Prediction of Incident Type 2 Diabetes: Modifiable and Non-Modifiable Risk Factors Among High-Risk Populations and Diverse Racial/Ethnic Groups

A DISSERTATION
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY

Julie Kristen Bower

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Pamela J. Schreiner, Advisor

July 2010

ACKNOWLEDGEMENTS

I would like to acknowledge my advisor (Pamela Schreiner) and my dissertation committee members (Aaron Folsom, James Pankow, and Elizabeth Seaquist) for their guidance and feedback throughout the dissertation process, and for a general willingness to share their wisdom so freely. I also would like to thank all of the teachers and mentors that have provided me with the tools to undertake such a large project. In addition, this research was completed with support from the Cardiovascular Disease Epidemiology and Prevention training grant T32 HL007779 from the National Heart, Lung, and Blood Institute. This training program provided me with invaluable learning opportunities and the chance to interact with amazing leaders in the field of cardiovascular epidemiology.

My family and friends have provided an abundance of instrumental and moral support for which I am so grateful. I would especially like to thank my parents who always encouraged me to challenge myself and to learn as much as I could in every experience, no matter how difficult it may be. Finally, I would like to acknowledge my husband, John. He is my partner in all things. His unconditional support and encouragement gives me the confidence to persevere and never give up. I am thankful to have someone by my side that is always cheering for me and always there to celebrate every milestone.

DEDICATION

This dissertation is dedicated to my first teacher, devoted friend, and mom, Julieta Sabater Cortese. From her I received my name, my love for learning new things, and a role model for how a strong and independent woman should charge through life.

ABSTRACT

Diabetes mellitus refers to a class of disorders resulting from defects in insulin secretion and/or insulin action. Nearly ten percent of adults in the United States aged twenty years and older live with type 2 diabetes, which accounts for approximately 90% to 95% of all diagnosed diabetes mellitus cases in adults. Improving risk stratification and risk prediction facilitates the identification of those most likely to develop diabetes and informs prevention efforts. The aim of this dissertation was to evaluate the predictive validity of currently established definitions of high risk across ethnically diverse population subgroups, and to examine the impact of modifiable and non-modifiable factors that predict type 2 diabetes risk, presented in three related manuscripts.

In the first manuscript, the predictive validity of fasting glucose and glycated hemoglobin A1c (HbA1c) measures were evaluated across population subgroups in the Multi-Ethnic Study of Atherosclerosis (MESA), a community-based observational longitudinal cohort study that enrolled 6,814 Caucasian Americans, Chinese Americans, African Americans, and Hispanic Americans aged 45-84 years and free of clinical cardiovascular disease (CVD). Over five years, 692 incident diabetes cases occurred; the association between fasting glucose levels at baseline and cumulative incidence of diabetes did not differ by self-reported ethnic group. However, the strength of association between HbA1c and diabetes incidence differed significantly by ethnicity. In multivariable models, the c-statistics (markers of goodness-of-fit) from logistic regression models ranged from 0.75 in African Americans to 0.87 in Caucasians.

The second manuscript examined predictors of incident diabetes among high-risk individuals with pre-diabetes in the Coronary Artery Risk Development in Young Adults (CARDIA) study, a community-based longitudinal observational study that enrolled 5,115 African Americans and Caucasian Americans aged 18-30 years and free of clinical CVD. In the CARDIA cohort, body mass index (BMI) was the strongest predictor of diabetes among individuals with pre-diabetes; other predictors included race/ethnicity, diastolic blood pressure, HDL-cholesterol, and alcohol consumption.

A final manuscript explored genetic predisposition to diabetes in addition to potential interactions between single nucleotide polymorphisms (SNPs), the trait of cardiorespiratory fitness, and diabetes-related outcomes in the CARDIA cohort. SNPs of rs8050136 of the fat mass and obesity associated (*FTO*) gene and rs3856806 of the peroxisome proliferator-activated receptor gamma (*PPARG*) gene were significantly associated with HbA1c levels; the association between cardiorespiratory fitness level and HbA1c differed by genotype of rs8050136 in Caucasians and rs3856806 in African Americans.

