. M., Carroll, M. D., Ogden, C. L., & Johnson, C. L. (2002). Prevalence and trends in obesity among US adults, 1999-2000. *Journal of the American Medical Association, 288*, 1723-1727.

Foster, K. R., & Lukaski, H. C. (1996). Whole-body impedance - what does it measure? *American Journal of Clinical Nutrition, 64 (suppl)*, 388S-396S.

Gong, Z., Agalliu, I., Lin, D., Stanford, J., & Kristal, A. (2007). Obesity is associated with increased risks of prostate cancer metastasis and death after initial cancer diagnosis in middle-aged men. *Cancer, 109*, 1192-1202.

Graham, J. W., Cumsille, P. E., & Elek-Fisk, E. (2003). Methods for handling missing data. In J. A. Schinka & W. F. Velicer (Eds.), *Research methods in psychology: Vol. 2 Handbook of psychology* (pp. 87-114). New York: Wiley.

Graham, J. W., Hofer, S. M., & Piccinin, A. M. (1994). Analysis with missing data in drug prevention research. In L. M. Collins & L. Seitz (Eds.), *Advances in data analysis for prevention intervention research* (pp. 13-63). Washington, D.C.: National Institute on Drug Abuse.

Graham, J. W., Taylor, B. J., Olchowski, A. E., & Cumsille, P. E. (2006). Planned missing data designs in psychological research. *Psychological Methods, 11*(4), 323-343.

Gregg, E. W., Cheng, Y. J., Cadwell, B. L., Imperatore, G., Williams, D. E., Flegal, K. M., et al. (2005). Secular trends in cardiovascular disease risk factors according to body mass index in US adults. *Journal of the American Medical Association, 293*, 1868-1874.

Guo, S. S., Chumlea, W. C., & Cockram, D. B. (1996). Use of statistical methods to estimate body composition. *American Journal of Clinical Nutrition, 64(suppl)*, 428S-435S.

Haider, A. W., Larson, M. G., Franklin, S. S., & Levy, D. (2003). Systolic blood pressure, diastolic blood pressure, and pulse pressure as predictors of risk for congestive heart failure in the Framingham Heart Study. *Annals of Internal Medicine, 138*, 10-16.

Hartung, G. H., Krock, L. P., Crandall, C. G., Bisson, R. U., & Myhre, L. G. (1993). Prediction of maximal oxygen uptake from submaximal exercise testing in aerobically fit and nonfit men. *Aviation, Space, and Environmental Medicine, 64*, 735-740.

Heyward, V. H., & Wagner, D. R. (2004). *Applied body composition assessment* (Second ed.). Champaign, IL: Human Kinetics.

Houtkooper, L. B., Lohman, T. G., Going, S. B., & Howell, W. H. (1996). Why bioelectrical impedance analysis should be used for estimating adiposity. *American Journal of Clinical Nutrition, 64(suppl)*, 436S-448S.

Hu, X., Juneja, S., Maihle, N., & Cleary, M. (2002). Leptin -- a growth factor in normal and malignant breast cells and for normal mammary gland development. *Journal of the National Cancer Institute, 94*, 1704-1711.

Ishikawa-Takata, K., Ohta, T., & Tanaka, H. (2003). How much exercise is required to reduce blood pressure in essential hypertensives: A dose-response study. *American Journal of Hypertension, 16*, 629-633.

Jakulj, F., Zernicke, K., Bacon, S. L., van Wielingen, L. E., Key, B. L., West, S. G., et al. (2007). A high-fat meal increases cardiovascular reactivity to psychological stress in healthy young adults. *The Journal of Nutrition, 137*, 935-939.

Johnson, J. G., Cohen, P., Kasen, S., & Brook, J. S. (2002). Eating disorders during adolescence and the risk for physical and mental disorders during early adulthood. *Archives of General Psychiatry, 59*, 545-552.

Kant, A. K., & Graubard, B. I. (2005). Energy density of diets reported by American adults: Association with food group intake, nutrient intake, and body weight. *International Journal of Obesity, 29*, 950-956.

