THE ASSOCIATION OF OVERWEIGHT AND OBESITY WITH INCIDENT HEART FAILURE

Laura R. Loehr MD, MS

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Epidemiology from the School of Public Health.

Chapel Hill
2008

Approved by,

Advisor: Wayne D. Rosamond PhD, MS

Reader: Gerardo Heiss MD, PhD

Reader: Charles Poole ScD

Reader: Patricia P. Chang MD, MPH

Reader: Ann Marie McNeill PhD

ABSTRACT

Laura R. Loehr, MD MS: The association of overweight and obesity with incident heart failure
(under the direction of Wayne D. Rosamond, PhD MS)

Obesity has been identified as a risk factor for heart failure. The importance of a measure of central adiposity (waist-hip ratio) as compared to BMI has not been extensively studied. The increasing prevalence of both obesity and heart failure (HF) make this association an important topic for prevention.

The Atherosclerosis Risk in Communities (ARIC) study is an ongoing bi-racial population-based cohort (45-65 years) from 4 U.S. communities with 14 years median follow-up for incident hospitalized or fatal heart failure. Waist-hip ratio (WHR), waist circumference and BMI were measured at the baseline visit (1987-1989). After exclusion of prevalent HF, missing anthropometry measures, and poorly represented race groups, there were N=8,031 women and N=6,659 men at baseline. BMI was categorized as normal weight (BMI <25 kg/m^2), overweight (BMI 25-30 kg/m^2), and obese (BMI $\geq$ 30 kg/m^2). WHR and waist circumference were divided into gender-specific tertiles. Models were adjusted for demographics, alcohol use, smoking, age, and educational level using Cox proportional hazards modeling. Because HF is a syndrome without clear objective diagnostic criteria, we evaluated the effect of outcome misclassification with Monte Carlo sensitivity analysis. The impact of a hypothetical reduction in obesity and overweight on heart failure incidence was estimated by calculation of the generalized impact fraction.

With stratification by race and gender, the adjusted hazard ratios for the comparison of the highest category of each anthropometric measure (obese) to the lowest were similar and ≥ 1.0. Results from the sensitivity analysis estimated the effect of outcome misclassification was to bias our findings toward the null. Calculation of the generalized impact fraction estimated that a hypothetical 30 % reduction in the prevalence of obesity would reduce incident HF by 6.7 % in the population.

Our findings do not support the superiority of a measure of central adiposity (WHR) over BMI for the prediction of incident HF. A 6.7 % reduction in heart failure, the estimated impact of a 30 % reduction in obesity, would result in 44,220 fewer incident heart failure cases per year. Such evaluations are vital to prioritize and inform future prevention programs regarding the possible impact of their efforts.

ACKNOWLEDGEMENTS

I would like to express my sincere appreciation to Dr. Wayne D. Rosamond, my committee chair, for his enduring patience and excellent mentoring over the last few years. This dissertation would not have been possible without his guidance. I would like to thank my committee members for their many contributions: Drs. Gerardo Heiss, Charles Poole, Ann Marie McNeill, and Patricia P. Chang. In addition, thanks to co-authors Drs. Aaron Folsom and Lloyd Chambless for their contributions to manuscript 1. I have been very fortunate to have such a great team of people to work with. I much appreciate the friendship of everyone in the CVD epidemiology group, and especially Phyllis Johnson.

I want to thank my mother, Doris Ross, my mother-in-law, Elizabeth Loehr, and all my brothers and sisters for their support. I wish to recognize Jennifer Wermig and Tina Muhtar for taking excellent care of my daughter while I worked. But most of all, I want to thank my husband, Steve, and my daughter, Millicent, for their sense of humor, patience, and constant love.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AF	Attributable fraction
ANOVA	Analysis of variance
ARIC	Atherosclerosis Risk in Communities
BMI	Body mass index
BNP	B-type natriuretic peptide
BP	Blood pressure
CAD	Coronary artery disease
CHD	Coronary heart disease
CHS	Cardiovascular Health Study
CI	Confidence interval
CIE	Change in estimate
CLR	Confidence limit ratio
CT	Computed tomography
CVD	Cardiovascular disease
DAG	Directed acyclic graph
FEV_1	Forced expiratory volume in 1 second
GIF	Generalized impact fraction
HERS	Heart and Estrogen/Progestin Replacement Study
HF	Heart failure
HDL-C	High-density lipoprotein cholesterol
HOPE	Heart Outcomes Prevention Evaluation
HR	Hazard ratio

ICD	International Classification of Disease
ICD-9-CM	International Classification of Disease, 9th revision, clinical modification
ICD-10-CM	International Classification of Disease, 10th revision
ICR	Interaction contrast ratio
IR	Incidence rate
JAMA	Journal of the American Medical Association
LDL-C	Low-density lipoprotein cholesterol
LRT	Likelihood ratio test
LVH	Left ventricular hypertrophy
MD	Maryland
MI	Myocardial infarction
MN	Minnesota
MS	Mississippi
MSE	Mean square error
N	Number
NIH	National Institute of Health
NC	North Carolina
NHANES (I)	National Health and Nutrition Examination Survey Epidemiologic Follow-up Study
NPV	Negative predictive value
PPV	Positive predictive value
RCT	Randomized controlled trial
ROC	Receiver operating curve

RR	Relative risk
P	Proportion
PHA	Proportional hazards assumption
PY	Person-years
SD	Standard deviation
SE	Standard error
T1	First tertile
T2	Second tertile
T3	Third tertile
WC	Waist circumference
WHI	Women's Health Initiative
WHR	Waist-hip ratio

CHAPTER I

SPECIFIC AIMS

This study aims to fulfill the following objectives: 1) quantify the association between overweight and obesity as measured by waist/hip ratio, waist circumference, and BMI with incident hospitalized heart failure; 2) compare the magnitude of association for these three anthropometrical measures as to their ability to predict heart failure; 3) assess the probable magnitude and direction of systematic error due to outcome misclassification; 4) determine the effect of outcome misclassification on the association between obesity and heart failure using a semi-automated probabilistic sensitivity analysis; 5) determine the population burden of incident heart failure that could be prevented if there were a hypothetical reduction in the distribution of obesity and overweight.

Specific Aim 1

Assess obesity/overweight as risk factors for the development of incident hospitalized heart failure. This aim will be achieved by the following sub-aims:

1.1: Determine the hazard ratio for the association of obesity (as measured by BMI, waist/hip ratio, and waist circumference) with incident heart failure, while controlling for potential effect modifiers and confounders. Assess race as a potential effect modifier.

1.2: Compare three anthropometrical measures (e.g. BMI, waist circumference, and waist/hip ratio) as to their ability to predict incident hospitalized heart failure.

Specific Aim 2

Assess the probable magnitude and direction of systematic error due to misclassification of the outcome of heart failure for the association of obesity with incident heart failure. This aim will be achieved by the following sub-aims:

2.1: Determine a distribution for sensitivity and specificity as estimated from the literature for the definition of the outcome (incident HF) as defined by hospital discharge and death codes.

2.2: Perform a semi-automated probabilistic sensitivity analysis to estimate the degree of bias due to disease misclassification based on the chosen distribution of sensitivity and specificity. Separate multivariable estimates of the odds ratio and its distribution will be obtained that include systematic error (from disease misclassification), random error and both.

Specific Aim 3

Assess the burden of incident heart failure that could be prevented if there were a reduction in the proportion of obesity and/or overweight. This aim will be accomplished through the use of generalized impact fractions, a generalization of the attributable fraction. We will assume several hypothetical changes to BMI distribution to determine the effect on the burden of heart failure.

3.1 Determine a range for the feasible reduction in the prevalence of obesity and overweight (as measured by BMI), based on the findings from the literature.

3.2 Given several hypothetical scenarios based on the literature for the reduction in prevalence of obesity and overweight, and the magnitude of association between BMI and incident heart failure (as determined from the previous aims), the generalized impact fraction will be determined, overall and stratified race, gender, and age. This will estimate the potential population-level impact of weight reduction on the incidence of heart failure.

3.3 Determine the population attributable fraction for overweight/obese for the outcome of incident hospitalized heart failure. Compare this more commonly used measure to that of the generalized impact fraction.

Rationale

This study adds to the existing literature on obesity and heart failure in several ways. First, the ARIC study will be the largest population-based cohort study to evaluate the association of waist/hip ratio and waist circumference with incident heart failure. Also, since the ARIC cohort is a large biracial study, we will be able to evaluate race and gender as potential effect modifiers of these associations. Second, to address the limitation of disease misclassification, we will use a novel approach to evaluate the effect of systematic error from disease misclassification. Third, we will use the generalized impact fraction as an important tool to predict the population burden of disease due to obesity and overweight that could be prevented given an effective weight reduction intervention program.

CHAPTER II

BACKGROUND AND SIGNIFICANCE

A. Public Health Burden of Heart Failure and Obesity

Heart failure is responsible for more hospitalizations than any other condition in those 65 and older. The temporal trend of heart failure (HF) indicates a steadily increasing population burden (Masoudi, Havranek et al. 2002). Furthermore, the prevalence of heart failure is higher in U.S. blacks than in any other race group in the United States(Brown, Haldeman et al. 2005). The primary risk factor for heart failure is coronary heart disease (CHD). Advances in medical care for CHD has resulted in longer survival time with CHD. Therefore, this emerging heart failure epidemic may partially be the result of the enlarging population with a history of CHD, in addition to the aging of the population in general(Braunwald 1997). Unfortunately, there is a high rate of hospital re-admission(Krumholz, Parent et al. 1997), and a poor prognosis associated with HF, such that between 25-33 % die within the first year following an incident HF hospitalization (Croft, Giles et al. 1999; Schellenbaum, Rea et al. 2004). In 2006, the estimated cost of HF in the United States is projected to be $29.6 billion(Thom, Haase et al. 2006).

Study of the epidemiology of heart failure has been hindered by the lack of a standard definition to define it(Goldberg and Konstam 1999). When present, the symptoms of heart

failure may include shortness of breath, fatigue and lower extremity edema which are not specific to heart failure. Signs of heart failure include reduced left ventricular function which can be assessed in most hospitals by echocardiography; however, the cut-point at which one defines HF is not well standardized. Originally, reduced LV function was thought to be a key component of the definition of HF, now it is understood that many people have HF with preserved LV function(Gottdiener, McClelland et al. 2002).

Risk factors for HF include advancing age, history of coronary heart disease, hypertension, male gender, and valvular heart disease(He, Ogden et al. 2001). Overweight and obesity and its associated conditions, diabetes and even insulin resistance, have also recently been implicated as risk factors for HF(He, Ogden et al. 2001; Kenchaiah, Evans et al. 2002; Ingelsson, Sundstrom et al. 2005; Murphy, Macintyre et al. 2005). This is alarming considering the rapidly increasing prevalence of obesity(2005). Over the last few decades, the prevalence of obesity (BMI $\geq$ 30) in the United States has doubled. Diabetes has shown a similar increase in prevalence which is largely attributed to the obesity trends(Ford, Williamson et al. 1997; Mokdad, Ford et al. 2000). Current estimates from the National Health and Nutrition examination survey (NHANES) show that 31 % of U.S. adults are obese and 65 % are overweight; however these estimates vary across race-gender groups. The prevalence of obesity is highest (49.6 %) in African-American women, followed by 31.3 % in white women, 28.9 % in black men and 28.7 % in white men(2005).

Obesity and overweight as measured by BMI have been identified as risk factors for heart failure from the Framingham Heart Study and other studies (Kenchaiah, Evans et al. 2002; Murphy, Macintyre et al. 2005). Further support for the association of obesity with HF comes from echocardiographic studies in which obesity is an even stronger predictor of left

ventricular mass, a known HF predictor, than hypertension (Drazner, Rame et al. 2004). BMI is the only ponderosity metric to date to be assessed as a HF risk factor in a large population based study; the importance of measures of central adiposity (such as waist circumference and waist-hip ratio) as compared to BMI has yet to be studied in this setting. The increasing prevalence of both obesity and heart failure make this association an important topic for further investigation.

B. Risk Factors for Heart Failure

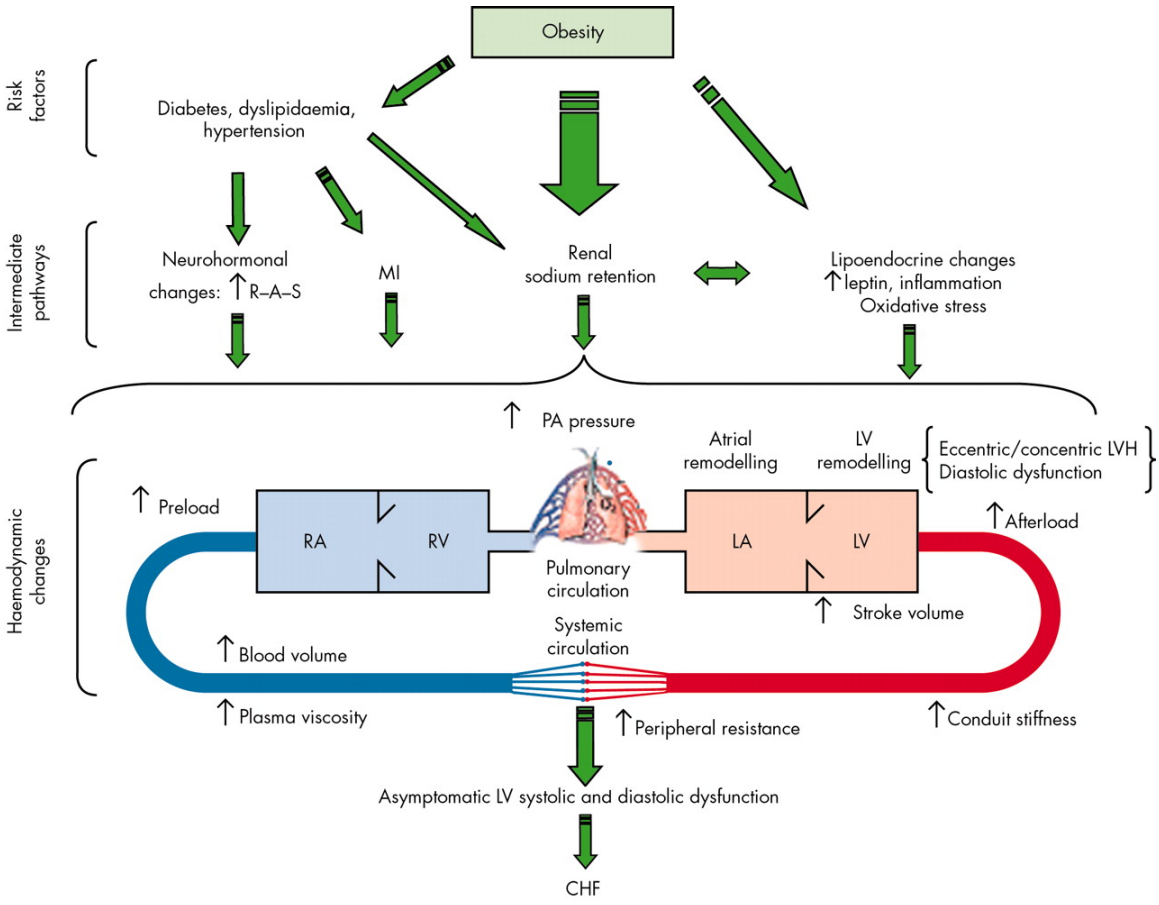
Heart failure is a chronic disease with multiple co-morbid conditions and underlying risk factors. Many of the risk factors associated with HF are preventable. Risk factors identified in the Framingham Heart Study include advanced age, CHD, left ventricular hypertrophy, hypertension, valvular heart disease, diabetes and obesity; with hypertension having the strongest influence on the development of HF(Kannel, D'Agostino et al. 1999; Kenchaiah, Evans et al. 2002). Heart failure risk factors identified from the New Haven, Connecticut cohort for the Established Population for Epidemiologic Studies of the Elderly program were previous myocardial infarction, male gender, older age, diabetes, pulse pressure and BMI $\geq$ 28(Chen, Vaccarino et al. 1999). More recent studies have found that insulin resistance, and retinopathy increase the risk of HF(Ingelsson, Sundstrom et al. 2005; Wong, Rosamond et al. 2005). The First National Health and Nutrition Examination Survey Epidemiologic Follow-up Study (NHANES) estimated the population attributable fraction for the following heart failure risk factors: CHD 61.1 %, smoking 17.1 %, hypertension 10.1 %, low physical activity 9.2 %, male gender 8.9 %, less than high school education 8.9, overweight 8.0, diabetes 3.1 % and valvular heart disease 2.2%(He, Ogden et al. 2001). Studies have been

consistent in finding that many of the HF risk factors are also risk factor for atherosclerotic disease. It is likely that risk factor profiles differ depending on whether the etiology is ischemic or non-ischemic, although there is little data to support this. Furthermore, the distribution of risk factors for HF is believed to vary by race. Results from the National Heart Failure Project found that hypertension and diabetes were more common co-morbid conditions in blacks hospitalized with HF than whites, whereas CHD was a more common co-morbidity in whites(Rathore, Foody et al. 2003).

C. Obesity as a Risk Factor for Heart Failure

Obesity may cause heart failure through both direct and indirect mechanisms. These are displayed in Figure 1 below by Vasan(Vasan 2003). The indirect mechanisms are those in which obesity causes other diseases, i.e. diabetes, hypertension, hyperlipidemia, and renal disease that then result in atherosclerotic disease, ie myocardial infarction or neurohormonal changes thus leading to heart failure. The direct pathway refers to the direct effects of increased adipose tissue on vascular structures and is described in the figure as lipoendocrine changes. Both direct and indirect mechanisms will be discussed below.

Figure 1. Direct and indirect mechanisms through which obesity can lead to heart failure, ***adapted from RS Vasan(2003)(Vasan 2003)***



Direct effects of excess adipose tissue

Independent of co-existing CHD risk factors, cardiac adaptation to excess body fat may result in decreased cardiac function(Poirier, Giles et al. 2006). This has been termed obesity cardiomyopathy. Possibly the first case study of obesity cardiomyopathy was in 1847 by William Harvey; he describes fatty adherences to the heart in an obese man with symptoms of orthopnea before his death(Ford 1950). In 1933, Smith and Willius described the association between obesity and increased heart weight(Smith and Willius 1933). This led to the next discovery that a mechanism through which obesity causes HF is via the development

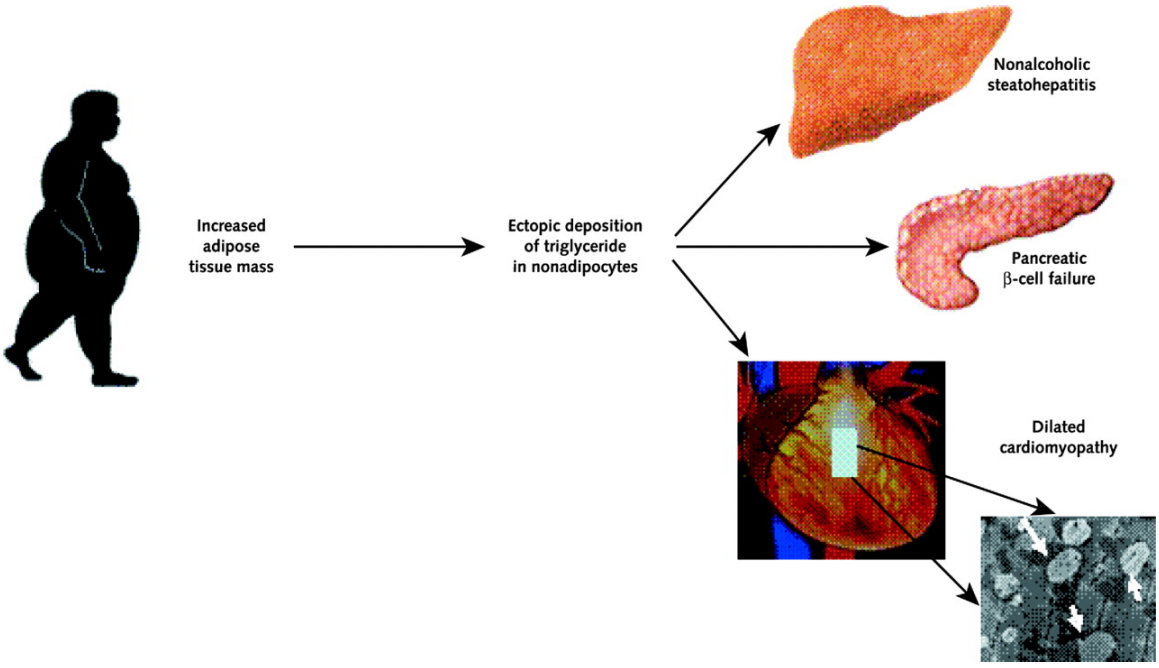
of pulmonary hypertension, and right heart failure(Auchincloss, Cook et al. 1955). The combination of obesity, right heart failure, and sleep apnea was termed the Pickwickian syndrome(Burwell, Robin et al. 1956). For many years, pulmonary hypertension and right heart failure (cor pulmonale) was believed the only mechanism by which obesity causes HF(Alexander 1998); however in 1965 Amad et al found increased heart weight in morbid obesity was primarily due to left ventricular hypertrophy(Amad, Brennan et al. 1965). Hemodynamic studies followed in which obesity was linked to increased plasma volume, increased cardiac output(Alexander, Dennis et al. 1962), increased LV filling pressures and decreased LV compliance (de Divitiis, Fazio et al. 1981); furthermore, improvements in these hemodynamic changes were noted after weight loss(Alexander and Peterson 1972). These findings were then confirmed in echocardiographic studies(Alpert, Terry et al. 1985; Alaud-din, Meterissian et al. 1990; Lauer, Anderson et al. 1991). Specifically, results from the Framingham Heart Study found a positive correlation between left ventricular chamber size as measured by echocardiography and severity of obesity as measured by BMI. Another study reported positive correlations of left ventricular mass with waist/hip ratio and waist circumference (Rasooly, Sasson et al. 1993).

More recently, research from animal models have lent support to the cardiotoxic effects of fat cells(Zhou, Grayburn et al. 2000). A process termed lipotoxicity has been described; it is the disruption of the usual mechanism that regulates triglyceride storage. Normally triglycerides are stored in adipose cells, however when this process is disrupted then triglycerides may deposit in non-adipose cells causing cell death or dysfunction(McGavock, Victor et al. 2006). This process, depicted below in **Error! Reference source not found.**, is also theorized to also cause type 2 diabetes mellitus and nonalcoholic hepatic steatosis.

Another novel hypothesis is that an increase in inflammatory cytokines from excess adipocytes may increase risk of HF(Vasan, Sullivan et al. 2003; McGavock, Victor et al. 2006).

Clinical manifestations of true obesity cardiomyopathy are not that common. It is thought to occur most frequently amongst those with extreme obesity (BMI $\geq$ 40 kg/m^2) of greater than 10 years duration(Kaltman and Goldring 1976). Amongst this population, it is estimated that approximately 10 % develop circulatory congestion(Alexander, Amad et al. 1962). Therefore, the majority of heart failure associated with pre-existing obesity may be through indirect pathways.

Figure 2. Depiction of lipotoxicity, ***adapted from McGavock JM (2006)(McGavock, Victor et al. 2006)***



Indirect effects of excess adipose tissue

More commonly, however, obesity leads to co-morbid conditions such as diabetes and hypertension that are also risk factors for heart failure. These co-morbid conditions are also known CVD risk factors; therefore, it is likely that a common pathway from obesity to heart failure is through these intermediary diseases which then cause coronary heart disease and ischemic heart failure. However, several of these secondary diseases have been found to cause heart failure independent of CHD. For example, obstructive sleep apnea causes pulmonary hypertension and resulting right heart disease which can lead to heart failure. Also, diabetes mellitus, insulin resistance and the metabolic syndrome have all been identified as independent causes of heart failure. Hypertension is known to cause left ventricular hypertrophy which can lead to heart failure.

D. Studies of Obesity as a Risk Factor for Heart Failure

Studies investigating the association of obesity with HF are reviewed in the following section. See Table 1, for a summary of these studies. Most of these studies are in primarily white populations, or specific populations (ie, clinical trial participants with specific exclusion criteria), and BMI is the most commonly used anthropometric measure.

The Framingham Heart Study

In a study from the Framingham Heart Study, Kenchaiah et al found that overweight and obesity was associated with an increased risk for the development of HF(Kenchaiah, Evans et al. 2002). Weight was measured at the baseline visit by BMI. Diagnosis of heart failure during follow-up was based on continuous surveillance for cardiovascular events followed by

committee adjudication using the Framingham Heart Study criteria for HF. There were 5,881 participants with a median follow-up of 14 years (mean age 55 years) and 496 cases of incident HF. BMI was analyzed as both continuous and categorical. After adjustment for potential confounders, the hazard ratio for incident heart failure was 1.06 (1.04-1.09 kg/m^2) for a one unit change in BMI. The multi-variable adjusted HR for HF for overweight (BMI 25 – 29.9 kg/m^2) and obese (BMI ≥30 kg/m^2) as compared to normal weight (BMI 18.5 – 24.9 kg/m^2) were 1.34 (1.08-1.67 kg/m^2) and 2.04 (1.59-2.63 kg/m^2), respectively. The results were similar for models with BMI and covariates treated as time-varying variables.

Effect measure modification was observed for both hypertension and myocardial infarction at baseline. They found that the HR for the trend across categories of BMI was lower in those with hypertension (HR = 1.30, 1.11-1.52) than in those without hypertension, (HR = 1.66, 1.33-2.07). For those with myocardial infarction at baseline (N=148), there was no effect across categories of BMI (0.80, 0.5-1.30) for incident HF, although they were underpowered to detect a HR of 1.5 or less. For those without myocardial infarction, the HR was 1.5 (95 % CI of 1.31- 1.71). No effect measure modification was found for the following variables: age, smoking status, gender, alcohol use, diabetes mellitus, or valvular heart disease.

The strength of this study is that it is extremely well characterized; such that the definition of heart failure is as close to a gold standard as there exists. This cohort is almost exclusively white and upper class. The size of the study is large, however may not be large enough for the inferences regarding effect modification. Effect modification was observed for hypertension and myocardial infarction. Unfortunately, BMI was the only metric of adiposity in this study.

First National Health and Nutrition Examination Survey (NHANES I) Epidemiologic Follow-up Study

The First National Health and Nutrition Examination Survey (NHANES I) studied the association of overweight as measured by BMI and incident heart failure(He, Ogden et al. 2001). There were 1,382 heart failure cases from 13,643 participants, age 25 to 74 years, followed for a mean of 19 years. Heart failure was defined by the presence of an ICD-9 code '428.0' or '428.9' from a hospitalization or death certificate. Overweight was defined as BMI $\geq$ 27.8 for men and $\geq$ 27.3 for women. The adjusted HR for overweight was 1.23 (1.00-1.52) in men and 1.34 (1.10 – 1.64) in women. These models were adjusted for age, race and time-dependent history of CHD. They determined that the population attributable fraction of HF due to overweight is 8 % (5.6 % for men and 9.6 % for women).

This strength of this study is the large size; however, it is limited by the single type of anthropometric measurement and the definition of heart failure by ICD codes. Potential effect modification by gender is an important finding.

Health, Aging and Body Composition Study

The Health, Aging and Body Composition Study (Health ABC) is a longitudinal cohort study of those aged 70-79 without CHD(Nicklas, Cesari et al. 2006). The association of multiple metrics of overweight/obesity and body composition with future HF events was studied. Of 2,435 participants with 6.1 years of follow-up, there were 166 cases of HF. A diagnosis of heart failure was based on surveillance for heart failure followed by physician adjudication based on specific HF criteria. Criteria for HF were met if any of the following

were present: 1) physician diagnosis of HF and treatment with diuretics and digoxin or a vasodilator, 2) pulmonary edema and cardiomegaly on chest x-ray, or 3) decreased systolic function by echocardiography or ventriculography. This study had multiple measures of body composition with dual-energy x-ray absorptiometry and computed tomography to measure visceral adipose tissue; in addition to traditional anthropometric measures of BMI, waist circumference and waist/thigh ratio were measured. All of the adiposity measures were significant predictors of heart failure in adjusted Cox models. In addition, models for waist circumference included BMI and fat mass still found waist circumference an independent predictor of HF. They also found a significant interaction for gender in terms of the waist to thigh ratio. In men the waist to thigh ratio was significantly associated with HF (HR = 1.33, 1.11-1.60), whereas in women it was not (HR = 1.06, 0.86-1.30).

Additional adjusted models considered both overweight/obese (BMI $\geq$ 25 kg/m^2 or not) and high waist circumference ($\geq$ 102 cm for men, $\geq$ 88 cm for women) as dichotomous variables. There was no significant association of overweight/obese with heart failure with BMI measured with this cut-point, however there was a significant association of high waist circumference with heart failure (HR = 1.91, 95 % CI = 1.32 – 2.75). This relationship remained significant after adjustment for obesity status. The authors concluded that it is the location of fat tissue that is important in predicting future HF; specifically, waist circumference was the most robust predictor of HF.

This is a very well characterized cohort in terms of the variety of anthropometric and body fat distribution measurements, however this study is not generalizable to those with coronary heart disease, as they were excluded at baseline. Unfortunately, a large proportion of those who develop heart failure have pre-existing CHD; this may be one reason for the limited

sample size (N = 166). Although approximately 50 % of the participants were black, it is unlikely that the study was adequately powered to make inferences by race and gender. Also, this study only included those between 70-79 years of age, therefore these associations may not be generalizable to a younger age group.

The Renfrew-Paisley Study

A community based study from Renfrew and Paisley, Scotland observed 15,402 participants aged 45-64 years beginning in 1972(Murphy, Macintyre et al. 2005). Heart failure was defined by ICD codes (ICD 425.4, 425.5, 428, 402) from hospital and death records, and was not necessarily incident. BMI was the only anthropometric measure. They found 641 HF cases over 20 years of follow-up. The crude HR for a one unit increase in BMI is 1.06 (1.04-1.08). In an adjusted model including diabetes, cholesterol level, and hypertension, obesity (BMI $\geq$ 30 kg/m^2 as compared to BMI 18.5-24.9 kg/m^2) was an independent risk factor for men 2.16 (1.57-2.57) and less so for women 1.37 (1.00-1.88).

This study is limited by the single type of anthropometric measurement, and the homogeneity of this Scottish population. In addition, HF was necessarily incident. As in the NHANES study above, they found effect modification by gender.

Uppsala Longitudinal Study of Adult Men

This community based cohort from Uppsala, Sweden investigated both insulin resistance and obesity as predictors of incident HF(Ingelsson, Sundstrom et al. 2005). There were 104 incident hospitalized heart failure cases from 1,187 elderly men (70 years and older) over a median follow-up of 8.9 years. Heart failure cases were initially identified by an ICD code

('428', 'I50' or '111.0') and then validated with physician adjudication using the classification scheme from the European Society of Cardiology(1995). Insulin sensitivity was measured with the euglycemic insulin clamp technique and anthropometric variables were BMI and waist circumference. Hazard ratios were estimated for a one standard deviation change. Multivariable models were adjusted for diabetes, prior acute MI, hypertension, smoking, left ventricular hypertrophy by electrocardiography, and serum cholesterol. They found that BMI (HR 1. 37, 95 % CI = 1.12-1.68) and waist circumference (HR 1.40, 95 % CI = 1.13-1.74) were important predictors of HF in a multivariable model, however this relationship was no longer significant when clamp glucose disposal rate was included in the model (HR for BMI = 1.17, 95 % CI = 0.92-1.50) and (HR for waist circumference = 1.18, 95 % CI = 0.88-1.53). They found similar results sub-samples without diabetes and without obesity and in models that included interim MI.

This study was in a very specific population, elderly Swedish men, therefore it is uncertain if these results are generalizable to women, younger age groups or other ethnic groups. This is the largest study to date to report on insulin sensitivity as measured by a euglycemic clamp and heart failure. The results of this study imply that the relationship between obesity and heart failure is largely mediated by insulin resistance.

Other studies

The New Haven, Connecticut cohort of the Established Population for Epidemiologic Studies of the Elderly program reported on the association of obesity in 173 heart failure cases from 1,749 participants free of CHD; however BMI was based on self-report of height and weight(Chen, Vaccarino et al. 1999). Heart failure was based on applying specific

criteria to the review of medical records with HF code. The HR for categorized BMI with baseline risk factors in the model was of borderline significance. This study is limited by the self-reported BMI.

Predictors of HF were studied amongst women (N =2,391) participating in the Heart and Estrogen/Progestin Replacement Study (HERS)(Bibbins-Domingo, Lin et al.). The HERS is a randomized controlled trial of hormone therapy in women with known CHD. HF was defined by hospitalization or death as discovered from routine surveillance and then validated by committee adjudication. There were 237 cases over a mean follow-up of 6.3 years. The adjusted HR for obese participants (BMI >35 kg/m^2) compared to those of normal weight (BMI 18.5 - 25 kg/m^2) was 1.9 (95 % CI = 1.1-3.0). This study is small and includes only women with known CHD. Also, this is a population participating in a randomized controlled trial which is likely healthier than the general population.

The Heart Outcomes Prevention Evaluation (HOPE) study is a randomized controlled trial assessing the use of an angiotensin-converting enzyme inhibitor (ramipril), vitamin E, or their combination in reducing CVD events(Dagenais, Yi et al. 2005). This study only included high risk patients for CVD events; participants must have had CVD or diabetes plus one other CVD risk factor. The population (N=8,802) was 75 % male. There were 297 HF cases over a median follow-up of 4.5 years. HF defined as a hospitalization with clinical or radiographic signs of congestion. They did not find that BMI or waist/hip ratio were significant predictors of heart failure in a fully adjusted model; however when stratified by gender, waist hip ratio was a significant predictor of heart failure in women only (HR = 2.30, 95 % CI = 1.25-4.21). Also, the hazard ratio for the highest tertile of waist circumference compared to the lowest tertile was 1.38 (95 % CI = 1.03- 1.85); however when BMI was

added to the model these findings were no longer significant. There are two main limitations of this study which reduce it's generalizability: 1) the use of tertiles to define cut-points for the main exposures are not recommended especially when recommended category boundaries are available(Greenland and Rothman 1998), and 2) the population is high-risk male participants in an randomized controlled trial (RCT).

Summary

Nearly all of these studies find a significant association between obesity and heart failure. Only three studies have waist circumference in addition to BMI; however, these particular studies are not population-based(Dagenais, Yi et al. 2005; Ingelsson, Sundstrom et al. 2005; Nicklas, Cesari et al. 2006), and instead are in very specifically defined populations. Waist/hip ratio was an anthropometric measure in only one study(Dagenais, Yi et al. 2005), although a similar measure of waist/thigh ratio was used in another study(Nicklas, Cesari et al. 2006). Future studies should assess BMI and WHR and compare their predictive ability. Also, most of these populations are primarily white, such that inferences by race were not possible. Gender (He, Ogden et al. 2001; Murphy, Macintyre et al. 2005; Nicklas, Cesari et al. 2006) was found to be an effect modifier; therefore, we have stratified by gender in this dissertation.

Table 1. Summary table of studies of obesity as a predictor of heart failure

Author, year	Study Population	Anthropometric Measures	Outcome Definition	Limitations
Kenchaiah, 2002	491 cases from 5,881 participants, cohort study from Framingham, Mass., mostly white, mean age 55	BMI	Adjudicated incident heart failure	BMI only metric, largely white cohort
Chen, 1999	173 cases from 1,749 participants from the New Haven, Connecticut cohort in the Established Population for Epidemiologic Studies of the Elderly, community based cohort, aged 65 years and older, free of CHD, 18 % black participants	BMI, based on self report of height and weight	Incident hospitalized heart failure based on chart review	Self reported BMI, mostly white cohort
Nicklas, 2006	166 cases from 2,435 participants in the Health ABC (Aging and Body Composition) study, cohort study of those 70-79 years without CHD, biracial (black/white)	BMI, waist circumference, waist/thigh ratio, dual-energy x-ray absorptiometry, CT scans	Adjudicated incident heart failure	Limited sample size, free of CHD at baseline
He, 2001	1,382 cases from13,643 participants in First National Health and Nutrition Examination Survey (NHANES), 25 to 74 years of age	BMI	First listing of ICD-9 code '428' from a hospitalization or UCOD on death certificate	BMI only metric, non-adjudicated definition for HF
Murphy, 2005	641 cases from 15,402 participants in the Renfrew-Paisley study, 45-64 years of age from Scotland	BMI	Heart failure ICD codes from a hospitalization or death certificate	BMI only metric, non-adjudicated definition for HF
Bibbins-Domingo, 2004	237 cases from 2,391 women with known CHD participating in the Heart and Estrogen Replacement Study (HERS), randomized controlled trial of estrogen replacement therapy	BMI	Adjudicated incident heart failure	BMI only metric, only women, RCT, small sample
Dagenais, 2005	297 cases from 8,802 participants in the Heart Outcomes Prevention Evaluation (HOPE) study, high risk patients for CVD events, 75 % male, randomized controlled trial of ACE inhibitor (ramipril) and vitamin E	BMI, waist circumference, waist/hip ratio	Adjudicated incident heart failure	RCT, included only high risk patients
Ingelsson, 2005	104 cases from 1,187 participants in the Uppsala longitudinal study of adult men, a community based observational cohort from Sweden, 70 years and older	BMI, waist circumference	First hospitalized heart failure, adjudicated	All men, largely white cohort

E. Validation Studies of ICD codes to Define Heart Failure

Hospital discharge codes are based on the *International Classification of Diseases (ICD)* system which was created primarily for administrative rather than epidemiologic purposes. Despite this, ICD codes are often used in large epidemiologic studies to define disease, when the gold standard may not be feasible(He, Ogden et al. 2001; Murphy, Macintyre et al. 2005). Unfortunately, the gold standard of committee adjudication is expensive and time consuming; also, there are not clearly agreed upon criteria to define heart failure(Goldberg and Konstam 1999). One reason for misclassification from the use of ICD codes are their linkage to hospital reimbursement. In particular, heart failure as a complication during a hospitalization can significantly increase the Medicare reimbursement associated with that hospitalization(Psaty, Boineau et al. 1999). Therefore, there is a financial incentive for hospitals to "up-code" or to list a heart failure ICD codes. As a result, studies have tried to determine the amount of misclassification of heart failure when diagnosed by ICD codes. These validation studies are reviewed below (see summary in Table 2).

Corpus Christi Heart Project

The Corpus Christi Heart Project is a population-based study with the main goal of surveillance for hospitalized coronary heart disease; therefore, these hospitalizations included those for definite and possible myocardial infarction, aortocoronary bypass surgery, and transluminal coronary angioplasty(Goff, Pandey et al. 2000). This data was used to determine the frequency and validity of ICD codes for heart failure. Surveillance took place in Nueces County, Texas with a population of 291,145 in 1990 for which there were 5,083 hospitalizations with possible heart failure. Heart failure cases were validated based on

criteria of a physician diagnosis of acute heart failure or evidence of pulmonary edema on chest x-ray.