This dissertation served to evaluate the predictive value of current screening tools and markers of risk for type 2 diabetes, and supported previous findings that modifiable risk factors significantly contribute to diabetes risk prediction. Improving diabetes prevention in population subgroups that are disproportionately affected — particularly those with pre-diabetes and in certain race/ethnic groups — is dependent upon effective screening tools validated across populations, and a more comprehensive understanding of

the role of modifiable and non-modifiable risk factors in influencing one's diabetes risk profile.

1. TABLE OF CONTENTS

1.1 LIST OF TABLES

1.2 LIST OF FIGURES

2. INTRODUCTION

Nearly ten percent of adults in the United States aged twenty years and older live with diabetes mellitus, a number that continues to increase over time.[1] Management of diabetes and associated comorbidities translate into an enormous burden on the medical and public health systems. Cardiovascular morbidity among people with type 2 diabetes is estimated at two to four-fold greater compared to non-diabetic individuals;[2,3] in 2007, the direct and indirect costs attributable to diabetes was estimated at $174 billion.[4] Thus, preventing diabetes, as well as managing hyperglycemia and cardiovascular risk factors among those living with diabetes, is critical to both the control of cardiovascular disease (CVD) and the broader management of health care costs worldwide.[5]

This chapter reviews the relevant literature regarding the epidemiology and etiology of diabetes, modifiable and non-modifiable risk factors, and the impact of diabetes on cardiovascular disease and mortality. The manuscripts that follow will address specific gaps in the literature and serve to inform diabetes prevention efforts. Manuscript #1 will evaluate and explore definitions of high risk in an ethnically-diverse population-based cohort study. Manuscript #2 will examine risk factors that predict progression from the pre-diabetic state to frank diabetes. Finally, manuscript #3 will explore genetic predisposition to diabetes in addition to potential interactions between single nucleotide polymorphisms (SNPs), the trait of cardiorespiratory fitness (which has both modifiable and non-modifiable components), and diabetes risk. This dissertation serves to add to the existing body of knowledge regarding the early identification of high-

risk individuals in diverse populations and modifiable risk factors that might influence diabetes development.

3. BACKGROUND AND SIGNIFICANCE

3.1 *Introduction to Diabetes*

3.1.1 *Terms and Definitions*

Diabetes mellitus refers to a class of disorders resulting from defects in insulin secretion, insulin action, or both.[6] Polyuric states reminiscent of diabetes were described as early as 1550 BC.[7] The term "diabetes mellitus" itself originated from an allusion to a condition characterized by urine that had a sweet, honeyed taste; was sticky to the touch; and attracted ants. At that time, two primary types of diabetes were observed: one that tended to affect older, more overweight people and the other that tended to affect younger people with short life expectancies. The two observed classes later developed into what we currently term type 1 and type 2 diabetes mellitus.

A vast majority of diabetes mellitus cases are either of type 1 or type 2. Type 1 diabetes mellitus is caused by the autoimmune destruction of the pancreatic beta cell (which ultimately leads to absolute insulin deficiency), while type 2 diabetes mellitus is caused by a combination of insulin resistance and insufficient insulin secretory response. Type 2 diabetics might experience disturbances in insulin secretion; insulin resistance in the muscle, liver, and adipocytes; and abnormal splanchnic glucose uptake.[8] Analyses for this project include data from prospective cohort studies of adults, are concerned with predicting incident diabetes, and therefore exclude both prevalent cases of diabetes at the baseline examination as well as women who report diabetes during pregnancy (gestational diabetes). The intent for this dissertation is to include only incident cases of type 2 diabetes mellitus (and thus, henceforth in this document the generic term

"diabetes" will simply be used). However, a negligible number of diabetes of other types may exist in the analytic samples. These will not be discussed and are expected to have no impact on study findings.