Kawano, Y., Matsuoka, H., Takishita, S., & Omae, T. (1998). Effects of magnesium supplementation in hypertensive patients: Assessment by office, home, and ambulatory blood pressures. *Hypertension, 32*, 260-265.

Kohlmeier, L., Mendez, M., McDuffie, J., & Miller, M. (1997). Computer-assisted self-interviewing: a multimedia approach to dietary assessment. *American Journal of Clinical Nutrition, 65, suppl.*, 1275S-1281S.

Kotler, D. P., Burastero, S., Wang, J., & Pierson Jr., R. N. (1996). Prediction of body cell mass, fat-free mass, and total body water with bioelectrical impedance analysis: Effects of race, sex, and disease. *American Journal of Clinical Nutrition, 64(suppl)*, 489S-497S.

Kyle, U. G., Genton, L., Karsegard, L., Slosman, D. O., & Pichard, C. (2001). Single prediction equation for bioelectrical impedance analysis in adults aged 20-94 years. *Nutrition, 17*, 248-253.

Lapu-Bula, R., & Ofili, E. (2007). From hypertension to heart failure: Role of nitric oxide-mediated endothelial dysfunction and emerging insighs from myocardial contrast echocardiography. *American Journal of Cardiology, 99*, 7D-14D.

Lee, C. D., Blair, S. N., & Jackson, A. S. (1999). Cardiorespiratory fitness, body composition, and all-cause and cardiovascular disease mortality in men. *American Journal of Clinical Nutrition, 69*, 373-380.

Lewington, S., Clarke, R., Qizilbash, N., Peto, R., Collins, R., & Collaboration, P. S. (2002). Age-specific relevance of usual blood pressure to vascular mortality: A meta-analysis of individual data for one million adults in 61 prospective studies. *The Lancet, 360*, 1903-1913.

Lind, P., Engstrom, G., Stavenow, L., Janzon, L., Lindgarde, F., & Hedblad, B. (2004). Risk of myocardial infarction and stroke in smokers is related to plasma levels of inflammation-sensitive proteins. *Aterioscleroses, Thrombosis, and Vascular Biology, 24*, 577-582.

Liszka, H. A., Mainous, A. G., King, D. E., Everett, C. J., & Egan, B. M. (2005). Prehypertension and cardiovascular morbidity. *Annals of Family Medicine, 3*, 294-299.

Mahmud, A., & Feely, J. (2003). Effect of smoking on arterial stiffness and pulse pressure amplification. *Hypertension, 41*, 183-187.

Mangels, A., Messina, V., & Melina, V. (2003). Position of the American Dietetic Assosication and Dieticians of Canada: Vegetarian diets. *Journal of the American Dietetic Association, 103*, 748-765.

Marshall, J. R. (2003). Methodologic and statistical considerations regarding use of biomarkers of nutritional exposure in epidemiology. *Journal of Nutrition, 133*, 881S-887S.

McKay, D. L., Perrone, G., Rasmussen, H., Dallal, G., & Blumberg, J. B. (2000). Multivitamin/mineral supplementation improves plasma B-vitamin status and homocysteine concentration in healthy older adults consuming a folate-fortified diet. *Journal of Nutrition, 130*, 3090-3096.

Mei, Z., Grummer-Strawn, L. M., Pietrobelli, A., Goulding, A., Goran, M. I., & Dietz, W. H. (2002). Validity of body mass index compared with other body-composition screening

indexes for the assessment of body fatness in children and adolescents. *American Journal of Clinical Nutrition, 75*, 978-985.

Melanson, E. L., Knoll, J. R., Bell, M. L., Donahoo, W. T., Hill, J. O., Nysse, L. J., et al. (2004). Commercially available pedometers: considerations for accurate step counting. *Preventive Medicine, 39*, 361-368.

Mensah, G. A., & Brown, D. W. (2007). An overview of cardiovascular disease burden in the United States. *Health Affairs, 26*, 38-48.

Mensah, G. A., Hokdad, A. H., Ford, E. S., Greenlund, K. J., & Croft, J. B. (2005). State of disparities in cardiovascular health in the United States. *Circulation, 111*, 1233-1241.