The overall prevalence of a heart failure ICD code was (1197/5083) 23.5 %, whereas the most frequent heart failure ICD code was '428' with a prevalence of 20.4 % (1,035/5,083), this was followed by 402.x1 (hypertensive heart disease with congestive heart failure) with a prevalence of 2.6 %. All other heart failure ICD codes had a prevalence of less than 1 %. The test characteristics of code '428' were as follows: sensitivity = 62.8 % (864/1376), specificity = 95.4 % (3536/3707), positive predictive value = 83.5 % and negative predictive value = 87.4 %; whereas, the test characteristics for any HF ICD code were: sensitivity = 67.1 % (923/1376), specificity = 92.6 % (3433/3707), positive predictive value = 77.1 %, and a negative predictive value = 88.3%. The sensitivity was higher, but the specificity was lower when all of heart failure ICD codes were used.

This study is a large population-based validation study of ICD codes for heart failure in those with a co-existing ICD code for coronary heart disease; therefore, it is likely that the resulting test characteristics are not generalizable those without CHD. It is notable that ICD code '428' is the most frequently used HF code and it is more specific than when combined with all other HF codes.

The Cardiovascular Health Study

In the Cardiovascular Health Study (CHS) 5,888 adults 65 years and older were followed from 1989 to 2000 during which there were 1,209 incident heart failure hospitalizations (1,072 had an ICD-9 diagnostic code for HF)(Schellenbaum, Heckbert et al. 2006). They used the following ICD-9 discharge codes to identify potential cases: 428 (heart failure),

398.91 (rheumatic heart failure), 425 (cardiomyopathy), 402.01, 402.11, 402.91 (hypertension with heart failure), and 997.1 (cardiac failure postoperatively). They only included those that survived hospitalization. Of these discharge codes, code '428' was the most frequent at 70.6 percent of the events followed by code 997.1 for "cardiac failure postoperatively" (11.6 %). These potential cases were then validated with committee adjudication using criteria created by the CHS. The CHS criteria required a physician diagnosis of heart failure in addition to supporting evidence of any of the following: heart failure signs and symptoms, pulmonary edema on chest x-ray or evidence of treatment for heart failure.

Using only the ICD-9 code '428', there were 523 cases that were adjudicated as heart failure, leaving 364 individuals with a code '428' that were not considered heart failure after adjudication. The test characteristics of a code '428' by the CHS criteria are as follows: sensitivity = 0.71, specificity = 0.925, positive predictive value = 0.59 and negative predictive value = 0.96. Mortality was selected as a surrogate endpoint for predictive validity of these HF identification methods. Mortality was not significantly different across these two methods of HF event definition. Also notable considering the proposed dissertation, BMI did not vary across the type of event definition (ie adjudication vs. discharge diagnosis).

This is a well done validation study from a large population-based cohort. The only issues that decrease generalizability to other population-based studies are the older age group (65 years and older) and the exclusion of those who died during hospitalization. It is interesting that mortality did not vary by method of case definition. This implies that the distinction between validated and non-validated groups is not clinically relevant.

Olmsted County, Minnesota

A study from Olmsted County, Minnesota primarily on incidence and survival also tested the accuracy of ICD codes for a heart failure by validation with 2 sets of criteria, Framingham and specific clinical criteria.(Roger, Weston et al. 2004) This is community-based open cohort with 22 years of follow-up. There were a large proportion of potential cases (26 %) that were outpatient only. They found 7,298 (80%) of those identified as potential cases had an ICD code '428', whereas only 1,877 potential cases were identified by other HF codes without an accompanying '428' (these other codes were 402.01, 'hypertensive heart disease malignant with congestive heart failure', 402.11, 'hypertensive heart disease benign with congestive heart failure', 425, 'cardiomyopathy', 429.3, 'cardiomegaly', and 514, 'pulmonary congestion'). First, they validated these potential cases by applying the Framingham criteria using committee adjudication. They found that 82 % of potential cases with a '428' code met Framingham criteria, whereas only 14-30 % of the other HF ICD codes (without an associated '428') met Framingham HF criteria. Next, they validated with clinical criteria of a physician diagnosis of heart failure. In this case, 90% of code '428' had a physician diagnosis of heart failure, whereas only 14-36 % of the other codes (without a '428') met clinical criteria. They found that secular trends for incidence and survival were similar between the Framingham and clinical criteria.

This study provided estimates of the positive predictive value for two sets of criteria from an open cohort. The Framingham criteria identified fewer cases than the clinical criteria. Positive predictive value was estimated for code '428' and all codes combined. The inclusion of outpatient cases is unique amongst these validation studies.

Women's Health Initiative (WHI)

A study from the observational cohort of the Women's Health Initiative (WHI) tested the accuracy of diagnostic codes to local adjudication(Heckbert, Kooperberg et al.). Heart failure was defined by the following criteria: symptoms and signs of HF and either pulmonary edema by chest x-ray, documented dilated ventricle or decreased ventricular function, or physician diagnosis of HF and receiving medical treatment(Curb, McTiernan et al. 2003). There were 93,676 post-menopausal women in the observational cohort of which 1,241 were hospitalized with an ICD 9 code '428'; of these codes, only 603 were validated as HF by local adjudication using a standard definition developed by the WHI. The positive predictive value of an ICD code 428 was estimated as 48.6 %. They also validated ICD code 425 (N = 134) and found the PPV was a little lower at 45 %. The main limitation of this study was that participants were all female; in addition, participants were mostly upper class and from urban areas. Compared to other validation studies, this one was not population-based.

National Registry for Atrial Fibrillation

A study from the National Registry for Atrial Fibrillation II evaluated the accuracy of ICD codes for co-morbid conditions including HF(Birman-Deych, Waterman et al. 2005). This registry includes patient information from 23,657 Medicare beneficiaries representing 3,586 hospitals in all 50 states. Mean age was 79 years of age. There were 11,014 heart failure cases based on medical record review for both chronic and current heart failure. They included any of the following ICD-9 codes in any position in the hospital claims data: 428.x, 398.91, 402.01, 402.11, 402.91, 404.01, 404.11, 404.91, 404.03, 404.13, and 404.93. They

estimated the following test characteristics based on current or past HF events: sensitivity = 0.76, specificity = 0.97, positive predictive value = 0.97 and negative predictive value = 0.74. When past HF was excluded, then the specificity (0.86) and PPV (0.85) were lower, and the sensitivity was higher (0.83).

One would expect a higher positive predictive value for a heart failure ICD code in a population with atrial fibrillation, since this arrhythmia is even a component of some HF diagnostic criteria(Eriksson, Caidhal et al. 1987). Unfortunately, test characteristics were not presented for ICD code '428' only.

Summary

These five validation studies provide insight into misclassification rates of ICD codes for heart failure. Unfortunately, there is no gold standard to define heart failure; therefore, the diagnostic criteria vary by study. In the proposed dissertation, there are no internal validation studies from which to develop these estimates. Four of the above studies found that ICD code '428' was by far the most frequently documented ICD code for heart failure(Goff, Pandey et al. 2000; Heckbert, Kooperberg et al. 2004; Roger, Weston et al. 2004; Schellenbaum, Rea et al. 2004; Schellenbaum, Heckbert et al. 2006). In fact, most of the other heart failure ICD codes occurred with a co-existing '428'. In addition, an ICD code '428' was more specific than when all of the ICD codes were included; therefore, the proposed study only included ICD code '428' as diagnostic criteria for heart failure. Estimates of sensitivity and specificity from these validation studies will be used in a sensitivity analysis for disease misclassification.

The proposed study will use the ARIC population, which is a large population based cohort of men and women between 45 and 64 years of age. The above validation studies differ from the ARIC cohort in the following ways: 1) the Corpus Christi Heart Study and the National Registry for Atrial Fibrillation are high-risk populations; 2) the WHI include only women; 3) Olmsted County includes a proportion of outpatient cases; and 4) the CHS is an older population. The CHS followed by the Corpus Christi Heart Study has the most similarities to the ARIC study, and provides estimates of sensitivity and specificity for an ICD code '428'; therefore, the results of this studies would receive the highest priority in selecting estimates for a sensitivity analysis.

Table 2. Summary table of validation studies of ICD codes to define heart failure

First author, year	Study Population	Validation criteria	Results
Goff, 2000	Corpus Christi Heart Project, included hospitalizations for definite or possible MI, aortocoronary bypass surgery, and transluminal coronary angioplasty, population-based, 5,083 hospitalizations with possible heart failure	Physician diagnosis of acute HF or pulmonary edema on chest x-ray	For ICD code '428': Sensitivity: 62.8 % (864/1376) Specificity: 95.4 % (3536/3707) PPV: 83.5 % (1035) NPV: 87.4 % (4048)
Schellenbaum, 2005	Cardiovascular Health Study, population-based study, 65 years and older, 1,209 with possible incident heart failure	Physician diagnosis of HF, in addition to any of the following: HF signs or symptoms, pulmonary edema on chest x-ray or evidence of treatment for HF	For ICD code '428': Sensitivity: 71 % Specificity: 93 % PPV: 59 % NPV: 96 %
Roger, 2004	Olmsted county, Minnesota, population-based open cohort, included 26 % outpatient only cases, 4,537 possible incident heart failure cases	Tested 2 criteria Framingham and Physician diagnosis of HF	For ICD code '428': Framingham criteria: PPV: 82% Clinical criteria: PPV: 90%
Heckbert, 2004	Women's Health Initiative, 93,657 post-menopausal women in the observational cohort, 1,241 participants with a code '428'	Symptoms/signs of HF and either pulmonary edema by chest x-ray, documented dilated ventricle/decreased ventricular function, or physician diagnosis of HF and receiving medical treatment	For ICD code '428': PPV: 49 % For all ICD codes: Sensitivity: 79 % PPV: 45 %
Birman-Deych, 2005	National Registry for Atrial Fibrillation II, 23,657 Medicare beneficiaries with an ICD code for atrial fibrillation, incident and recurrent HF	Chart review for mention of current or history of heart failure	For all HF ICD codes: Sensitivity: 76 % Specificity: 97 % PPV: 97 % NPV: 74 %

F. Methods to Adjust for Bias

Some degree of misclassification is common in observational studies; however, bias due to misclassification is rarely quantified. The conventional approach to the presentation of uncertainty in the scientific literature is the 95 % confidence interval; although, confidence intervals only quantify residual random error(Poole 2001). Unfortunately, the confidence interval is often misinterpreted to represent all sources of error (Greenland 2001). Currently, the standards for publishing in most journals do not include the need to go beyond the uncertainty of random error. Sensitivity analyses attempt to quantify the effect of bias on the results of a study. These techniques have not yet reached the mainstream; however, several epidemiologists have called for a more thorough presentation of the uncertainty inherent in scientific research(Maclure and Schneeweiss 2001; Greenland 2005). In fact, a journal named 'Epidemiologic Perspectives & Innovations' was created in 2004 inviting the exploration of such issues(Phillips, Goodman et al. 2004). The editors request submission of research that expresses "full and proper disclosure of uncertainty in study results" and explores "decision making in the face of this full disclosure"(Maldonado and Phillips 2004).

A simple sensitivity analysis for misclassification is to back-calculate the expected results given a plausible set of specific estimates for sensitivity and specificity(Greenland 1998). This method does not consider the likelihood of each set of values(Phillips 2003). In addition, the presentation of such results requires a table rather than a more succinct summary. Recently computer intensive methods have been created using Monte Carlo type sensitivity analyses (Lash and Fink 2003; Fox, Lash et al. 2005). One such software program uses SAS software to iteratively change individual level data for the biased parameter, while maintaining covariate data such the conventional regression analysis can be performed. This

type of incorporation of uncertainty into regression analysis is called a Monte Carlo Risk Assessment (MCRA) or Monte Carlo uncertainty analysis. Such an analysis requires the statement of a prior distribution for each sensitivity parameter which should include an explicit rationale for the priors(Greenland 2001). More specifically, an MCRA can iteratively sample for multiple pairs of sensitivity parameters from the input distribution, then create modified datasets for each set of parameters, and then re-run the conventional regression for each modified dataset. A summary of the distribution of the effect estimates obtained from each modified dataset can be presented graphically, or with 95 % "uncertainty intervals" by taking the 2.5th and 97.5th percentiles from this distribution.

This is an improvement over older sensitivity analysis methods for the following reasons: 1) a distribution of sensitivities and specificities can be specified, such that the likelihood of each set of values is taken into account; 2) the distribution of sensitivities and specificities are used to create modified datasets for analysis with conventional regression, such that a summary distribution of effect estimates is created; 3) concise summarization of results can be presented that account for random error, systematic error and both; 3) the magnitude of multiple sources of bias (i.e., from the exposure, outcome or covariates) can be considered in one analysis that still results in concise summary statistics and graphs. This technique is useful for the proposed study as we would like to consider the effect of misclassification of the outcome (incident heart failure). After specifying a distribution for the bias due to outcome misclassification, then we will have a new summary statistics to compare with the conventional analysis.

In an article by Lash and Fink(Lash and Fink 2003) entitled, 'semi-automatic probabilistic sensitivity analysis to assess systematic errors in observational studies', the authors introduce a methodology for sensitivity analysis. An example is given in which the conventional survival analysis results are compared with the results of sensitivity analyses of

multiple biases. They evaluated 3 types of systematic error (selection bias, misclassification, and confounding). These three types of systematic error are analyzed independently and all together; next random error is incorporated into each analysis (See below). The end result is the inclusion of all sources of error in the bottom right three columns such that the simulated median is 1.52 with 95 % uncertainty intervals of 0.81-2.81; which can be compared to the conventional hazard ratio of 2.0 with 95 % CI of 1.2-3.4. The relative impact of each source of error can be visualized in the below in **Error! Reference source not found.**. In this figure, the dotted/dashed line represents the results from the sensitivity analysis including all 3 sources of systematic error, the dotted line represents the results of the conventional analysis, and the solid line represents the combination of systematic and random error from the bootstrapped samples. The availability of high performance software has facilitated this advancement in multiple bias modeling, however such techniques are not often used. We suggest the use of this methodology for the analysis of systematic error in the proposed study. From **Error! Reference source not found.** and

Table 3, one can envision how the results of such an analysis might appear.

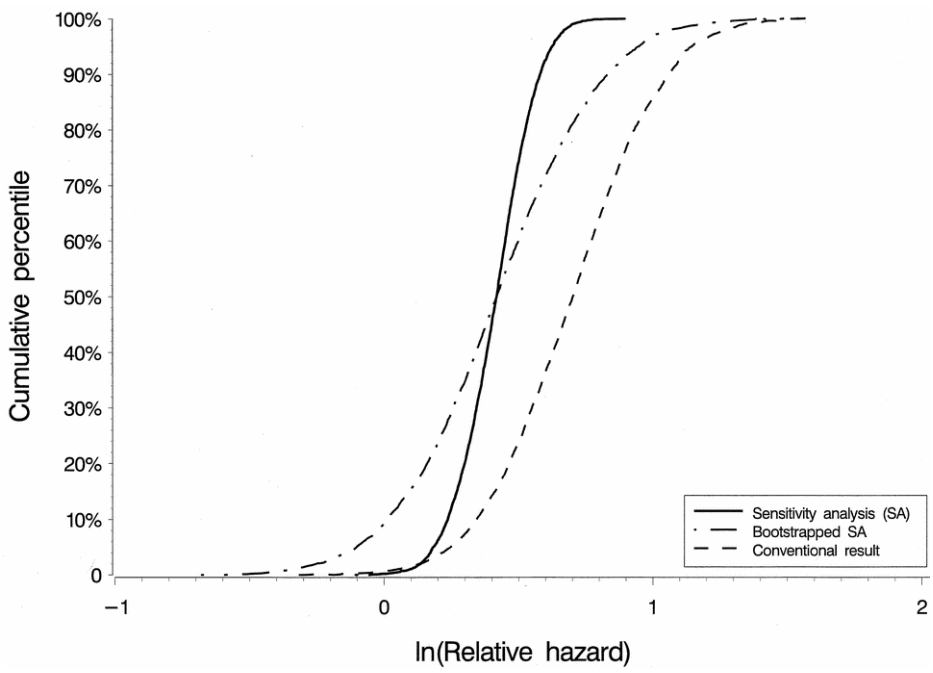
Table 3. An example of the succinct summarization of results from a semi-automated probabilistic sensitivity analysis, from Lash and Fink(Lash and Fink 2003)

Summary of relative hazards from the cumulative probability distributions yielded by the sensitivity analysis of the effect of less-than-definitive therapy on breast cancer mortality over five years of follow-up. The conventional survival analysis yielded a relative hazard of 2.0, with 95% confidence interval of 1.2–3.4.

	Systematic Error			Systematic and Random Error		
Analysis	**2.5%**	**50%**	**97.5%**	**2.5%**	**50%**	**97.5%**
Selection bias	1.51	1.81	2.19	1.06	1.82	3.10
Stage misclassification	1.59	1.83	2.08	1.01	1.84	3.45
Unknown confounder	1.60	1.85	2.01	1.03	1.84	3.32
All three sources	1.16	1.52	1.94	0.81	1.52	2.81

Figure 3. Graphic representation of the results summarized in

Table 3. An example of the succinct summarization of results from a semi-automated, From Lash and Fink(Lash and Fink 2003)


100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0
Cumulative percentile
−1
0
1
2
ln(Relative hazard)
Sensitivity analysis (SA)
Bootstrapped SA
Conventional result

G. Impact measures of risk factor-disease associations

In epidemiology we most often study the association of a given exposure with a disease. Once a potentially causative association is found with a modifiable exposure then we want to determine the burden of disease that could be prevented given the implementation of available interventions to modify the exposure. This latter part allows prediction of the possible public health impact of programs to reduce or eliminate a given exposure. The primary impact measure used in practice is the population attributable fraction (also known as, attributable risk). The attributable fraction, introduced by Levin in 1953(Levin 1953), is often misused and/or misinterpreted(Rockhill, Newman et al. 1998). It estimates the proportional reduction in disease given complete elimination of an exposure. For many exposures complete elimination is impossible or highly unlikely(Benichou 2007). Obesity and overweight is an exposure that is unlikely to be eliminated. Despite this, the attributable fraction has been used extensively in the recent articles attempting to estimate the burden of deaths attributable to overweight and obesity(Flegal, Graubard et al. 2004; Flegal, Graubard et al. 2005). A more useful calculation for risk factors such as obesity is the generalized impact fraction, also known as the potential impact fraction and the generalized attributable fraction(Benichou 2001; Rodgers, Ezzati et al. 2004). It is a generalization of the attributable fraction which estimates the proportional reduction in disease incidence given a reduction in prevalence of a given risk factor. These calculations are often stratified by important confounders, allowing the identification of subgroups with particularly high impact that can be targeted for intervention. For common risk factors such as obesity and overweight, the

impact of intervention will likely reveal important changes in the burden of disease even when risk factor-disease associations are relatively weak(Morgenstern and Bursic 1982).

Generalized impact fraction

The generalized impact fraction was originally described by Walter in 1980(Walter 1980), then further elucidated and coined by Morgenstern and Bursic in 1982(Morgenstern and Bursic 1982). It is "the proportional reduction in the total number of new (incident) cases of a certain disease, resulting from a specific change in the distribution of a risk factor in the population at risk."(Morgenstern and Bursic 1982). Despite its introduction over 25 years ago, it hasn't caught on as either a replacement for or as an additional measure to report with the attributable fraction.

Morgenstern and Bursic (1982) illustrate the use of the generalized impact fraction with an example of the impact of a hypothetical weight loss programs on the incidence of coronary heart disease(Morgenstern and Bursic 1982). Here they reveal the above mentioned advantages of the identification of high risk subgroups and the likely unexpected impact of modest changes in the prevalence of obesity and overweight. Morgenstern and Bursic discuss several key assumptions of the potential impact fraction. A relative risk of the relationship under study will be needed to calculate the GIF. It is assumed that it is a valid estimate of the relationship between exposure and disease for the desired target population, and is therefore, not biased by confounding or misclassification. Also, we assume that those who are hypothetically shifted to a different risk group post-intervention, now have the same risk that this group had pre-intervention. When this condition is not met then the potential impact fraction may overestimate the impact of the intervention. Also, it is assumed there

are not secular trends in risk of disease that are not due to the intervention; and there are not secular trends in intermediary diseases, such as diabetes for obesity, that would effect the relationship between the exposure and disease.

Inherent in estimating the GIF, one must consider a counterfactual approach(Rodgers 2002). The theoretical minimum risk distribution would be the complete elimination of obesity, or for a continuous measure it would be the point of lowest risk on the distribution. However, it is more likely that there could be partial eradication of obesity. There are multiple counterfactual situations that could be considered between the current and minimum distributions and these are called distributional transitions by the World Health Organization. Most likely to occur are small distributional shifts, such as 10 % or 20 % change. Murray and Lopez introduce four types of counterfactual exposure distributions: theoretical minimum risk, plausible minimum risk, cost-effective minimum risk, and feasible minimum risk. In our case, we are interested in feasible minimum risk(Murray and Lopez 1999). Feasible means that the distribution change has been achieved in some population and is also possible for the current population. A plausible distribution is imaginable, but rather may be possible for some society during some time period.

Feasible goals for weight reduction

To determine feasible goals for weight reduction, we will consider results from previous weight-loss intervention studies and the goals set by the US Department Health and Human services in the Healthy People 2010 report. Healthy People 2010 are a set of recommended health objectives to be achieved by the year 2010(2000). Obesity and overweight are listed as high priority public health issues. These goals are meant to serve as a basis from which to

develop US national, state and community plans for health improvement. Specifically, their objective is to reduce the proportion of obese adults aged 20 years and older to 15 % from 23% (25 % in females and 20 % in males, data from 1988-1994, age adjusted to year 2000). A second objective is to increase the proportion of adults at a healthy weight to 60 % from 42 % (45 % in women, 38 % in men, data from 1988-1994, age-adjusted to year 2000).

Obesity and overweight are unfortunately increasing rather than decreasing(Rosamond, Flegal et al. 2007). Many strategies for weight reduction exist, however all have limited effectiveness. A meta-analysis by Franz et al categorized weight loss clinical trials into 8 types of intervention: diet and exercise, exercise alone, advice alone, meal replacements, very-low-energy diets, and weight-loss medications (orlistat and sibutramine)(Franz, VanWormer et al. 2007). Results from this meta-analysis, which includes clinical trial completers with at least one year of follow-up, found that the mean weight loss was 5-8.5 kg (5-9%) at 6 months and stabilized at 4.5-7.5 kg (4.8-8 %) at 12 months. A diet and exercise intervention would be most applicable for a population wide strategy. Those in the diet and exercise group lost 7.6 kg at 12 months, whereas those in the diet-alone group lost 4.6 kg. A systematic review of diet and exercise trials reported that a mean of 6.7 kg of weight loss was maintained after one year(Curioni and Lourenco 2005). Based on these findings of those who completed one year of follow-up, it is apparent that at best only a modest weight reduction could be expected from a population wide intervention.

H. Summary and public health significance

Heart failure and obesity are important healthcare issues that have increasing prevalence. Previous studies have found an association between obesity and heart failure; although most

of these studies used BMI as the sole metric of adiposity and were in primarily white populations, specific clinical trial populations or isolated community populations. Insulin resistance is likely an important intermediary factor in the association between obesity and HF. Waist hip ratio, a metric of central adiposity, is more closely associated with insulin resistance than BMI. Selecting the best anthropometric for the prediction of HF could have public health implications for the screening and prevention of HF. Currently, screening for obesity and overweight is primarily assessed with BMI, rather than waist circumference or WHR. Further evidence supporting measures of central adiposity, rather than BMI, could eventually result in a shift in current practice patterns.

The magnitude of misclassification of heart failure due to the use of ICD codes has been quantified in several studies. For large observational studies without committee validation, one method to adjust for disease misclassification is to perform a sensitivity analysis to show the effect of misclassification. Approaches called multiple bias modeling allow the incorporation of various types of bias, either from the main exposure, outcome or covariates, into multivariate modeling using Monte Carlo techniques. This study provides further example of how Monte Carlo Risk Assessment techniques can be used to succinctly summarize the possible effects of systematic error. The application of this method has public health importance in providing further example of how the interpretation of results can easily include a sensitivity analysis which incorporates systematic error in addition to random error. For studies in which results will potentially change policy or alter patient care, it is vital that such estimates of uncertainty be considered in the interpretation of results.

We estimated the public health impact on the incidence of heart failure given a hypothetical and feasible reduction in the prevalence of obesity and/or overweight. This will

provide further example of how this important measure, the generalized impact fraction, should be routinely used in practice for risk factors that are unlikely to be completely eradicated. For common risk factors such as obesity and overweight, the impact of intervention will likely reveal important changes in the burden of disease even when risk factor-disease associations are relatively weak. Such evaluations are vital to prioritize and inform future prevention programs regarding the possible impact of their efforts.

CHAPTER III

METHODS

A. Study population

The ARIC cohort was recruited using probability sampling of those aged 45-64 from the following four US communities: Forsyth County, North Carolina (includes Winston-Salem); the city of Jackson, Mississippi; the northwestern suburbs of Minneapolis, Minnesota; and Washington County, Maryland (includes Hagerstown). The distribution of blacks and whites from each county is representative of the area (mostly white in Minneapolis and Washington County), except for Forsyth County in which blacks were over-sampled (15 %) and in Jackson where only blacks were sampled. Response rates were 46 % in Jackson and between 65-67 % for the other communities. The design and rationale of the ARIC study (Investigators 1989) and the comparison between responders and non-responders (Jackson, Chambless et al. 1996) have been previously published. The institutional review boards from each site approved the ARIC study; also, the institutional review board at UNC-Chapel Hill approved this dissertation. All participants provided written informed consent.

Initially, a home interview was administered to all potential cohort members. Those meeting criteria and interested in participating were provided an ARIC brochure and scheduled for their first examination. Informed consent was obtained on the arrival of

participants to the local field center. Staff members were certified in the appropriate method for obtaining consent. The first portion of the examination was performed following a 12 hour fast. Cohort examinations took place every 3 years for 4 visits, beginning with the baseline visit in 1987-1989. Between cohort examinations, a telephone questionnaire was administered yearly to identify intervening hospitalizations and deaths. In addition, community-wide surveillance was performed to identify all cohort hospitalizations and deaths(Investigators 1989).

B. Exclusion criteria

Racial groups not classified as white or black (N=48) and those missing anthropometry (N=33) were excluded. In addition, those with prevalent heart failure at baseline were excluded from this analysis by the following criteria: 1) those answering "yes" to the following question: "Were any of the medications you took during the last two weeks for heart failure?" (N = 83), or else 2) those with stage 3 or 'manifest heart failure' by applying Gothenburg criteria (N= 699), or 3) those who did not meet one of these 2 criteria, but were unresponsive to the HF medication question or any component of the Gothenburg criteria (N = 289). The Gothenburg criteria are based on a study from Gothenburg, Sweden of men born in 1913(Eriksson, Caidahl et al. 1987). It is composed of three scores: 1) cardiac, 2) pulmonary, and 3) therapy. In order to have stage 3 heart failure, one must have a point from each category. See Table 4, for a description of the Gothenburg criteria. All current medications (taken within the last two weeks) were brought into the clinic and documented. Use of digoxin and diuretics was determined from these medication lists. Atrial fibrillation was diagnosed by visual inspection of a 2 minute rhythm strip from leads V_1, II and V_5 using

standardized methodology(Vitelli, Crow et al. 1998). All other components were determined by participant self-report. After these exclusions, the total sample size was 14,690.

Table 4. Description of Gothenburg score

	Gothenburg components
Cardiac	Coronary Heart Disease -1 point if ever, 2 points if within the last year
	Angina - 1 point if ever, 2 points if within the last year
	Leg edema – 1 point
	Shortness of breath at night – 1 point
	Rales on lung exam – 1 point
	Atrial fibrillation on ECG – 1 point
Pulmonary	History of bronchitis – 1 point
	History of asthma – 1 point
	Cough, phlegm or wheezing – 1 point
	Rhonchi on lung exam – 1 point
Therapy	Treatment with digoxin – 1 point
	Treatment with diuretics – 1 point

C. Ascertainment of heart failure events

The following three methods were used for ascertainment of events during the follow-up period: 1) Participants were contacted annually by phone and interviewed about interim hospitalizations; 2) local hospitals provided lists of discharges with cardiovascular diagnoses and these were reviewed to identify cohort hospitalizations; and 3) health department death certificate files were continuously surveyed. All discharge codes for cohort hospitalizations and listed causes of death from death certificates were recorded. We used any listing of an ICD code for HF on a death certificate rather than HF listed as an underlying cause of death, because usually the cause of HF is listed as the underlying cause of death, rather than just

HF. Therefore, many deaths due to HF would be missed using only a definition of HF as underlying cause of death.

D. Incident heart failure event criteria

Incident HF was defined as the first occurrence of either: 1) a hospitalization which included an ICD-9-CM (International Classification of Diseases, 9th revision, clinical modification) discharge code beginning with '428' in any position (N = 1,329) or else 2) a death certificate with an ICD-9 code beginning with '428' (HF) or ICD-10 code 'I50' (HF) in any position (N = 76). Follow-up time for those with an incident HF event was defined as the time from the date of their baseline examination (1987-1989) until the incident event (follow-up through Dec. 31st, 2003). The end of follow-up time for those without HF was the first occurrence of either: 1) December 31st, 2003, 2) date of last contact for those lost to follow-up, or 3) date of death.

E. Anthropometric measures

Anthropometric indices were measured with participants in standard scrub suits and lightweight non-constricting underwear. Participants had been instructed to fast overnight and to empty their bladder prior to these measures. Measures were taken by either two technicians or one technician with a full length mirror for assistance. Technicians measured height with participants barefoot using a wall mounted ruler. An anthropometric measuring tape was applied horizontally to measure hip and waist girth; participants stood upright with weight evenly distributed between both feet and breathing quietly. Waist girth was measured at level of the umbilicus and hip girth at the level of maximal protrusion of the gluteal

muscles. Values were rounded down and were recorded to the centimeter. Weight measurements were performed using a scale (Detecto model 437) that was zeroed daily and calibrated quarterly. Body mass index (BMI) was calculated as weight divided by height squared (kilograms/meters^2), whereas waist/hip ratio is the waist girth divided by the hip girth. Inter-technician reliability coefficients for waist and hip girth and WHR were > 0.91.(Ferrario, Carpenter et al. 1995)

F. Baseline covariate definitions

All covariates were collected from the baseline visit. Race, gender, educational level, current alcohol use, and smoking status were obtained by participant self-report. Race categories included white, black, American Indian, Asian/Pacific Islander, and Hispanic; however, as mentioned earlier, racial groups not classified as white or black were excluded due to their limited numbers (N= 48). Prior history of myocardial infarction (MI) was defined as a reported history of physician-diagnosed MI or silent MI as identified by electrocardiography. History of coronary heart disease includes those with a history of MI, and those with a prior coronary revascularization procedure or coronary artery bypass surgery. History of stroke was defined by self-report of physician diagnosis of stroke.

Blood pressure (BP) measurements were taken using a standardized protocol. The last two of three sitting BP measurements were averaged. Hypertension was defined by either a diastolic blood pressure ≥ 90 mm Hg or a systolic blood pressure ≥ 140 mm Hg, or anti-hypertensive medication use during the previous 2 weeks. Left ventricular hypertrophy was identified by electrocardiography using Cornell criteria (Crow, Prineas et al. 1995). Diabetes mellitus was defined as any of the following: self-reported history of physician diagnosed

diabetes, medication use for diabetes over the last 2 weeks, a blood glucose ≥ 126 mg/dl fasting or a blood glucose ≥ 200 mg/dl non-fasting. Forced expiratory volume (FEV1) was obtained from pulmonary spirometry performed by trained technicians with computer assistance. The FEV_1 measurements have been adjusted for age, race, gender and height(Shahar, Boland et al.). Methods for the measurement of blood levels of albumin, creatinine, and glucose have been previously described (Eckfeldt, Chambless et al. 1994). Low-density lipoprotein cholesterol (LDL-C), high-density lipoprotein cholesterol (HDL-C), triglycerides, and total cholesterol were measured in fasting state using standardized methods (Investigators 1989).

G. Data quality

Data was directly entered into a computer-assisted data collection system. Suspicious values were immediately detected while the participant was onsite, such that the value could be confirmed or corrected. The main study data was held at the coordinating center; the main data was updated weekly by diskettes of study data mailed from the centers. Reports on data were generated routinely for review by all study sites. Quality control measures were undertaken to assure uniform data collection across time and between centers. All staff received training and certification in data collection procedures. Staff performance was monitored by supervisors, both directly and indirectly with video-recordings of participant interviews. When deemed necessary, re-training and re-certification was performed. Also as part of their training, staff was instructed on the ethical conduct of research involving human subjects, with an emphasis on confidentiality of the study data and its protection. All ARIC staff signed a pledge of confidentiality. Other data monitoring steps were undertaken with

included: 1) assessment of participant effort during spirometry; 2) assessment for digit preference in blood pressure data; 3) measurement of blood flow rate during venipuncture; 4) repeat anthropometric measures by the same and different technicians; and 5) blind analysis of duplicate blood samples and electrocardiograms(Investigators 1989).

H. Statistical power analysis

Since there are accepted cut-points for BMI, we used BMI as the main exposure for the statistical power analysis. The data for this study has already been collected and therefore the numbers per groups has already been defined. Power analyses for incident HF were estimated stratified by gender for the comparison between the highest category of BMI (BMI >30) as compared to the referent group (BMI < 25). Power analysis assumes no confounding, no missing values, and no bias. Preliminary data from the ARIC study was used to determine the exponential survival parameters for each BMI group, whereas we assumed a dropout rate of 5 % per year. The group with the smallest number was provided as "n per group" and was held constant across the table. The highlighted numbers were estimated by the software.

Table **5** provides an estimate of power given different hazard ratios for men. The proportion used for the exponential parameter for λ_1 was based on the proportion of men in group BMI > 30 with HF by the end of follow-up. In the first column, λ_2 is the proportion of men in the BMI <25 group, with HF by the end of follow-up. However, for columns 2-4, the HR varies and therefore λ_2 varies depending on the HR that was input. The power based on a HR of 2.1 is 99 % if there were 133 cases (example 1), whereas when the hazard ratio is decreased to 1.5, the power decreases to 89 % if there were 248 cases (example 2). As for the women (see

power **Error! Reference source not found.**) based on a HR of 2.9 is 99 % if there were 66 cases (example 1), whereas as when the HR decreases to 1.5, the power decreases to 91 % if there were 275 cases (example 3). The actual total number of events for these two BMI categories combined is 372 for men and 383 for women. Since we have over 350 cases for both men and women, this tables shows that we have adequate power even after stratification by gender. Power estimates and tables were created using *NQuery Advisor 5.0*.

Table 5. Two group test of equal exponential survival, with exponential dropout, only men included

	1	2	3	4
Test significance level, α	0.050	0.050	0.050	0.050
1 or 2 sided test?	2	2	2	2
Length of accrual period	3.00	3.00	3.00	3.00
Maximum length of follow-up	16.00	16.00	16.00	16.00
Common exponential dropout rate, d	0.0500	0.0500	0.0500	0.0500
BMI >30, exponential parameter, λ_1	0.0101	0.0101	0.0101	0.0101
BMI <25, exponential parameter, λ_2	0.0048	0.0067	0.0078	0.0072
Hazard ratio, h= λ_1 / λ_2	2.104	1.500	1.300	1.400
Power (%)	99	89	57	77
n per group	1567	1567	1567	1567
Total number of events required, E	133	248	268	258

Table 6. Two group test of equal exponential survival, with exponential dropout, women only

	1	2	3	4
Test significance level, α	0.050	0.050	0.050	0.050
1 or 2 sided test?	2	2	2	2
Length of accrual period	3.00	3.00	3.00	3.00
Maximum length of follow-up	16.00	16.00	16.00	16.00
Common exponential dropout rate, d	0.0500	0.0500	0.0500	0.0500
BMI >30, exponential parameter, λ_1	0.0072	0.0072	0.0072	0.0072
BMI <25, exponential parameter, λ_2	0.0025	0.0036	0.0048	0.0055
Hazard ratio, h= λ_1 / λ_2	2.880	2.000	1.500	1.300
Power (%)	99	99	91	62
n per group	2405	2405	2405	2405
Total number of events required, E	66	153	275	298

In addition, a power analysis was performed to determine the power to assess for effect modification by race, hypertension, diabetes, and CHD using Power Program v.3.0.0 (downloaded from http://dceg.cancer.gov/POWER/). Given that the analysis is already stratified by gender, there is inadequate power to detect effect modification for any of the potential effect modifiers, given a theta of 1.5 (See Table 7). All of these effect modifiers except race could be considered causal intermediates, therefore we will only further consider race as a potential effect modifier. There are no race- and gender-stratified estimates in the literature for the association of obesity with HF. Despite the estimated lack of power to find a significant difference between these groups, we will investigate these stratified estimates to gain some insight as to how this association might vary across these race and gender groups.

Table 7. Estimate of power* to assess for multiplicative effect measure modification by the following variables, given probabilities and sample size in the ARIC

	Power in Men	Power in Women
Age	0.31	0.34
Race	0.49	0.35
Hypertension	0.40	0.32
Diabetes	0.25	0.16
CHD	0.19	0.07

I. Statistical Analysis

The distributions of all three exposures (BMI, WHR, and WC) were inspected for outliers. The distribution of men and women were compared for each exposure. Pearson correlation coefficients were estimated to determine the correlation between BMI, waist-hip ratio and waist circumference. Categorization of BMI was defined as established in the literature(NIH 1998): 1) BMI < 25 kg/m^2 (normal weight); 2) BMI between 25-30 kg/m^2 (overweight); and 3) BMI >30 kg/m^2 (obese). No well accepted epidemiologic standard exists for the categorization of WHR and WC, therefore both of these measures were categorized into approximate tertiles by gender. Although a sex-specific dichotomous cut-point has been suggested for WC (obese defined as men > 102 cm and women > 88 cm)(NIH 1998), we preferred approximate sex-specific tertiles of WC for better specification. Categorized variables were represented as indicator variables with comparison to the lowest group (normal weight) as referent. We evaluated BMI in classes of weight as represented in the clinical guidelines from the National Institutes of Health(1998). In addition, we characterized obesity within BMI category by low or high waist circumference as seen in the NIH's clinical guidelines on the identification, evaluation, and treatment of overweight and obesity in adults(NIH 1998). The cut-point for low and high waist circumference is 102 cm. in men and 88 cm. women are recommended to further define those at high risk among

individuals with a BMI between 25 kg/m^2 and 34.9 kg/m^2. Furthermore, BMI categories of normal weight, overweight and obese, were stratified by low and high waist-hip ratio using gender-specific cutpoints from the U.S. Department of Agriculture (WHR cutpoints: women, 0.88; men, 0.95)(US Department of Agriculture 1990).