3.1.2 *Pathogenesis*

Within the pancreatic exocrine tissues are the islets of Langerhans, which contain four distinct cell types that secrete different hormones: beta cells secrete insulin, alpha cells secrete glucagon, delta cells produce somatostatin, and PP cells produce pancreatic polypeptides. The key hormone that regulates blood glucose is insulin and the ratio of glucagon to insulin plays an important role in regulating metabolism; a balance between insulin secretion and efficiency in insulin action maintains normoglycemia in a non-diseased individual. Impaired insulin secretion is a consistent finding in type 2 diabetes patients.[9] In the fasting state, most glucose is produced in the liver; about half of the glucose is used for brain glucose metabolism and the remainder is utilized in other tissues (most notably the muscle and adipose tissue). When insulin levels are low, glucose uptake into the fat and muscle is also low and glucose production from the liver is high. After a meal, insulin is secreted in larger amounts, resulting in diminished liver glucose production in addition to enhancement of muscle and adipose tissue glucose uptake.

A normal pancreatic beta cell is adaptable to changes in insulin action. For example, as insulin action is decreased, insulin secretion increases to meet the demand imposed by the change in insulin action. When adaption of the beta cell are insufficient for a given degree of insulin sensitivity, impaired glucose tolerance or type 2 diabetes develops. Insulin resistance—where the muscle, adipose tissue, and liver cells do not use insulin properly—develops when the biological effects of insulin are abnormal for

glucose disposal in skeletal muscle as well as endogenous glucose production suppression. [10] When endogenous glucose production is accelerated, impaired fasting glucose and insulin resistance can be observed as the pancreas loses compensatory effectiveness. Hypertension in addition to hyperglycemia is common in diabetes patients; these two conditions eventually can lead to complications and decreased quality of life (particularly when disease management is poor). Other potential contributors to diabetes development include increased amounts of abdominal fat, infection, and genetic factors through mechanisms that are not yet fully understood.[11-13]

3.2 *Diagnosis of Diabetes*

A diagnosis of diabetes, as defined by the American Diabetes Association (ADA), requires one of the following four criteria to be met: fasting plasma glucose (FPG) ≥ 126 mg/l (7.0 mmol/L) after a minimum of eight hours with no caloric intake, symptoms of hyperglycemia and a casual (random) plasma glucose ≥ 200 mg/dL (11.1 mmol/L), two-hour plasma glucose ≥ 200 mg/dl (11.1 mmol/L) during a 75-gram oral glucose tolerance test (OGTT), or glycated hemoglobin A1c (HbA1c) ≥ 6.5% using a method that is National Glycohemoglobin Standardization Program (NGSP) certified and standardized to the Diabetes Control and Complications Trial (DCCT) assay.[14] The ADA also notes that if results are abnormal, a repeat test on the subsequent day should be performed to confirm the diagnosis.[14]

The FPG has been the preferred test for diagnosis due to convenience and reliability when measured in the true fasting state. For the same reason, this test is commonly used in epidemiological studies to characterize study populations and identify

undiagnosed diabetes cases.[11,15] The OGTT uses a standardized glucose load containing the equivalent of 75 g of anhydrous glucose that is dissolved in water, wherein a patient's plasma glucose is measured two hours following consumption. Other diabetes symptoms include polyuria (abnormally large urine volume), polydipsia (abnormally high level of thirst), unexplained weight loss, and blurred vision. HbA1c cut points were a recent addition to the diagnostic criteria and offers some important advantages over FPG and the OGTT (these will be discussed in further detail in a later section of this document).

The type of diabetes assigned at diagnosis often depends upon symptoms at presentation to the clinic.[14] Complicating this issue is the fact that many diabetes cases do not easily fall into a single class.[14]. For example, a woman with gestational diabetes might remain hyperglycemic after delivery and be later diagnosed with type 2 diabetes. Despite the assigned type, all diabetes cases share the common characteristic of blood glucose levels that are abnormally high.

3.3 *Management of Diabetes*

Management of diabetes involves a combination of monitoring (both by the patient and the provider), lifestyle modification, and use of medication in some patients. Additionally, control of cardiovascular risk factors and obesity are paramount in diabetic populations, since diabetes patients are at markedly high risk for cardiovascular morbidity and mortality.[3,5,16,17] Because diabetes is considered a progressive disease, careful and comprehensive management of cases has significant effects in minimizing and/or delaying complications. Details on management strategies and tools are summarized in the subsequent subsections.