Mizia-Stec, K. (2006). Cytokines and adhesive molecules in detection of endothelial dysfunction. *Pharmacological Reports, 58, suppl.*, 21-32.

Mohamed-Ali, V., Goodrick, S., Bulmer, K., Holly, J., Yudkin, J. S., & Coppack, S. W. (1999). Production of soluble tumor necrosis factor receptors by human subcutaneous adipose tissue in vivo. *American Journal of Physiology 277*, E971-E975.

Munter, P., He, J., Cutler, J. A., Wildman, R. P., & Whelton, P. K. (2004). Trends in blood pressure among children and adolescents. *Journal of the American Medical Association, 291*, 2107-2113.

Muthen, B., Kaplan, D., & Hollis, M. (1987). On structural equation modeling with data that are not missing completely at random. *Psychometrika, 52*, 431-462.

Myers, J., Kaykha, A., George, S., Abella, J., Zaheer, N., Lear, S., et al. (2004). Fitness versus physical activity patterns in predicting mortality in men. *The American Journal of Medicine, 117*, 912-918.

Natarajan, S., & Nietert, P. J. (2003). National trends in screening, prevalence, and treatment of cardiovascular risk factors. *Preventive Medicine, 36*, 389-397.

NCHS. (2000). *NHANES: Physician Examination Procedures Manual*. Hyattsville: MD: National Center for Health Statistics.

NCHS. (2001a). *NHANES laboratory procedures manual: Part I*. Hyattsville, MD: National Center for Health Statistics.

NCHS. (2001b). *NHANES laboratory procedures manual: Part II*. Hyattsville, MD: National Center for Health Statistics.

NCHS. (2002a). *NHANES anthropometry procedures manual*. Hyattsville, MD: National Center for Health Statistics.

NCHS. (2002b). *NHANES body composition procedures manual*. Hyattsville, MD: National Center for Health Statistics.

NCHS. (2002c). *NHANES MEC in-person dietary interview procedures manual*. Hyattsville, MD: National Center for Health Statistics.

NCHS. (2003). *NHANES 1999-2000 Data Release: Dietary Interview Component, Total Nutrients Intakes File*. Hyattsville: MD: National Center for Health Statistics.

NCHS. (2004a). *NHANES 2001-2002 Data Release: Dietary Interview Component, Total Nutrient Intakes Files*. Hyattsville: MD: National Center for Health Statistics.

NCHS. (2004b). *NHANES cardiovascular fitness procedures manual*. Hyattsville, MD: National Center for Health Statistics.

Nicholls, S., Lundman, P., Harmer, J., Cutri, B., Griffiths, K., Rye, K., et al. (2006). Consumption of saturated fat impairs the anti-inflammatory properties of high-density

lipoproteins and endothelial function. *Journal of the American College of Cardiology, 48*, 715-720.

Nichols, W. W. (2005). Clinical measurement of arterial stiffness obtained from noninvasive pressure waveforms. *American Journal of Hypertension, 18*, 3S-10S.

Noonan, V., & Dean, E. (2001). Submaximal exercise testing: Clinical application and interpretation. *Physical Therapy, 80*, 782-807.

Osei, K., Rhinesmith, S., Gaillard, T., & Schuster, D. (2004). Impaired insulin sensitivity, insulin secretion, and glucose effectiveness predict future development of impaired glucose tolerance and type 2 diabetes in pre-diabetic African Americans. *Diabetes Care, 27*, 1439-1446.

Pastor-Barriuso, R., Banegas, J. R., Damian, J., Appel, L. J., & Guallar, E. (2003). Systolic blood pressure, diastolic blood pressure, and pulse pressure: An evaluation of their joint effect on mortality. *Annals of Internal Medicine, 139*, 731-739.

Pate, R., Wang, C., Dowda, M., Farrell, S., & O'Neill, J. (2006). Cardiorespiratory fitness levels among US youth 12 to 19 years of age: Findings from the 1999-2002 National Health and Nutrition Examination Survey. *Archives of Pediatric and Adolescent Medicine, 160*, 1005-1012.