Univariate associations by gender were determined across categories of the three exposures for each categorical covariate using chi-square and for each continuous covariate using analysis of variance (ANOVA). Poisson models were used to age-adjust (to the mean age at baseline, 54 years) incidence rates for heart failure for each weight category, stratified by race-gender.

Analysis to address aim 1.1

> *1.1: Determine the hazard ratio for the association of obesity (as measured by BMI, waist/hip ratio, and waist circumference) with incident heart failure, while controlling for potential effect modifiers and confounders. Assess race as a potential effect modifier.*

Multivariable Cox proportional hazard regression was used to model the association between each anthropometric variable (BMI, WHR and waist circumference) and time to incident HF. Log (- log) survival curves and time interaction terms for the main exposures and all covariates were performed to validate the proportional hazards assumption. In the multivariable model, ties were handled with the method suggested by Efron. All statistical analyses were performed using SAS software v 9.1 (Cary, NC). Multivariable hazard ratios were estimated for the main exposure variables (BMI, WHR, and waist circumference) in both continuous (per one standard deviation change) and categorized form.

Assessment of confounding

A directed acyclic graph (DAG) was drawn representing a conceptual framework of the relationship between obesity and HF (Weinberg 1993; Greenland, Pearl et al. 1999), see Figure 4. This technique has been recommended as a formal method to assess whether possible confounders truly meet the definition of a confounder. There are three characteristics that define a covariate as a confounder: 1) the covariate should be associated with the disease, 2) and the exposure, and 3) not affected by the exposure(Rothman and Greenland 1998). These relationships are based on findings from the literature and on biologic plausibility. Possible causal relationships are represented by arrows. Since a confounder should not be affected by the exposure, then all arrows emanating from the exposure (obesity) are removed (see Figure 5. Same as Figure 4, except all arrows emanating from the main exposure (obesity); therefore, now there are unblocked backdoor paths between age, gender, education, race, center, smoking and alcohol use. These unblocked backdoor paths indicate potential confounders. All the other factors, such as CVD risk factors and CHD, are along the causal pathway between obesity and HF. These factors are caused by the exposure and adjustment for or stratification by any of these intermediates can create bias(Weinberg 1993). Therefore, only these factors with unblocked backdoor paths were considered as potential confounders.

Figure 4. A Directed Acyclic Graph (DAG) for the association between obesity and incident heart failure

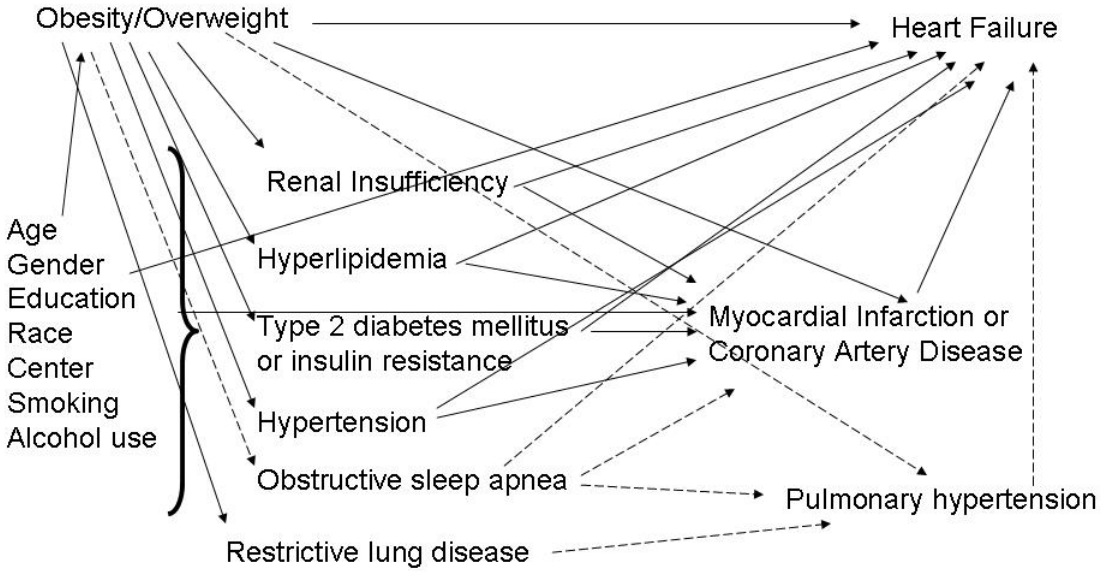
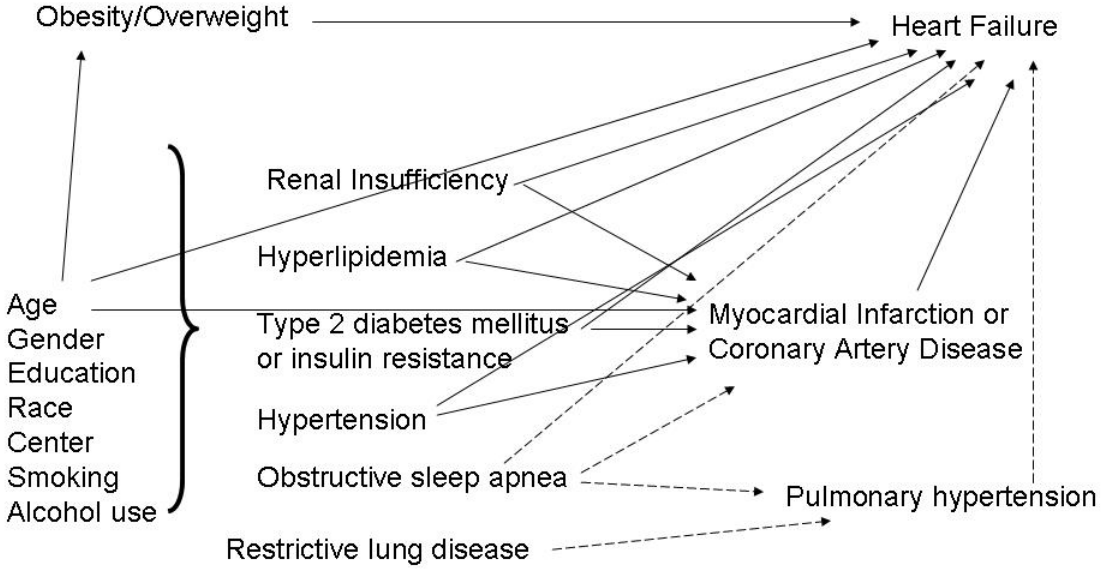


Figure 5. Same as Figure 4, except all arrows emanating from the main exposure (obesity) have been removed



As an alternative to the full model, a change in estimate approach was used to further select confounders to be in a more parsimonious model. The following formula was used in which the full model was compared to the reduced model:

$$\text{Change in estimate} = | \ln(HR_{full} / HR_{reduced}) |,$$

Where, HR_{full} = the hazard ratio for obesity or overweight with all covariates in the model, and $HR_{reduced}$ = the hazard ratio for obesity or overweight with all covariates except the covariate of interest.

The cut-point for inclusion in the model was a change in estimate of > 0.1. In addition, precision was assessed for all models with the confidence limit ratio (upper confidence limit/lower confidence limit). Furthermore, in order to quantify the trade-off between precision and validity for the inclusion versus exclusion of a given covariate, the following variation on the estimate of mean square error ("MSE") was calculated:

$$\text{"MSE"full} = \text{variance}(\ln HR)$$

$$\text{"MSE"reduced} = \text{variance}(\ln HR_{reduced}) + (\ln HR_{full} - \ln HR_{reduced})^2$$

If "MSE" for the full model is < the "MSE" for the reduced model then the assessment of the precision/validity trade-off favors adjustment for the covariate of interest. These same methods, change in estimate, confidence limit ratio and "MSE" where used to assess the different forms of certain covariates.

Covariate specification

The appropriate form of each covariate (age, education, race, center, smoking, alcohol use) was investigated. For the continuous covariate (age in years) linearity was assessed, separately for men and women: 1) Age groups were created by dividing age into 3 year categories and then creating indicator variables with the youngest group as referent, 2) The HR for the association of age with HF was plotted to assess for linearity. Age in continuous form was re-centered to an age of 45 years. All other variables were considered as indicator variables. Specifically, the form of alcohol use and smoking status were evaluated as never, current and former and also in more detailed categories of never, former, current occasional, current moderate, and current heavy. Race and center were considered jointly with indicator variables because the distribution race and center were dependent. One center (Jackson, MS) was all black, two centers (Minneapolis, MN and Washington County, MD) were almost exclusively white, and the other center (Forsyth County, NC) contained both race groups; therefore, blacks not from Jackson or Forsyth (N= 22 women, N=26 men) were excluded from all models. To control for both race and center, indicator variables were created for blacks from Jackson, whites form Minnesota, whites from Washington County, blacks from Forsyth, and whites from Forsyth that compared each with a referent group of whites from Minnesota. The Minnesota group was selected as referent because it was the largest group. Furthermore, a complete case analysis was performed, therefore, participants missing values for any covariates were excluded from modeling (N = 7967 for women, N = 6,603 for men, after deletion of 120 participants that were either missing values for any of these 6 covariates or were blacks not from Jackson or Forsyth centers). Complete case analysis is a good choice since there is relatively little missing information on covariates.

Assessment for effect measure modification

To assess for additive effect measure modification by race, interaction contrast ratios (ICR) with 95 % CI were calculated. Specifically, ICR's were calculated separately for men and women for the comparison of obese (BMI $\geq$ 30) with normal weight (BMI < 25). The formula for the ICR is as follows:

$$ICR = HR_{11} - HR_{10} - HR_{01} + 1$$

Where, HR_{11} = hazard for those obese and black/ hazard for those normal weight and white, jointly exposed,

HR_{10} = hazard for those obese and black/ hazard for those normal weight and white, single exposure to black race, and

HR_{01} = hazard for those normal weight and black/ hazard for those normal weight and white, single exposure to obesity

The formula for 95 % CI for ICR is as follows:

$$(-HR_{01})^2 \text{ X Variance}(\beta_1) + (-HR_{10})^2 \text{ X Variance}(\beta_2) + (HR_{11})^2 \text{ X Variance}(\beta_3) +$$

$$((-HR_{01}) \text{ X } (-HR_{10}) \text{ X } 2(\text{Covariance}(\beta_1\ \beta_2)) + ((-HR_{01}) \text{ X } (HR_{11}) \text{X } 2(\text{Covariance}(\beta_1\ \beta_3)) +$$

$$((HR_{21}) \text{ X } (-HR_{01}) \text{ X } 2(\text{Covariance}(\beta_2\ \beta_3))$$

Where $\beta_1 = \beta_{\text{normalweight/black}}$, $\beta_2 = \beta_{\text{obese/white}}$, and $\beta_3 = \beta_{\text{obese/black}}$

A value for the interactive contrast ratio (and it's 95 % CI) different from zero suggests an additive interaction.

Multivariable models assume constancy of effect between covariates. To assess whether there were statistical interactions between covariates, we calculated likelihood ratio tests by gender for models with and without interaction terms (p value < 0.05 considered significant). The final model with significant interaction terms were compared with the reduced model using the above mentioned calculations of change in estimate, confidence limit ratio and "MSE".

Analysis to address aim 1.2

1.2 Compare three anthropometrical measures (e.g. BMI, waist circumference, and waist/hip ratio) as to their ability to predict incident hospitalized heart failure.

The predictive ability of each anthropometric variable for incident heart failure was compared using time-dependent receiver operating curves (ROC) for estimation of area under the curve at 10 years (AUC(10)), as described by Chambless and Diao(Chambless and Diao 2006). We determined the AUC (10) for each categorized anthropometric measure based on predicted probabilities estimated from gender-stratified Cox proportional hazards model with adjustment for age, educational level, race, smoking status, and alcohol use. Predicted probabilities at 10 years were divided into deciles and each compared to the lowest group. Goodness of fit was assessed using Hosmer-Lemeshow tests. To test for the differences in AUC(10) between waist-hip ratio, waist circumference, BMI, and BMI stratified by high and low waist circumference, bootstrapping (sampling with replacement) was used to estimate a distribution of AUC(10) differences such that 95 % confidence intervals were determined. Furthermore, we adjusted for optimism due to use of the same

dataset for determination of AUC(10) for both models being compared(Harrell, Lee et al. 1996).

Analysis to address aim 2.1

2.1 Determine a distribution for sensitivity and specificity as estimated from the literature for the definition of the outcome (incident HF) as defined by hospital discharge and death codes.

Estimates of the range of values for sensitivity (62.8 %) and specificity (95.4 %) of '428.0 – 428.9' ICD code to define HF were based on findings from Goff et. al.(Goff, Pandey et al. 2000). Although there are other studies which validate HF when defined by ICD codes, these other studies used additional codes besides '428.0 – 428.9'.

Analysis to address aim 2.2

2.2 Perform a semi-automated probabilistic sensitivity analysis to estimate the degree of bias due to disease misclassification based on the chosen distribution of sensitivity and specificity. Separate multivariable estimates of the odds ratio and its distribution will be obtained that incorporate systematic error (from disease misclassification), random error and both.

A sensitivity analysis was performed to assess the impact of systematic error due to misclassification of the outcome (heart failure). The *SAS* software macro developed by Lash and Fink(Lash and Fink 2003) was used. Since this macro is written for use with logistic regression, the odds ratios for the association of obesity with HF from multivariable logistic regression models were compared to the hazard ratios from multivariable Cox proportional

hazards regression models. Since the results were similar then logistic regression was used for the sensitivity analysis, to avoid reclassifying survival time for those with reclassified outcomes.

A distribution of values of sensitivity and specificity were input based on the finding from aim 2.1 for non-differential outcome misclassification. The program deleted all iterations that are not plausible combinations of sensitivity and specificity. Initial parameters for the trapezoidal distribution of sensitivity (minimum = 0.75, maximum = 1) and specificity (minimum = 0.75, maximum = 1) resulted in multiple deleted iterations (>25,000), therefore a broader range of sensitivity and specificity values (0.25 – 1 with modes of 0.5 and 0.75) were entered, such that the plausible range of values from the non-deleted iterations could be determined. The plausible range of values for the trapezoidal distribution of sensitivity were: minimum =0.6, mode 1 = 0.7, mode 2 = 0.85, maximum = 1, and for specificity: minimum = 0.94, mode 1 = 0.96, mode 2=0.96, maximum = 1. Therefore, this was the range used for the analysis.

Analysis to address aim 3.1

> *3.1 Determine a range for the feasible reduction in the prevalence of obesity and overweight (as measured by BMI categories), based on findings from weight reduction studies.*

Based on above review of the literature on weight-loss reduction intervention studies only modest weight-loss (about 5-7 kgs) was achieved. However, these studies do not tell us the population-level impact of an intervention in terms of the reduced proportion obese or overweight. Instead we relied on goals set by the US Department Health and Human

services in the Healthy People 2010 report. Healthy People 2010 are a set of recommended health objectives meant to serve as a basis from which to develop US national, state and community plans for health improvement by the year 2010(2000) . Obesity and overweight are listed as high priority public health issues. Specifically, their objective is to reduce the proportion of obese adults aged 20 years and older to 15 % from 23% (25 % in females and 20 % in males, data from 1988-1994, age adjusted to year 2000), which is approximately a 35 % reduction in the proportion of obese. We chose several scenarios for the reduction in the proportion of obese and overweight, for which the maximum achievable goal selected was a 30 % reduction in both obesity and overweight.

Analysis to address aim 3.2

3.2 Given several estimates for the reduction in prevalence of obesity and overweight, and the magnitude of association between BMI and incident heart failure, the generalized impact fraction will be determined, stratified by gender. This will estimate the potential impact of weight reduction on the incidence of heart failure.

The generalized impact fraction (GIF) for the association of overweight/obesity with the incidence of HF will be estimated using the following equation:

$$GIF = (Pr(D) - Pr(D/E^*)) / Pr(D)$$

Where Pr (D) = probability of disease, and
Pr (D/E*) = probability of disease given modified exposure distribution.

The generalized impact fraction (GIF) for heart failure given a reduction in overweight and or obesity was estimated using equations 1 and 3 (see box, below). Specifically, for this

study, the GIF was calculated for each age-, race- and gender- stratified group (see equation 1) and then combined using a case-load weighted-sum method to obtain an overall GIF (see equation 3). For each stratum, the components of GIF equation were defined as follows: 1) P_i is the prevalence of the exposure (obesity, overweight and normal weight) in the ARIC population at baseline (1987-1989); 2) P_i* is the reduced prevalence of obesity, overweight and normal weight (reduced as compared to P_i) after a hypothetical change in the distribution of BMI; and 3) RR_i, in this case HR_i, is the crude hazard ratio of incident heart failure for those overweight, obese or normal weight as compared to the referent group (normal weight), for that stratum. Cox proportional hazard regression was used to model the association between obesity (BMI $\geq$ 30 kg/m^2), overweight (BMI $\geq$ 25 and $\leq$ 30, kg/m^2) with normal weight (BMI < 25 kg/m^2) and time to incident HF. BMI was categorized as represented in the clinical guidelines from the National Institutes of Health(NIH 1998). In the multivariable model, ties were handled with the method suggested by Efron. All statistical analyses were performed using SAS software v 9.1 (Cary, NC).

Since calculation of the GIF is based on a single study population (the ARIC cohort), we added further uncertainty to these calculations by bootstrapping to obtain 95 % simulation intervals for the GIF and attributable fraction (see Figure 6)(Greenland 2004). Ten thousand bootstrapped samples (with replacement) were performed for each subgroup of race, gender and age (therefore, there were 8 subgroups each with 10,000 bootstrapped datasets). Each sampled dataset was the same size as the subgroup in the original dataset. The unadjusted hazard ratios (HR_i) for obesity and overweight (the HR for normal weight is always 1), and the prevalence (P_i) of normal weight, overweight and obesity were determined within each stratum for each of the 10,000 bootstrapped dataset. The GIF for each stratum was

calculated given these values and a specified alternative distribution for the reduced prevalence of obesity, overweight and normal weight (P_i*).

The overall GIF for each hypothetical scenario of weight reduction was determined using the case-load weighted-sum approach for each bootstrapped sample using equation 2 (Benichou 2001). From the distribution of the 10,000 overall generalized impact fractions, the median (along with 2.5 % and 97.5 % simulation intervals) were calculated for each hypothetical scenario.

(Equation 1)

$$\text{GIF for a given bootstrap sample, and stratum a,b,c} = \frac{\sum P_{i,\,abc}(RR_{i,\,abc}) - \sum P_i^{*}{}_{,abc}\,(RR_{i,\,abc})}{\sum P_{i,\,abc}(RR_{i,abc})}$$

Where,
GIF = generalized attributable fraction for one bootstrap sample, from a given stratum of abc,
P_i = proportion of the population in exposure category, i,
P_i*= proportion of the population in exposure category i after an intervention or other change,
RR_i = Relative risk at exposure category i compared to the reference level,
i = normal weight, overweight, or obese categories as defined by BMI,
abc = stratum of age (age < 55 years, age ≥ 55 years), race (black or white) and gender (male or female)

Specifically,

$$\text{Numerator } GIF_{a,b,c} = [P_{overwt,\,abc}(RR_{overwt,\,abc}) + P_{obese,\,abc}(RR_{obese,\,abc}) + P_{normal\ wt,\,abc}(1)] - [(P^{*}_{overwt,\,abc})(RR_{overwt,\,abc}) + (P^{*}_{obese,\,abc})(RR_{obese,\,abc}) + (P^{*}_{normal\ wt,\,abc})(1)]$$

$$\text{Denominator } GIF_{a,b,c} = [(P_{overwt,\,abc})(RR_{overwt,\,abc}) + (P_{obese,\,abc})(RR_{obese,\,abc}) + (P_{normal\ wt,\,abc})(1)]$$

(Equation 2)

$$\text{AF (or maximum GIF) for a given bootstrap sample from stratum a, b, c} = [\sum(pd_i(RR_i) - 1)] \,/\, [\sum(pd_i(RR_i)]$$

where pd_i = proportion of total cases arising from the ith exposure category
RR_i = the unadjusted RR (in this case HR) for the ith exposure category compared to referent (i = overweight, or obese categories compared to normal)

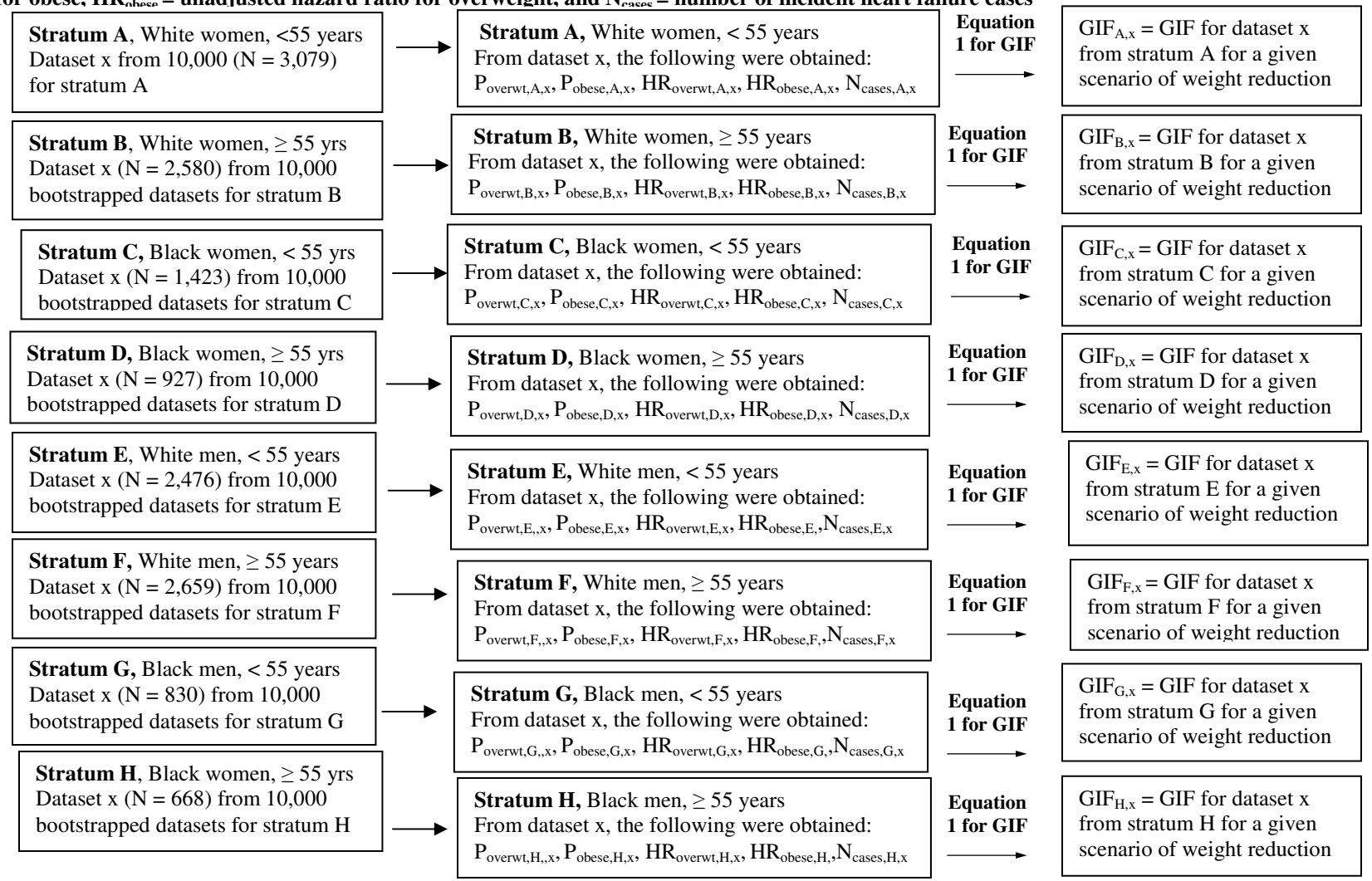
(Equation 3)

Case-load weighted sum method for $GIF_{overall}$

$$\text{Numerator } GIF_{overall} \text{ (or } AF_{overall}\text{) for a one bootstrap sample} = (\#cases_{a1b1c1})(GIF_{a1b1c1}) + (\#cases_{a1b1c0})(GIF_{a1b1c0}) + (\#cases_{a0b1c1})(GIF_{a0b1c1}) + (\#cases_{a0b1c0})(GIF_{a0b1c0}) + (\#cases_{a0b1c1})(GIF_{a0b1c1}) + (\#cases_{a0b1c0})(GIF_{a0b1c0}) + (\#cases_{a1b0c1})(GIF_{a1b0c1}) + (\#cases_{a1b0c0})(GIF_{a1b0c0}) + (\#cases_{a1b0c1})(GIF_{a0b0c1}) + (\#cases_{a0b0c0})(GIF_{a0b0c0}) + (\#cases_{a0b0c1})(GIF_{a0b0c1}) + (\#cases_{a0b0c0})(GIF_{a0b0c0})$$

$$\text{Denominator } GIF_{overall} \text{ (or } AF_{overall}\text{) for a given boostrap sample} = (\#cases_{a1b1c1}) + (\#cases_{a1b1c0}) + (\#cases_{a0b1c1}) + (\#cases_{a0b1c0}) + (\#cases_{a0b1c1}) + (\#cases_{a0b1c0}) + (\#cases_{a1b0c1}) + (\#cases_{a1b0c0}) + (\#cases_{a1b0c1}) + (\#cases_{a0b0c0}) + (\#cases_{a0b0c1}) + (\#cases_{a0b0c0})$$

Figure 6. Diagram of intermediate steps in the estimation of the distribution of the generalized impact fraction (GIF) using bootstrapped datasets (sampled with replacement), x = dataset number out of 10,000 bootstrapped datasets. Each dataset contains the same name number (N) in each stratum as found in the original data. P_{overwt}= proportion overweight, P_{obese} = proportion obese, HR_{overwt} = unadjusted hazard ratio for obese, HR_{obese} = unadjusted hazard ratio for overweight, and N_{cases} = number of incident heart failure cases



Analysis to address aim 3.3

3.3 Determine the population attributable fraction for overweight/obese for the outcome of incident hospitalized heart failure. Compare this more commonly used measure to that of the generalized impact fraction.

The attributable fraction, or in this case the maximum GIF, was calculated with the same case-load weight-sum method using equation 2(Morgenstern and Bursic 1982). Based on the distribution of values from the bootstrapped samples, the median attributable fraction along with 2.5 and 97.5 percentiles were reported.

J. Methodological strengths and limitations

Strengths of the analysis for manuscript 1 included the large sample size of this well characterized cohort and the extended period of follow-up. Measures were obtained using standardized techniques. Anthropometric variables were analyzed in both continuous form and using cut-points. Since there are three anthropometric measures available from the baseline visit, we were able to compare the predictive ability of the three measures using a newly described method for time-dependent ROC analysis.

The main limitation of this analysis is the misclassification of the outcome measure, however the uncertainty due to this bias was evaluated with the sensitivity analysis from the second aim in which we show the estimated effect of outcome misclassification on our results. In this aim we will use a novel technique called multiple bias modeling to allow the incorporation of this bias, into multivariate modeling using Monte Carlo techniques. This study will provide further example of how Monte Carlo Risk Assessment techniques can be used to succinctly summarize the effect of systematic error.

Strengths of the analysis for manuscript 2 are that we re-introduced the generalized impact fraction, as an alternative to the attributable fraction, for exposures that are unlikely to be eliminated, such as obesity. We show how confidence intervals can be obtained from bootstrapped datasets. In addition, we use these findings to extrapolate the public health impact in terms of the predicted decrease in heart failure hospitalizations and healthcare costs. Limitations of the analysis for manuscript 2 include the basic assumptions of the GIF, which are the same as for the attributable fraction, and the need to stratify on important confounders, unlike the attributable fraction.

CHAPTER IV

RESULTS

A. Manuscript 1: The Association of Overweight and Obesity with Incident Heart Failure: the Atherosclerosis Risk in Communities (ARIC) Study

ABSTRACT

Introduction: Obesity, as measured by BMI, has been identified as a risk factor for heart failure. This association has not been studied in non-white racial groups. Furthermore, the importance of a measure of central adiposity (waist-hip ratio) as compared to BMI has yet to be studied in a large population-based study. Methods: The Atherosclerosis Risk in Communities (ARIC) study is an ongoing bi-racial population-based cohort of those aged 45-65 years from 4 U.S. communities with 14 years median follow-up for incident, hospitalized or fatal heart failure. Waist-hip ratio, waist circumference and BMI were measured at the baseline visit (1987-1989). After exclusion of prevalent heart failure, missing anthropometry measures, and poorly represented race groups, the sample size was 14,642. BMI was categorized as normal weight (BMI <25 kg/m^2), overweight (BMI 25-30 kg/m^2), and obese (BMI $\geq$ 30 kg/m^2). WHR and waist circumference were divided into gender-specific tertiles. A first occurrence of ICD-9-CM codes of heart failure, either hospital discharge (428.0-428.9, N=1,329), or on a death certificate (428.0-428.9 or I50.0-I50.9, N=76) was considered

an incident case. Analyses were adjusted for alcohol use, smoking, age, and educational level using Cox proportional hazards modeling. Results: Even with stratification by race and gender, the adjusted hazard ratios for the comparison of the highest category of each measure (obese) to the lowest were similar and well above 1.0 for all three anthropometric measures (HR with 95 % CI for 3rd vs. 1st tertile of waist-hip ratio: 2.55 (1.88, 3.44) for white women; 3.10 (2.13, 4.51) for black women, 2.69 (2.10, 3.45) for white men, and 2.45 (1.75, 3.43) for black men. Hazard ratios for the middle category for all three anthropometric measures tended to be intermediate, thus suggesting a graded response between body size and heart failure.

Conclusion: Generalized obesity, overweight, and central adiposity, as measured by three different anthropometric measures, were associated with incident heart failure over 14 years median follow-up of the ARIC cohort. The current study does not support the superiority of waist-hip ratio and waist circumference over BMI for the prediction of incident heart failure.

1. Introduction

Obesity, as measured by body mass index (BMI), has been identified as a risk factor for heart failure from the Framingham Heart Study(Kenchaiah, Evans et al. 2002) and other studies(He, Ogden et al. 2001; Murphy, Macintyre et al. 2005). The increasing prevalence of both obesity and heart failure (HF) in the U.S. make this association an important topic for prevention(Rosamond, Flegal et al. 2007). Replication of the association of obesity with heart failure in non-white racial groups has been mentioned as an important area for future research.(Kenchaiah, Gaziano et al. 2004).

Furthermore, research on the association of measures of central adiposity with heart failure as has been mentioned as a priority research area(O'brien G 2002). Because diabetes and insulin resistance are risk factors for heart failure(Bertoni, Hundley et al. 2004; Ingelsson, Sundstrom et al. 2005), one would expect that a measure of central adiposity, a correlate of impaired insulin sensitivity, would have a stronger association with incident heart failure than a measure of generalized adiposity. Waist circumference is highly correlated with BMI and is therefore considered a measure of both central and generalized adiposity, whereas waist-hip ratio is a measure of central adiposity only(Folsom, Kushi et al. 2000). Existing studies imply that waist circumference and waist-hip ratio are also associated with incident heart failure(Dagenais, Yi et al. 2005; Ingelsson, Sundstrom et al. 2005; Nicklas, Cesari et al. 2006), however results vary depending on the population under study.

The Atherosclerosis Risk in Communities (ARIC) study is a bi-racial longitudinal cohort study of cardiovascular disease and its risk factors in those aged 45-65 years. It has anthropometric measures of BMI, waist-hip ratio, and waist circumference from the baseline visit along with 14 years median follow-up for incident heart failure. In this study, we evaluated the race- and gender-specific associations of overweight and obesity with incident heart failure. Furthermore, we determined whether a measure of central adiposity (waist-hip ratio) would be more closely associated with incident heart failure than BMI.

2. Methods

Study Population

The ARIC cohort (N=15,792) was recruited between 1987-1989 using probability sampling of those aged 45-64 years from the following four US communities: Forsyth County, North Carolina; Jackson, Mississippi; suburbs of Minneapolis, Minnesota; and Washington County, Maryland. The distribution of blacks and whites from each field center was representative of the area (mostly white in Minneapolis and Washington County, 15 % black in Forsyth County), except in Jackson where only blacks were sampled. Response rates at baseline were 46 % in Jackson and between 65-67 % for the other communities. Three subsequent visits were conducted at approximately three-year intervals through 1996-98. The design and rationale of the ARIC study (The ARIC Investigators 1989) and a comparison of responders and nonresponders (Jackson, Chambless et al. 1996) have been previously published. The institutional review boards from each site approved the ARIC study. All participants provided written informed consent.

Racial groups not classified as white or black (N = 48), and blacks not from Jackson or Forsyth County (N = 120) were excluded in this study due to their limited numbers. In addition, those with missing anthropometry (N = 33), prevalent heart failure at baseline (N=751), or missing criteria used to define prevalent heart failure (N=289) were excluded from this analysis. Criteria to define prevalent heart failure were as follows: 1) those answering "yes" to the following question: "Were any of the medications you took during the last two weeks for heart failure?" (N = 83), or 2) those with stage 3 heart failure by applying Gothenburg criteria (N = 699)(Eriksson, Caidahl et al. 1987). Gothenburg criteria were defined by self-reported medical history, medication lists and electrocardiography(Loehr, Rosamond et al. 2008). After these exclusions, the total sample size was 14,642.

Anthropometric Measures

Anthropometric indices were measured after an overnight fast with participants in standard scrub attire. Technicians measured height with participants barefoot using a wall mounted ruler. An anthropometric measuring tape was applied horizontally to measure hip and abdominal girth; participants stood upright with weight evenly distributed between both feet and breathing quietly. Abdominal girth was measured at the level of the umbilicus and hip girth at the level of maximal protrusion of the gluteal muscles. Weight was measured using a scale (Detecto model 437) that was zeroed daily and calibrated quarterly. Body mass index (BMI) was calculated as weight divided by height squared (kilograms/meters^2), whereas waist/hip ratio (WHR) was the waist girth divided by the hip girth. Inter-technician reliability coefficients for waist and hip girth and waist-hip ratio were >0.91(Ferrario, Carpenter et al. 1995)

Ascertainment of heart failure events

The following methods were used for ascertainment of heart failure events: 1) participants were contacted annually by phone and interviewed about interim hospitalizations (93-96 % response); 2) local hospitals provided lists of hospital discharges with cardiovascular diagnoses and these were reviewed to identify cohort hospitalizations; and 3) health department death certificate files were continuously surveyed. All discharge diagnosis codes for cohort hospitalizations and underlying or contributory causes of death from death certificates were recorded.

Incident heart failure event criteria

Incident heart failure was defined as the first occurrence of either: 1) a hospitalization which included an ICD-9-CM (International Classification of Diseases, 9th revision, clinical modification) discharge diagnosis code for heart failure beginning with '428' (i.e. 428.0 to 428.9) in any position (N = 1,329) or else 2) a death certificate ICD-9 code beginning with '428' (heart failure) or ICD-10 code 'I50' (heart failure or I50.0 to I50.9) in any position (N = 76). Follow-up time for those with an incident heart failure event was defined as the time from the date of their baseline examination until the incident event. The end of follow-up time for those without heart failure was the first occurrence of either: 1) December 31st, 2003, 2) date of last contact for those lost to follow-up, or 3) date of death.

Baseline covariate definitions

All covariates were collected from the baseline visit. Race was self-reported; educational level, alcohol use, and smoking status were obtained with interviewer-administered questionnaires. Prior history of myocardial infarction (MI) was defined as a reported history of physician-diagnosed MI or silent MI as identified by electrocardiography. History of coronary heart disease included those with a history of MI, and those with a prior coronary revascularization procedure or coronary artery bypass surgery.

Hypertension was defined by either a diastolic blood pressure ≥ 90 mm Hg or a systolic blood pressure ≥ 140 mm Hg measured with random-zero mercury manometers, or anti-hypertensive medication use during the previous 2 weeks. Left ventricular hypertrophy was identified by electrocardiography using Cornell criteria (Crow, Prineas et al. 1995). Diabetes mellitus was defined as any of the following: self-reported history of physician diagnosed

diabetes, medication use for diabetes over the last 2 weeks, a blood glucose ≥ 126 mg/dl fasting or a blood glucose ≥ 200 mg/dl non-fasting. Low-density lipoprotein cholesterol (LDL-C), high-density lipoprotein cholesterol (HDL-C), triglycerides, and total cholesterol were measured in fasting state using standardized methods(The ARIC Investigators 1989).

Statistical Analysis

Pearson correlation coefficients were estimated to determine the correlation between BMI, waist-hip ratio and waist circumference. Multivariable Cox proportional hazard regression was used to model the association between each anthropometric variable (BMI, waist-hip ratio and waist circumference) and time to incident heart failure, after stratification by race and gender. Waist circumference, BMI and waist-hip ratio were analyzed as both continuous and categorized variables. BMI was categorized as normal weight (< 25 kg/m^2), overweight (25 – 30 kg/m^2) and obese (≥ 30 kg/m^2), as represented in the clinical guidelines from the National Institutes of Health(NIH 1998). Waist-hip ratio and waist circumference were categorized into approximate gender-specific tertiles (**Table 9** (MS. 1, Table 2)) details the specific cutpoints used)(Dagenais, Yi et al. 2005). Furthermore, BMI categories of normal weight, overweight and obese, were stratified by low and high waist-hip ratio using gender-specific cutpoints from the U.S. Department of Agriculture (WHR cutpoints: women, 0.88; men, 0.95)(US Department of Agriculture 1990). Poisson models were used to age-adjust incidence rates for heart failure to the mean age at baseline (54 years). Log (- log) survival curves and time interaction terms for the main exposures and all covariates were used to evaluate the proportional hazards assumption. In the multivariable model, ties were handled with the method suggested by Efron(Efron 1977). A directed acyclic graph (DAG) was

drawn (DAG not shown) for the identification of confounders which represents our conceptual framework of the relationship between obesity and heart failure, distinguishing confounders from causal intermediates(Weinberg 1993; Greenland, Pearl et al. 1999). In addition, a backward selection change-in-estimate approach was used to further reduce confounders.

The predictive ability of each anthropometric variable for incident heart failure was compared using time-dependent receiver operating curves (ROC) for estimation of area under the curve at 10 years (AUC(10)), as described by Chambless and Diao(Chambless and Diao 2006). We determined the AUC (10) for each anthropometric measure (using categories shown in (Table 8 (MS. 1, Table 2) and Table 11 (MS. 1, Table 4)) based on predicted probabilities estimated from gender-stratified Cox proportional hazards model with adjustment for age, educational level, race, smoking status, and alcohol use. Predicted probabilities at 10 years were divided into deciles and each compared to the lowest group. Goodness of fit was assessed using Hosmer-Lemeshow tests. To test for the differences in AUC(10) between waist-hip ratio, waist circumference and BMI, bootstrapping (sampling with replacement) was used to estimate a distribution of AUC(10) differences such that 95 % confidence intervals were determined. Furthermore, we adjusted for optimism due to use of the same dataset for determination of AUC(10) for both models being compared(Harrell, Lee et al. 1996).