3.3.1 Pharmacological Management of Diabetes

Figure 1 summarizes the general strategy for managing type 2 diabetes, as summarized by Nathan et al. [18] In brief, the approach includes a combination of lifestyle change along with pharmacological approaches. A major factor in selecting a class of drugs, or a specific medication within a class of drugs, is the ambient level of glycemic control.[18] When HbA1c levels are high, classes with high potency are typically chosen.[18] When HbA1c levels are near target levels, medications less potent medications may be selected.[18]

Table 1. Summary of glucose-lowering intervention recommendations.

Intervention	**Expected decrease in HbA1c with monotherapy (%)**	**Advantages**	**Disadvantage**
Tier 1: Well validated core			
Step 1: Initial therapy			
Lifestyle to decrease weight and increase activity	1.0-2.0	Broad benefits	Insufficient for most within first year
Metformin	1.0-2.0	Weight neutral	GI side effects, contraindicated with renal insufficiency
Step 2: Additional therapy			
Insulin	1.5-3.5	No dose limit, rapidly effective, improved lipid profile	One to four injections daily, monitoring, weight gain, hypoglycemia, analogues are expensive

Intervention	Expected decrease in HbA1c with monotherapy (%)	Advantages	Disadvantage
Sulfonylurea	1.0-2.0	Rapidly effective	Weight gain, hypoglycemia (especially with glibenclamide or chlorpropamide)
Tier 2: Less well validated			
TZDs	0.5-1.4	Improved lipid profile (pioglitazone), potential decrease in MI (pioglitazone)	Fluid retention, CHF, weight gain, bone fractures, expensive, potential increase in MI (rosiglitazone)
GLP-1 agonist	0.5-1.0	Weight loss	1-2 injections daily, frequent GI side effects, long-term safety not established, expensive
Other therapy			
α-Glucosidase inhibitor	0.5-0.8	Weight neutral	Frequent GI side effects, three times/day dosing, expensive
Glinide	0.5-1.5[a]	Rapidly effective	Weight gain, three times/day dosing, hypoglycemia, expensive
Pramlintide	0.5-1.0	Weight loss	Three injections daily, frequent GI side effects, long-tem safety not established, expensive

Intervention	Expected decrease in HbA1c with monotherapy (%)	Advantages	Disadvantage
DPP-4 inhibitor	0.5-0.8	Weight neutral	Long-term safety not established, expensive

[a]Repaglinide more effective in lowering A1c than nateglinide; CHF, congestive heart failure; GI, gastro intestinal; MI, myocardial infarction.
Adapted from: Nathan, et al., 2009. [18]

While the above guidelines provide a common starting point for initiation of glucose-lowering interventions, ultimately the treatment approach needs to be individualized to the patient. Choice of treatment is also highly dependent upon side effects, cost, and patient compliance. Achievement and maintenance of normal glycemic levels is the primary goal of pharmacological approaches. Early addition of insulin therapy should also be considered for patients who do not meet target goals.[18] Finally, control of blood pressure and lipid levels reduces the risk for micro- and macrovascular complications that disproportionately affect this population.[19]

3.3.2 *Lifestyle Modification*

Lifestyle modification plays a crucial role in the management of diabetes. While debate continues about the specific type of diet and exercise that is most beneficial for those with diabetes, weight loss is clearly indicated in those who are overweight.[18,20] Decreased energy intake almost always improves glycemic levels while concurrently improving CVD risk factors. [21] Hence, weight reduction is the focus of many behavioral

interventions in diabetes patients. The primarily challenge with this approach, unsurprisingly, is that long-term maintenance of weight loss is difficult to achieve.[19]