Patterson, P. (2000). Reliability, validity, and methodological responses to the assessment of physical activity via self-report. *Research Quarterly for Exercise and Sport, 71*, 15-20.

Perez-Stable, E. J., Benowitz, N. L., & Marin, G. (1995). Is serum cotinine a better measure of cigarette smoking than self-report? *Preventive Medicine, 24*, 171-179.

Peterson, M. J., Czerwinski, S. A., & Siervogel, R. M. (2003). Development and validation of skinfold-thickness prediction equations with a 4-compartment model. *American Journal of Clinical Nutrition, 77*, 1186-1191.

Piers, L. S., Soares, M. J., Frandsen, S. L., & O'Dea, K. (2000). Indirect estimates of body composition are useful for groups but unreliable in individuals *International Journal of Obesity, 24*, 1145-1152.

Plantinga, Y., Ghiadoni, L., Magagna, A., Giannarelli, C., Franzone, F., Taddei, S., et al. (2007). Supplementation with vitamins C and E improves arterial stiffness and endothelial function in essential hypertensive patients. *American Journal of Hypertension, 20*, 392-297.

Qureshi, A. I., Suri, M. F., Kirmani, J. F., & Divani, A. A. (2005). Prevalence and trends of prehypertension and hypertension in the United States: National Health and Nutrition Examination Surveys 1976 to 2000. *Medical Science Monitor, 11*, CR403-409.

Rogers, W. M., & Schmitt, N. (2004). Parameter recovery and model fit using multidimensional composites: A comparison of four empirical parceling algorithms. *Multivariate Behavioral Research, 3*, 379-412.

Sallis, J. F., & Saelens, B. E. (2000). Assessment of physical activity by self-report: Status, limitations, and future directions. *Research Quarterly for Exercise and Sport, 71*, 1-14.

Sampei, M. A., Novo, N. F., Juliano, Y., & Sigulem, D. M. (2001). Comparison of the body mass index to other methods of body fat evaluation in ethnic Japanese and Caucasian adolescent girls. *International Journal of Obesity, 25*, 400-408.

Silverthorn, D. U. (2001). *Human physiology: An integrated approach, 2nd ed.* . Upper Saddle River, NJ: Prentice-Hall.

Singhal, A. (2005). Endothelial dysfunction: Role in obesity-related disorders and the early origins of CVD. *Proceedings of the Nutrition Society, 64*, 15-22.

Staessen, J. A., Gasowski, J., Wang, J. G., Thijs, L., Hond, E. D., Boissel, J., et al. (2000). Risks of untreated and treated isolated hypertension in the elderly: meta-analysis of outcome trials. *The Lancet, 355*, 865-872.

Staffileno, B. A., Braun, L. T., & Rosenson, R. S. (2001). The accumulative effects of physical activity in hypertensive, post-menopausal women. *Journal of Cardiovascular Risk, 8*, 283-290.

Steiger, J. H., & Lind, J. M. (1980). *Statistically based tests for the number of common factors*. Paper presented at the the annual meeting of the Psychometric Society.

Steinberger, J., Jacobs, D. R., Raatz, S., Moran, A., Hong, C., & Sinaiko, A. R. (2005). Comparison of body fatness measurements by BMI and skinfolds vs dual energy x-ray absorptiometry and their relation to cardiovascular risk factors in adolescents. *International Journal of Obesity, 29*, 1346-1352.

Sternfeld, B., Sidney, S., Jacobs, D. R., Sadler, M. C., Haskell, W. L., & Schreiner, P. J. (1999). Seven-year changes in physical fitness, physical activity, and lipid profile in the CARDIA study. *Annals of Epidemiology, 9*, 25-33.

Streppel, M., Arends, L., van't Veer, P., Grobbee, D., & Geleijnse, J. (2005). Dietary fiber and blood pressure: A meta-analysis of randomized placebo-controlled trials. *Archives of Internal Medicine, 165*, 150-156.

Streppel, M., Boshuizen, H. C., Ocke, M. C., Kok, F. J., & Kromhout, D. (2007). Mortality and life expectancy in relation to long-term cigarette, cigar, and pipe smoking: The Zutphen Study. *Tobacco Control, 16*, 107-113.