To assess for additive effect measure modification by race, interaction contrast ratios(Rothman and Greenland 1998) (ICR) with 95 % CI's were calculated. To assess for improved model fit, alternative models with covariate interactions were tested using likelihood ratio tests.

A sensitivity analysis was performed to assess the impact of systematic error due to misclassification of the outcome (heart failure) on our results. A SAS macro (available at http://sph.bu.edu/index.php?option=com_content&task=view&id=405&Itemid=508#epidemiologic) developed by Lash and Fink(Lash and Fink 2003) was used to incorporate uncertainty due to systematic bias from outcome misclassification into traditional regression using Monte Carlo uncertainty analysis. All statistical analyses were performed using SAS software v 9.1 (Cary, NC).

3. Results

Those who developed heart failure were older at baseline, less well educated, and were more likely to be Black and male (Table 8 (MS. 1, Table 1)). As would be expected, there was a higher percentage of coronary heart disease and cardiovascular risk factors, such as diabetes and hypertension, among heart failure cases. Waist circumference and waist-hip ratio cutpoints for men were higher than they were for women (**Table 9** (MS. 1, Table 2)). The mean value for each anthropometric measure for each category was surprisingly similar by gender and across race. Unlike the other race and gender groups, black women had a consistently higher percentage of person-time in the highest group (obese) of each measure compared to the lowest. For all race and gender groups, age-adjusted heart failure incidence rates were two to three fold higher when comparing the lowest group to the highest for all three anthropometric measures (**Table 9** (MS. 1, Table 2)). Across all three anthropometric exposures, heart failure incidence rates were higher in blacks compared to whites for both men and women. The heart failure incidence rates in black women compared to white women were nearly two-fold higher for all exposures categories.

Potential confounders available from the ARIC study and identified using the directed acyclic graph (DAG) were educational level, alcohol use, age, center, and smoking status. The following were identified as causal intermediates and therefore were not considered as confounders: coronary heart disease, hypertension, diabetes mellitus, and kidney disease. For all three anthropometric exposures, reduced models (data not shown) determined using a backward selection change-in-estimate approach (with 10 % change as criteria) resulted in hazard ratios of similar magnitude and precision to the models including all identified confounders. For comparability across the three exposures, models with all identified confounders are reported here. In addition, there was no observed improvement in model fit with the inclusion of significant ($P < 0.05$) covariate interactions.

Adjusted hazard ratios for the comparison of the highest anthropometric category (obese) to the lowest for all three measures were well above 1.0 for all race and gender groups (Table 10 (MS. 1, Table 3) and Figure 7 (MS. 1, Figure 1), Panel B). In fact, when combined across race, the adjusted hazard ratios for heart failure with obesity by all three measures were well above 2.0. Hazard ratios for the middle category (overweight) for all three anthropometric measures tended to be intermediate, thus suggesting a graded response between body size and heart failure (Table 10 (MS. 1, Table 3) and Figure 7 (MS. 1, Figure 1), Panel A). Calculation of interaction contrast ratios and interaction term analysis ($P < 0.05$) did not provide evidence for additive or multiplicative effect measure modification by race. Although we did not find such support for differences by race, for completeness, we present race- and gender- stratified estimates of the association of heart failure and adiposity.

Multivariable models for a standard deviation (gender-specific) change in the three exposures showed positive associations for all race and gender groups ((MS. 1, Table 3)).

After adjustment for confounding factors, women with a 1 SD (0.08) higher waist-hip ratio were 62 % more likely to develop incident heart failure over the next 14 years, whereas men with a 1 SD higher waist-hip ratio (0.05) were 50 % more likely to develop incident heart failure. Results were similar for BMI and waist circumference. In addition, to refute the hypothesis that obesity at baseline was due to fluid retention from heart failure, we excluded the first three years of follow-up time from the analysis; there was little change in the associations examined here (data not shown).

BMI and waist circumference were highly correlated (r = 0.88), waist-hip ratio and waist circumference were moderately correlated (r= 0.73), and BMI and waist-hip ratio were poorly correlated (r = 0.43). Stratification of BMI by waist circumference resulted in small sample sizes for certain categories, such as those with normal BMI and high waist circumference; therefore, we describe heart failure incidence rates for BMI categories, stratified by high and low waist hip ratio. Age-adjusted heart failure incidence rates for BMI (normal weight, overweight and obese), stratified by high and low waist-hip ratio show an increasing trend across these categories, with particularly a high heart failure incident rate (10.1 per 1,000 person-years) in those who have a high BMI and waist-hip ratio (Table 11 (MS. 1, Table 4) and Figure 8 (MS. 1, Figure 2)).

Estimation of AUC (10) from predicted probabilities from multivariable models for each anthropometric were as follows: BMI (0.73 for men, 0.79 for women); waist-hip ratio (0.74 for men, 0.79 for women); waist circumference (0.73 for men, 0.79 for women); and BMI stratified by high and low waist-hip ratio (0.74 for men, 0.79 for women). By gender, there were no significant differences in AUC (10) between the three measures, or when compared to models with cross-classified categories of BMI and waist-hip ratio ($P > 0.05$ for all

comparisons). The inferences were the same from this analysis when stratified by race and gender groups. Of note, in some cases, the Hosmer-Lemeshow tests did not support good model fit.

The sensitivity analysis assessed the effect of misclassification of the outcome (incident heart failure) on the multivariable association of obesity (BMI $\geq$ 30) and incident heart failure. The distribution of sensitivity parameters selected were based on findings from the literature (sensitivity: minimum =0.6, mode 1 =0.7, mode 2 =0.85, maximum =1; for specificity, minimum =0.94, mode 1 =0.96, maximum =1)(Goff, Pandey et al. 2000). The results of the sensitivity analysis show the median odds ratio with 2.5th and 97.5th percentiles was higher (OR = 4.54, 95 % CI = 2.93, 14.83) after incorporating uncertainty due to outcome misclassification compared to the conventional analysis, incorporating only random error (OR = 2.89, 95 % CI =2.47, 3.42). As would be expected, this suggests that the effect of outcome misclassification on our findings, given the chosen distribution of sensitivity and specificity, was to bias them towards the null.

4. Discussion

Generalized obesity and central adiposity, as measured by three different anthropometric measures, were associated with incident heart failure over 14 years median follow-up of the ARIC cohort. The magnitudes and patterns of the associations were similar for all three measures and there was evidence of a graded relation for all race, gender groups. Furthermore, an adverse association existed between adiposity and incident heart failure even for those who were overweight compared to normal weight, although less consistent across race- and gender-stratified analyses. Analyses of receiver operating curves found no

significant differences in the prediction of incident heart failure for the three measures. This implies that measures of waist-hip ratio and waist circumference are not superior to BMI in the prediction of heart failure. This is contrary to our hypothesis that measures of central adiposity will be more closely associated with heart failure due to their closer association with diabetes, a known heart failure risk factor.

Several previous studies have found similar associations to those observed here between obesity as measured by BMI and incident heart failure. The Framingham Heart Study found that overweight (BMI 25 – 29.9 kg/m^2) and obesity (BMI ≥30 kg/m^2) as measured by BMI were associated with an increased risk for the development of heart failure over 14 years median follow-up; multivariable adjusted hazard ratios were 1.34 (95 % CI = 1.08-1.67) and 2.04 (95 % CI = 1.59-2.63), respectively (Kenchaiah, Evans et al. 2002). In these models, the Framingham study adjusted for diseases, such as diabetes, which are considered along the causal pathway between obesity to heart failure. As they note, adjustment for factors along the causal pathway may underestimate the effect of overweight or obesity with heart failure. The Framingham Heart Study is a primarily white population from a single community that is approximately a third the size of the ARIC study. Since the ARIC study included blacks and whites from 4 US communities, one might expect to find more heterogeneity in the findings between these two studies. Instead the robustness of these findings supports a lack of differences in this association by race. Further study differences between the Framingham study and ours include that the outcome of heart failure from the Framingham study was validated using Framingham criteria. Heart failure was not one of the original main outcomes of the ARIC cohort and therefore validation of heart failure events was not

performed. Again despite these study differences, the findings are strikingly similar for the association of obesity with heart failure.

A community based study from Scotland, the Renfrew-Paisley study, observed a stronger association of heart failure with obesity, as defined by BMI, in men (HR = 2.16, 95 % CI = 1.57-2.57) than in women (HR = 1.37, 95 % CI = 1.00-1.88)(Murphy, Macintyre et al. 2005). In contrast, we observed a similar magnitude of association between BMI and heart failure for men and women, even after additional stratification by race. The First National Health and Nutrition Examination Survey (NHANES I) studied the association of excess BMI and incident heart failure(He, Ogden et al. 2001), using a dichotomous cut-point for BMI (27.8 kg/m^2 in men and 27.3 kg/m^2 in women). The adjusted hazard ratio was 1.23 (95 % CI =1.00-1.52) in men and 1.34 (95 % CI = 1.10 – 1.64) in women. This dichotomous cutpoint was based on the 85th percentile from NHANES I(Must, Spadano et al. 1999), and it occurs in the middle of the overweight group by the current classification scheme. Yet, their findings are very similar to the hazard ratios observed for our overweight group.

To date, studies of central obesity and heart failure have involved less generalizable populations. For example, the Health, Aging and Body Composition Study (Health ABC) is a longitudinal cohort study of those aged 70-79 without CHD, in which multiple metrics of overweight/obesity and body composition were measured(Nicklas, Cesari et al. 2006). This study had a limited number of heart failure events (N = 166), in part due to the exclusion of those with CHD, a common heart failure precursor(Kannel, D'Agostino et al. 1999). Despite the small sample size, BMI, waist circumference, and waist/thigh ratio were all positively associated with heart failure incidence in adjusted multivariable models. In contrast, a study of participants with CVD from the Heart Outcomes Prevention Evaluation (HOPE) study did

not find an association between obesity and incident heart failure (N=297 HF events) for BMI or waist/hip ratio; however when stratified by gender, waist-hip ratio was a significant predictor of heart failure in women only (HR = 2.30, 95 % CI = 1.25-4.21) (Dagenais, Yi et al. 2005).

We also describe risk groups using the cross-classification of two anthropometric measures, BMI and waist-hip ratio. Unlike waist circumference which is highly correlated with BMI, waist hip ratio seems a distinct measure of central adiposity.

We find that incidence rates increase across categories of BMI stratified by waist-hip ratio, however the ROC analysis does not support better prediction of incident heart failure with this cross-classification of BMI and waist-hip ratio as compared to these measures alone.

Further support for the association of obesity with heart failure comes from echocardiographic studies. Specifically, results from the Framingham Heart Study found positive correlations between obesity (BMI > 30 kg/m^2) and echocardiographic measures of left ventricular mass, which were also associated with increase in left ventricular internal dimensions, and wall thickness(Lauer, Anderson et al. 1991). A small study in normotensive men reported stronger positive correlations of left ventricular mass with waist/hip ratio and waist circumference as compared to BMI (Rasooly, Sasson et al. 1993). Recently, McGavock et. al. found that cardiac steatosis as seen with magnetic resonance spectroscopy occurs with impaired glucose tolerance even before the development of type 2 diabetes mellitus and/or left ventricular dysfunction(McGavock, Lingvay et al. 2007).

Obesity's role in the development of heart failure may be through either direct and/or indirect mechanisms. The indirect mechanisms are those in which obesity causes other diseases, such as diabetes(Ford, Williamson et al. 1997), hypertension(Stamler 1991) or

coronary heart disease, which are themselves risk factors for heart failure(Kannel, D'Agostino et al. 1999). A direct mechanism might be that cardiac adaptation to excess body fat can result in decreased cardiac function(Poirier, Giles et al. 2006). This has been termed obesity cardiomyopathy. Several mechanisms have been proposed for the cardiotoxic effect of fat cells. Research from animal models has lent support to a process termed lipotoxicity, which is the disruption of the usual mechanism that regulates triglyceride storage(Zhou, Grayburn et al. 2000) (McGavock, Victor et al. 2006). Another novel hypothesis is that an increase in inflammatory cytokines from excess adipocytes may increase risk of heart failure(Vasan, Sullivan et al. 2003; McGavock, Victor et al. 2006). The incidence of true obesity cardiomyopathy is low. Obesity cardiomyopathy is thought to occur most frequently amongst those with extreme obesity (BMI $\geq$ 40 kg/m^2) of greater than 10 years duration(Kaltman and Goldring 1976). In any given case of heart failure preceded by obesity, multiple mechanisms may be contributing factors.

The main limitation of this study is our definition of heart failure. We included hospitalized and fatal heart failure, as we did not have data on outpatient heart failure; however, community surveillance reports have indicated that 74 % of outpatient heart failure cases are hospitalized within 1.7 years(Roger, Weston et al. 2004). To address this limitation, we performed a sensitivity analysis to explore the effect of outcome misclassification on our findings. As would be expected with outcome misclassification, we found it should bias our findings toward the null.

This study is relevant and important to the understanding of the etiology of heart failure for three reasons. First, the ARIC study is the largest population-based cohort study to evaluate the association between waist/hip ratio and waist circumference and incident heart

failure. Furthermore, this is a well characterized cohort with a long period of follow-up for which standardized methodology was used for the measurement of BMI, waist-hip ratio and waist circumference. Because it is a large biracial study, we were able to describe this association stratified by race and gender.

In conclusion, we observed that obesity was associated with incident heart failure and there was a graded relation with body size. This association did not vary by race or gender. The consistency of findings for the association of overweight with incident heart failure is in striking contrast with the inconsistencies regarding the association for overweight with coronary heart disease and mortality (Barrett-Conner 1985; Wilson, D'Agostino et al. 2002; Flegal, Graubard et al. 2007). Selecting the best anthropometric for the prediction of heart failure could have implications for the screening and prevention of heart failure. The current study does not support the superiority of waist-hip ratio or waist circumference over BMI for the prediction of heart failure.

Acknowledgements

The Atherosclerosis Risk in Communities Study is carried out as a collaborative study supported by National Heart, Lung, and Blood Institute contracts N01-HC-55015, N01-HC-55016, N01-HC-55018, N01-HC-55019, N01-HC-55020, N01-HC-55021, and N01-HC-55022. The authors thank the staff and participants of the ARIC study for their important contributions.

Table 8. (MS. 1, Table1) Characteristics at baseline(1987-1989) of those who did or did not develop heart failure, ARIC

Characteristics	Incident Heart Failure (N = 1,405)	Non-cases (N = 13,237)
Age, mean (SD), years	56.8 (5.4)	53.8 (5.7)
Men, %	54	44
Black, %	35	25
Center, %		
Jackson, MS	31	23
Forsyth County, NC	24	26
Minneapolis, MN	17	27
Washington County, MD	28	25
Educational level, %		
Less than high school	40	21
High school graduate	30	33
Greater than high school	30	46
Smoking status, %		
Never smoker	29	43
Former smoker	34	32
Current smoker	37	25
Alcohol use, %		
Never drinker	27	25
Former drinker	29	17
Current drinker	44	58
Diabetes, %	32	9
Hypertension, %	55	30
Left ventricular hypertrophy by ECG, %	7	2
History of coronary heart disease, %	14	3
History of myocardial infarction, %	13	2
Plasma lipids, mean (SD), mg/dl		
Total cholesterol	220 (45)	214 (41)
LDL-C	143 (41)	137 (39)
HDL-C	46 (16)	52 (17)
Triglycerides	159 (115)	127 (85)
Anthropometric variables, mean (SD)		
Body mass index, kg/m^2	29.6 (6.1)	27.3 (5.1)
Waist circumference, cm	103 (15)	96 (13)
Waist/Hip ratio	0.97 (0.07)	0.92 (0.08)

Abbreviation: LDL-C, low density lipoprotein cholesterol; HDL-C, High density lipoprotein cholesterol

SI conversion: to convert total cholesterol values to g/L, multiply by 0.01; LDL-C, HDL-C to mmol/L, multiply by 0.0259; triglyceride to mmol/L, multiply by 0.113.

Table 9. (MS. 1, Table 2) Number of heart failure cases, total person-years of follow-up and age-adjusted* incidence rates (IR) for heart failure by category of each anthropometric measure (BMI, waist circumference, and waist-hip ratio), stratified by race and gender, ARIC, 1987-2003

	White Women (N = 5,659)				Black Women (N = 2,350)			
	Mean Value	Total person-years N (%)	Heart failure Events (N)	Age-adjusted* heart failure IR (95 % CI) per 1,000 person years	Mean Value	Total person-years	Heart failure events (N)	Age-adjusted* heart failure IR (95 % CI) per 1,000 person years
BMI, kg/m^2								
< 25	22.2	39,557 (48)	104	2.27 (1.86, 2.77)	22.7	5,742 (18)	39	5.42 (3.99, 7.37)
25-30	27.2	25,259 (31)	109	3.96 (3.48, 4.50)	27.5	11,410 (35)	78	7.69 (6.54, 9.03)
≥ 30	34.5	17,788 (21)	146	6.89 (5.77, 8.23)	35.8	15,135 (47)	175	10.91 (9.23, 12.89)
Waist circumference, cm								
1st Tertile, < 87	78.9	31,814 (39)	73	1.92 (1.50, 2.46)	79.4	6,762 (21)	31	4.17 (2.96, 5.88)
2nd Tertile, 87 ≤ 100	92.6	28,110 (34)	94	3.57 (3.10, 4.13)	93.2	10,667 (33)	68	7.08 (5.92, 8.46)
3rd Tertile, ≥ 100	111.2	22,680 (27)	192	6.64 (5.56, 7.93)	113.4	14,857 (46)	193	12.01 (10.15, 14.22)
Waist-hip ratio								
1st Tertile, < 0.86	0.80	30,449 (37)	58	1.95 (1.57, 2.42)	0.80	10,136 (31)	35	4.28 (3.15, 5.81)
2nd Tertile, 86 ≤ 0.93	0.89	25,910 (31)	97	3.48 (3.08, 3.94)	0.89	9,353 (29)	79	7.48 (6.35, 8.81)
3rd Tertile, ≥ 0.93	0.98	26,245 (32)	204	6.22 (5.36, 7.22)	0.98	12,797 (40)	178	13.07 (10.99, 15.53)
	White Men (N = 5,135)				**Black Men (N = 1,498)**			
BMI, kg/m^2								
<25	23.1	19,561 (27)	113	3.92 (3.30, 4.67)	22.4	5,787 (30)	49	6.85 (5.23, 8.99)
25-30	27.3	36,665 (51)	256	6.12 (5.51, 6.79)	27.4	8,434 (43)	70	9.47 (8.11, 11.05)
≥ 30	33.0	15,010 (21)	190	9.54 (8.25, 11.03)	33.6	5,271 (27)	76	13.07 (10.44, 16.37)
Waist circumference, cm								
1st Tertile, <87	88.9	24,020 (34)	108	4.01 (3.36, 4.78)	86.2	8,972 (46)	66	6.84 (5.39, 8.67)
2nd Tertile, 87 ≤ 100	98.4	23,715 (33)	193	5.92 (5.30, 6.61)	98.5	5,108 (26)	44	9.81 (8.44, 11.41)
3rd Tertile, ≥100	110	23,501 (33)	258	8.75 (7.66, 10.01)	111.8	5,413 (28)	85	14.01 (11.34, 17.46)
Waist-hip ratio								
1st Tertile, <0.86	0.91	20,995 (29)	78	3.07 92.53, 3.72)	0.90	10,197 (52)	66	6.39 (4.97, 8.21)
2nd Tertile, 86 ≤ 0.93	0.96	21,479 (30)	131	5.33 (4.77, 5.97)	0.96	4,866 (25)	46	10.20 (8.68, 11.98)
3rd Tertile, ≥0.93	1.02	28,761 (40)	350	9.28 (8.25, 10.44)	1.01	4,429 (23)	83	16.27 (12.83, 20.6)

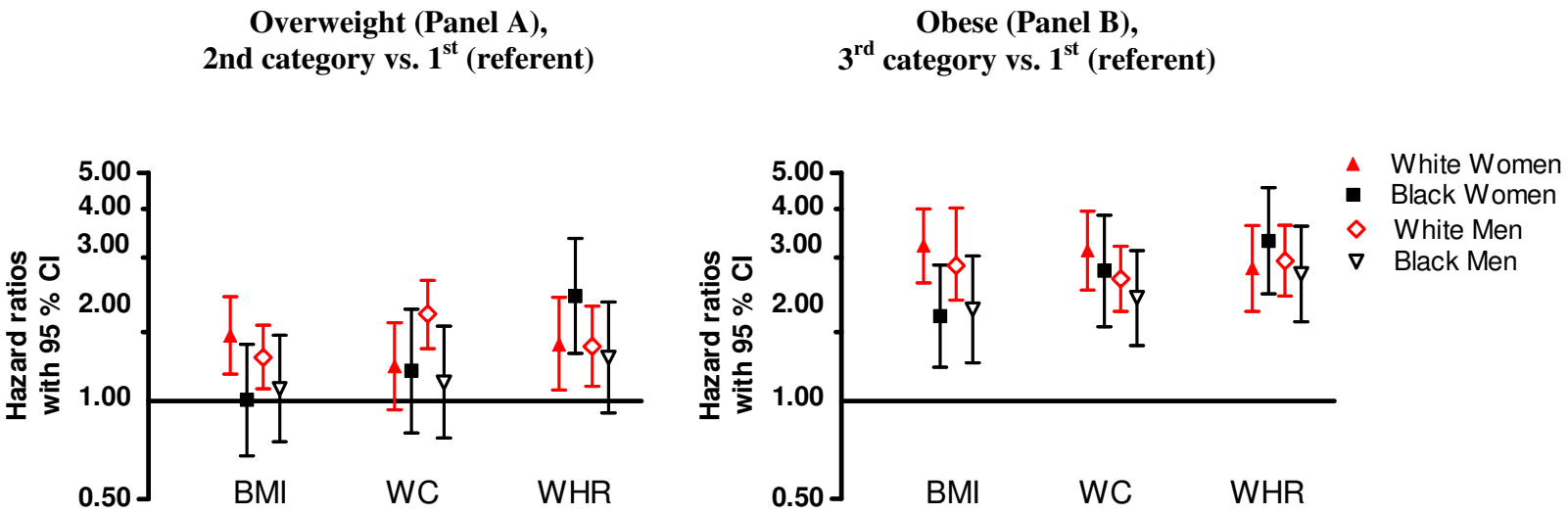
*age-adjusted to mean age at baseline, 54 years

Table 10. (MS. 1, Table 3) Unadjusted and adjusted* hazard ratios (with 95 % CI) for incident heart failure by BMI, waist circumference and waist-hip ratio, stratified by race and gender, ARIC, 1987-2003

	Women			Men		
	White women HR (95 % CI)	**Black women HR (95 % CI)**	**All women HR (95 % CI)**	**White men HR (95 % CI)**	**Black men HR (95 % CI)**	**All men HR (95 % CI)**
BMI						
Overweight vs. normal, crude	1.62 (1.25, 2.15)	1.00 (0.68, 1.47)	1.62 (1.30, 2.01)	1.21 (0.97, 1.51)	0.97 (0.67, 1.40)	1.13 (0.93, 1.36)
Overweight vs. normal, adjusted	1.59 (1.21, 2.09)	1.01 (0.68, 1.49)	1.38 (1.10, 1.72)	1.36 (1.09, 1.71)	1.09 (0.75, 1.59)	1.31 (1.08, 1.58)
Obese vs. normal, crude	3.15 (2.45, 4.05)	1.70 (1.19, 2.40)	3.13 (2.57, 3.81)	2.22 (1.76, 2.80)	1.69 (1.18, 2.42)	2.07 (1.70, 2.51)
Obese vs. normal, adjusted	2.99 (2.30, 3.88)	1.82 (1.27, 2.61)	2.52 (2.03, 3.12)	2.60 (2.04, 3.30)	1.91 (1.31, 2.78)	2.43 (1.98, 2.98)
Continuous, 1 SD change†, adjusted	1.64 (1.48, 1.81)	1.35 (1.22, 1.49)	1.47 (1.37, 1.57)	1.51 (1.39, 1.63)	1.34 (1.20, 1.49)	1.45 (1.36, 1.54)
Waist Circumference						
2nd Tertile vs. 1st tertile, crude	1.46 (1.07, 1.98)	1.38 (0.90, 2.10)	1.55 (1.21, 1.98)	1.78 (1.41, 2.25)	1.17 (0.80, 1.71)	1.54 (1.27, 1.87)
2nd Tertile vs. 1st tertile, adjusted	1.28 (0.94, 1.74)	1.24 (0.80, 1.91)	1.27 (0.99, 1.63)	1.85 (1.45, 2.34)	1.14 (0.77, 1.70)	1.63 (1.34, 1.99)
3rd tertile vs. 1st tertile, crude	3.73 (2.85, 4.88)	2.86 (1.95, 4.17)	3.86 (3.11, 4.78)	2.38 (1.90, 2.97)	2.16 (1.56, 2.97)	2.20 (1.84, 2.64)
3rd tertile vs. 1st tertile, adjusted	2.89 (2.19, 3.82)	2.51 (1.69, 3.71)	2.70 (2.15, 3.39)	2.37 (1.88, 2.98)	2.07 (1.48, 2.89)	2.26 (1.88, 2.73)
Continuous, 1 SD change†, adjusted	1.61 (1.46, 1.78)	1.46 (1.32, 1.63)	1.54 (1.43, 1.65)	1.56 (1.44, 1.69)	1.40 (1.26, 1.56)	1.50 (1.40, 1.60)
Waist-hip Ratio						
2nd Tertile vs. 1st tertile, crude	1.98 (1.43, 2.73)	2.45 (1.65, 3.65)	2.19 (1.70, 2.81)	1.65 (1.25, 2.19)	1.47 (1.01, 2.14)	1.46 (1.17, 1.82)
2nd Tertile vs. 1st tertile, adjusted	1.49 (1.08, 2.08)	2.10 (1.40, 3.15)	1.73 (1.34, 2.24)	1.47 (1.11, 1.95)	1.36 (0.92, 2.01)	1.41 (1.13, 1.77)
3rd tertile vs. 1st tertile, crude	4.12 (3.08, 5.52)	4.08 (2.84, 5.86)	4.32 (3.45, 5.42)	3.34 (2.61, 4.27)	2.98 (2.16, 4.12)	2.88 (2.38, 3.47)
3rd tertile vs. 1st tertile, adjusted	2.55 (1.88, 3.44)	3.10 (2.13, 4.51)	2.76 (2.18, 3.49)	2.69 (2.10, 3.45)	2.45 (1.75, 3.43)	2.58 (2.12, 3.14)
Continuous, 1 SD change†, adjusted	1.57 (1.40, 1.76)	1.67 (1.47, 1.89)	1.62 (1.48, 1.76)	1.51 (1.40, 1.63)	1.47 (1.29, 1.67)	1.50 (1.41, 1.60)

*Adjusted models included age, alcohol use, educational level, smoking status. Models of all men and women also adjust for race and center.
Adjusted models exclude 243 participants with missing covariates: smoking status (N=9), educational level (N=23), and alcohol use (N=218).
†1 SD change for BMI = 6 kg/m^2 for women, 4.2 kg/m^2 for men; 1 SD change for waist circumference = 15.4 cm for women and 10.9 cm for men; 1 SD change for waist-hip ratio = 0.08 for women and 0.05 for men.

Figure 7. (MS. 1, Figure 1) Adjusted hazard ratios (with 95 % CI) for the association of body mass index (BMI), waist circumference (WC) and waist-hip ratio (WHR) with incident heart failure, stratified by gender and race, ARIC, 1987-2003. Shown are overweight (Panel A) and obesity (panel B) as compared with the lowest category for that respective measure.



Models adjusted for educational level, age, smoking status, and alcohol use. Adjusted models exclude 243 participants with missing covariates: smoking status (N=9), educational level (N=23), and alcohol use (N=218)

Table 11. (MS. 1, Table 4) Numbers of heart failure events, person-years of follow-up, age-adjusted heart failure incidence rates and multivariable adjusted hazard ratios of BMI stratified by high or low waist-hip ratio (WHR)[Ω], ARIC 1987-2003

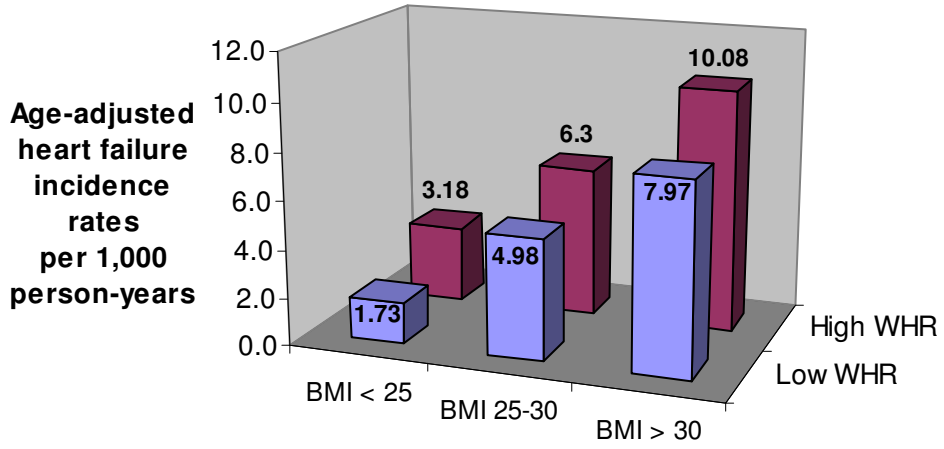
	Heart failure events, N	Person-years Of follow-up	*Age-adjusted heart failure incidence rates per 1,000 person-years (95 % CI)	†Multivariable adjusted HR (95 % CI)
BMI < 25 kg/m^2, Low WHR	164	47,838	3.11 (2.77, 3.50)	1.00
BMI< 25 kg/m^2, High WHR	137	21,722	3.94 (3.59, 4.31)	1.48 (1.18, 1.86)
BMI 25-30 kg/m^2, Low WHR	119	30,795	4.98 (4.64, 5.35)	1.14 (0.90, 1.45)
BMI 25-30 kg/m^2, High WHR	382	49,679	6.30 (5.93, 6.70)	1.86 (1.54, 2.23)
BMI > 30 kg/m^2, Low WHR	48	9,169	7.97 (7.46, 8.52))	1.73 (1.25, 2.41)
BMI > 30 kg/m^2, High WHR	524	43,060	10.08 (9.27, 10.97)	3.18 (2.66, 3.81)

*Age-adjusted to mean age at baseline, 54 years.

†Multivariable Cox proportional hazards models adjusted for age, educational level, race and center, gender, smoking status and alcohol use: excludes 243 participants with missing covariates: smoking status (N=9), educational level (N=23), and alcohol use (N=218)

[Ω]Waist-hip ratio cutpoints were: 0.88 for women and 0.95 for men

Figure 8. (MS. 1, Figure 2) *Age-adjusted heart failure incidence rate (per 1,000) person-years by categories of BMI, stratified by category of waist-hip ratio (cutpoints were = 0.88 for women and 0.95 for men), ARIC 1987-2003



*Age-adjusted to mean age at baseline, 54 years

B. Manuscript 2: The preventable burden of heart failure due to obesity: The Atherosclerosis Risk in Communities (ARIC) Study

ABSTRACT

Prevention of heart failure is a priority as it is associated with significant personal suffering, healthcare utilization and costs. We estimated the proportion of incident heart failure that would be prevented from a range of feasible reductions in obesity and/or overweight, by calculation of the generalized impact fractions (GIF). Methods: The Atherosclerosis Risk in Communities (ARIC) study is an ongoing bi-racial population-based cohort of those aged 45-65 years from 4 U.S. communities with 14 years median follow-up for incident, hospitalized or fatal heart failure. BMI was measured at the baseline visit (1987-1989) and categorized as normal weight (BMI <25 kg/m^2), overweight (BMI 25-30 kg/m^2), and obese (BMI $\geq$ 30 kg/m^2). After exclusion of baseline prevalent heart failure, missing anthropometry measures, and poorly represented race groups, the sample size was 14,642. The GIF was calculated for each age-, race- and gender- stratified group and then combined overall groups using a case-load weighted-sum method. A 95 % distribution of the GIF was estimated from 10,000 bootstrapped datasets. Results: A 30% hypothetical reduction in both obesity and overweight would prevent 8.5 % (95 % intervals = 6.1%, 10.7%) of incident heart failure cases, whereas a 30 % reduction in obesity only would prevent 6.7 % (95 % intervals = 5.4 %, 8.0 %) of incident heart failure. However, a 6.7 % reduction in heart failure would result in 44,220 fewer incident heart failure cases with approximately $477,674,356 cost savings per year. The attributable fraction (28 % [95 % intervals = 20 %, 36 %]), which assumes complete elimination of obesity and overweight, is approximately 3 times larger than the highest amount of weight reduction (30 % reduction in obesity) assumed feasible here. Conclusion: Calculation of the GIF was enlightening as to the actual impact of an aggressive

reduction in obesity and overweight. The overall GIF number appears small, however when applied to the actual impact on healthcare utilization and cost, it has a large impact since HF is a common condition that is expensive to manage. We propose studies should consider estimation of the generalized impact fraction in addition to the attributable fraction, in order to avoid estimates of population impact based on unachievable goals, such as eradication of obesity.

1. Introduction

Heart failure (HF) is common, with an estimated 5,300,000 American adults living with HF, and it is increasing in prevalence(Rosamond, Flegal et al. 2008). Hospital discharges for heart failure in the U.S. have increased 171 % since 1979, with 660,000 incident heart failure cases expected in the U.S. for 2008 with nearly 35 billion in U.S healthcare costs{Rosamond, 2008 #201}. Clearly the prevention of heart failure must be a priority. Obesity is a known risk factor for heart failure(Kenchaiah, Narula et al. 2004). In fact, obesity has been recognized as a qualifying risk factor for stage A heart failure, a pre-heart failure stage, in the latest staging scheme for heart failure from the American Heart Association (AHA)(Hunt, Abraham et al. 2005). A recent scientific statement from the AHA was released specifically on the prevention of heart failure; it recommends maintenance of normal weight as one way to prevent heart failure(Schocken, Benjamin et al. 2008). The actual impact of a population-level reduction in obesity on the potential prevention of heart failure has been assessed only in terms of the attributable fraction, which estimates the proportional reduction in disease given complete elimination of an exposure(He, Ogden et al. 2001; Kenchaiah, Narula et al. 2004). For many exposures, complete elimination is impossible or highly unlikely, a

circumstance applicable to obesity and overweight as exposures that are unlikely to be eliminated from society(Benichou 2007). Despite this, the attributable fraction has been used extensively in the obesity literature, including recent articles estimating the burden of all-cause mortality attributable to overweight and obesity(Flegal, Graubard et al. 2004; Flegal, Graubard et al. 2005).

A more realistic and meaningful calculation for estimating the impact of weight reduction on disease incidence is the generalized impact fraction (GIF), also known as the potential impact fraction and the generalized attributable fraction(Benichou 2001; Vander Hoorn, Ezzati et al. 2004). It is a generalization of the attributable fraction which estimates the proportional reduction in disease incidence given a graded reduction in the prevalence of a risk factor. For common risk factors and diseases, the impact of a hypothetical reduction in the exposure may reveal an important effect on disease incidence even when risk factor-disease associations are relatively weak(Morgenstern and Bursic 1982). The GIF was originally described by Walter in 1980(Walter 1980), then further elucidated and coined by Morgenstern and Bursic in 1982(Morgenstern and Bursic 1982). Specifically, it is "the proportional reduction in the total number of new (incident) cases of a certain disease, resulting from a specific change in the distribution of a risk factor in the population at risk."(Morgenstern and Bursic 1982). Despite its introduction over 25 years ago, it hasn't caught on as either a replacement for, or complement to, the attributable fraction. Despite it's limited adoption in the literature as compared to the attributable fraction, the generalized impact fraction has been adopted recently in the literature on the global burden of disease(Murray and Lopez 1999), health policy(Nilunger, Diderichsen et al.

2004),environmental epidemiology(Semenza, Rubin et al. 1996) and nutritional epidemiology(Graham 2000).

The goal of this study is to estimate the population burden of incident heart failure that would be prevented from a feasible reduction in the prevalence of obesity and overweight, by calculation of the generalized impact fraction. Several scenarios of reduced prevalence of obesity and overweight are considered. In addition, we calculate the attributable fraction to compare this more commonly used measure with the GIF.

2. Methods

Study population

The ARIC cohort (N=15,792) was recruited between 1987-1989 using probability sampling of those aged 45-64 from the following four US communities: Forsyth County, North Carolina; Jackson, Mississippi; suburbs of Minneapolis, Minnesota; and Washington County, Maryland. The distribution of blacks and whites from each field center was representative of the area (mostly white in Minneapolis and Washington County, 15 % black in Forsyth County) except in Jackson where only blacks were sampled. Response to the baseline examination was 46 % in Jackson and between 65-67 % for the other communities. Three subsequent examinations were conducted at approximately three-year intervals through 1996-98. The design and rationale of the ARIC study (The ARIC Investigators 1989) and the comparison between responders and nonresponders (Jackson, Chambless et al. 1996) have been previously published. The institutional review boards from each study site approved the ARIC study. Participants provided written informed consent at each examination.

Racial groups not classified as White or Black (N = 48) and Blacks not from the city of Jackson or Forsyth County (N = 120) were excluded from analyses due to their limited numbers. In addition, those with missing anthropometry (N = 33) or those with prevalent heart failure (HF) at baseline were excluded from this analysis. Criteria to define prevalent HF were as follows: 1) those answering "yes" to the following question: "Were any of the medications you took during the last two weeks for heart failure?" (N = 83), or 2) those with stage 3 or 'manifest HF' by applying Gothenburg criteria (N = 699)(Eriksson, Caidahl et al. 1987). Gothenburg criteria were defined by self-reported medical history, medication lists and electrocardiography. After these exclusions, the total sample size was 14,399.

Anthropometric and covariate measurement

Body mass index (BMI) was calculated as weight (kg) divided by height (meters^2) squared. Height and weight were measured after an overnight fast with participants in standard scrub attire. Technicians measured height with participants barefoot using a wall mounted ruler. Weight measurements were performed using a scale (Detecto model 437) that was zeroed daily and calibrated quarterly.

All covariates are drawn from the baseline visit. Race, gender, educational level, alcohol use, and smoking status were obtained by participant self-report. History of coronary heart disease included those with a history of myocardial infarction (MI) (either self-report of physician-diagnosed MI or silent MI as identified by electrocardiography), and those with a prior coronary revascularization procedure or coronary artery bypass surgery. Hypertension was defined by either a diastolic blood pressure ≥ 90 mm Hg or a systolic blood pressure ≥ 140 mm Hg measured with random-zero mercury manometers, or anti-hypertensive

medication use during the previous 2 weeks. Diabetes mellitus was defined as any of the following: self-reported history of physician diagnosed diabetes, medication use for diabetes over the last 2 weeks, a blood glucose ≥ 126 mg/dl fasting or a blood glucose ≥ 200 mg/dl non-fasting.