The composition of the diet is important for diabetes management, yet specific recommendations are less clear. Diabetes patients are advised to pay particular attention to the type and amount of carbohydrates consumed, due to the effects on postprandial glucose responses; high-fiber diets are generally recommended, as well as close monitoring of blood glucose levels after meals.[22,23] Regulation of glucose metabolism is impaired in diabetes patients. Particularly in the postprandial period, patients are more susceptible to the influence of diet on plasma glucose compared to the non-diabetic population that can better accommodate a variety of types and amounts of certain foods.[23] Other dietary recommendations include limiting fat intake (particularly saturated and trans fats) and alcohol consumption.[19,23]

Moderate to vigorous-intensity exercise facilitates weight loss and management, and also improves insulin sensitivity independent of weight loss.[21,22,24] The benefits of regular exercise in managing hyperglycemia are apparent not only among diabetes patients, but also among those with pre-diabetes.[25,26] In a review of controlled trials of individuals with IGT, increases in physical activity were associated with weight loss, lower incidence of diabetes, and decreases in plasma glucose levels.[26-28] Other intervention studies have demonstrated decreases in HbA1c levels in response to a variety of exercise interventions.[29]

3.3.3 *Hemoglobin A1c*

Close monitoring of two measures, glucose and HbA1c, primarily assesses the efficacy of diabetes therapy. The development of the HbA1c assay provided an important

tool for clinicians in the monitoring of patients with diabetes. Found in red blood cells formed when glucose attaches to hemoglobin, HbA1c was first separated from other forms of hemoglobin (HbA1a and HbA1b) using a chromatographic column.[30] HbA1c is typically monitored in diabetic patients to assess glycemic control.[31,32] Levels of HbA1c represent the average level of blood glucose to which a red blood cell has been exposed during its lifetime, which is approximately 120 days.[6,33]

Clinical guidelines state that the goal of diabetes therapy for most patients is to maintain HbA1c levels at less than 7%.[34] In individuals with poorly controlled diabetes, measured HbA1c levels tend to be high. Even in the normal range, CVD risk rises as a function of HbA1c.[3,35] In the Atherosclerosis Risk in Communities (ARIC) study, a positive and graded association was found between HbA1c level and peripheral arterial disease (PAD) risk (independent of known risk factors for PAD).[36] It should be noted, however, that the principle of "the lower, the better" does not universally hold for all diabetes patients,[17,37,38] and therefore aggressive and continuous lowering of HbA1c is not always warranted. Consequences of overly aggressive lowering of HbA1c levels include hypoglycemia, a dangerous condition that highlights the importance of individualizing goals for optimal levels.

The recent establishment of guidelines for the use of HbA1c in the diagnosis of diabetes and as a measure of increased risk among non-diabetic individuals offers some advantages over fasting glucose.[35,39,40] First, measurement of HbA1c does not require patient fasting for accurate measurement. Secondly, HbA1c reflects longer-term glycemia compared to fasting glucose measures and can be thought of as an integrated measure of hyperglycemia. Reliability and standardization of measurement of HbA1c has improved

greatly in recent years, and measurement error due to non-glycemic factors are infrequent.[40] Further, HbA1c correlates well with undiagnosed diabetes in non-treated populations and has lower intraindividual variability compared to measures such as fasting glucose.[39] Thus, efforts to establish standards for HbA1c levels in initially non-diabetic populations are well justified.

While the concept of utilizing measurement of HbA1c as a screening tool is not new, challenges in standardizing the HbA1c assay has limited its use.[41] Further, factors that affect individual variation in HbA1c levels and the precise meaning of inter-individual variations in HbA1c (both within the "normal" range and even in diabetics) are incomplete. For example, evidence suggests that observed ethnic differences in mean HbA1c levels may reflect underlying population differences in diabetes risk or disparities in the efficacy of diabetes therapy between populations.[42,43] A recent meta-analysis evaluating differences in HbA1c levels between Hispanic and non-Hispanic White adults found that unadjusted HbA1c values were approximately one-half percent higher in Hispanic patients with diabetes than in non-Hispanic patients.[42] Despite this preliminary evidence, the data thus far appear inconclusive; it may be the case that a given HbA1c level might confer a different degree of risk between individuals in a multi-ethnic population, or it may reflect differences in the quality of care amongst population subgroups. Further research should clarify the utility of HbA1c as a screening tool.