Subar, A. F., Kipnis, V., Troiano, R. P., Midthune, D., Schoeller, D. A., Bingham, S., et al. (2003). Using intake biomarkers to evaluate the extent of dietary misreporting in a large sample of adults: The OPEN Study. *American Journal of Epidemiology, 158*, 1-13.

Sun, S. S., Chumlea, W. C., Heymsfield, S. B., Lukaski, H. C., Schoeller, D., Friedl, K., et al. (2003). Development of bioelectrical impedance analysis prediction equations for body composition with the use of a multicomponent model for use in epidemioloigc surveys. *American Journal of Clinical Nutrition, 77*, 331-340.

Symons, J. D., Rutledge, J. C., Simonsen, U., & Pattathu, R. A. (2006). Vascular dysfunction produced by hyperhomocysteinemia is more severe in the presence of low folate. *American Journal of Physiology: Heart and Circulatory Physiology, 290*, H181-H191.

Talbot, L. A., Metter, E. J., & Fleg, J. L. (2000). Leisure-time physical activities and their relationship to cardiorespiratory fitness in healthy men and women 18-95 years old. *Medicine and Science in Sports and Exercise, 32*, 417-425.

Tchernof, A., Nolan, A., Sites, C. K., Ades, P. A., & Poehlman, E. T. (2002). Weight loss reduces C-reactive protein levels in obese postmenopausal women. *Circulation, 105*, 564-569.

Teo, K. K., Ounpuu, S., Hawken, S., Pandey, M. R., Valentin, V., Hunt, D., et al. (2006). Tobacco use and risk of myocardial infarction in 52 countries in the INTERHEART study: A case-control study. *The Lancet, 368*, 647-658.

Thompson, D. R., Obarzanek, E., Franko, D. L., Barton, B. A., Morrison, J., Biro, F. M., et al. (2007). Childhood overweight and cardiovascular disease risk factors: The National Heart, Lung, and Blood Institute Growth and Health Study. *Journal of Pediatrics, 150*, 18-25.

Tooze, J. A., Subar, A. F., Thompson, F. E., Troiano, R. P., Schatzkin, A., & Kipnis, V. (2004). Psychosocial predictors of energy underreporting in a large doubly labeled water study. *American Journal of Clinical Nutrition, 79*, 795-804.

Tucker, L. R., & Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika, 38*, 1-10.

Tudor-Locke, C. E., & Myers, A. M. (2001). Challenges and opportunities for measuring physical activity in sedentary adults. *Sports Medicine, 31*, 91-100.

USDHHS. (2002). Physical activity and fitness. In *Healthy People 2010: Objectives for improving health, 2nd ed.* . Washington, DC: U.S. Government Printing Office

Vartiainen, E., Seppala, T., Lillsunde, P., & Puska, P. (2002). Validation of self reported smoking by serum cotinine measurement in a community-based study. *Journal of Epidemiology and Community Health, 56*, 167-170.

Vasan, R. S., Beiser, A., Seshadri, S., Larson, M. G., Kannel, W. B., D'Agostino, R. B., et al. (2002). Residual lifetime risk for developing hypertension in middle-aged women and men: The Framingham Heart Study. *Journal of the American Medical Association, 287*(8), 1003-1010.

Vogel, R. A., Corretti, M. C., & Plotnick, G. D. (1997). Effect of a single high-fat meal on endothelial function in healthy subjects. *American Journal of Cardiology, 79*, 350-354.

Wagenknecht, L. E., Burke, G. L., Perkins, L. L., Haley, N. J., & Friedman, G. D. (1992). Misclassification of smoking status in the CARDIA study: A comparison of self-report with serum cotinine levels. *American Journal of Public Health, 82*(1), 33-36.

Wallenfeldt, K., Hulthe, J., Bokemark, L., Wikstrand, J., & Fagerberg, B. (2001). Carotid and femoral atherosclerosis, cardiovascular risk factors, and C-reactive protein in relation to smokeless tobacco use or smoking in 58-year-old men. *Journal of Internal Medicine, 250*, 492-501.