Ascertainment of heart failure events

The following methods were used for ascertainment of HF events: 1) participants were contacted annually by phone and interviewed about interim hospitalizations; 2) local hospitals provided lists of hospital discharges with cardiovascular diagnoses and these were reviewed to identify cohort hospitalizations; and 3) health department death certificate files were continuously surveyed. All discharge codes for cohort hospitalizations and listed causes of death from death certificates were recorded (underlying cause of death or contributory causes).

Incident heart failure event criteria

Incident HF was defined as the first occurrence of either: 1) a hospitalization which included an ICD-9-CM (International Classification of Diseases, 9th revision, clinical modification) discharge code for heart failure beginning with '428' (428.0 - 428.9) in any position (N = 1,329) or else 2) a death certificate beginning with ICD-9 code '428' or ICD-10 code 'I50' (I50.0 – I50.9) in any position (N = 76). Follow-up time for those with an incident HF event was defined as the time from the date of their baseline examination until the incident event with follow-up through December 31st, 2003. The end of follow-up time

for those without HF was the first occurrence of either: 1) December 31st, 2003, 2) date of last contact for those lost to follow-up, or 3) date of death.

Estimates for the reduced prevalence of obesity and overweight

To determine feasible goals for weight reduction, we considered the goals set by the US Department Health and Human services in the Healthy People 2010 report. Healthy People 2010 is a set of recommended health objectives meant to serve as a basis from which to develop US national, state and community plans for health improvement by the year 2010(2000) . Obesity and overweight are listed as high priority public health issues. Specifically, Healthy People 2010 aim to reduce the proportion of obese adults aged 20 years and older to 15 % from 23% by the year 2010 (25 % in females and 20 % in males, data from 1988-1994, age adjusted to year 2000), which is approximately a 35 % reduction in the proportion of obese. We chose several scenarios for the reduction in the proportion of obese and overweight, for which the maximum achievable goal selected was a 30 % reduction in both obesity and overweight. We chose a 30 % reduction in each subgroup, rather than targeting 15 % prevalence of obesity for all subgroups, as certain subgroups, such as Black women in which 47-48 % were obese at study baseline, would require an impractical amount of widespread weight reduction to meet this goal.

Statistical methods

The generalized impact fraction for heart failure given a reduction in overweight and or obesity was estimated using equations 1 and 3 (see box). Specifically, for this study, the GIF was calculated for each age-, race- and gender- stratified group (see equation 1) and then

combined using a case-load weighted sum method to obtain an overall GIF (see equation 3). For each stratum, the components of GIF equation were defined as follows: 1) P_i is the prevalence of the exposure (obesity, overweight and normal weight) in the ARIC population at baseline (1987-1989); 2) P_i* is the reduced prevalence of obesity, overweight and normal weight (reduced as compared to P_i) after a hypothetical change in the distribution of BMI; and 3) RR_i, in this case, is the crude hazard ratio of incident heart failure for those overweight, obese or normal weight as compared to the referent group (normal weight), for that stratum. Cox proportional hazard regression was used to model the association between obesity (BMI $\geq$ 30 kg/m^2), overweight (BMI $\geq$ 25 and $\leq$ 30, kg/m^2) with normal weight (BMI $<$ 25 kg/m^2) and time to incident HF. BMI was categorized as represented in the clinical guidelines from the National Institutes of Health(NIH 1998). In the multivariable model, ties were handled with the method suggested by Efron. All statistical analyses were performed using SAS software v 9.1 (Cary, NC).

Since calculation of the GIF is based on a single study population (the ARIC cohort), we added further uncertainty to these calculations by bootstrapping to obtain 95 % simulation intervals for the GIF and attributable fraction (see Figure 9 (MS. 2, Figure 1))(Greenland 2004). Ten thousand bootstrapped samples (with replacement) were performed for each subgroup of race, gender and age (therefore, there were 8 subgroups each with 10,000 bootstrapped samples). Each sampled dataset was the same size as the subgroup in the original dataset. The unadjusted hazard ratios (HR_i) for obesity and overweight (the HR for normal weight is always 1), and the prevalence (P_i) of normal weight, overweight and obesity were determined within each stratum for each of the 10,000 bootstrapped dataset. The GIF

for each stratum was calculated given these values and a specified alternative distribution for the reduced prevalence of obesity, overweight and normal weight (P_i*).

The overall GIF for each hypothetical scenario of weight reduction was determined using the case-load weighted-sum approach for each bootstrapped sample using equation 2 (Benichou 2001). From the distribution of the 10,000 overall generalized impact fractions, the median (along with 2.5 % and 97.5 % simulation intervals) were calculated for each hypothetical scenario. The attributable fraction, or in this case the maximum GIF, was calculated using the same case-load weight-sum method using equation 2(Morgenstern and Bursic 1982). Based on the distribution of values from the bootstrapped samples, the median attributable fraction along with 2.5 and 97.5 percentiles are reported.

(Equation 1)

$$\text{GIF for a given bootstrap sample, and stratum a,b,c} = \frac{\sum P_{i,\,abc}(RR_{i,\,abc}) - \sum P_i^{*}{}_{,abc}\,(RR_{i,\,abc})}{\sum P_{i,\,abc}(RR_{i,abc})}$$

Where,
GIF = generalized attributable fraction for one bootstrap sample, from a given stratum of abc,
P_i = proportion of the population in exposure category, i,
P_i^*= proportion of the population in exposure category i after an intervention or other change,
RR_i = Relative risk at exposure category i compared to the reference level,
i = normal weight, overweight, or obese categories as defined by BMI,
abc = stratum of age (age < 55 years, age ≥ 55 years), race (black or white) and gender (male or female)

Specifically,

$$\text{Numerator } GIF_{a,b,c} = [P_{overwt,\,abc}(RR_{overwt,\,abc}) + P_{obese,\,abc}(RR_{obese,\,abc}) + P_{normal\ wt,\,abc}(1)] - [(P^*_{overwt,\,abc})(RR_{overwt,\,abc}) + (P^*_{obese,\,abc})(RR_{obese,\,abc}) + (P^*_{normal\ wt,\,abc})(1)]$$

$$\text{Denominator } GIF_{a,b,c} = [(P_{overwt,\,abc})(RR_{overwt,\,abc}) + (P_{obese,\,abc})(RR_{obese,\,abc}) + (P_{normal\ wt,\,abc})(1)]$$

(Equation 2)

AF (or maximum GIF) for a given bootstrap sample from stratum a, b, c =

$$[\sum(pd_i(RR_i) - 1)] / [\sum(pd_i(RR_i)]$$

where pd_i = proportion of total cases arising from the ith exposure category
RR_i = the unadjusted RR (in this case HR) for the ith exposure category compared to referent (i = overweight, or obese categories compared to normal)

(Equation 3)

Case-load weighted sum method for $GIF_{overall}$

Numerator $GIF_{overall}$ (or $AF_{overall}$) for a one bootstrap sample =
$(\#cases_{a1b1c1})(GIF_{a1b1c1}) + (\#cases_{a1b1c0})(GIF_{a1b1c0}) +$
$(\#cases_{a0b1c1})(GIF_{a0b1c1}) + (\#cases_{a0b1c0})(GIF_{a0b1c0}) +$
$(\#cases_{a0b1c1})(GIF_{a0b1c1}) + (\#cases_{a0b1c0})(GIF_{a0b1c0}) +$
$(\#cases_{a1b0c1})(GIF_{a1b0c1}) + (\#cases_{a1b0c0})(\ GIF_{a1b0c0}) +$
$(\#cases_{a1b0c1})(GIF_{a0b0c1}) + (\#cases_{a0b0c0})(\ GIF_{a0b0c0}) +$
$(\#cases_{a0b0c1})(\ GIF_{a0b0c1}) + (\#cases_{a0b0c0})(GIF_{a0b0c0})$

Denominator $GIF_{overall}$ (or $AF_{overall}$) for a given boostrap sample =
$(\#cases_{a1b1c1}) + (\#cases_{a1b1c0}) + (\#cases_{a0b1c1}) + (\#cases_{a0b1c0}) + (\#cases_{a0b1c1}) +$
$(\#cases_{a0b1c0}) + (\#cases_{a1b0c1}) + (\#cases_{a1b0c0}) + (\#cases_{a1b0c1}) + (\#cases_{a0b0c0}) +$
$(\#cases_{a0b0c1}) + (\#cases_{a0b0c0})$

3. Results

There were 3,887 obese and 5,787 overweight participants in the ARIC cohort at baseline (Table 12, (MS 2. Table 1). Obese participants were more often Black, female, with less than high school education, and more likely to have diabetes, or hypertension. Furthermore, they were less likely to be smokers, and current drinkers of alcohol. There were a larger percentage of obese participants from Jackson than any other study site; only African-Americans were represented from this site. Prevalent coronary heart disease was less common among normal weight participants compared to those overweight or obese.

The main components of the GIF formula are shown by subgroup in Table 13 (MS. 2, Table 2), namely the number of heart failure cases, the proportion of participants in each category of the exposure variable (normal weight, overweight, or obese), and the unadjusted hazard ratios for the involved comparisons (overweight and obese as compared to referent). This table also shows incidence rates of HF by subgroup. Across all race and gender groups, the incidence rates of HF are at least two times larger for older compared to the younger subgroups. Older Black women and Black men have particularly high HF incidence rates (14.0 and 15.2 per 1,000 person-years, for women and men, respectively). Although there are fewer Blacks than Whites across all age and gender groups, the GIF formula weights by the number of heart failure cases in each subgroup. For example, while the number of older Black men (N = 668) is over four times smaller than younger White women (N = 3,079), older Black men have nearly the same weight as young White women in the overall GIF formula since they have a similar number of heart failure cases (120 and 119, respectively). Approximately half of White women were in the normal weight category whereas only 18 % of Black women were normal weight; nearly half of Black women were obese. There was a higher proportion of overweight among White men (over

50 %) compared to Black men, whereas there was a higher proportion obese among Black men compared to White men. For all strata, the unadjusted hazard ratios and their confidence intervals for obesity as compared to referent were above 1.0 representing a harmful effect of obesity on the incidence of heart failure. Hazard ratios for heart failure were particularly high among young White women (HR = 4.14 (95 % CI = 2.70, 6.36) although the confidence intervals overlap for most other subgroups. The hazard ratios for overweight are lower in magnitude with confidence intervals crossing 1.0 in most cases. In fact, two of the hazard ratios for the comparison of overweight with referent were less than 1.0, although the confidence intervals also cross 1.0 (young Black men, HR = 0.93 (95 % CI, 0.49, 1.76); young Black women, HR = 0.94 (0.52, 1.69)).

The goal of healthy people 2010 for a reduction in obesity to 15 % among the U.S. adult population is similar to our goal here of a 30 % reduction in the proportion obese. Table 14 (MS 2, Table 3) shows the proportion obese before and after a hypothetical intervention resulting in a 30 % reduction in the proportion obese (shifting all who lost weight from the obese category to normal weight). In most subgroups, a hypothetical 30 % reduction in obesity results in a prevalence of 15-19 % obesity in the population of interest. Black women were the exception since there would be a prevalence of 33-34% of obesity for Black women even after a 30 %. The median GIF (with 95 % simulation intervals) from 10,000 bootstrapped datasets for a 30 % reduction in obesity range from 4% to 11% for the eight subgroups. The 2.5 % tail of the distribution of median GIF is negative for older Black men. The overall median GIF for the scenario of a 30 % reduction in the proportion obese (scenario 6 from Table 15 (MS. 2, Table 4)) was 6.7 % (95 % intervals = 5.4, 8.0). This can be interpreted to mean that 6.7% of incident

heart failure could be prevented in this population with a 30% reduction in obesity, assuming weight loss down to the normal weight category.

The overall median GIF with 95 % simulation intervals for 10 different scenarios of weight reduction are shown in Table 15 (MS. 2, Table 4) and Figure 10 (MS. 2, Figure 2). The most aggressive weight reduction scenario assumed here, other than complete elimination as required in calculation of the attributable fraction, is a 30% reduction of obese and overweight down to the normal weight category (scenario 8, see Table 15 (MS. 2, Table 4)). We predict that scenario 8 would result in the prevention of 8.5 % (95 % intervals = 6.1%, 10.7%) of incident heart failure cases. The last two scenarios represent the attributable fraction which assumes complete elimination of obesity and overweight in the case of the 10^{th} scenario, whereas the 9^{th} scenario assumes complete elimination of obesity only. The attributable fraction (scenario 10) is approximately 3 times larger than our most aggressive scenario of weight reduction presented here (scenario 8). There is only a small change in the GIF when those who are obese and have lost weight are shifted to normal weight (scenarios 2 and 6), rather than overweight (scenarios 1 and 5). This implies that weight loss among those who are obese down to the overweight group would have a similar population-level impact on heart failure incidence as weight loss down to the normal weight group.

4. Discussion

We examined the possible public health impact on heart failure incidence given multiple hypothetical and feasible reductions in obesity and/or overweight, and found that the practical level of impact appears small, but may be potentially important. We found that a 30 % reduction in obesity with presumed weight loss to the normal weight category, the goal most similar to

Healthy People 2010 goal of 15 % obese, would hypothetically result in 6.7 % fewer cases of incident heart failure. It is estimated there will be 660,000 incident heart failure events in 2008. The average cost per hospital discharge for HF (estimated from 2003) was $6,577(Rosamond, Flegal et al. 2008). To extrapolate our findings to these estimates, a 6.7 % reduction in heart failure would have resulted in 44,220 fewer incident heart failure cases with an estimated $477,674,356 cost savings for that year. Due to the high prevalence of heart failure and frequent need for re-hospitalization, such an impact would offer important reduction in disease burden and healthcare utilization. Although not accounted for in these calculations, population-wide weight reduction would also have a preventive effect on common heart failure risk factors, such as hypertension(Levy, Larson et al. 1996) and diabetes(Masoudi and Inzucchi 2007), resulting in further reductions in healthcare cost and utilization.

Another goal from healthy people 2010 is an increase in the percentage of normal weight up to 60 %(US Department of Health and Human Services 2000). Even with a 30 % reduction in overweight and obese, we estimate we still will not have achieved this goal (data not shown); instead the percentage of normal weight after a 30 % reduction in obesity and overweight would be between 43–54 % per subgroup, except for among White women in which it would be 61 - 65%. More widespread weight reduction would be needed to accomplish this goal among race and gender groups other than White women.

Our findings for the generalized impact fraction are relatively small compared to our much higher estimate for the attributable fraction of 28 %. As compared to the attributable fraction, the generalized impact fraction is a measure of preventable burden of disease that can be based on multiple hypothetical changes in the exposure distribution. In fact, although not shown here, the generalized impact fraction can be used to predict the effect of a given shift in the

distribution of a continuous exposure measurement. To date, the generalized impact fraction has not been used elsewhere to describe the preventable burden of heart failure due to obesity. However, two other studies have calculated the attributable fraction. First, the Framingham Heart Study estimated the attributable fraction for heart failure was 14 % for overweight women, 13.9 % for obese women, 8.8 % for overweight men and 10.9 % for obese men(Kenchaiah, Evans et al. 2002). They used an equation for the attributable fraction appropriate when the relative risk has been adjusted for confounders. The Framingham population has a lower prevalence of obesity compared to the ARIC study and they adjusted for causal intermediates which lowers the magnitude of the effect estimate. Secondly, the First National Health and Nutrition Examination Survey (NHANES I) calculated the attributable fraction of heart failure for overweight (5.6 % in men; 9.6 % in women), however they dichotomized BMI at 27.8 kg/m^2 for men and 27.3 kg/m^2 for women(He, Ogden et al. 2001). In both studies, confidence intervals for attributable fraction were not estimated. These estimates of attributable fraction are lower than those estimated here. This is due to differences in the prevalence of obesity and overweight in the populations, and the magnitude of the relative estimate of effect.

There are several assumptions of the generalized impact fraction and the attributable fraction(Morgenstern and Bursic 1982). The relative risk, or in this case hazard ratio, must be a valid estimate for a causal relationship between exposure and disease for the desired target population, and not biased by confounding or misclassification. In this case, we have extrapolated from our study population, the ARIC cohort, to the U.S. population, our target. Also, we assume that those who are hypothetically shifted (obese) to a different risk group (normal weight) post-intervention, now have the same risk that their new risk group had pre-intervention. We have assumed those who lose weight will have the same risk of heart failure as

those who were normal weight without intervention. When this condition is not met then the generalized impact fraction may overestimate the impact of the intervention. Also, it is assumed there are not secular trends in risk of disease that are not due to the intervention; and there are not secular trends in intermediary diseases, such as diabetes for obesity, that would affect the relationship between the exposure and disease. There are secular trends showing increasing incidence of heart failure, obesity, and diabetes(Rosamond, Flegal et al. 2008), which is one reason that prevention of disease has become vital. Although all of these criteria are not easily met, as in this case, these are the same assumptions made when estimating the often used attributable fraction.

Limitations of this analysis and the generalized impact fraction in general, include that unlike the attributable fraction, one must stratify by important confounders rather use an adjusted relative measure of effect such that small sample sizes can be a problem. Although in the current study, there were few confounders identified using a 10 % change in estimate criterion. Inherent in estimating the generalized impact fraction, one must consider a counterfactual approach(Rodgers 2002), and in so doing we wanted to use feasible goals for population-wide weight reduction. It is difficult to determine feasible goals for population-wide weight reduction; however, we considered a spectrum of weight loss with the goals recommended in the Healthy People 2010 report as our most optimistic goal.

In 1982, Morgenstern and Bursic illustrate the use of the generalized impact fraction with an example of the impact of a hypothetical weight loss programs on the incidence of coronary heart disease(Morgenstern and Bursic 1982). Similarly, we estimated the effect of a feasible reduction in obesity on the burden of heart failure. These examples show how this important measure, the generalized impact fraction, should be routinely used in practice for risk factors that are unlikely

to be completely eradicated. For common risk factors such as obesity and overweight, the impact of intervention may reveal important changes in the burden of disease even when risk factor-disease associations are relatively weak. Reduction in the incidence of heart failure of 6.7%, the predicted impact of a 30% reduction in the prevalence of obesity, would result in an impressive reduction in the number of incident for heart failure events. We propose studies should consider estimation of the generalized impact fraction instead of the attributable fraction, in order to avoid estimates of population impact based on unachievable goals, such as eradication of obesity. Such evaluations are vital to prioritize and inform future prevention programs regarding the possible impact of their efforts.

Acknowledgements

The Atherosclerosis Risk in Communities Study is carried out as a collaborative study supported by National Heart, Lung, and Blood Institute contracts N01-HC-55015, N01-HC-55016, N01-HC-55018, N01-HC-55019, N01-HC-55020, N01-HC-55021, and N01-HC-55022. The authors thank the staff and participants of the ARIC study for their important contributions.

Figure 9. (MS. 2, Figure 1) Diagram of intermediate steps in the estimation of the distribution of the generalized impact fraction (GIF) using bootstrapped datasets (sampled with replacement), x = dataset number out of 10,000 bootstrapped datasets. Each dataset contains the same name number (N) in each stratum as found in the original data. P_{overwt}= proportion overweight, P_{obese} = proportion obese, HR_{overwt} = unadjusted hazard ratio for obese, HR_{obese} = unadjusted hazard ratio for overweight, and N_{cases} = number of incident heart failure cases

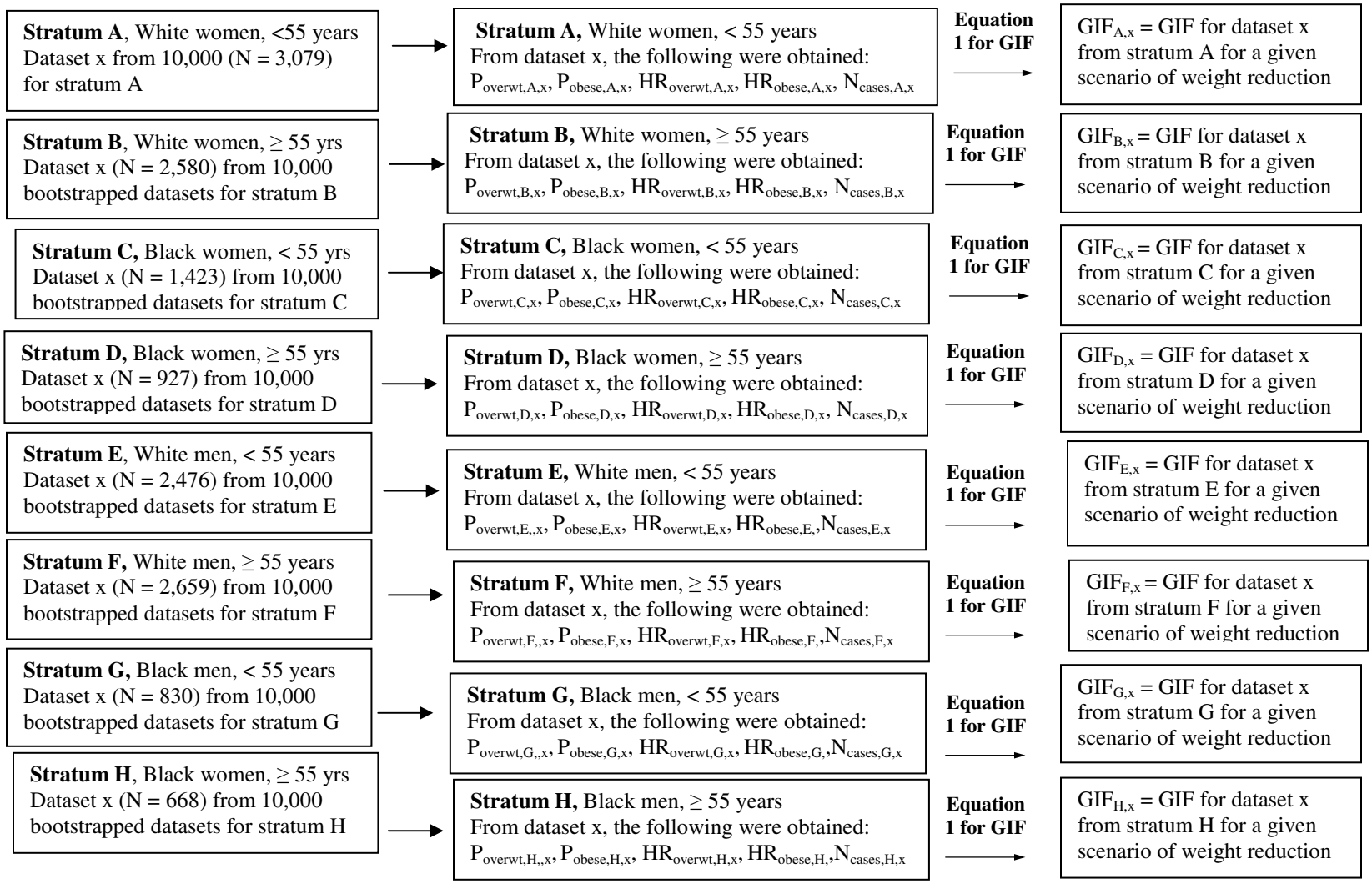


Table 12. (MS. 2, Table 1) Characteristics of the population (N = 14,642) at baseline by category of BMI, ARIC, 1987-1989

Characteristics	Normal weight BMI < 25 kg/m^2 (N = 4,968)	Overweight BMI 25-30 kg/m^2 (N = 5,787)	Obese BMI > 30 kg/m^2 (N = 3,887)
Body mass index, kg/m^2, mean (SD)	22.5 (1.8)	27 (1.4)	34 (4.3)
Age, years, mean (SD)	54 (6)	54 (6)	54 (6)
Men, %	37	56	39
Black, %	18	25	39
Center, %			
Jackson, MS	15	22	36
Forsyth County, NC	34	25	17
Minneapolis, MN	27	28	21
Washington County, MD	24	25	26
Educational level, %			
Less than high school	18	23	30
High school graduate	34	32	32
Greater than high school	48	46	38
Smoking status, %			
Never smoker	39	40	48
Former smoker	28	36	33
Current smoker	34	24	19
Alcohol use, %			
Never drinker	21	23	34
Former drinker	16	18	21
Current drinker	63	59	45
Diabetes, %	5	10	21
Hypertension, %	21	32	49
History of coronary heart disease, %	3	5	5

Table 13. (MS. 2, Table 2) Total number of participants, number of heart failure events, follow-up time, proportion in each BMI category, and unadjusted hazard ratios for incident heart failure, stratified by race, gender and age, ARIC 1987-2003

	HF events, N	Person years of Follow-up	Heart failure IR per 1,000 person-years	Proportion normal weight	Proportion overweight	Proportion obese	Unadjusted HR for overweight (95 % CI)	Unadjusted HR for obese (95 % CI)
White women, age < 55 years, N= 3,079	119	45,887	2.59	0.50	0.29	0.21	1.52 (0.92, 2.51)	4.14 (2.70, 6.36)
White women, age ≥ 55 years, N=2,580	240	36,717	6.54	0.45	0.32	0.23	1.56 (1.31, 2.14)	2.58 (1.89, 3.53)
Black women, age < 55 years, N=1,423	123	20,247	6.07	0.18	0.35	0.47	0.94 (0.52, 1.69)	1.63 (0.96, 2.76)
Black women, age ≥ 55 years, N = 927	169	12,039	14.04	0.18	0.34	0.48	1.04 (0.62, 1.72)	1.75 (1.10, 2.77)
White men, age < 55 years, N= 2,476	151	36,034	4.19	0.27	0.51	0.22	1.65 (1.03, 2.64)	2.64 (2.01, 3.25)
White men, age ≥ 55 years, N=2,659	408	35,202	11.59	0.27	0.51	0.22	1.09 (0.85, 1.40)	1.98 (1.51, 2.59)
Black men, age < 55 years, N=830	75	11,611	6.46	0.28	0.44	0.28	0.93 (0.49, 1.73)	2.10 (1.16, 3.80)
Black men, age ≥ 55 years, N=669	120	7,881	15.23	0.34	0.39	0.26	1.12 (0.72, 1.76)	1.58 (1.00, 2.49)

Table 14. (MS. 2, Table 3) Median proportion in each BMI category along with median GIF from 10,000 bootstrapped datasets, after 2 scenarios of weight reduction, a 30 % reduction in obesity to normal weight, and a 30 % reduction in obesity and overweight with reduction to normal weight, stratified by race, gender and age, ARIC 1987-2003

	Median Proportion Overweight	Median Proportion Obese	Median Proportion Obese* (after a 30 % Reduction)	Median GIF*, % (95 % intervals)	Median GIF**, % (95 % intervals)
White women, age < 55 years	0.21	0.21	0.15	11 (8, 15)	13.5 (8.6, 18.3)
White women, age ≥ 55 years	0.23	0.23	0.16	7 (5, 10)	10.5 (6.7, 14.4)
Black women, age < 55 years	0.47	0.47	0.33	7 (0.5, 13)	6.7 (-4.4, 16.4)
Black women, age ≥ 55 years	0.48	0.48	0.34	8 (2, 13)	8.2 (-1.1, 16.1)
White men, age < 55 years	0.22	0.22	0.16	8 (5, 11)	13.6 (7.2, 20.0)
White men, age ≥ 55 years	0.21	0.21	0.15	5 (3, 7)	6.1 (1.6, 10.4)
Black men, age < 55 years	0.27	0.27	0.19	7 (2, 13)	6.6 (-4.4, 16.0)
Black men, age ≥ 55 years	0.26	0.26	0.18	4 (-0.1, 8)	5.0 (-2.6, 11.9)

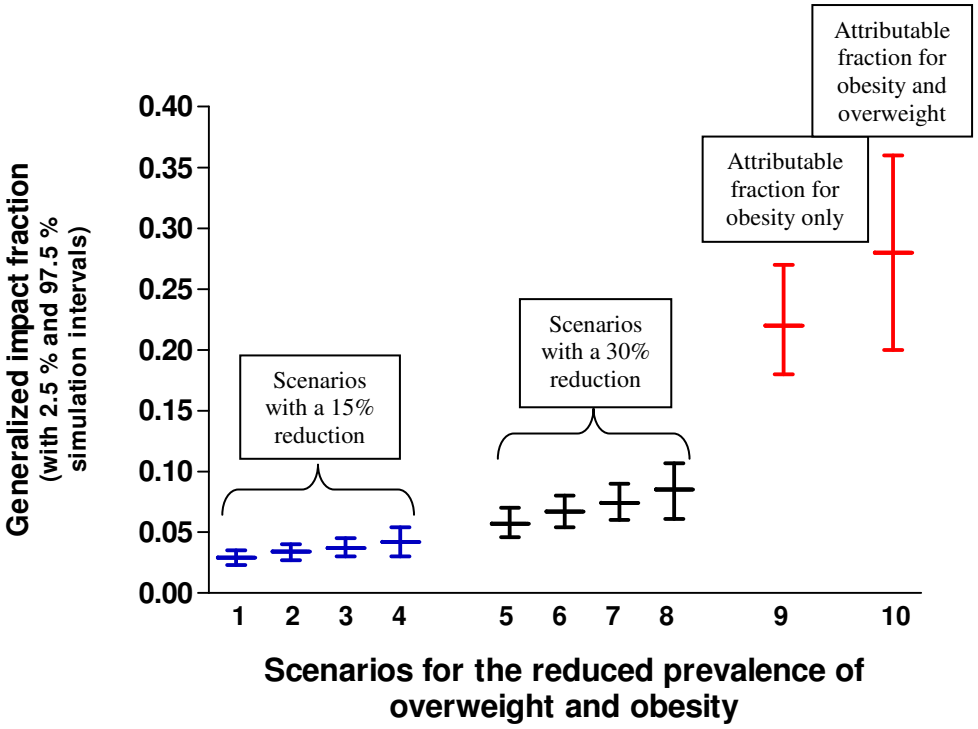
*30 % reduction in obesity, shift all to normal weight

**30 % reduction in obesity and overweight, shift all to normal weight

Table 15. (MS. 2, Table 4) The median generalized impact fraction and attributable fraction (with 2.5 % and 97.5 % simulation intervals) from 10,000 bootstrap datasets using the case-load weighted-sum method, given 10 scenarios of reduced prevalence of obesity and overweight, ARIC, 1987-2003

Hypothetical scenarios of weight reduction	**Median GIF (X 100) (95 % simulation intervals)**
Scenario 1: 15 % reduction in obesity, shift obese to overweight	2.9 (2.3, 3.5)
Scenario 2: 15 % reduction in obesity, shift obese to normal weight	3.4 (2.7, 4.0)
Scenario 3: 15 % reduction in obesity and overweight, shift both down one category	3.7 (3.0, 4.5)
Scenario 4: 15 % reduction in obesity and overweight, shift both to normal weight	4.2 (3.0, 5.4)
Scenario 5: 30 % reduction in obesity, shift obese to overweight	5.7 (4.6, 7.0)
Scenario 6: 30 % reduction in obesity, shift obese to normal weight	6.7 (5.4, 8.0)
Scenario 7: 30 % reduction in obesity and overweight, shift both down one category	7.4 (6.0, 9.0)
Scenario 8: 30 % reduction in obesity and overweight, shift both to normal weight	8.5 (6.1, 10.7)
Scenario 9: Complete elimination of obesity, shift to normal weight	22.0 (18.0, 27.0)
Scenario 10: Complete elimination of obesity and overweight, shift to normal weight	28.0 (20.0, 36.0)

Figure 10. (MS. 2, Figure 2) The median generalized impact fraction and attributable fraction (with 2.5 % and 97.5 % simulation intervals) from 10,000 bootstrap datasets using the case-load weighted-sum method, given 10 scenarios of reduced prevalence of obesity and overweight, ARIC, 1987-2003



1: 15 %reduction in obesity, shift obese to overweight
2: 15 %reduction in obesity, shift obese to normal weight
3: 15 %reduction in obesity and overweight, shift down 1 category
4: 15 %reduction in obesity and overweight, down to normal wt.

5: 30 %reduction in obesity, shift obese to overweight
6: 30 %reduction in obesity, shift obese to normal weight.
7: 30 %reduction in obesity and overweight, shift down one category
8: 30 %reduction in obesity and overweight, shift to normal weight

9: Attributable fraction, complete elimination of obesity

10: Attributable fraction, complete elimination of obesity and overweight

CHAPTER V

CONCLUSIONS

A. Recapitulation of overall specific aims

Our overall goal was to characterize the association between 3 different anthropometric measures to define obesity and overweight and their association with incident heart failure, stratified by race and gender. To better interpret our findings, we also determined the probable effect of outcome misclassification on our results. To put our findings in a public health context, we determined the preventable burden of heart failure due to overweight and obesity. Briefly these are recapitulated below in specific aims 1 and 2 which are addressed in manuscript 1, and specific aim 3 which is addressed in manuscript 2.

Specific Aim 1

Assess obesity/overweight as risk factors for the development of incident hospitalized heart failure. Assess race as a potential effect modifier. Compare three anthropometrical measures (e.g. BMI, waist circumference, and waist/hip ratio) as to their association with incident hospitalized heart failure.

Research question: Is there a positive association for obesity and overweight, as defined by each anthropometric measure (BMI, WHR and waist circumference), with incident heart

failure? Does this association vary by race- and gender-group? Is one measure superior to the other two in the prediction of incident heart failure?

Specific Aim 2

Assess the probable magnitude and direction of systematic error due to misclassification of the outcome of hospitalized heart failure for the association of obesity with incident hospitalized heart failure. This aim was achieved by performing a semi-automated probabilistic sensitivity analysis to estimate the degree of bias due to disease misclassification.

Research question: What is the estimated magnitude and direction of bias due to outcome misclassification in this analysis?

Specific Aim 3

Assess the burden of incident heart failure that could be prevented if there were a reduction in the proportion of obesity and/or overweight persons in the population. This aim was accomplished through the use of generalized impact fractions, a generalization of the attributable fraction. The results for the generalized impact fraction were compared to the more commonly used attributable fraction.

Research question: What is the reduced burden of incident heart failure that would be expected if there were a feasible reduction in the proportion of overweight and obese persons in the population? By how much does the attributable fraction differ from the generalized impact fraction in its estimate of preventable heart failure burden due to obesity?

B. Discussion of results

We found there was a positive association for obesity as defined by all three measures for incident HF. Even with stratification by race and gender, the adjusted hazard ratios for the comparison of the highest category of each measure (obese) to the lowest were similar and well above 1.0 for all three anthropometric measures (For example, HR with 95 % CI for 3rd vs. 1st tertile of waist-hip ratio: 2.55 (1.88, 3.44) for white women; 3.10 (2.13, 4.51) for black women, 2.69 (2.10, 3.45) for white men, and 2.45 (1.75, 3.43) for black men. Hazard ratios for the middle category for all three anthropometric measures tended to be intermediate, thus suggesting a graded response between body size and heart failure (For example, HR with 95 % CI for 2nd vs. 1st tertile of waist-hip ratio: 1.49 (1.08, 2.08) for white women; 2.10 (1.40, 3.15) for black women, 1.47 (1.11, 1.95) for white men, and 1.36 (0.92, 2.01) for black men). These associations were of similar magnitude for all four race and gender groups studied here. We further defined anthropometric categories by cross-classifying BMI categories by high or low waist-hip ratio. Age-adjusted heart failure incidence rates for BMI (normal weight, overweight and obese), stratified by high and low WHR show an increasing trend across these categories, with particularly a high heart failure incident rate (10.1 per 1,000 person-years) in those who have a high BMI and WHR. Other forms and combinations of these three anthropometrics were explored although not included in manuscript 1, largely due to limitations in sample size. These included further stratification of BMI into 6 categories, as defined by the six clinical action levels by the National Institutes of Health(1998). There were small numbers for certain groups (for example, BMI < 18.5 kg/m^2), even when men and women were combined**.** This analysis (see shown Figure 20), shows an increasing linear trend for overweight group, and the 3 levels of

obesity with a peak hazard ratio for those with BMI > 40 (HR= 4.76, 95 % CI = 3.68, 6.15); however the underweight group (BMI < 18.5 kg/m^2) actually shows a higher risk compared to the referent group (BMI 18.5 – 25 kg/m^2), although imprecise due to small sample size for this group (HR = 1.77, 95 % CI = 1.05, 2.98). Similarly, we divided BMI, waist circumference and WHR into quintiles (see Table 30), in which we show there is an increasing linear trend for all three measures. The adjusted HR for incident HF for the highest group with BMI > 40 as compared to the referent group is 4.76 (95 % CI = 3.68, 6.15).

We considered cross-classification of BMI by high and low waist circumference (see Figure 21), as these are clinical categories created by the National Institute of Health (NIH)(NIH 1998). The cut-point that defines high or low waist circumference is 102 cm for men and 88 cm for women; since these two measures are highly correlated (r = 0.88) there were small sample sizes for the discordant groups (especially those with normal weight by BMI and high waist circumference and those with obesity by BMI and low waist circumference). Even with men and women combined, the confidence intervals were imprecise for these discordant groups. For those in the category of obese BMI ($\geq$ 30 kg/m^2) and high waist circumference (adjusted HR = 2.62, 95 % CI = 2.25, 3.06), the findings are similar to the hazard ratio reported in manuscript 1 for the comparison of obese BMI and high WHR (adjusted HR = 3.18 95 % CI = 2.66, 3.81).

We used several methods to evaluate the effect of potential confounders. Utilizing a change-in-estimate approach, we found that only a few of our potential confounders met a 10 % change criterion. For women, change-in-estimate criterion were met by two variables for BMI (smoking and race and center), whereas only age met the criterion for WHR, and none

of the variables met the criterion for waist circumference. For men, age and smoking meet criterion for BMI, whereas none of the covariates meet criterion for waist circumference or WHR. For all three exposures and in both men and women, there is minimal change in the confidence limit ratio with a reduced as compared to full model (precision). Also, calculation of the "MSE", as described in the methods section, indicated that the precision/validity trade-off favored adjustment for the excluded variable in all cases. Therefore, we chose to present the models inclusive of all potential confounders for consistency across models.

We also assessed the effect on model fit of the inclusion of significant covariate interactions in the final model, as shown in Table 28 of the appendix. Likelihood ratio tests (LRT) were used to compare nested models (significance criterion of $p < 0.05$). Calculation of change-in-estimate and "mean square error" were also considered here. For both men and women and across all three exposures, when significant covariate interaction terms are included in the model, the HR and their associated confidence intervals changed minimally; for simplicity of interpretation, covariate interaction terms were excluded from the final model. We conclude that our findings were fairly robust to changes in model form.

With time-dependent ROC analysis, we did not find that any one anthropometric measure, or even the cross-classified measures of BMI and WHR, as superior in the prediction of incident heart failure. Stratification by race and gender did not change our inferences in this respect.