3.3.4 *Self-monitoring in diabetes management*

Self-monitoring of blood glucose is often recommended and advised by clinicians for a comprehensive diabetes management program. For one, the measure is quick and provides immediate feedback. With a small finger prick and minimal amount of blood,

patients can measure blood glucose at any time.[44] This practice provides immediate feedback about blood glucose responses to foods eaten or other behaviors, and can facilitate self-assessment of the effectiveness of medications. However, the utility of self-monitoring should be weighed against issues including poor compliance, cost, evidence that suggests only a minimal impact of self-monitoring on glucose control, and the current lack of strong evidence showing that self-monitoring actually improves HbA1c levels. [44,45]

3.4 *Pre-diabetes: Pre-disease state or distinct clinical entity?*

Because the population burden of diabetes is so great, prevention efforts are an important public health priority. Identification of high-risk individuals is paramount for efficient targeting of lifestyle interventions and is thus the basis for the rationale behind the establishment of the distinct entity, pre-diabetes. Pre-diabetes is considered a state of abnormal glucose homeostasis in individuals without diabetes, where a deficiency or resistance to insulin is apparent.[46,47] The benefit of creating the pre-diabetes entity is that progression to diabetes can be prevented or delayed if individuals are identified early enough; once disease develops, the goal becomes management and is considered irreversible.[48]

Pre-diabetes, until recently, was considered present when fasting plasma glucose levels are ≥ 100 mg/L and < 126 mg/dL, or levels are ≥ 140 mg/dL but < 200 mg/dl following an oral glucose tolerance test (OGTT).[6] Elevated fasting glucose is termed "impaired fasting glucose" (IFG) and elevated post-OGTT glucose is termed "impaired glucose tolerance" (IGT). In 2010, cut points for HbA1c (5.7-6.4%) were also established

as an alternative tool for identifying individuals at high risk.[14] From the data thus far, this range appears to be less sensitive than fasting glucose but more specific, and with a higher positive predictive value for later development of diabetes.[14]

In 2003, the American Diabetes Association decreased the lower cut point for impaired fasting glucose from 110 to 100 mg/dL to increase sensitivity in detecting individuals at increased risk for developing diabetes, albeit reducing specificity.[33] In this study, the maximization of sensitivity and specificity occurred at 103 mg/dL.[33] A 2006 report utilizing data from the Screening for Impaired Glucose Tolerance Study 1 (SIGT 1) compared the current cut point of 100 mg/dL for IFG (IFG 100) to the previous 110 mg/dL cut point (IFG 110).[49] While lowering the cut point decreased specificity for detecting risk of developing diabetes, sensitivity was improved for predicting a variety of outcomes. For example, IFG 110 carried an increased risk of the metabolic syndrome [odds ratio 7.10 (95% CI: 4.39–11.46) vs. 10.33 (95% CI: 4.87–21.88)].[49] Other outcomes examined included LDL cholesterol > 130 mg/dl, high urine albumin-to-creatinine ratio, and higher likelihood of post-challenge hyperglycemia.[49] A retrospective cohort study of members of Kaiser Permanente Northwest followed 5,452 individuals who met the criteria for IFG (fasting glucose 100–125 mg/dL) and with no prior history of diabetes.[50] Over a mean follow-up of 6.3 years, 8.1% of the cohort progressed to diabetes, at an annual rate of 1.34%.

While pre-diabetes cut points were developed and are used for the purpose of identifying pre-diseased, high-risk individuals, up to 25% of individuals with new diagnosis of diabetes actually present with established diabetic retinopathy or microalbuminuria.[51,52] Findings of this nature suggest an apparent gap between actual

disease onset and diagnosis. This is suggestive of the presence of undiagnosed overt diabetes cases, though one could also argue that pre-diabetes might be more than simply a marker of risk.[53] Recent studies, such as one described by Hu and colleagues utilizing data from the Nurses' Health Study,[54] have also observed substantially increased risk of myocardial infarction and stroke prior to clinical diagnosis of type 2 diabetes among women. Both macrovascular and microvascular complications have been observed in pre-diabetes, highlighting the importance of early identification of high-risk persons (regardless of whether pre-diabetes should be considered a clinical entity).[55,56]