Warnecke, R. B., Johnson, T. P., Chavez, N., Sudman, S., O'Rourke, D. P., Lacey, L., et al. (1997). Improving question wording in surveys of culturally diverse populations. *Annals of Epidemiology, 7*, 334-342.

Weickert, M. O., Mohlig, M., Schofl, C., Arafat, A. M., Otto, B., Viehoff, H., et al. (2006). Cereal fiber improves whole-body insulin sensitivity in overweight and obese women. *Diabetes Care, 29*, 775-780.

West, S. G. (2001). Effect of diet on vascular reactivity : An emerging marker for vascular risk. *Current Atherosclerosis Reports, 3*, 446-455.

Wilsgaard, T., Schirmer, H., & Arnesen, E. (2000). Impact of body weight on blood pressure with a focus on sex differences. *Archives of Internal Medicine, 160*, 2847-2853.

Wolever, T., Campbell, J. E., Geleva, D., & Anderson, G. H. (2004). High-fiber cereal reduces postprandial insulin responses in hyperinsulinemic but not normoinsulinemic subjects. *Diabetes Care, 27*, 1281-1285.

Wong, N. D., Pio, J. R., Franklin, S. S., L'Italien, G. J., Kamath, T. V., & Williams, G. R. (2003). Preventing coronary events by optimal control of blood pressure and lipids in patients with the metabolic syndrome. *American Journal of Cardiology, 91*, 1421-1426.

Wood, F. G. (2002). Ethnic differences in exercise among adults with diabetes. *Western Journal of Nursing Research, 24*, 502-515.

Yamani, M. H., Wells, L., Massie, B. M., Ammon, S., Der, E., & Prouty, K. (1995). Relation of the nine-minute self-powered treadmill test to maximal exercise capacity and skeletal muscle function in patients with congestive heart failure. *American Journal of Cardiology, 76*, 788-792.

Zevin, S., Saunders, S., Gourlay, S. G., Jacob III, P., & Benowitz, N. L. (2001). Cardiovascular effects of carbon monoxide and cigarette smoking. *Journal of the American College of Cardiology, 38*, 1633-1638.

Zhou, M., Kosaka, H., & Yoneyama, H. (2000). Potassium augments vascular relaxation mediated by nitric oxide in the carotid arteries of hypertensive Dahl rats. *American Journal of Hypertension, 13*, 666-672.

**Vita**

**Allison Elizabeth Olchowski**

Department of Biobehavioral Health
The Pennsylvania State University
315 East Health and Human Development
University Park, Pennsylvania 16802-6509

## EDUCATION

• **Ph.D.,** Biobehavioral Health, 2007
The Pennsylvania State University, State College, PA
Research mentor: Dr. John W. Graham

• **M.S.,** Health Policy & Administration, 2004
The Pennsylvania State University, State College, PA
Research mentor: Dr. E. Michael Foster

• **B.S.,** Biology, 2002 (*Cum laude)*
Mary Washington College, Fredericksburg, VA

## PEER-REVIEWED PUBLICATIONS

• **Olchowski, A.E.**, Foster, E.M., & Webster-Stratton, C. (2007). Examining the differential cost-effectiveness of behavioral health interventions: An assessment of the Incredible Years Series. *Journal of Early and Intensive Behavioral Intervention, 4*(1), 284-304.

• Graham, J.W., Taylor, B.J., **Olchowski, A.E.**, & Cumsille, P.E. (2006). Planned missing data designs in psychological research. *Psychological Methods*, *11*(4), 323-343.

• **Olchowski, A.E.,** Graham, J.W., Beverly, E.A., Dupkanick, C. Cigarette smoking, physical activity, and the health status of college students. Forthcoming in the *Journal of Applied Social Psychology*.

• Foster, E.M., **Olchowski, A.E.,** & Webster-Stratton, C. Is stacking intervention components cost-effective? An Analysis of the Incredible Years Program. Forthcoming in the *Journal of the American Academy of Child and Adolescent Psychiatry.*

• Graham, J.W., **Olchowski, A.E.**, & Gilreath, T. How many imputations are really needed? Some fine tuning of MI theory. Forthcoming in *Prevention Science*.