Calculation of the GIF in manuscript 2 was enlightening as to the actual impact of an aggressive reduction in obesity and overweight. The overall GIF number appears small; however, when applied to the actual impact on healthcare utilization and cost, then there

seems to be a large impact, since HF is a common condition. A 30% hypothetical reduction in both obesity and overweight would prevent 8.5 % (95 % intervals = 6.1%, 10.7%) of incident heart failure cases, whereas a 30 % reduction in obesity only would prevent 6.7 % (95 % intervals = 5.4 %, 8.0 %) of incident heart failure. However, a 6.7 % reduction in heart failure would result in an estimated 72,628 fewer heart failure hospitalizations with approximately $477,674,356 cost savings per year. The attributable fraction (22 % (95 % intervals = 18 %, 27 %)), which assumes complete elimination of obesity and overweight, is approximately 3 times larger than the highest amount of weight reduction (30 %) assumed feasible here.

C. Degree to which doctoral goals have been met

This dissertation is the product of my work as the lead investigator in the design, analysis and writing, under the primary supervision of my dissertation chair, Wayne Rosamond, PhD, MS. Suggestions provided by this dissertation committee, and other co-authors have been incorporated. This dissertation topic and its scope were approved at the time of my dissertation proposal defense and further at an interim dissertation meeting. I believe that the expected goals of doctoral research work will have been met with these dissertation committee meetings, the preparation and submission of this document the graduate school, submission for publication of manuscript 1 to the ARIC publications committee, and the final dissertation defense itself.

There are four areas of expectations and standards for a doctoral dissertation specified in the "Academic Policies" manual for 2007-2008, by the department of epidemiology at UNC-Chapel Hill; these areas include originality, depth, scholarship, and writing skills. I'll discuss

each area as to my achievement of these expectations, however it is the ultimately the decision of this dissertation committee to determine if these four expectations have been shown during my time in training here, as well as in my final doctoral dissertation document. There are several aspects of my dissertation that are original in substantive content and in creative application of existing and relatively new methods. Substantively, the association of obesity with incident heart failure has not previously been evaluated stratified by race and gender. The ability of these three measures to predict heart failure has not been previously compared with ROC analysis. In addition, the second paper on the impact of reduction in obesity and overweight on the burden of heart failure is novel. Methodologically, we use several techniques that are original in their application to this topic area including multiple-bias modeling, time-dependent ROC analysis (with adjustment for optimism), and calculation of the generalized impact fraction. These methods are also technically sophisticated and add considerable depth and scholarship to the analysis. Depth and scholarship are the second and third expectations for a doctoral dissertation. Depth is shown in the review of the literature on this topic, as well as the methods used to explore the best form for the main exposures and covariates, evaluation of potential confounders and effect measure modifiers. Similarly, scholarship is shown in the synthesis of all of this work succinctly into the two enclosed manuscripts.

The fourth expectation of a doctoral dissertation is that it shows competence in scientific writing. I have worked on my writing skills during my 5 years at UNC, with both required and optional courses. I attended an elective 5-day seminar by George Gopen JD, PhD, offered by the UNC office of postdoctoral services called “Writing from the Reader’s

Perspective". In addition to this dissertation as an example of my scientific writing skills, I have also published two additional articles (as the first author) in peer-reviewed journals.

In summary, during my training at UNC Chapel Hill, I have gained considerable substantive knowledge in the area of cardiovascular epidemiology, improved scientific writing skills, as well as knowledge in how to critically appraise the literature. I have learned skills in advanced epidemiologic methods. This document shows the synthesis of all of these learned skills.

D. Strengths

Strengths of the analysis for manuscript 1 include the large sample size of this well characterized cohort and the extended period of follow-up. Measures were obtained using standardized techniques. Anthropometric variables were analyzed in both continuous form and using cut-points. Results are presented stratified by race and gender. Since there are three anthropometric measures available from the baseline visit, we were able to compare the predictive ability of the three measures using a newly described method for time-dependent ROC analysis. Furthermore, we used a novel technique for multiple bias modeling which incorporates the bias due to outcome misclassification, into multivariate modeling with Monte Carlo techniques. This study provides further example of how Monte Carlo Risk Assessment techniques can be used to succinctly summarize the effect of systematic error.

Strengths of the analysis for manuscript 2 are the re-introduction of the generalized impact fraction as an alternative to the attributable fraction for exposures that are unlikely to be eliminated, such as obesity. We show how confidence intervals can be obtained from

bootstrapped datasets. In addition, we use these findings to extrapolate the public health impact in terms of the estimated decrease in heart failure hospitalizations and healthcare costs.

As compared to the attributable fraction, the GIF is a measure of preventable burden of disease that can be based on multiple hypothetical changes in the exposure distribution. The utility of the GIF shown here should encourage others in its use. As example of it's limited use compared to the attributable fraction, in the New England Journal of Medicine, there is only one paper(Semenza, Rubin et al. 1996) using impact fractions between 1995 and 11/2007, and yet there are 79 papers which mention the attributable risk or fraction between 1993 and 2007. Despite this relatively high number for the use of the attributable fraction, it is even considered underutilized(Nakayama 2000). The limited application of the generalized impact fraction is likely due to lack of familiarity by most researchers; papers such as this one may provide an impetus for its increased use in the practice of epidemiology.

E. Limitations

Limitations of these two studies include the definition of heart failure. We included hospitalized and fatal heart failure, as we did not have data on outpatient heart failure; however, community surveillance reports have indicated that 74 % of outpatient heart failure cases are hospitalized within 1.7 years(Roger, Weston et al. 2004). Furthermore, we do not know the etiology of the heart failure (i.e., ischemic, viral, or idiopathic) events or have echocardiographic data to define the type of heart failure (diastolic versus systolic). Heart failure events were defined solely by ICD codes and were not validated by committee adjudication. Because HF is a clinical syndrome without clear objective diagnostic criteria,

we were concerned with the possibility of misclassification of the discharge code based definition. To address this potential limitation, we performed a sensitivity analysis to explore the effect of outcome misclassification on our findings. We found that outcome misclassification would bias our results toward the null, given the parameters for sensitivity and specificity that we provided.

Limitations of the analysis (from manuscript 2) using the generalized impact fraction, include that unlike the attributable fraction, one must stratify by important confounders rather use an adjusted relative measure of effect such that small sample sizes can be a problem. There was not adequate sample size to stratify by all potential confounders that were adjusted for in manuscript 1 (age, race, gender, education, smoking, alcohol use); therefore we stratified based on age, race and gender only. There were few confounders identified using a 10 % change-in-estimate criterion. Therefore, I suspect that further stratification with a larger sample would not have had much of an impact on the results for GIF. An additional issue with stratification for multiple variables is that it makes the creation of bootstrapped datasets time-consuming, as each cross-classified group (race by gender by age group) must be bootstrapped independently.

Inherent in estimating the generalized impact fraction, one must consider a counterfactual approach(Rodgers 2002), and in so doing we wanted to use feasible goals for population-wide reduction in obesity and overweight. It is difficult to determine feasible goals for population-wide weight reduction; however, we considered a spectrum of weight loss with the goals recommended in the Healthy People 2010 report as our most optimistic goal. Clinical trial studies on interventions for weight loss have provided the mean number of kilograms lost(Franz, VanWormer et al. 2007), but these studies do not provide the actual

number we sought for our calculation of the GIF, which is the feasible proportion of obese whom could lose weight to shift categories to the normal weight group.

F. Future Directions

Further improvements to the multiple-bias modeling sensitivity analysis used here would be to expand this technique to work with outcome misclassification in Cox proportional hazards models. In this case, one would have to create a distribution to re-define person-time for those whom the outcome is reclassified.

Assessment of the potential impact of exposure reduction on disease incidence, such as done here with the generalized impact fraction, could be extended to include simulation techniques. For example, bootstrapped datasets could be created with reduced prevalence of obesity (by subgroup) from which a distribution of heart failure incident rates could be calculated from this altered population as compared to the original population.

Future research in the association of obesity with heart failure should consider whether one type of adiposity, either visceral or subcutaneous, is more predictive of heart failure. Unfortunately, definitive measurements to distinguish visceral and subcutaneous fat require expensive imaging with computed tomography. In addition, it would be interesting to assess the type of heart failure (diastolic or systolic dysfunction) most often associated with obesity. Currently, obesity is thought to most often result in diastolic dysfunction. The ARIC study has begun detailed abstraction of medical records for all heart failure hospitalizations, along with physician validation; such data will aid in the determination of the type of heart failure associated with obesity. The Multi-Ethnic Study of Atherosclerosis (MESA) study recently began in 2000-2002 will eventually have enough follow-up time for incident heart failure

outcomes, along with measures of chest computed tomography and cardiac MRI. In addition, other studies should consider the age of onset of obesity and/or the amount of time obese as it relates to the age of onset of incident heart failure. The epidemic of childhood obesity and metabolic complications may contribute to an earlier age of onset for heart failure.

APPENDICES

A. IRB certification

B. Supplemental results for manuscript 1

C. Supplemental results for manuscript 2

A. IRB certification

THE UNIVERSITY of NORTH CAROLINA at CHAPEL HILL

OFFICE OF HUMAN RESEARCH ETHICS
Medical School Building 52
Mason Farm Road
CB #7097
Chapel Hill, NC 27599-4105
(919) 966-3113
Web site: ohre.unc.edu
https://my.research.unc.edu for IRB status
Federalwide Assurance (FWA) #4801

To: Laura Loehr
Epidemiology
CB: 8050

From: Public Health-Nursing IRB

Ruth Humphrey for
Authorized signature on behalf of IRB
Approval Date: 12/14/2007
Expiration Date of Approval: 12/12/2008
RE: Notice of IRB Approval by Expedited Review (under 45 CFR 46.110)
Submission Type: Renewal
Expedited Category: 5.Existing or non-research data
Study #: 07-0013
Study Title: The Association of Overweight and Obesity with Incident Hospitalized Heart Failure

This submission has been approved by the above IRB for the period indicated.

Study Description:
The goal of this study is to evaluate obesity as an independent predictor of incident heart failure (HF) using three different anthropometric measures. This study will also assess whether one measure is superior to others in the prediction of HF and determine the burden of HF disease that might be prevented if a downward shift in the distribution of body mass index occurred. Study participants will consist of the ARIC study, which is a bi-racial cohort with an average of 13 years of follow-up. Study procedures include adjusting for several types of systematic error using semi-automated probabilistic sensitivity analysis.

Regulatory and other findings:

This research meets criteria for a waiver of consent entirely according to 45 CFR 46.116(d).

Investigator's Responsibilities
Federal regulations require that all research be reviewed at least annually. It is the Principal Investigator's responsibility to submit for renewal and obtain approval before the expiration date. You may not continue any research activity beyond the expiration date without IRB approval. Failure to receive approval for continuation before the expiration date will result in automatic termination of the approval for this study on the expiration date.

When applicable, enclosed are stamped copies of approved consent documents and other recruitment materials. You must copy the stamped consent forms for use with subjects unless you have approval to do otherwise.

You are required to obtain IRB approval for any changes to any aspect of this study before they can

be implemented (use the modification form at ohre.unc.edu/forms). Should any adverse event or unanticipated problem involving risks to subjects or others occur it must be reported immediately to the IRB using the adverse event form at the same web site.

CC: Wayne Rosamond, Epidemiology

I received approval from the ARIC study for this work. I have also completed the collaborative institutional training initiative (CITI) program in research ethics.

B. Supplemental Results, Manuscript 1

Figure 11. (MS. 1, supplemental results) Distribution of measures of BMI, among men and women (red dotted line represents category cutpoints), ARIC 1987-1989

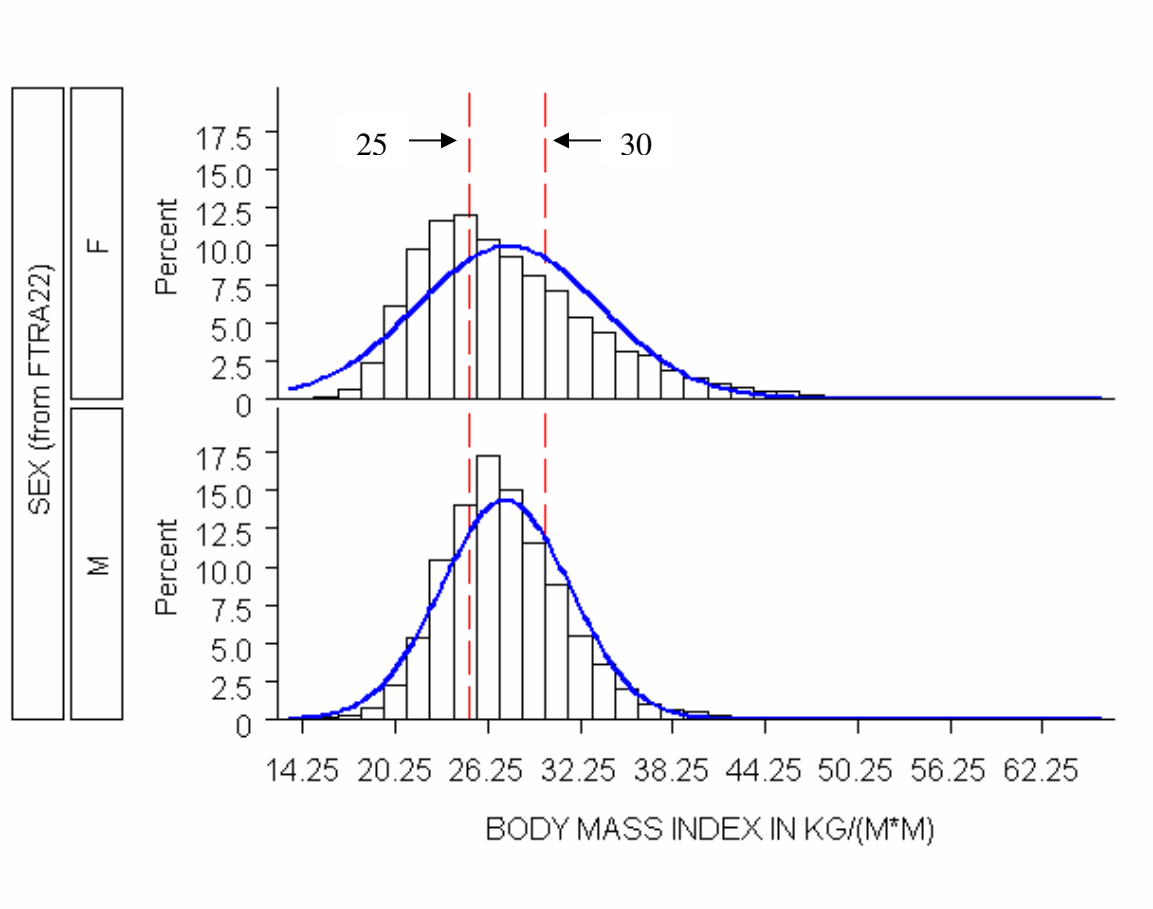


Figure 12. (MS. 1, supplemental results) Distribution of measures of waist circumference, among men and women (red dotted line represents tertile cutpoints), ARIC 1987-1989

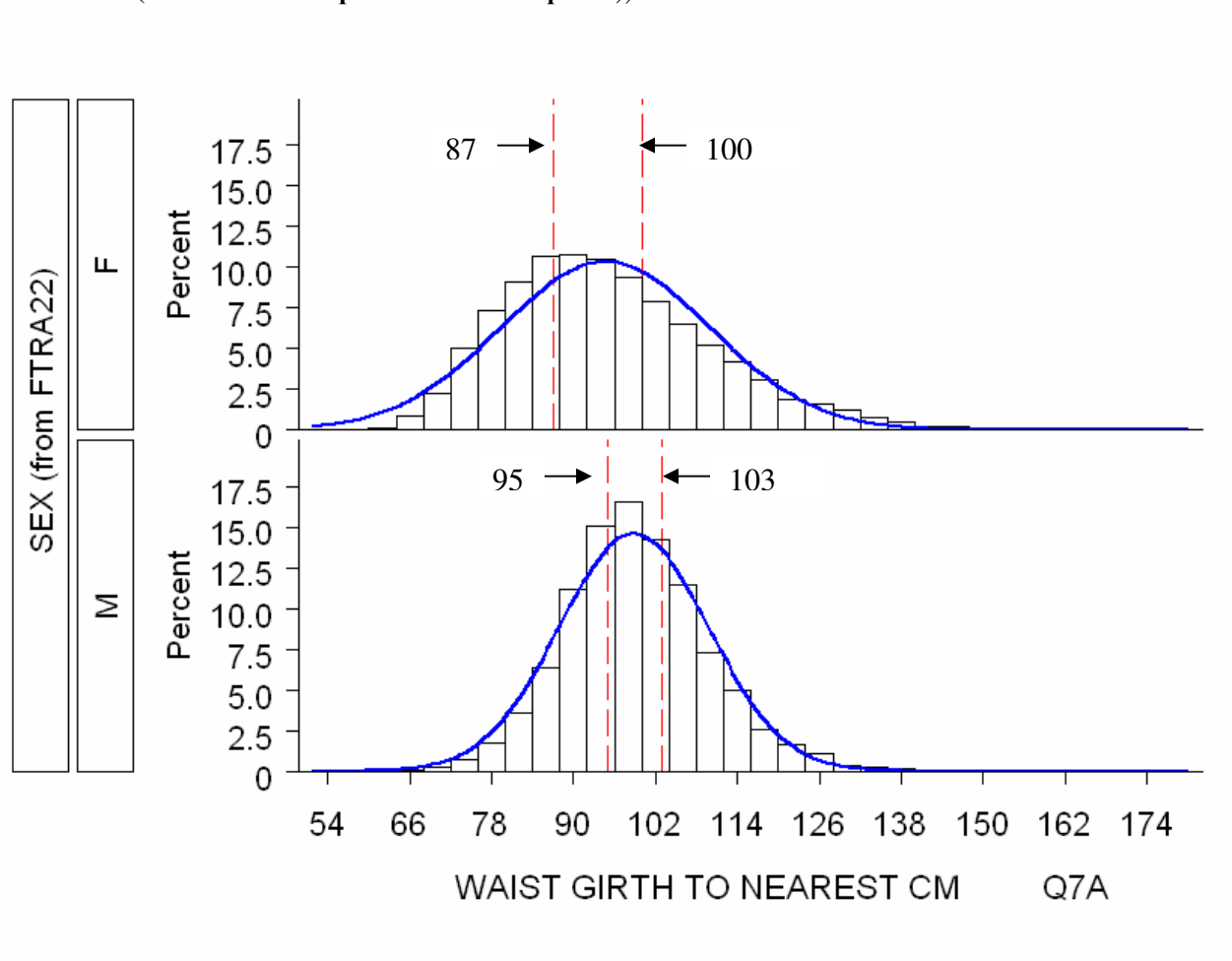


Figure 13. (MS. 1, supplemental results) Distribution of measures of WHR, among men and women, (red dotted line represents tertile cutpoints), ARIC 1987-1989

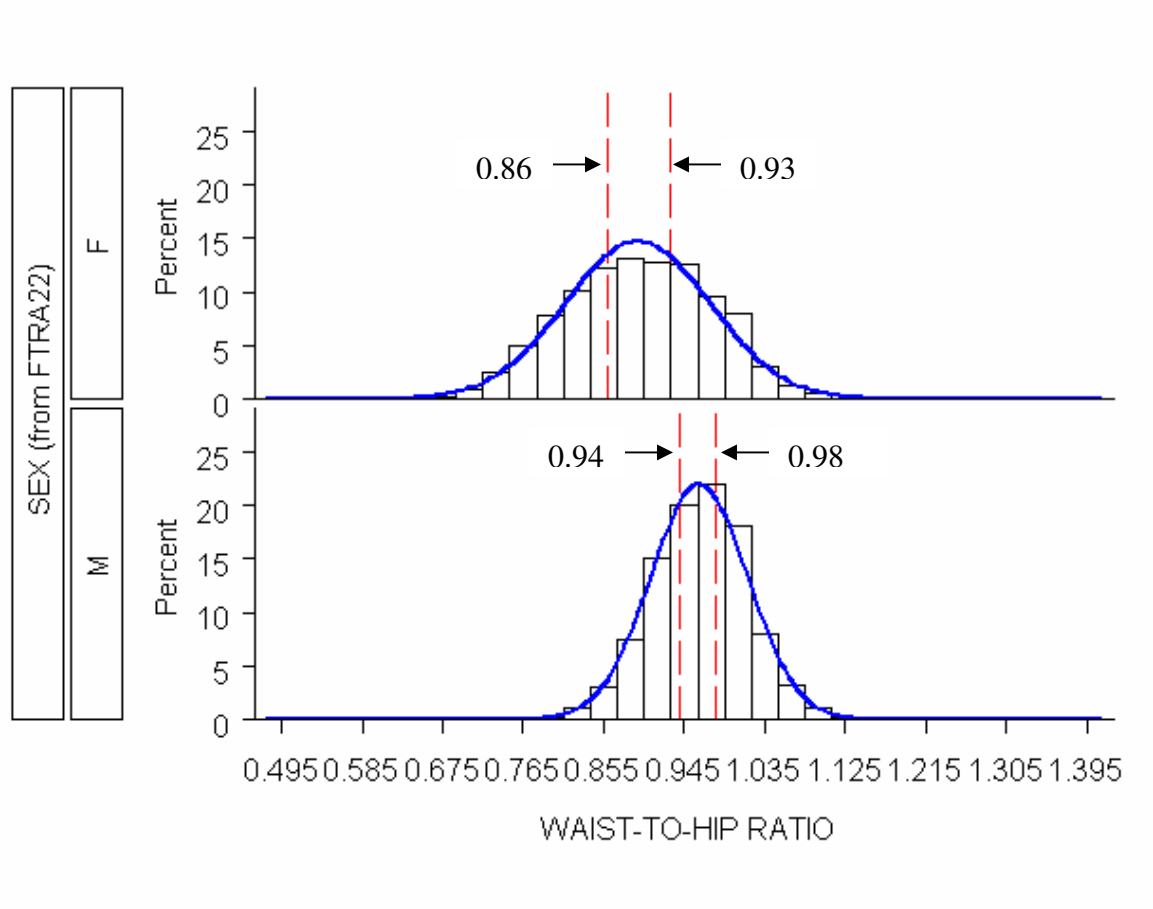


Table 16. (MS. 1, supplemental results) Distribution of each anthropometric measure within category of each, for women, ARIC 1987-1989

	Mean	Median	Standard Error	Interquartile Range
BMI, kg/m^2				
<25, N = 3,113	22.3	22.5	0.03	2.7
25-30, N = 2,545	27.3	27.2	0.03	2.5
≥ 30, N = 2,373	35.1	33.8	0.10	5.7
Waist Circumference, cm				
1st Tertile < 87, N = 2,643	79.0	80.0	0.10	8.0
2nd Tertile 87 - <100, N = 2,668	92.8	93.0	0.07	6.0
3rd Tertile > 100, N = 2,720	112.1	109.0	0.21	13.0
Waist Hip Ratio				
1st Tertile <0.86, N = 2,762	0.80	0.81	0.0008	0.06
2nd Tertile 0.86 - <0.93, N= 2,449	0.89	0.89	0.0004	0.03
3rd Tertile ≥ 0.93, N = 2,820	0.98	0.97	0.0007	0.05

Range of BMI for women is 14.2-65.9
Range of waist circumference for women is 52-178
Range of WHR for women is 0.49-1.29

Table 17. (MS. 1, supplemental results) Distribution of each anthropometric measure within category of each, for men, ARIC 1987-1989

	Mean	Median	Standard Error	Interquartile Range
BMI, kg/m^2				
<25, N = 1,865	23.0	23.3	0.04	2.1
25-30, N = 3,261	27.3	27.2	0.02	2.3
≥ 30, N = 1,533	33.2	32.2	0.09	3.2
Waist Circumference, cm				
1st Tertile < 95, N = 2,349	88.1	89.0	0.11	7.0
2nd Tertile 95 - < 103, N = 2,099	98.4	98.0	0.05	4.0
3rd Tertile > 103, N = 2,211	110.7	108.0	0.17	9.0
Waist Hip Ratio				
1st Tertile < 0.94, N = 2,228	0.90	0.91	0.0007	0.04
2nd Tertile 0.94 - <0.98, N = 1,900	0.96	0.96	0.0003	0.02
3rd Tertile ≥ 0.98, N = 2,531	1.02	1.01	0.0006	0.04

Range of BMI for men is 15.4-65.9
Range of waist circumference for men is 52-178
Range of WHR for men is 0.55-1.39

Table 18. (MS. 1, supplemental results) Comparison of characteristics at baseline by category of BMI, waist hip ratio and waist circumference, ARIC, 1987-1989, for women

	BMI			Waist Circumference			Waist-hip ratio		
	< 25	25-30	≥ 30	T1	T2	T3	T1	T2	T3
Age, years, mean (SD)	54(6)	54(6)	54(6)	53(6)	54(6)	54 (6)	52(5)	54(6)	55(6)
Black, %	14	32	47	18	28	42	26	28	35
Center, %									
Jackson	12	28	43	17	25	37	24	24	30
Forsyth	34	23	16	30	26	19	26	28	22
Minnesota	30	25	16	34	24	15	34	23	17
Maryland	24	24	25	19	25	29	17	25	31
Education, %									
< high school	14	23	33	13	22	33	13	22	32
HS graduate	39	37	35	38	38	36	37	38	36
> high school	47	40	32	50	40	32	50	40	32
Smoking status, %									
Never	46	55	61	48	54	58	53	53	53
Former	23	22	21	23	22	22	23	26	25
Current	31	24	18	29	24	20	23	26	25
Alcohol use, %									
Never	26	36	46	27	35	44	31	34	41
Former	13	16	20	12	14	21	11	17	20
Occasional	23	21	17	23	22	17	23	21	18
Light	29	23	15	30	23	16	29	22	17
Heavy	8	5	3	8	6	3	6	6	5
CHD, %	1.1	1.9	1.9	1.0	1.7	2.2	0.6	1.8	2.5
MI, %	1.0	1.7	1.7	0.9	1.5	1.9	0.6	1.5	2.2
Diabetes, %	4	10	22	3	7	22	3	7	21
Hypertension, %	19	11	50	18	31	48	21	30	46
LVH by ECG, %	1.3	2.1	3.4	1.3	2.0	3.2	1.2	2.5	2.9
Lipids, mg/dl, mean (SD)									
Total	214 (40)	222 (45)	220 (44)	211 (40)	221 (44)	222 (45)	209 (40)	219 (43)	226 (45)
LDL-C	129(39)	141 (42)	141 (40)	126 (39)	140(41)	142 (40)	126(38)	138 (41)	145 (41)
HDL-C	64(18)	57 (16)	51 (14)	65 (18)	58 (16)	51(14)	64 (18)	57 (16)	52(15)
Triglycerides	104(55)	124 (83)	141 (92)	97 (50)	121 (74)	146 (95)	96(55)	117(65)	150 (97)
Creatnine, g/dl, mean (SD)	0.99 (0.3)	1.02 (0.4)	1.04 (0.6)	0.99 (0.2)	1.01 (0.4)	1.03 (0.6)	0.99(0.2)	1.00 (0.3)	1.03 (0.6)
Albumin, g/dl, mean (SD)	3.88 (0.3)	3.84 (0.3)	3.75 (0.3)	3.87 (0.3)	3.85 (0.3)	3.77 (0.3)	3.85 (0.3)	3.84 (0.3)	3.81 (0.3)
FEV_1, L/sec, mean (SD)	2.57 (0.3)	2.57 (0.3)	2.57 (0.3)	2.55 (0.3)	2.51 (0.3)	2.47 (0.3)	2.58(0.3)	2.52(0.3)	2.44(0.3)

LDL-C = low density cholesterol; HDL-C = high density cholesterol

Table 19. (MS. 1, supplemental results) Comparison of characteristics at baseline by category of BMI, waist-hip ratio and waist circumference, ARIC, 1987-1989, for men,

	BMI			Waist Circumference			Waist-hip ratio		
	< 25	25-30	≥ 30	T1	T2	T3	T1	T2	T3
Age, years, mean (SD)	55 (6)	54 (6)	54 (6)	54 (6)	55 (6)	55 (6)	54 (5)	54 (6)	55 (6)
Black, %	25	20	27	30	18	20	35	20	15
Center, %									
Jackson	21	17	23	26	16	17	30	17	12
Forsyth	32	26	23	26	27	26	26	27	26
Minnesota	23	30	29	26	28	29	26	28	29
Maryland	24	26	28	23	28	28	18	27	33
Education, %									
< high school	26	22	25	25	22	24	21	23	26
HS graduate	26	27	27	26	28	27	25	28	28
> high school	49	51	48	49	50	48	54	49	45
Smoking status, %									
Never	26	29	29	30	29	25	33	28	24
Former	35	46	50	36	46	51	38	45	49
Current	40	24	21	34	25	23	29	27	27
Alcohol use, %									
Never	13	12	14	13	12	13	15	13	11
Former	22	21	25	22	21	24	21	22	24
Occasional	12	13	14	13	12	14	13	13	13
Light	41	44	38	42	44	40	43	42	41
Heavy	11	9	9	10	10	10	9	10	11
CHD, %	6.3	6.8	8.7	6.1	7.2	8.1	5.3	6.7	9.0
MI, %	5.2	5.6	7.3	5.0	6.0	6.8	4.6	5.5	7.4
Diabetes, %	6	10	21	6	10	18	6	10	18
Hypertension, %	25	31	47	26	31	42	27	31	40
LVH by ECG, %	1.6	2.1	2.4	1.5	2.7	1.9	1.6	2.4	2.1
Lipids, mg/dl, mean (SD)									
Total	206 (39)	213 (39)	213 (41)	208 (39)	213 (40)	213 (40)	207 (39)	212 (40)	212 (40)
LDL-C	134 (38)	142 (36)	141 (37)	136 (38)	142 (36)	140 (36)	136 (38)	141 (37)	141 (36)
HDL-C	50 (17)	43 (12)	40 (11)	50 (16)	43 (12)	40 (11)	50 (16)	44 (13)	41 (11)
Triglycerides	111 (67)	143 (89)	171 (134)	111 (65)	144 (91)	168 (124)	108 (58)	143 (98)	166 (118)
Creatnine, g/dl, mean (SD)	1.19 (0.2)	1.24 (0.4)	1.23 (0.2)	1.21 (0.2)	1.24 (0.5)	1.22 (0.2)	1.22 (0.3)	1.22 (0.4)	1.22 (0.2)
Albumin, g/dl, mean (SD)	3.92 (0.3)	3.94 (0.3)	3.90 (0.3)	3.93 (0.3)	3.93 (0.3)	3.91 (0.3)	3.93 (0.3)	3.94 (0.3)	3.91 (0.3)
FEV_1, L/sec, mean (SD)	3.66 (0.4)	3.68 (0.4)	3.64 (0.4)	3.58 (0.4)	3.69 (0.4)	2.73 (0.4)	3.64 (0.4)	3.69 (0.4)	3.67 (0.4)

LDL-C = low density cholesterol; HDL-C = high density cholesterol

Table 20. (MS. 1, supplemental results) Person-Time contribution in years (%) by covariates and category of each exposure, for women (only those non-missing for all covariates included), ARIC, 1987-2003

	BMI < 25 44,714 P-Y	BMI 25-30 36,094 P-Y	BMI ≥ 30 32,432 P-Y	WC -T1 38,045 P-Y	WC - T2 38,177 P-Y	WC - T3 37,018 P-Y	WHR - T1 40,010 P-Y	WHR - T2 34,760 P-Y	WHR - T3 38,470 P-Y
Race and center									
Whites, Minnesota	13,734 (31 %)	9,119 (25 %)	5,487 (17 %)	13,088 (34 %)	9,357 (25 %)	5,897 (16 %)	13,500 (34 %)	8,142 (23 %)	6,699 (17 %)
Whites, Forsyth	14,745 (33 %)	6,975 (19 %)	3,892 (12 %)	10,939 (29 %)	8,991 (24 %)	5,683 (15 %)	9,750 (24 %)	8,680 (25 %)	7,183 (19 %)
Blacks, Forsyth	762 (2 %)	1,308 (4 %)	1,146 (4 %)	583 (2 %)	1,109 (3 %)	1,524 (4 %)	726 (2 %)	955 (3 %)	1,534 (4 %)
Blacks, Jackson	4,877 (11 %)	9,841 (27 %)	13,665 (42 %)	6,068 (16 %)	9,332 (24 %)	12,984 (35 %)	9,265 (23 %)	8,228 (24 %)	10,890 (28 %)
Whites, Maryland	10,596 (24 %)	8,850 (25 %)	8,241 (25 %)	7,368 (19 %)	9,389 (25 %)	10,931 (30 %)	6,769 (17 %)	8,754 (25 %)	12,164 (32 %)
Education									
< HS educ.	5,838 (13 %)	8,041 (22 %)	10,137 (31 %)	4,533 (12 %)	7,965 (21 %)	11,518 (31 %)	5,061 (13 %)	7,324 (21 %)	11,632 (30 %)
HS grad.	17,447 (39 %)	13,507 (37 %)	11,484 (35 %)	14,391 (38 %)	14,655 (38 %)	13,392 (36 %)	14,830 (37 %)	13,317 (38 %)	14,291 (37 %)
> HS educat.	21,429 (48 %)	14,546 (40%)	10,811 (33 %)	19,121 (50 %)	15,558 (41 %)	12,108 (33 %)	20,119 (50 %)	14,119 (41 %)	12,547 (33 %)
Smoking status									
Never smoker	21,230 (47 %)	20,072 (56 %)	20,133 (62 %)	18,557 (49 %)	21,075 (55 %)	21,802 (59 %)	21,692 (54 %)	18,826 (54 %)	20,917 (33 %)
Former smoker	10,537 (24 %)	7,893 (22 %)	6,919 (21 %)	8,861 (23 %)	8,378 (22 %)	8,110 (22 %)	9,332 (23 %)	7,447 (21 %)	8,570 (22 %)
Current smoker	12,947 (29 %)	8,129 (23 %)	5,380 (17 %)	10,626 (28 %)	8,724 (23 %)	7,106 (19 %)	8,986 (22 %)	8,487 (24 %)	8,983 (23 %)
Alcohol use									
Never drinker	11,592 (26 %)	12,971 (36 %)	14,967 (46 %)	10,140 (27 %)	13,298 (35 %)	16,092 (43 %)	12,272 (31 %)	11,845 (34 %)	15,413 (40 %)
Former drinker	5,637 (13 %)	5,384 (15 %)	6,038 (19 %)	4,540 (12 %)	5,327 (14 %)	7,193 (19 %)	4,412 (11 %)	5,528 (16 %)	7,119 (19 %)
Occasional drinker	10,578 (24 %)	7,663 (21 %)	5,675 (17 %)	8,848 (23 %)	8,637 (23 %)	6,431 (17 %)	9,457 (24 %)	7,344 (21 %)	7,116 (18 %)
Light/mod drinker	13,228 (30 %)	8,259 (23 %)	4,853 (15 %)	11,623 (31 %)	8,685 (23 %)	6,032 (16 %)	11,498 (29 %)	7,861 (23 %)	6,980 (18 %)
Heavy drinker	3,679(8%)	1,817(5%)	899(3%)	2,894(8%)	2,230(6%)	1,270(3%)	2,371(6%)	2,181(6%)	1,815(5%)

Table 21. (MS. 1, supplemental results) Person-Time contribution in years (%) by covariates and category of each exposure, for men (only those non-missing for all covariates included), ARIC, 1987-2003

	BMI < 25 24,846 P-Y	BMI 25-30 44,380 P-Y	BMI ≥ 30 19,797 P-Y	WC -T1 31,813 P-Y	WC - T2 28,443 P-Y	WC - T3 28,768 P-Y	WHR - T1 30,683 P-Y	WHR - T2 25,804 P-Y	WHR - T3 32,537 P-Y
Race by Center									
Whites, Minnesota	5,905 (24 %)	13,667 (31 %)	5,882 (59 %)	8,709 (27 %)	8,226 (29 %)	8,520 (30 %)	8,430 (27 %)	7,402 (29 %)	9,623 (30 %)
Whites, Forsyth	7,191 (29 %)	11,031 (25 %)	3,362 (17 %)	7,237 (23 %)	7,270 (26 %)	7,077 (25 %)	7,065 (23 %)	6,589 (26 %)	7,930 (24 %)
Blacks,Forsyth	823(3 %)	875(2 %)	581(3 %)	991(3 %)	593(2 %)	695(2 %)	1,043(3 %)	520(2 %)	715(2 %)
Blacks, Jackson	4,656 (19 %)	7,330 (17 %)	4,508 (23 %)	7,545 (24 %)	4,385 (15 %)	4,565 (16 %)	8,784 (29 %)	4,098 (16 %)	3,613 (11 %)
Whites, Maryland	6,273 (25 %)	11,476 (26%)	5,463 (28 %)	7,331 (23 %)	7,969 (28 %)	7,911 (27 %)	5,362 (17 %)	7,194 (28 %)	10,656 (33 %)
Education									
< HS education	5,640 (23 %)	9,223 (21 %)	4,548 (23 %)	7,184 (23 %)	5,726 (20 %)	6,500 (23 %)	5,897 (19 %)	5,628 (22 %)	7,885 (24 %)
High school grad.	6,479 (26 %)	12,044 (27 %)	5,319 (27 %)	8,158 (26 %)	7,926 (28 %)	7,757 (27 %)	7,477 (24 %)	7,188 (28 %)	9,177 (28 %)
> HS education	12,728 (51 %)	23,114 (52 %)	9,930 (50 %)	16,470 (52 %)	14,791 (52 %)	11,511 (40 %)	17,309 (56 %)	12,988 (50 %)	15,475 (48 %)
Smoking status									
Never smoker	7,153 (29 %)	13,555 (31 %)	5,886 (30 %)	10,336 (32 %)	8,729 (31 %)	7,529 (26 %)	10,655 (35 %)	7,743 (30 %)	8,196 (25 %)
Former smoker	9,054 (36 %)	20,766 (47 %)	9,905 (50 %)	11,718 (37 %)	13,225 (46 %)	14,782 (51 %)	11,878 (39 %)	11,789 (46 %)	16,058 (50 %)
Current smoker	8,640 (35 %)	10,059 (23%)	4,006 (20 %)	9,759 (31 %)	6,489 (23 %)	6,457 (22 %)	8,151 (27 %)	6,272 (24 %)	8,282 (25 %)
Alcohol use									
Never drinker	3,242 (13 %)	5,558 (13 %)	2,837 (14 %)	4,265 (13 %)	3,637 (13 %)	3,736 (13 %)	4,490 (15 %)	3,482 (13 %)	3,666 (11 %)
Former drinker	5,221 (21 %)	8,840 (20 %)	4,687 (24 %)	6,689 (21 %)	5,705 (20 %)	6,354 (22 %)	6,053 (20 %)	5,524 (21 %)	7,171 (22 %)
Occasional drinker	3,110 (13 %)	5,691 (13 %)	2,791 (14 %)	4,103 (13 %)	3,480 (12 %)	4,008 (14 %)	3,880 (13 %)	3,361 (13 %)	4,351 (13 %)
Light/mod drinker	10,581 (43 %)	20,081 (45%)	7,662 (39 %)	13,811 (43 %)	12,801 (45 %)	11,712 (41 %)	13,754 (45 %)	10,914 (42 %)	13,656 (42 %)
Heavy drinker	2,693(11%)	4,210(9%)	1,820(9%)	2,946(9%)	2,820(10%)	2,957(10 %)	2,506(8%)	2,523(10%)	3,693(1%)