Some research suggests that individuals who present with only IFG (isolated IFG) or only IGT (isolated IGT) may have different aberrations in glucose metabolism.[47,57-59] For instance, blood glucose levels should be lower several hours after eating. Blood glucose levels should rise after drinking a sugar solution, but should still be below 140 mg/dL; higher than normal levels of this measure indicates an abnormality in utilizing glucose or mounting an adequate insulin response. On the other hand, high levels of fasting glucose suggest insulin resistance. IFG may more reflect impaired early-late phase insulin secretion, while IGT might more reflect impaired late-phase insulin secretion.[57] Therefore, effective prevention strategies for thwarting progression to diabetes might differ among individuals with isolated IFG compared to those with isolated IGT (though these conditions are not commonly mutually exclusive).

3.5 *Diabetes and Cardiovascular Disease*

The complications associated with diabetes are primarily microvascular and macrovascular diseases that result from accelerated atherogenesis.[10] Type 2 diabetes

patients die of macrovascular CVD at rates that are two to four times higher than patients of similar demographic characteristics without diabetes.[17,60-64] With such a large proportion of patients developing cardiovascular complications, close CVD risk factor management is crucial in patients with diabetes.

3.5.1 *Microvascular Complications*

Microvascular complications of diabetes include retinopathy, nephropathy, and peripheral or autonomic neuropathy.[65] The structure and function of the eye tends to be severely affected by diabetes.[65] The primary microvascular complication is diabetic retinopathy, which is the leading cause of blindness in adults aged 20-64 years in the United States.[66] In fact, up to four million Americans aged 40 years and older with diabetes also have retinopathy, with nearly one million of those having sight-threatening retinopathy.[67] Also of interest is that new studies in large populations show retinal vascular changes even in pre-diabetes cases prior to overt diabetes.[55]

Microvascular complications of diabetes are primarily due to glucose, and the effect of hyperglycemia on the vasculature. Hyperglycemia leads to high glucose flux across endothelial cell membranes, which overwhelms the mitochondrial electron transport system (leading to the release of reactive oxygen species).[68] Increase oxidative stress reduces the bioavailability of nitric oxide by converting available nitric oxide into peroxynitrate, eventually resulting in endothelial dysfunction.[69]

3.5.2 *Macrovascular Complications*

Heart disease was noted in 68% of diabetes-related death certificates in 2004 among adults aged 65 and older; stroke was noted in 16% of diabetes-related death certificates.[4] Longitudinal data from the Framingham Study found that the risk factor-

adjusted relative risk of coronary heart disease (CHD) was 1.38 (95% CI: 0.99-1.92) times higher and the risk for CHD death was 1.86 times greater (95% CI: 1.17-2.93) for each 10-year increase in duration of diabetes.[70] Diabetes also increases risk of stroke with a relative risk ranging from 1.8 to close to 6.0 compared to non-diabetic patients.[71] Insulin resistance is believed to be the primary cause of atherosclerotic damage in diabetics; according to this hypothesis, hyperinsulinemia leads to increased atherogenesis both directly and indirectly through effects of insulin on blood lipids and blood pressure.[72]

3.6 *Epidemiology and Public Health Burden of Diabetes*

Nearly ten percent of Americans aged twenty years and older live with diabetes mellitus, a number that has increased in recent years.[1] In 2000, the prevalence of diabetes mellitus worldwide was estimated at 2.8% (171 million people); this number is expected to rise to approximately 4.4% (366 million people) by 2030 if recent trends continue.[73] Based on 1999-2002 NHANES data, the prevalence of diagnosed/undetected diabetes was 15.3% in adults aged ≥ 65 years, and the prevalence of undiagnosed diabetes was 6.9%.[74] Additionally, one and a half million new cases of diabetes were diagnosed in 2005 among adults ages twenty years and older; in 2005-2006, the prevalence of diagnosed diabetes was 7.7% in individuals aged 20 years and older.[75] In adults, type 2 diabetes accounts for an estimated ninety to ninety-five percent of all diagnosed cases of diabetes.[4]