Figure 14. (MS. 1, supplemental results) Plot of cumulative incidence for heart failure (person-years) by category of BMI (black line = normal weight, red line = overweight, green line = obese), stratified by race and gender, ARIC, 1987-2003

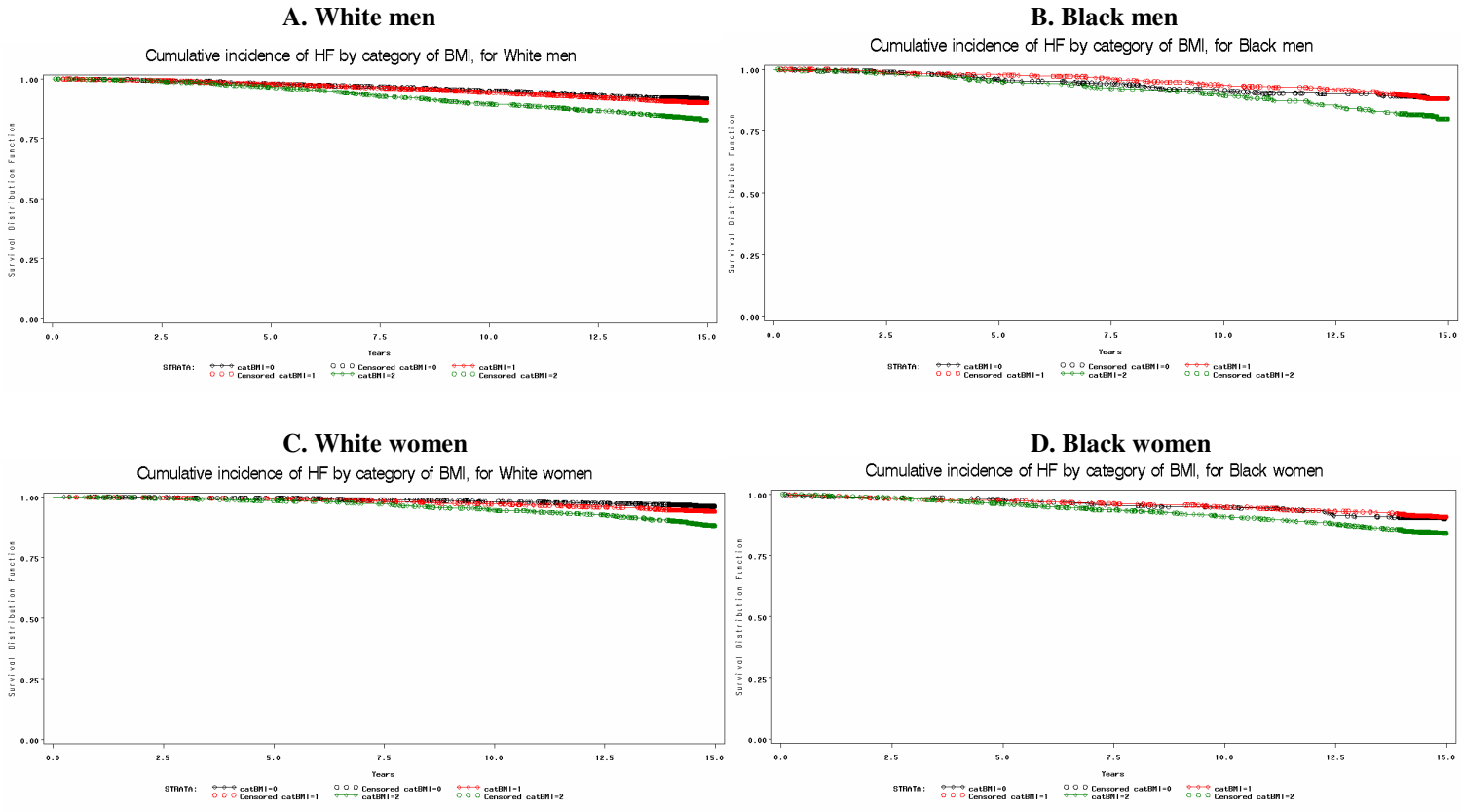


Figure 15. (MS. 1, supplemental results) Plot of cumulative incidence for heart failure (Person-years) by tertile of waist circumference (black line = first tertile, red line = second tertile, green line = third tertile), stratified by race and gender, ARIC, 1987-2003

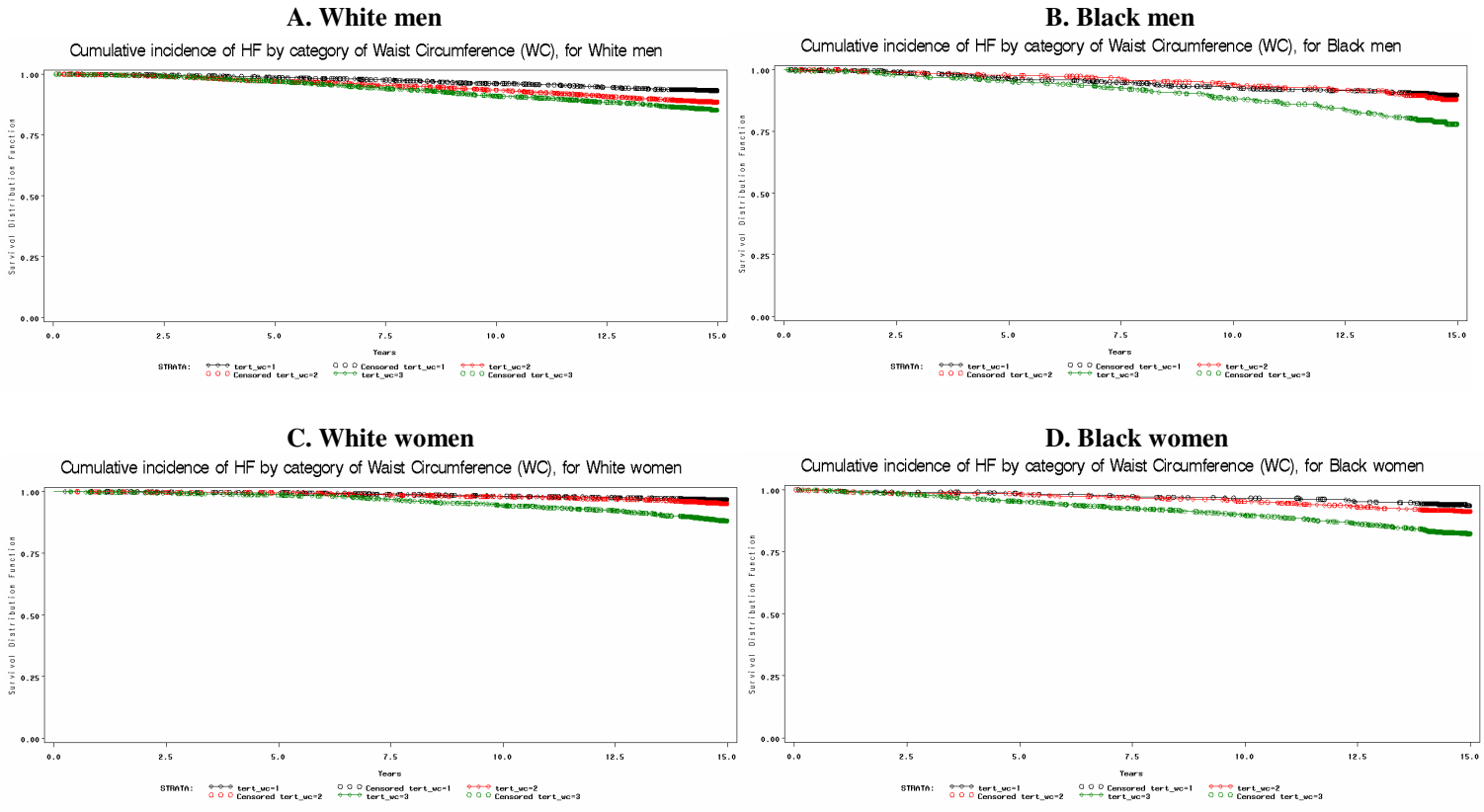
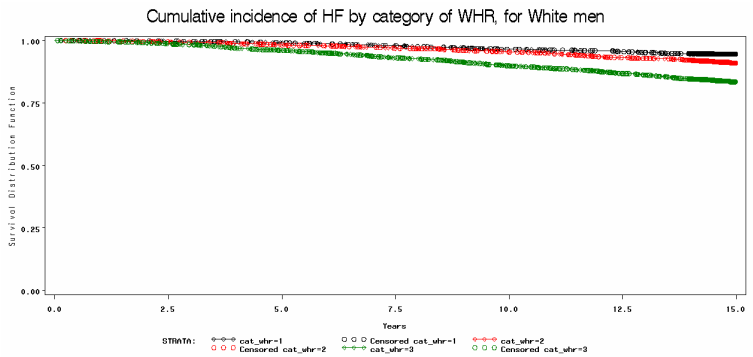
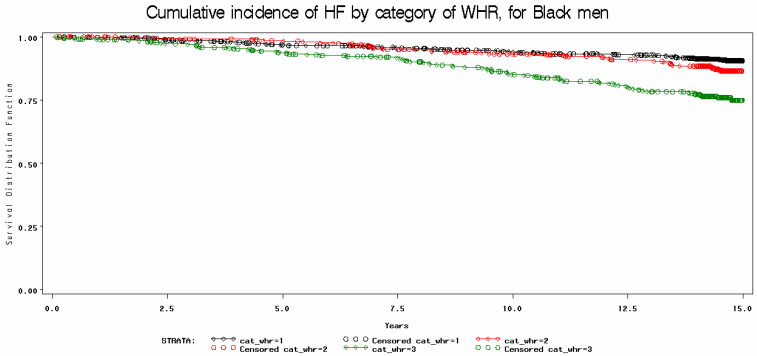


Figure 16. (MS. 1, supplemental results) Plots of cumulative incidence for heart failure (person-years) by tertile of waist-hip ratio (black line = first tertile, red line = second tertile, green line = third tertile), stratified by race and gender, ARIC, 1987-2003

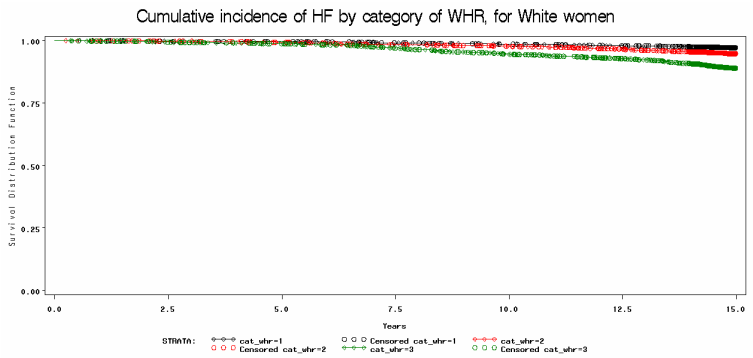
A. White men



B. Black men



C. White women



D. Black women

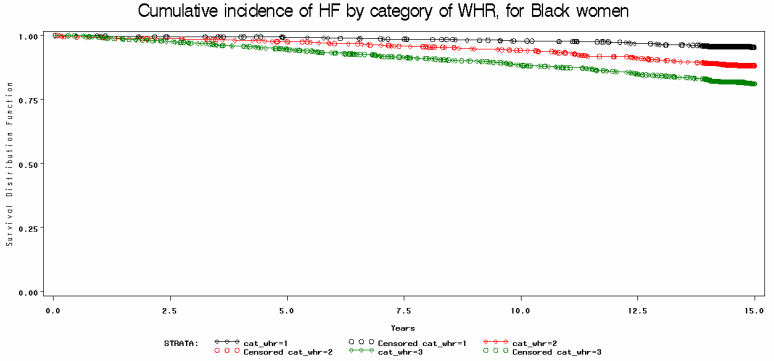


Table 22. (MS. 1, supplemental results) Assessment of proportional hazards assumption using log negative log survival plot and time interaction terms for all 3 exposures, as indicator and continuous variables, ARIC, 1987-2003

	Log negative log plot		Time interaction term	PHA Decision?
	Obese (3rd) vs. referent (1)	Overweight (2) vs. referent (1)		
BMI				
Women	1 cross @ 245[1]	Overlap then parallel (181)[2]	P = 0.4	No violation
Men	No cross	Overlap	P = 0.1	No violation
Waist-hip ratio				
Women	No cross	No cross	P = 0.09	No violation
Men	No cross	1 cross @ 403[3]	P = 0.1	No violation
Waist circumference				
Women	Cross @110[4]	Overlap then parallel (181)[5]	P = 0.5	No violation
Men	No cross	Overlap then parallel (403)[6]	P = 0.03	No violation

[1] --8 events before cross
[2] –9 events occur before lines become parallel
[3] --4 events before cross
[4] --3 events before cross
[5] --6 events before cross
[6] --8 events before cross

Table 23. (MS. 1, supplemental results) Results of modeling using a backward-selection change-in-estimate approach for categorized anthropometric measures, comparing highest to referent, for women, ARIC 1987-2003

	BMI				Waist Circumference				Waist Hip Ratio			
	Adjusted HR (95 % CI)	CLR	CIE	MSE	Adjusted HR (95 % CI)	CLR	CIE	MSE	Adjusted HR (95 % CI)	CLR	CIE	MSE
Full model 1	2.52 (2.03, 3.12)	1.54	--	0.0120	2.70 (2.15, 3.39)	1.58	--	0.0136	2.76 (2.18, 3.49)	1.60	--	0.0144
- Age	2.42 (1.95, 2.99)	1.53	0.04		2.96 (2.36, 3.72)	1.58	0.09		3.34 (2.65, 4.22)	1.59	0.19*	0.05
- Alcohol	2.66 (2.15, 3.29)	1.53	0.02		2.84 (2.26, 3.56)	1.58	0.05		2.85 (2.25, 3.60)	1.60	0.03	
- Smoking	2.18 (1.76, 2.70)	1.53	0.14*	0.03	2.46 (1.96, 3.09)	1.58	0.09		2.79 (2.20, 3.53)	1.60	0.01	
- Education	2.66 (2.15, 3.29)	1.53	0.02		2.86 (2.28, 3.58)	1.57	0.06		2.92 (2.31, 3.69)	1.60	0.06	
- Race/center	2.87 (2.33, 3.53)	1.52	0.13*	0.028	2.95 (2.36, 3.69)	1.56	0.09		2.79 (2.21, 3.52)	1.59	0	

*criterion met for change in estimate

CIE = change in estimate = | ln(crude HR/adjusted HR) |, MSE (full) = variance ($lnHR_{full}$); MSE (reduced) = variance ($lnHR_{red}$) + CIE^2

Full model, age in years centered at 45, alcohol (never, former, current occasional, current mod/light, and current heavy), smoking (never, former, ever), education (< high school, high school, > high school), race_center indicator variables

If MSE full is < MSE reduced then precision/validity trade off favors adjustment

Table 24. (MS. 1, supplemental results) Results of modeling using a backward-selection change-in-estimate approach for continuous measures per SD change for each anthropometric variable, for women, ARIC 1987-2003

	BMI			Waist Circumference			Waist Hip Ratio		
	Adjusted HR (95 % CI)	CLR	CIE	Adjusted HR (95 % CI)	CLR	CIE	Adjusted HR (95 % CI)	CLR	CIE
Full model 1	1.47 (1.37, 1.58)	1.15	--	1.54 (1.43, 1.65)	1.15	--	1.62 (1.48, 1.76)	1.19	--
- Age	1.41 (1.31, 1.51)	1.15	0.04	1.54 (1.43, 1.65)	1.15	0.006	1.74 (1.60, 1.89)	1.18	0.07
- Alcohol	1.49 (1.39, 1.60)	1.15	0.01	1.57 (1.46, 1.69)	1.16	0.02	1.63 (1.50, 1.78)	1.19	0.006
- Smoking	1.37 (1.27, 1.47)	1.16	0.07	1.47 (1.36, 1.58)	1.16	0.05	1.62 (1.49, 1.76)	1.18	0
- Education	1.50 (1.40, 1.61)	1.15	0.02	1.57 (1.46, 1.69)	1.16	0.03	1.65 (1.51, 1.79)	1.19	0.02
- Race/center	1.53 (1.43, 1.64)	1.15	0.04	1.59 (1.48, 1.71)	1.16	0.06	1.63 (1.50, 1.77)	1.18	0.02

For Table 23 and **Table 24**, complete case analysis, n= 7967 or 64 deleted after deletion of blacks not from Jackson or Forsyth and any missing variables

Table 25. (MS. 1, supplemental results) Results of modeling using a backward-selection change-in-estimate approach for categorized anthropometric measures, comparing highest to referent, for men, ARIC 1987-2003

	BMI				Waist Circumference				Waist Hip Ratio			
	Adjusted HR (95 % CI)	CLR	CIE	MSE	Adjusted HR (95 % CI)	CLR	CIE	MSE	Adjusted HR (95 % CI)	CLR	CIE	MSE
Full model	2.43 (1.98, 2.98)	1.51	--	0.011	2.26 (1.88, 2.73)	1.45	--	0.009	2.58 (2.12, 3.14)	1.48	--	0.010
- Age (cont)	2.17 (1.78, 2.66)	1.49	0.11*	0.022	2.23 (1.85, 2.69)	1.45	0.01		2.78 (2.29, 3.39)	1.48	0.07	
- Alcohol	2.46 (2.01, 3.01)	1.50	0.008		2.27 (1.89, 2.74)	1.45	0.004		2.57 (2.11, 3.13)	1.48	0.004	
- Smoking	2.16 (1.77, 2.64)	1.49	0.12*	0.025	2.15 (1.79, 2.59)	1.45	0.05		2.59 (2.13, 3.15)	1.48	0.008	
- Education	2.45 (2.00, 3.00)	1.5	0.01		2.28 (1.89, 2.75)	1.46	0.009		2.65 (2.18, 3.23)	1.48	0.03	
- Race/center	2.41 (1.97, 2.95)	1.5	0.01		2.22 (1.84, 2.68)	1.46	0.02		2.43 (2.00, 2.94)	1.47	0.07	

*criterion met for change in estimate, CIE = change in estimate = | ln(crude HR/adjusted HR) |

Full model, age in years, alcohol (never, former, current occasional, current mod/light, and current heavy), smoking (never, former, ever), education (< high school, high school, > high school), race_center indicator variables

If MSE full is < MSE reduced then precision/validity trade off favors adjustment

Table 26. (MS. 1, supplemental results) Results of modeling using a backward-selection change-in-estimate approach for continuous measures per SD change for each anthropometric variable, for men, ARIC 1987-2003

	BMI			Waist Circumference			Waist Hip Ratio		
	Adjusted HR (95 % CI)	CLR	CIE	Adjusted HR (95 % CI)	CLR	CIE	Adjusted HR (95 % CI)	CLR	CIE
Full model 1	1.45 (1.36, 1.56)	1.15	--	1.50 (1.40, 1.60)	1.14	--	1.50 (1.41, 1.60)	1.13	--
- Age	1.37 (1.29, 1.45)	1.12	0.06	1.45 (1.36, 1.54)	1.13	0.006	1.54 (1.44, 1.64)	1.14	0.03
- Alcohol	1.46 (1.37, 1.56)	1.14	0.007	1.51 (1.42, 1.61)	1.13	0.007	1.50 (1.41, 1.60)	1.13	0
- Smoking	1.40 (1.32, 1.49)	1.13	0.04	1.46 (1.37, 1.56)	1.14	0.03	1.49 (1.40, 1.59)	1.14	0.007
- Education	1.46 (1.36, 1.55)	1.14	0.007	1.50 (1.41, 1.60)	1.13	0	1.51 (1.42, 1.65)	1.16	0.007
- Race/center	1.46 (1.37, 1.55)	1.13	0.007	1.50 (1.41, 1.60)	1.13	0	1.47 (1.37, 1.56)	1.14	0.02

For Table 25 and Table 26, complete case analysis, n= 6603 or 56 deleted after deletion of blacks not from Jackson or Forsyth and any missing variables

Table 27 (MS. 1, supplemental results) Assessment for additive interaction using interaction contrast ratios (ICR) for race, stratified by gender

	MEN ICR*	WOMEN ICR*
BMI		
Overweight vs. normal	**-0.27**	**-0.65**
Obese vs. normal	**- 0.19**	**-0.32**
WHR		
2nd tert vs. 1st tert	**0.22**	**1.66**
3rd tert vs. 1st tert	**1.09**	**2.55**
WC		
2nd tert vs. 1st tert	**-0.51**	**0.3**
3rd tert vs. 1st tert	**0.49**	**1.05**

*all 95 % confidence intervals included 0

Table 28. (MS. 1, supplemental results) Assessment for improved model fit with covariate interactions for the association of categorized BMI, waist circumference and WHR with incident heart failure, ARIC, 1987-2003

	-2 log likelihood	Likelihood Ratio Test, P value	HR (95 % CI)	CLR	CIE	MSE
BMI -Women						
ME: Main effects	10706.916		2.52 (2.03, 3.12)	1.54	-	
M1: ME + smoke_dum1*age (P = 0.005), educ_dum1*age	10695.141	0.003 (M1 vs.ME)	2.48 (2.00, 3.08)	1.54	0.016	0.0123
M2: ME + smoke_dum1*age (P = 0.006)	10698.871	0.005 (M2 vs.ME) 0.05 (M2 vs.M1)	2.50 (2.01, 3.09)	1.54	0.008	0.0120
BMI - Men						
ME: Main effects	12056.961		2.43 (1.98, 2.98)	1.51	-	0.0107
M1: ME + smoke_dum1*etoh_dum1, educ_dum1*age	12055.395	0.5 (M1 vs.ME)	2.43 (1.98, 2.98)	1.51	0	
Waist Circumference - Women						
ME: Main effects	10681.349		2.70 (2.15, 3.39)	1.58	-	0.0137
M1: ME + Smoke_dum1*age (P = 0.006), educ_dum1*age	10669.844	0.003 (M1 vs.ME)	2.67 (2.13, 3.36)	1.58	0.01	0.0136
M2: ME + Smoke_dum1*age (P = 0.006)	10673.536	0.005 (M2 vs.ME) 0.06 (M2 vs.M1)	2.68 (2.13, 3.36)	1.58	0.007	
Waist Circumference – Men						
ME: Main effects	12062.006		2.26 (1.88, 2.73)	1.45	-	0.0091
M1: ME + smoke_dum1*etoh_dum1 (P = 0.03), smoke_dum2*etoh_dum1 (P = 0.04), educ_dum1*age	12056.275	0.1 (M1 vs.ME)	2.26 (1.87, 2.72)	1.45	0	0.0091
M2: ME + smoke_dum1*etoh_dum1 (P = 0.04), smoke_dum2*etoh_dum1 (P = 0.05)	12057.324	0.1 (M2 vs.ME) 0.3 (M2 vs.M1)	2.26 (1.87, 2.72)	1.45	0	
WHR - Women						
ME: Main effects	10703.239		2.76 (2.18, 3.49)	1.60	-	0.0144
M1: ME + smoke_dum1*age (P = 0.006), educ_dum2*age	10694.958	0.02 (M1 vs.ME)	2.74 (2.17, 3.47)	1.60	0.007	0.0145
M2: ME + smoke_dum1*age (P = 0.005)	10695.174	0.02 (M2 vs.ME) 0.6 (M2 vs.M1)	2.74 (2.16, 3.46)	1.60	0.007	
WHR - Men						
ME: Main effects	12028.931		2.58 (2.12, 3.14)	1.48		0.010
M1: ME + smoke_dum1*etoh_dum1, smoke_dum2*etoh_dum1, educ_dum2*age (P = 0.0007)	12013.411	0.002 (M1 vs.ME)	2.55 (2.09, 3.10)	1.48	0.01	0.010
M2: ME + educ_dum2*age (P = 0.0006)	12017.258	0.0006 (M2vs.ME) 0.2 (M2 vs.M1)	2.56 (2.10, 3.12)	1.49	0.008	0.010

Figure 17. (MS. 1, supplemental results) Fully adjusted hazard ratios (with 95 % CI) for incident heart failure by category of obesity and overweight as measured by BMI, waist circumference and WHR, for women

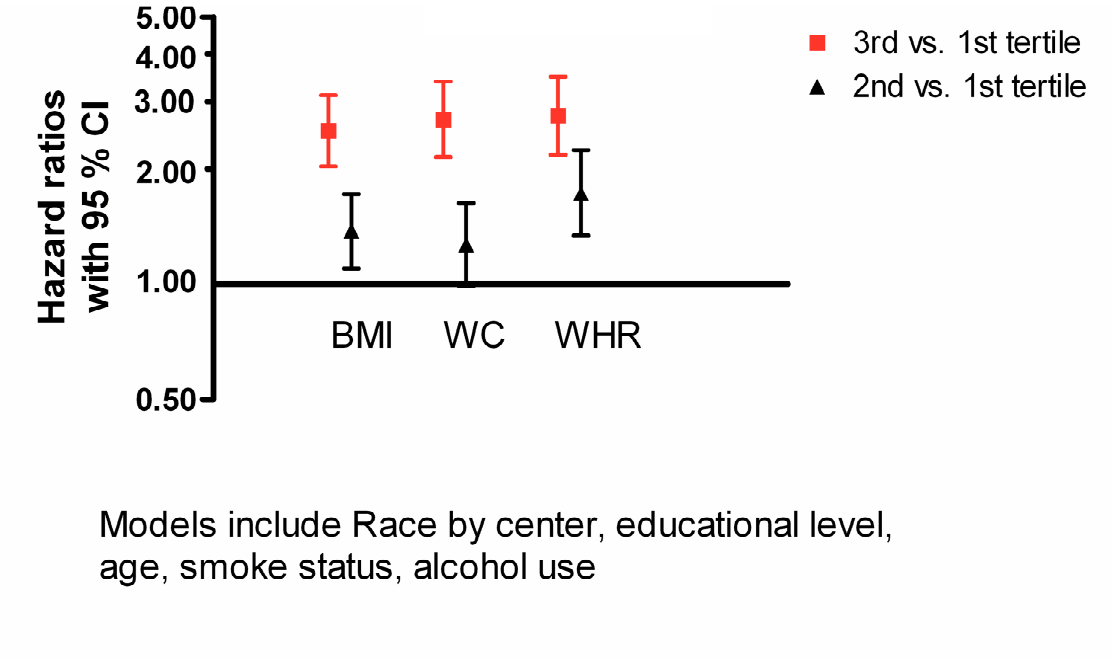


Figure 18. (MS. 1, supplemental results) Fully adjusted hazard ratios (with 95 % CI) for incident heart failure by category of obesity and overweight as measured by BMI, waist circumference and WHR, for men

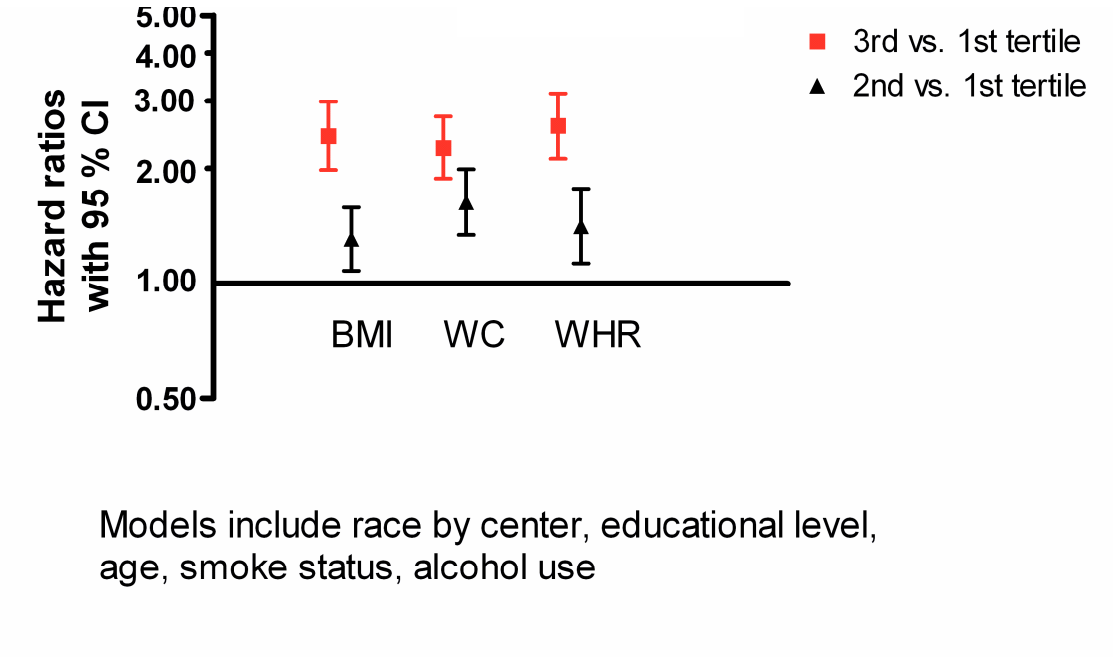


Table 29. (MS. 1, supplemental results) Unadjusted and adjusted* hazard ratios (with 95 % CI) for incident heart failure by continuous measures of BMI, waist circumference and waist hip ratio, stratified by gender, ARIC, 1987-2003

	Women HR (95 % CI)	Men HR (95 % CI)
BMI, 1 unit change, unadjusted	1.07 (1.06, 1.09)	1.08 (1.06, 1.09)
BMI, 1 unit change, adjusted	1.07 (1.05, 1.08)	1.09 (1.08, 1.11)
BMI, 1 SD change, unadjusted	1.53 (1.44, 1.63)	1.35 (1.27, 1.44)
BMI, 1 SD change, fully adjusted	1.47 (1.37, 1.57)	1.45 (1.36, 1.54)
Waist Circumference, 1 SD change, unadjusted	1.68 (1.58, 1.80)	1.44 (1.36, 1.53)
Waist Circumference, 1 SD change, fully adjusted	1.54 (1.43, 1.65)	1.50 (1.40, 1.60)
WHR, 1 SD change, unadjusted	1.89 (1.75, 2.04)	1.54 (1.45, 1.63)
WHR, 1 SD change, fully adjusted	1.62 (1.48, 1.76)	1.50 (1.41, 1.60)

SD for BMI in women = 6.0, SD for BMI in men = 4.2
SD for waist circumference in women = 15.4, SD for waist circumference in men = 10.9
SD for WHR in women = 0.08, SD for WHR in men = 0.05
All models including center have blacks not from Jackson or Forsyth excluded (n = 22 women, n= 26 men)
Adjusted for age, smoking, alcohol use, education and race and center
Complete case analysis for unadjusted models too
Same results estimate whether centered BMI or not

Figure 19. (MS. 1, supplemental results) Fully adjusted models for incident heart failure by standard deviation change in anthropometric measures of BMI, waist circumference and WHR, by gender

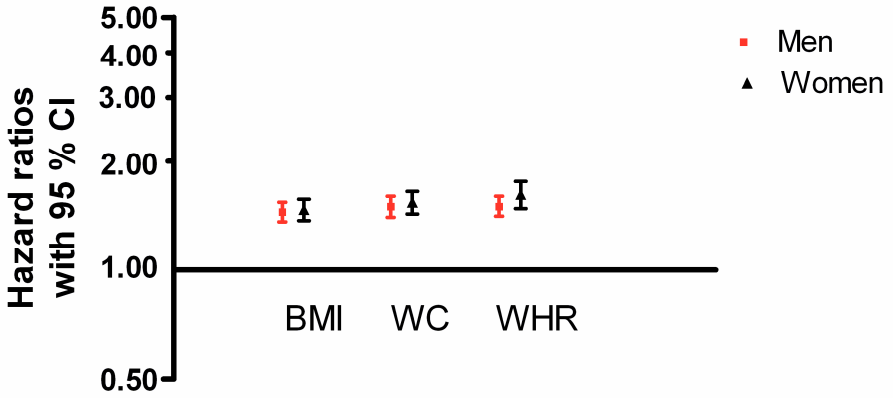
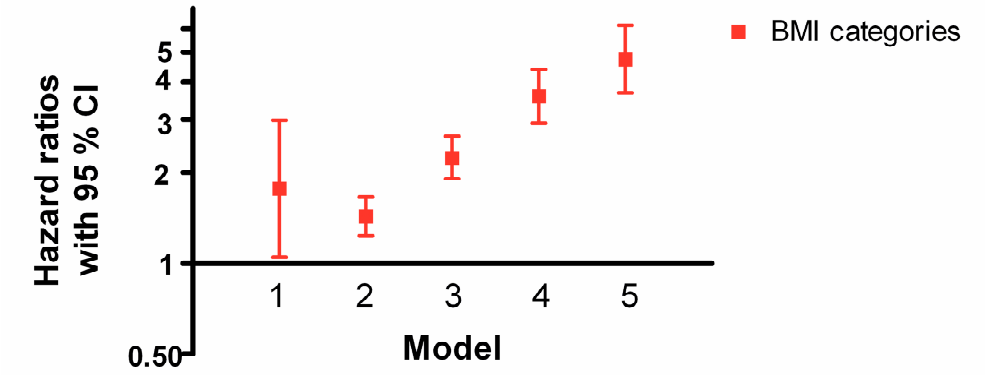


Figure 20. (MS. 1, supplemental results) Fully adjusted hazard ratios (with 95 % CI) for incident heart failure by action level categories of BMI, men and women combined, ARIC, 1987-2003



Referent: BMI >18.5 and < 24.9
1: BMI < 18.5
2: BMI > 25 and < 29.9
3: BMI > 30 and < 35.9
4: BMI > 35 and < 39.9
5: BMI > 40

Table 30. (MS. 1, supplemental results) Adjusted* hazard ratios (with 95 % CI) for incident heart failure by BMI, waist circumference and waist-hip ratio, each evenly divided into 5 groups (quintiles) to assess for a linear trend, stratified by gender, ARIC, 1987-2003

	Women HR (95 % CI)	P for trend	Men HR (95 % CI)	P for trend
BMI		<0.0001		<0.0001
1st quintile	1.00		1.00	
2nd quintile	0.99 (0.72, 1.39)		1.33 (1.02, 1.72)	
3rd quintile	1.30 (0.96, 1.77)		1.30 (1.00, 1.69)	
4th quintile	1.77 (1.32, 2.37)		1.59 (1.24, 2.03)	
5th quintile	2.82 (2.11, 3.76)		2.77 (2.20, 3.49)	
Waist Circumference		<0.0001		<0.0001
1st quintile	1.00		1.00	
2nd quintile	1.20 (0.79, 1.59)		1.28 (0.95, 1.73)	
3rd quintile	1.08 (0.76, 1.54)		1.70 (1.29, 2.30)	
4th quintile	1.73 (1.26, 2.38)		1.95 (1.49, 2.55)	
5th quintile	3.05 (2.25, 4.15)		2.86 (2.22, 3.67)	
WHR		<0.0001		<0.0001
1st quintile	1.00		1.00	
2nd quintile	1.39 (0.97, 1.99)		1.79 (1.32, 2.43)	
3rd quintile	1.76 (1.25, 2.49)		1.89 (1.39, 2.55)	
4th quintile	2.11 (1.51, 2.94)		2.75 (2.06, 3.67)	
5th quintile	3.40 (2.47, 4.68)		3.94 (2.98, 5.20)	

*Adjusted for age, smoking, alcohol use, education and race and center

Table 31. (MS. 1, supplemental results) Adjusted* hazard ratios (with 95 % CI) for incident heart failure by categories of BMI, waist circumference and waist hip ratio, stratified by gender, with the first three years of follow-up excluded, ARIC, 1987-2003

	Women HR (95 % CI)	Men HR (95 % CI)
BMI, Overweight vs. normal weight	1.39 (1.09, 1.77)	1.31 (1.07, 1.61)
BMI, Obese vs. normal weight	2.70 (2.15, 3.39)	2.49 (2.01, 3.08)
Waist Circumference, 2nd Tertile vs. 1st tertile	1.28 (0.98, 1.67)	1.61 (1.31, 1.99)
Waist Circumference, 3rd Tertile vs. 1st tertile	2.86 (2.24, 3.64)	2.30 (1.89, 2.80)
WHR, 2nd Tertile vs. 1st tertile	1.84 (1.41, 2.42)	2.62 (2.13, 3.22)
WHR, 3rd Tertile vs. 1st tertile	2.89 (2.25, 3.72)	1.46 (1.15, 1.85)

*Adjusted for age, smoking, alcohol use, education and race and center

Figure 21. (MS. 1, supplemental results) Fully adjusted hazard ratios (with 95 % CI) for incident heart failure by category of waist circumference within category of BMI, men and women combined, ARIC, 1987-2003

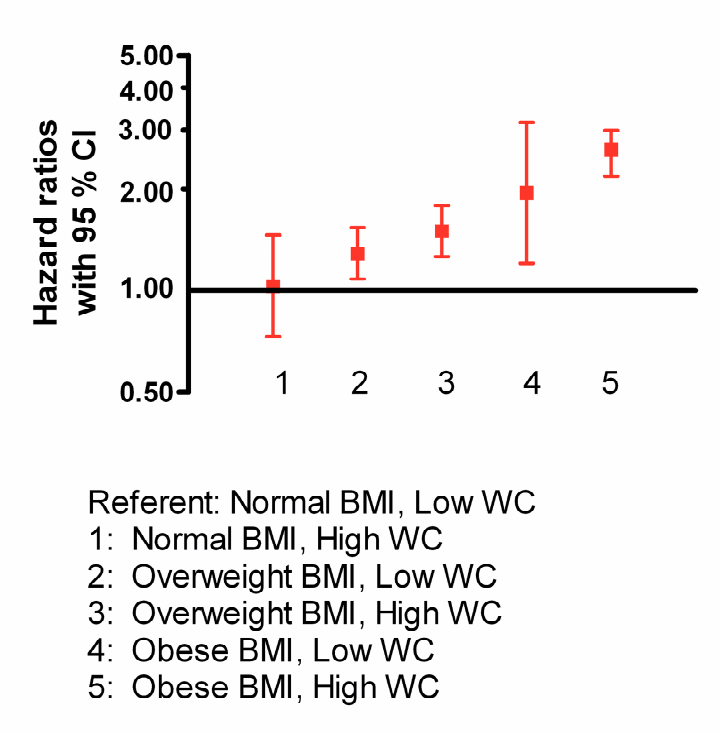
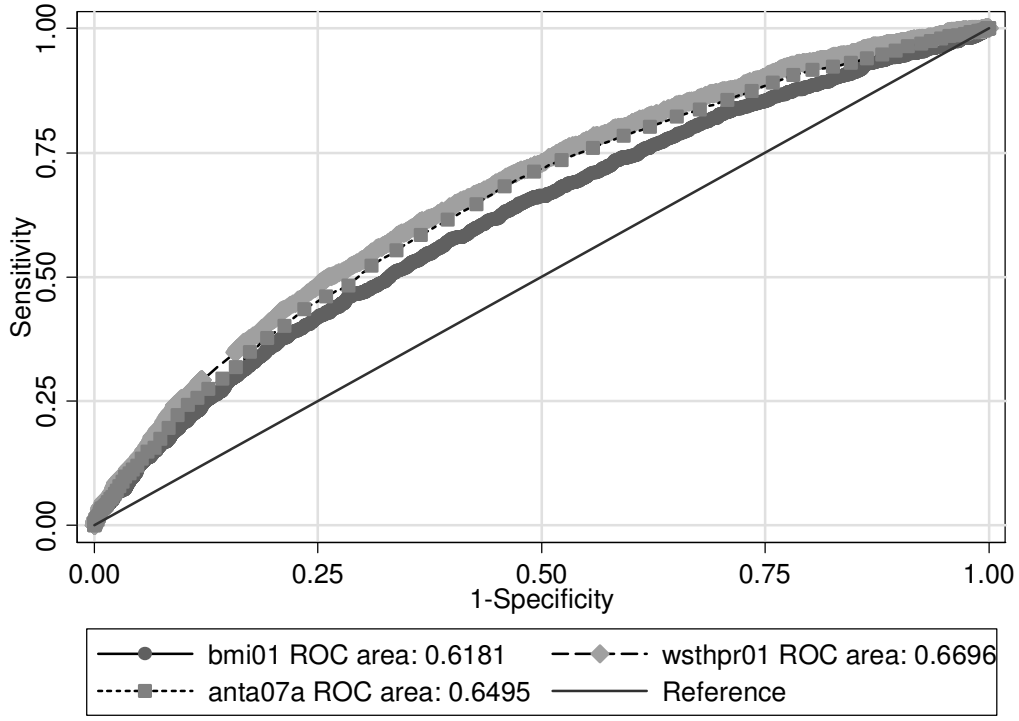


Figure 22. (MS. 1, supplemental results) Receiver operating curve (ROC) comparing BMI, WHR and WC for the prediction of incident heart failure, ARIC, 1987-2003



Anta07a is the variable for waist circumference
Wsthpr01 is the variable for waist hip ratio
Bmi01 is the variable for BMI

Table 32. (MS. 1, supplemental results) Summary of area under receiver operating curve at year 10 (AUC10), adjusted for optimism, for BMI, WHR and waist circumference, and BMI stratified by high and low WHR, and accompanying Hosmer-Lemeshow statistic (H-L), stratified by gender

	BMI	WHR	Waist circumference	BMI, stratified by WHR
All women				
AUC (10)[1], H-L	0.78910, 20.5	0.79171, 5.84	0.7913, 13.4	0.7935, 21.1
AUC (10)[2], H-L	0.79, 18.5	0.79, 5.3	0.79, 13.6	
All men				
AUC (10)[1], H-L	0.7375, 28.7	0.7384, 17.4	0.7335, 26.6	0.7416, 21
AUC (10)[2], H-L	0.74, 16.6	0.74, 11.6	0.73, 20.7	

[1]Model 1: In addition to anthropometric measure, adjusted for smoking status, drinking, age, educational status, and race

[2]Model 2: Model 1 + race and center

Conclusions were there was no difference in inferences based on AUC(10) with inclusion of race and center variable.

Table 33. (MS. 1, supplemental results) Summary of area under receiver operating curve at year 10 (AUC10) for BMI, WHR and waist circumference, stratified by race and gender

	AUC (10)[1], for BMI	AUC (10)[1], for WHR	AUC (10)[1], for WC
Black women	0.75	0.77	0.76
White women	0.79	0.79	0.79
Black men	0.69	0.71	0.70
White men	0.75	0.75	0.74

In addition to anthropometric measure, adjusted for smoking status, drinking, age, educational status

Figure 23. (MS. 1, supplemental results) Adjusted* hazard ratios (with 95 % CI) for incident heart failure by BMI, waist circumference and waist hip ratio, stratified by gender, ARIC, 1987-2003

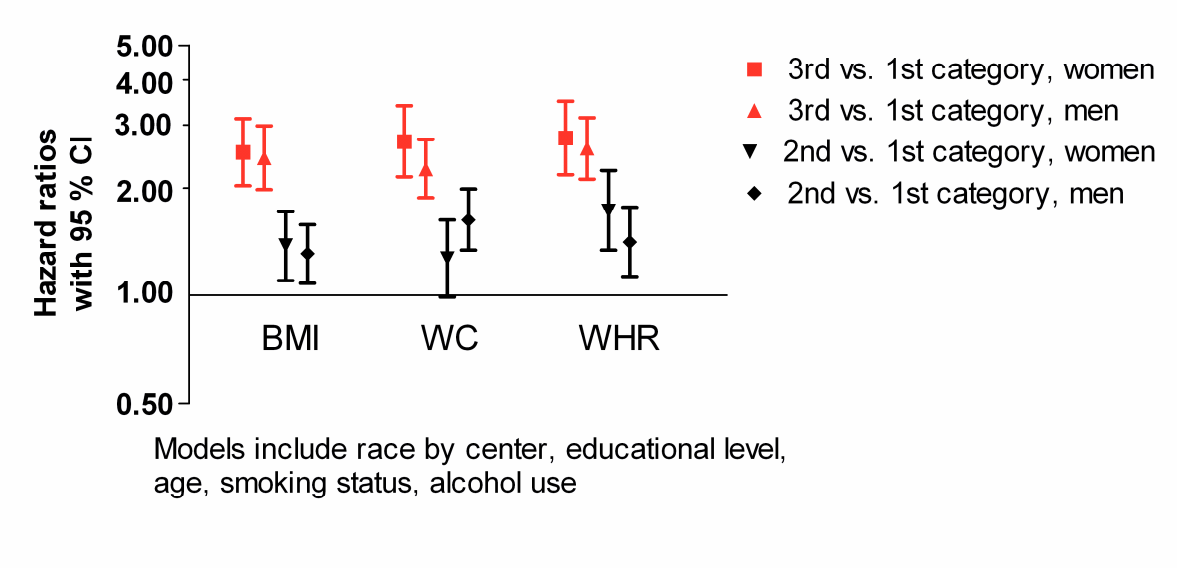


Table 34 (MS. 1, supplemental results) Comparison of adjusted* odds ratios (95 % CI) from logistic models with adjusted* hazard ratios (with 95 % CI) from Cox proportional hazards models for incident heart failure by categories of BMI, waist circumference (WC) and waist hip ratio (WHR), stratified by gender, ARIC, 1987-2003

	WOMEN		MEN	
	HR (95 % CI)	OR (95 % CI)	HR (95 % CI)	OR (95 % CI)
BMI, Overweight vs. normal	1.39 (1.09, 1.77)	1.46 (1.15, 1.85)	1.31 (1.07, 1.61)	1.39 (1.13, 1.71)
BMI, Obese vs. normal	2.70 (2.15, 3.39)	2.78 (2.21, 3.50)	2.49 (2.01, 3.08)	2.77 (2.21, 3.47)
WC, 2nd Tertile vs. 1st tertile	1.28 (0.98, 1.67)	1.34 (1.03, 1.74)	1.61 (1.31, 1.99)	1.71 (1.38, 2.12)
WC, 3rd Tertile vs. 1st tertile	2.86 (2.24, 3.64)	2.96 (2.33, 3.77)	2.30 (1.89, 2.80)	2.45 (2.00, 3.00)
WHR, 2nd Tertile vs. 1st tertile	1.84 (1.41, 2.42)	1.78 (1.36, 2.32)	1.46 (1.15, 1.85)	1.45 (1.14, 1.85)
WHR, 3rd Tertile vs. 1st tertile	2.89 (2.25, 3.72)	2.92 (2.28, 3.73)	2.62 (2.13, 3.22)	2.75 (2.23, 3.40)

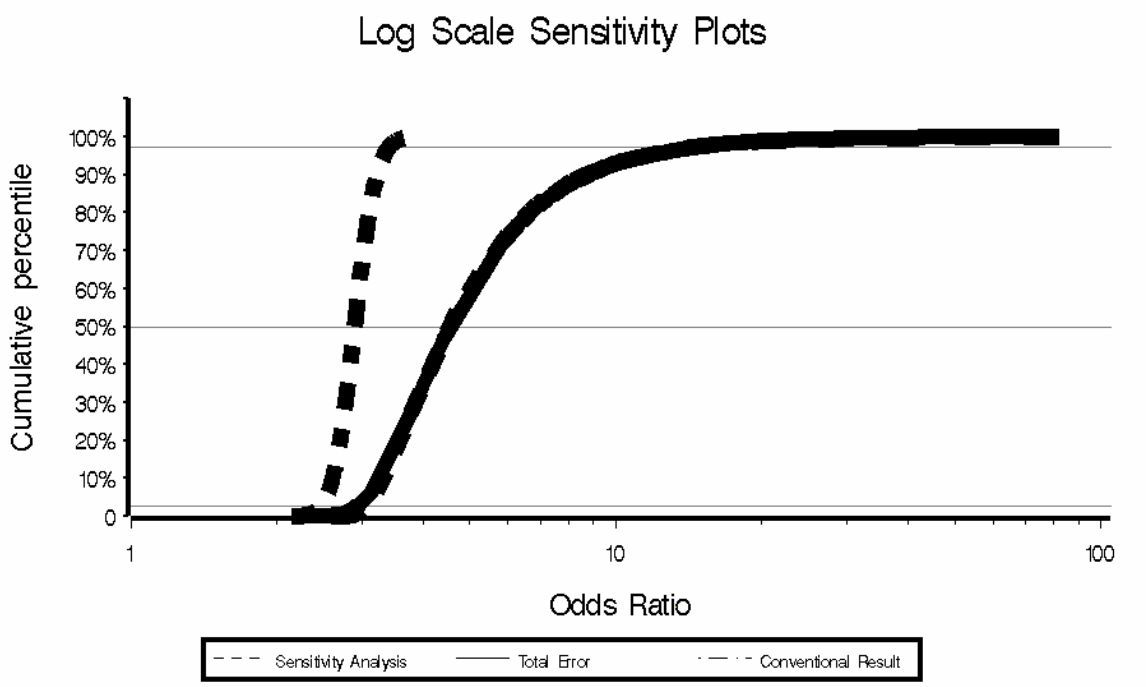
*adjusted for smoking status, drinking, age, educational status, and race and center

Table 35. (MS. 1, supplemental results) Summary of the odds ratios from the cumulative probability distributions from the sensitivity analysis of the effect of nondifferential misclassification of the outcome (heart failure) on the association of obesity as defined by BMI with incident heart failure

	median OR with 95 % uncertainty intervals	Width of intervals
Nondifferential outcome misclassification		
Conventional analysis (random error only)	**2.89 (2.47, 3.42)**	**1.39**
Sensitivity analysis (systematic error only)	**4.54 (3.06, 14.63)**	**4.79**
Total error analysis (systematic + random error)	**4.54 (2.93, 14.83)**	**5.05**

Sensitivity: min=0.6, mode 1 = 0.7, mode2 = 0.85, max. = 1
Specificity: min = 0.94, mode 1 = 0.96, mode 2 = 0.96, max = 1

Figure 24. (MS. 1, supplemental results) Graphical summary of the odds ratios from the cumulative probability distributions from the sensitivity analysis of the effect of nondifferential misclassification of the outcome (heart failure) on the association of obesity as defined by BMI with incident heart failure



C. Supplemental Results, Manuscript 2

Figure 25. (MS. 2, supplemental results) Percentage of normal weight, overweight and obese as defined by category of BMI, waist circumference and waist hip ratio, stratified by gender

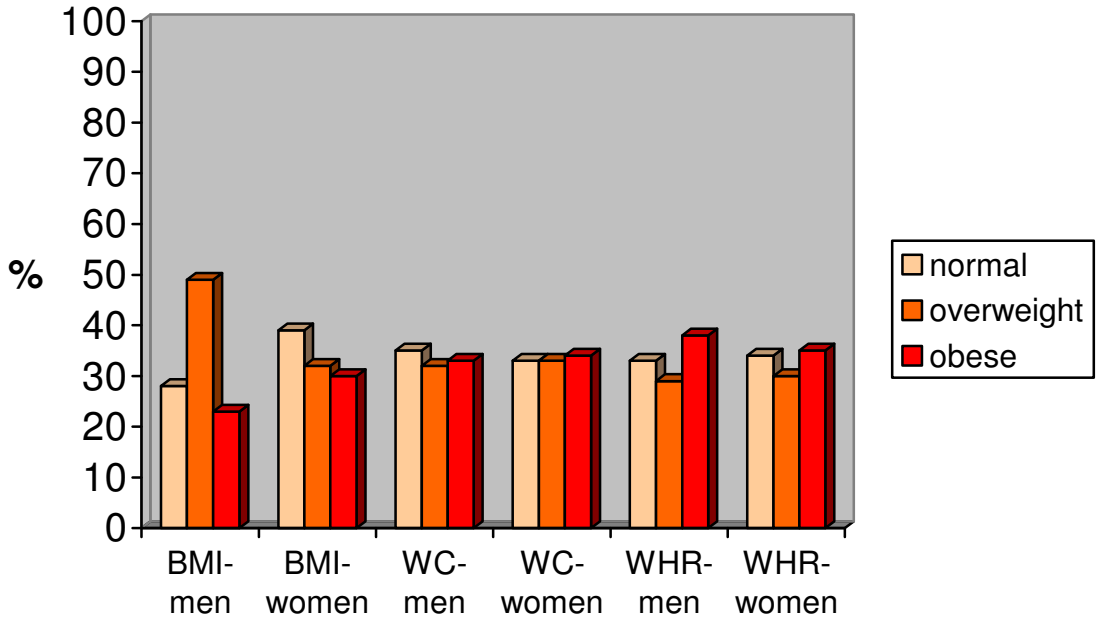
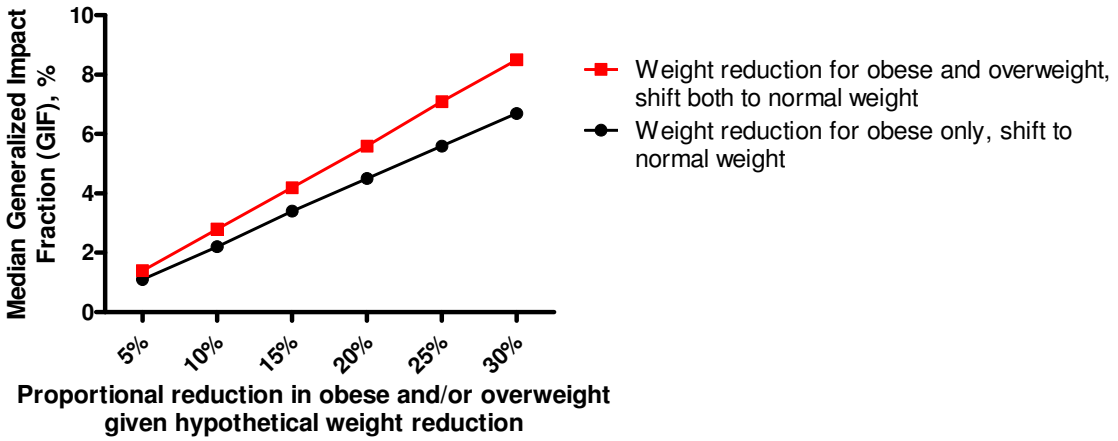


Table 36. (MS. 2, supplemental results) The median generalized impact fraction and attributable fraction (with 2.5 % and 97.5 % simulation intervals) from 10,000 bootstrap datasets using the case-load weighted-sum method, given 10 scenarios of reduced prevalence of obesity and overweight, ARIC, 1987-2003

Hypothetical scenarios of weight reduction	**Median GIF, % (95 % simulation intervals)**
Reduction in obesity, shift to normal weight	
Scenario 1: 5 % reduction in obesity	1.1 (0.9, 1.3)
Scenario 2: 10 % reduction in obesity	2.2 (1.8, 2.7)
Scenario 3: 15 % reduction in obesity	3.4 (2.7, 4.0)
Scenario 4: 20 % reduction in obesity	4.5 (3.6, 5.3)
Scenario 5: 25 % reduction in obesity	5.6 (4.5, 6.7)
Scenario 6: 30 % reduction in obesity	6.7 (5.4, 8.0)
Reduction in obesity and overweight, shift both to normal weight	
Scenario 1: 5 % reduction in obesity and overweight	1.4 (1.0, 1.8)
Scenario 2: 10 % reduction in obesity and overweight	2.8 (2.0, 3.6)
Scenario 3: 15 % reduction in obesity and overweight	4.2 (3.0, 5.4)
Scenario 4: 20 % reduction in obesity and overweight	5.6 (4.1, 7.1)
Scenario 5: 25 % reduction in obesity and overweight	7.1 (5.1, 8.9)
Scenario 6: 30 % reduction in obesity and overweight	8.5 (6.1, 10.7)

Figure 26. (MS. 2, supplemental results) Graph of the median generalized impact fraction for 5 % up to 30 % weight reduction with weight loss down to normal weight for obese (black) and for obese and overweight (red)



REFERENCES

Alaud-din, A., S. Meterissian, et al. (1990). "Assessment of cardiac function in patients who were morbidly obese." Surgery **108**(4): 809-18; discussion 818-20.

Alexander, J., Ed. (1998). Historical Notes. The Heart and Lung in Obesity. Armonk, New York, Futura.

Alexander, J. K., K. H. Amad, et al. (1962). "Observations on some clinical features of extreme obesity, with particular reference to cardiorespiratory effects." Am J Med **32**: 512-524.

Alexander, J. K., E. Dennis, et al. (1962). "Blood volume, cardiac output and distribution of systemic blood flow in extreme obesity." Cardiovasc Res Center Bull **1**: 39-44.

Alexander, J. K. and K. L. Peterson (1972). "Cardiovascular effects of weight reduction." Circulation **45**(2): 310-8.

Alpert, M. A., B. E. Terry, et al. (1985). "Effect of weight loss on cardiac chamber size, wall thickness and left ventricular function in morbid obesity." Am J Cardiol **55**(6): 783-6.

Amad, K. H., J. C. Brennan, et al. (1965). "The cardiac pathology of chronic exogenous obesity." Circulation **32**(5): 740-5.

Anonymous (1995). "Guidelines for the diagnosis of heart failure. The Task Force on Heart Failure of the European Society of Cardiology." Eur Heart J **16**(6): 741-51.

ARIC Investigators (1989). "The Atherosclerosis Risk in Communities (ARIC) Study: design and objectives." Am. J. Epidemiol. **129**: 687-702.

Auchincloss, J. J., E. Cook, et al. (1955). "Polycythemia of Unknown Cause with Alveolar Hypoventilation." Clin Res Proc **3**: 31-32.

Barrett-Conner, E. (1985). "Obesity, atherosclerosis, and coronary heart disease." Annals of Internal Medicine **103**: 1010-1019.

Benichou, J. (2001). "A review of adjusted estimators of attributable risk." Stat Methods Med Res **10**(3): 195-216.

Benichou, J. (2007). "Biostatistics and epidemiology: measuring the risk attributable to an environmental or genetic factor." C R Biol **330**(4): 281-98.

Bertoni, A. G., W. G. Hundley, et al. (2004). "Heart Failure Prevalence, Incidence, and Mortality in the Elderly With Diabetes." Diabetes Care **27**(3): 699-703.

Bibbins-Domingo, K., F. Lin, et al. "Predictors of Heart Failure Among Women With Coronary Disease." Circulation **110**(11): 1424-1430.

Birman-Deych, E., A. D. Waterman, et al. (2005). "Accuracy of ICD-9-CM codes for identifying cardiovascular and stroke risk factors." Med Care **43**(5): 480-5.

Braunwald, E. (1997). "Shattuck lecture--cardiovascular medicine at the turn of the millennium: triumphs, concerns, and opportunities." N Engl J Med **337**(19): 1360-9.

Brown, D. W., G. A. Haldeman, et al. (2005). "Racial or ethnic differences in hospitalization for heart failure among elderly adults: Medicare, 1990 to 2000." Am Heart J **150**(3): 448-54.

Burwell, C., E. Robin, et al. (1956). "Extreme obesity associated with hypoventilation - a Pickwickian syndrome." Am J Med **21**: 811-818.

Chambless, L. E. and G. Diao (2006). "Estimation of time-dependent area under the ROC curve for long-term risk prediction." Stat Med **25**(20): 3474-86.

Chen, Y. T., V. Vaccarino, et al. (1999). "Risk factors for heart failure in the elderly: a prospective community-based study." Am J Med **106**(6): 605-12.

Croft, J. B., W. H. Giles, et al. (1999). "Heart failure survival among older adults in the United States: a poor prognosis for an emerging epidemic in the Medicare population." Arch Intern Med **159**(5): 505-10.

Crow, R., R. Prineas, et al. (1995). "Relation between electrocardiography and echocardiography for left ventricular mass in mild systemic hypertension (results from Treatment of Mild Hypertension Study)." Am J Cardiol **75**: 1233-8.

Curb, J. D., A. McTiernan, et al. (2003). "Outcomes ascertainment and adjudication methods in the Women's Health Initiative." Ann Epidemiol **13**(9 Suppl): S122-8.

Curioni, C. C. and P. M. Lourenco (2005). "Long-term weight loss after diet and exercise: a systematic review." Int J Obes (Lond) **29**(10): 1168-74.

Dagenais, G. R., Q. Yi, et al. (2005). "Prognostic impact of body weight and abdominal obesity in women and men with cardiovascular disease." Am Heart J **149**(1): 54-60.

de Divitiis, O., S. Fazio, et al. (1981). "Obesity and cardiac function." Circulation **64**(3): 477-82.

Drazner, M. H., J. E. Rame, et al. (2004). "Increased left ventricular mass is a risk factor for the development of a depressed left ventricular ejection fraction within five years: the Cardiovascular Health Study." J Am Coll Cardiol **43**(12): 2207-15.

Eckfeldt, J., L. Chambless, et al. (1994). "Short-term, within-person variability in clinical chemistry test results. Experience from the Atherosclerosis Risk in Communities Study." Arch Pathol Lab Med **1994**(118): 496-500.

Efron, B. (1977). "The efficiency of Cox's likelihood function for censored data." Journal of the American Statistical Association **72**(359): 557-565.

Eriksson, H., K. Caidahl, et al. (1987). "Cardiac and pulmonary causes of dyspnoea--validation of a scoring test for clinical-epidemiological use: the Study of Men Born in 1913." Eur Heart J **8**(9): 1007-14.

Eriksson, H., K. Caidhal, et al. (1987). "Cardiac and pulmonary causes of dyspnea-validation of a scoring test for clinical-epidemiological use: the study of men born in 1913." Eur Heart J **8**: 1007-14.

Ferrario, M., M. Carpenter, et al. (1995). "Reliability of body fat distribution measurements. The ARIC Study baseline cohort results. Atherosclerosis Risk in Communities Study." Int J Obes Relat Metab Disord **19**(7): 449-57.

Flegal, K. M., B. I. Graubard, et al. (2004). "Methods of calculating deaths attributable to obesity." Am J Epidemiol **160**(4): 331-8.

Flegal, K. M., B. I. Graubard, et al. (2005). "Excess deaths associated with underweight, overweight, and obesity." Jama **293**(15): 1861-7.

Flegal, K. M., B. I. Graubard, et al. (2007). "Cause-specific excess deaths associated with underweight, overweight, and obesity." Jama **298**(17): 2028-37.

Folsom, A. R., L. H. Kushi, et al. (2000). "Associations of general and abdominal obesity with multiple health outcomes in older women: the Iowa Women's Health Study." Arch Intern Med **160**(14): 2117-28.

Ford, E. S., D. F. Williamson, et al. (1997). "Weight change and diabetes incidence: findings from a national cohort of US adults." Am J Epidemiol **146**(3): 214-22.

Ford, W. (1950). "Old Pharr." Bull Hist Med **24**: 219-226.

Fox, M. P., T. L. Lash, et al. (2005). "A method to automate probabilistic sensitivity analyses of misclassified binary variables." Int J Epidemiol **34**(6): 1370-6.

Franz, M. J., J. J. VanWormer, et al. (2007). "Weight-loss outcomes: a systematic review and meta-analysis of weight-loss clinical trials with a minimum 1-year follow-up." J Am Diet Assoc **107**(10): 1755-67.

Goff, D. C., Jr., D. K. Pandey, et al. (2000). "Congestive heart failure in the United States: is there more than meets the I(CD code)? The Corpus Christi Heart Project." Arch Intern Med **160**(2): 197-202.

Goldberg, R. J. and M. A. Konstam (1999). "Assessing the population burden from heart failure: need for sentinel population-based surveillance systems." Arch Intern Med **159**(1): 15-7.

Gottdiener, J. S., R. L. McClelland, et al. (2002). "Outcome of congestive heart failure in elderly persons: influence of left ventricular systolic function. The Cardiovascular Health Study." Ann Intern Med **137**(8): 631-9.

Graham, P. (2000). "Bayesian inference for a generalized population attributable fraction: the impact of early vitamin A levels on chronic lung disease in very low birthweight infants." Stat Med **19**(7): 937-56.

Greenland, S. (1998). Basic Methods for Sensitivity Analysis and External Adjustment. Modern Epidemiology. K. Rothman and S. Greenland. Philadelphia, PA, Lippincott-Raven**:** 343-357.

Greenland, S. (2001). "Sensitivity analysis, Monte Carlo risk analysis, and Bayesian uncertainty assessment." Risk Anal **21**(4): 579-83.

Greenland, S. (2004). "Interval estimation by simulation as an alternative to and extension of confidence intervals." Int J Epidemiol **33**(6): 1389-97.

Greenland, S. (2005). "Multiple-bias modelling for analysis of observational data." J.R. Statist. Soc. A **Part 2**(168): 267-306.

Greenland, S., J. Pearl, et al. (1999). "Causal diagrams for epidemiologic research." Epidemiology **10**(1): 37-48.

Greenland, S. and K. Rothman (1998). Modern Epidemiology. Philadelphia, Lippincott-Raven.

Harrell, F. E., Jr., K. L. Lee, et al. (1996). "Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors." Stat Med **15**(4): 361-87.

He, J., L. G. Ogden, et al. (2001). "Risk factors for congestive heart failure in US men and women: NHANES I epidemiologic follow-up study." Arch Intern Med **161**(7): 996-1002.

Heckbert, S. R., C. Kooperberg, et al. (2004). "Comparison of Self-Report, Hospital Discharge Codes, and Adjudication of Cardiovascular Events in the Women's Health Initiative." Am. J. Epidemiol. **160**(12): 1152-1158.

Hunt, S. A., W. T. Abraham, et al. (2005). "ACC/AHA 2005 Guideline Update for the Diagnosis and Management of Chronic Heart Failure in the Adult: A Report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines (Writing Committee to Update the 2001 Guidelines for the Evaluation and Management of Heart Failure): Developed in Collaboration With the American College of Chest Physicians and the International Society for Heart and Lung Transplantation: Endorsed by the Heart Rhythm Society." Circulation **112**(12): e154-235.

Ingelsson, E., J. Sundstrom, et al. (2005). "Insulin resistance and risk of congestive heart failure." JAMA **294**(3): 334-41.

Jackson, R., L. Chambless, et al. (1996). "Differences between respondents and nonrespondents in a multicenter community-based study vary by gender and ethnicity." J Clin Epidemiol **49**: 1441-1446.

Kaltman, A. J. and R. M. Goldring (1976). "Role of circulatory congestion in the cardiorespiratory failure of obesity." Am J Med **60**(5): 645-53.

Kannel, W. B., R. B. D'Agostino, et al. (1999). "Profile for estimating risk of heart failure." Arch Intern Med **159**(11): 1197-204.

Kenchaiah, S., J. C. Evans, et al. (2002). "Obesity and the risk of heart failure." N Engl J Med **347**(5): 305-13.

Kenchaiah, S., J. M. Gaziano, et al. (2004). "Impact of obesity on the risk of heart failure and survival after the onset of heart failure." Med Clin North Am **88**(5): 1273-94.

Kenchaiah, S., J. Narula, et al. (2004). "Risk factors for heart failure." Med Clin North Am **88**(5): 1145-72.

Krumholz, H. M., E. M. Parent, et al. (1997). "Readmission after hospitalization for congestive heart failure among Medicare beneficiaries." Arch Intern Med **157**(1): 99-104.

Lash, T. L. and A. K. Fink (2003). "Semi-automated sensitivity analysis to assess systematic errors in observational data." Epidemiology **14**(4): 451-8.

Lauer, M. S., K. M. Anderson, et al. (1991). "The impact of obesity on left ventricular mass and geometry. The Framingham Heart Study." Jama **266**(2): 231-6.

Levin, M. (1953). "The occurrence of lung canceer in man." Acta Unio Internationalis Contra Cancrum: 531-41.

Levy, D., M. G. Larson, et al. (1996). "The progression from hypertension to congestive heart failure." JAMA **275**(20): 1557-1562.

Loehr, L. R., W. D. Rosamond, et al. (2008). "Heart Failure Incidence and Survival (from the Atherosclerosis Risk in Communities Study)." Am J Cardiol **101**(7): 1016-22.

Maclure, M. and S. Schneeweiss (2001). "Causation of bias: the episcope." Epidemiology **12**(1): 114-22.

Maldonado, G. and C. V. Phillips (2004). "Editorial: Wishful thinking." Epidemiol Perspect Innov **1**(1): 2.

Masoudi, F. and S. Inzucchi (2007). "Diabetes mellitus and heart failure: epidemiology, mechanisms, and pharmacotherapy." American Journal of Cardiology **99**(4A): 113B-132B.

Masoudi, F. A., E. P. Havranek, et al. (2002). "The burden of chronic congestive heart failure in older persons: magnitude and implications for policy and research." Heart Fail Rev. **7**(1): 9-16.

McGavock, J. M., I. Lingvay, et al. (2007). "Cardiac steatosis in diabetes mellitus: a 1H-magnetic resonance spectroscopy study." Circulation **116**(10): 1170-5.

McGavock, J. M., R. G. Victor, et al. (2006). "Adiposity of the Heart*, Revisited." Ann Intern Med **144**(7): 517-524.

Mokdad, A. H., E. S. Ford, et al. (2000). "Diabetes trends in the U.S.: 1990-1998." Diabetes Care **23**(9): 1278-83.

Morgenstern, H. and E. S. Bursic (1982). "A method for using epidemiologic data to estimate the potential impact of an intervention on the health status of a target population." J Community Health **7**(4): 292-309.

Murphy, N., K. Macintyre, et al. (2005). "Long-term cardiovascular consequences of obesity: 20-year follow-up of more than 15,000 middle-aged men and women (the Renfrew-Paisley study)." European Heart Journal: 1-11.

Murray, C. J. and A. D. Lopez (1999). "On the comparable quantification of health risks: lessons from the Global Burden of Disease Study." Epidemiology **10**(5): 594-605.

Must, A., J. Spadano, et al. (1999). "The disease burden associated with overweight and obesity." Jama **282**(16): 1523-9.

Nakayama, T. (2000). "Under-reporting of attributable risk and reporting of the risk ratio in epidemiologic literature." Epidemiology **11**(3): 366-7.

National Center for Health Statistics (2005). Health, United States, 2005 with chartbook on trends in the health of americans with special feature on adults 55-64 years. Hyattsville, Maryland.

Nicklas, B. J., M. Cesari, et al. (2006). "Abdominal Obesity Is an Independent Risk Factor for Chronic Heart Failure in Older People." Journal of the American Geriatrics Society **54**(3): 413-420.

NIH (1998). "Clinical Guidelines on the Identification, Evaluation, and Treatment of Overweight and Obesity in Adults: The Evidence Report: National Institutes of Health." Obesity Research(supplement 2): 51S - 209S.

Nilunger, L., F. Diderichsen, et al. (2004). "Using risk analysis in Health Impact Assessment: the impact of different relative risks for men and women in different socio-economic groups." Health Policy **67**(2): 215-24.

O'brien G (2002). "Obesity and the Risk of Heart Failure." New England Journal of Medicine **347**(23): 1887-9 author reply.

Phillips, C. V. (2003). "Quantifying and reporting uncertainty from systematic errors." Epidemiology **14**(4): 459-66.

Phillips, C. V., K. J. Goodman, et al. (2004). "Lead editorial: The need for greater perspective and innovation in epidemiology." Epidemiol Perspect Innov **1**(1): 1.

Poirier, P., T. D. Giles, et al. (2006). "Obesity and cardiovascular disease: pathophysiology, evaluation, and effect of weight loss: an update of the 1997 American Heart Association Scientific Statement on Obesity and Heart Disease from the Obesity Committee of the Council on Nutrition, Physical Activity, and Metabolism." Circulation **113**(6): 898-918.

Poole, C. (2001). "Low P-values or narrow confidence intervals: which are more durable?" Epidemiology **12**(3): 291-4.

Psaty, B. M., R. Boineau, et al. (1999). "The potential costs of upcoding for heart failure in the United States." Am J Cardiol **84**(1): 108-9, A9.

Rasooly, Y., Z. Sasson, et al. (1993). "Relation between body fat distribution and left ventricular mass in men without structural heart disease or systemic hypertension." Am J Cardiol **71**(16): 1477-9.

Rathore, S. S., J. M. Foody, et al. (2003). "Race, Quality of Care, and Outcomes of Elderly Patients Hospitalized With Heart Failure." JAMA **289**(19): 2517-2524.

Rockhill, B., B. Newman, et al. (1998). "Use and misuse of population attributable fractions." Am J Public Health **88**(1): 15-9.

Rodgers, A. (2002). Defining and Assessing Risks to Health. The World Health Report: 2002 - Reducing Risks, Reporting Health Life. C. J. Murray and A. D. Lopez. Geneva, World Health Organization.

Rodgers, A., M. Ezzati, et al. (2004). Distribution of Major Health Risks: Findings from the Global Burden of Disease Study. **1:** e27.

Roger, V. L., S. A. Weston, et al. (2004). "Trends in Heart Failure Incidence and Survival in a Community-Based Population." JAMA **292**(3): 344-350.

Rosamond, W., K. Flegal, et al. (2007). Heart Disease and Stroke Statistics--2007 Update: A Report From the American Heart Association Statistics Committee and Stroke Statistics Subcommittee. **115:** e69-171.

Rosamond, W., K. Flegal, et al. (2008). Heart Disease and Stroke Statistics--2008 Update: A Report From the American Heart Association Statistics Committee and Stroke Statistics Subcommittee. **117:** e25-146.

Rothman, K. and S. Greenland (1998). Modern Epidemiology. Philadelphia, Lippincott Williams & Wilkins.

Schellenbaum, G. D., S. R. Heckbert, et al. (2006). "Congestive heart failure incidence and prognosis: case identification using central adjudication versus hospital discharge diagnoses." Ann Epidemiol **16**(2): 115-22.

Schellenbaum, G. D., T. D. Rea, et al. (2004). "Survival Associated with Two Sets of Diagnostic Criteria for Congestive Heart Failure." Am. J. Epidemiol. **160**(7): 628-635.

Schocken, D. D., E. J. Benjamin, et al. (2008). "Prevention of Heart Failure: A Scientific Statement From the American Heart Association Councils on Epidemiology and Prevention, Clinical Cardiology, Cardiovascular Nursing, and High Blood Pressure Research; Quality of Care and Outcomes Research Interdisciplinary Working Group; and Functional Genomics and Translational Biology Interdisciplinary Working Group." Circulation **117**(19): 2544-2565.

Semenza, J. C., C. H. Rubin, et al. (1996). "Heat-related deaths during the July 1995 heat wave in Chicago." N Engl J Med **335**(2): 84-90.

Shahar, E., Lori L. Boland, et al. "Docosahexaenoic Acid and Smoking-Related Chronic Obstructive Pulmonary Disease." Am. J. Respir. Crit. Care Med. **159**(6): 1780-1785.

Smith, H. and F. Willius (1933). "Adiposity of the Heart." Archives of Internal Medicine **52**: 911-931.

Stamler, J. (1991). "Epidemiologic findings on body mass and blood pressure in adults." Ann Epidemiol **1**(4): 347-62.

The ARIC Investigators (1989). "The Atherosclerosis Risk in Communities (ARIC) Study: design and objectives." Am J Epidemiol **129**: 687-702.

Thom, T., N. Haase, et al. (2006). "Heart Disease and Stroke Statistics--2006 Update. A Report From the American Heart Association Statistics Committee and Stroke Statistics Subcommittee." Circulation **113**(14): e85-151.

US Department of Agriculture (1990). Dietary Guidelines for Americans, Washington DC.

US Department of Health and Human Services (2000). Healthy People 2010: Understanding and Improving Health. Washington, DC, U.S. Government Printing Office. **II**.

Vander Hoorn, S., M. Ezzati, et al. (2004). Estimating Attributable Burden of Disease from Exposure and Hazard Data. Comparative Quantification of Health Risks, Global and Regional Burden of Disease Attributable to Selected Major Risk Factors. M. Ezzati, A. D. Lopez, A. Rodgers and C. J. Murray. Geneva, World Health Organization. **1:** 2129-2140.

Vasan, R. S. (2003). "Cardiac function and obesity." Heart **89**(10): 1127-9.

Vasan, R. S., L. M. Sullivan, et al. (2003). "Inflammatory markers and risk of heart failure in elderly subjects without prior myocardial infarction: the Framingham Heart Study." Circulation **107**(11): 1486-91.

Vitelli, L. L., R. S. Crow, et al. (1998). "Electrocardiographic findings in a healthy biracial population. Atherosclerosis Risk in Communities (ARIC) Study Investigators." Am J Cardiol **81**(4): 453-9.

Walter, S. D. (1980). "Prevention for multifactorial diseases." Am J Epidemiol **112**(3): 409-16.

Weinberg, C. R. (1993). "Toward a clearer definition of confounding." Am J Epidemiol **137**(1): 1-8.

Wilson, P. W., R. B. D'Agostino, et al. (2002). "Overweight and obesity as determinants of cardiovascular risk: the Framingham experience." Arch Intern Med **162**(16): 1867-72.

Wong, T. Y., W. Rosamond, et al. (2005). "Retinopathy and Risk of Congestive Heart Failure." JAMA **293**(1): 63-69.

Zhou, Y. T., P. Grayburn, et al. (2000). "Lipotoxic heart disease in obese rats: implications for human obesity." Proc Natl Acad Sci U S A **97**(4): 1784-9.