Diabetes is an important contributor to the global burden of chronic disease, most notably cardiovascular morbidity and mortality.[2,76] According to data from the Global

Burden of Disease 2000 study, the estimated excess global mortality attributable to diabetes in the year 2000 was 2.9 million deaths.[77] This number translates to approximately 5.2% of all deaths worldwide, and ranges from 2-3% in the poorest countries and is over 8% in the United States, Canada, and the Middle East. Further, in 2007 the direct cost attributable to diabetes totaled $116 billion while the indirect costs summed to $58 billion.[4] Overall, diabetes prevalence is higher in men, but there is a larger absolute number of women worldwide with diabetes.[78]

3.6.1 *Race/Ethnicity and Diabetes*

At the population level, there are marked disparities in the impact of diabetes between racial and ethnic groups. For example, NHANES data observed a disproportionately high prevalence of diabetes in non-Hispanic black individuals and Mexican Americans compared with non-Hispanic whites.[79]. While genetic heterogeneity may account for a proportion of population differences in risk for diabetes, at least part of those population differences is likely due to variations in lifestyle behaviors, access to health services, and environmental characteristics.

Ethnicity and race are continuously evolving concepts, and self-identified categories are often used to group individuals in research studies that are believed to share common genetic factors that influence disease risk through some shared etiology.[80] While the term "race" generally refers to ancestral background, "ethnicity" alludes more to cultural and social identity. This distinction is important to recognize in the interpretation of studies of race/ethnic differences. Most studies classify individuals into groups based on self-reported race/ethnic group. The assigned category may be based on

questions referring to genetic ancestry or the individual may be asked to pick a particular group to which he/she most strongly identifies.

The use of race as an etiologic quantity has been criticized in medical research.[81,82] However, clear racial/ethnic differences in the burden of disease are repeatedly observed. Therefore, we should not ignore this observation but rather use caution in interpretation. Additionally, it is important to clearly communicate how race and ethnicity are defined for a particular study. Finally, more effort should be placed in understanding and investigating the nature of non-genetic factors that contribute to racial/ethnic differences in the burden of diabetes and of disease states more broadly.

3.7 *Epidemiology and Public Health Burden of Pre-Diabetes*

Seven percent of U.S. adolescents aged twelve to nineteen years had impaired fasting glucose in 1999 to 2000;[4] almost one quarter of U.S. adults are known to have pre-diabetes.[25] Additionally, analyses of data from 1988-1994 in the Third National Health and Nutrition Examination Survey found that 17.1% of overweight adults aged 45-74 years had IGT, 11.9% had IFG, 22.6% had pre-diabetes (IGT or IFG), and 5.6% had both IGT and IFG.[79] Using NHANES 2005-2006 data and examining participants aged 20 years and older, Cowie et al. reported that 25.7% of participants had IFG, 13.8% had IGT, and the crude prevalence of IFG or IGT (i.e., pre-diabetes) was 29.5%.[75] Prevalence was also significantly higher in men compared to women (crude 36% versus 23.5% and age-standardized 35.7% versus 22.8%; both $P = 0.0002$).[75]

As mentioned previously, a notable benefit of identifying individuals at high risk prior to disease diagnosis is that non-pharmacological treatments can delay or prevent the

need for medication; these intervention strategies also tend to be more cost effective in the long-term (with residual benefits on CVD risk).[4] In a cost-effectiveness analysis based on the Diabetes Prevention Program (DPP), for example, investigators found that screening for pre-diabetes as well as treating those identified as having both IGT and IFG with the DPP lifestyle intervention (weight loss and physical activity) had a cost effectiveness ratio of $8,181 per quality-adjusted life year (QALY) relative to no screening. Treating those with either IGT *or* IFG increased the cost-effectiveness ratio to $9,511 per QALY.

3.8 *Diabetes Risk Factors*

3.8.1 *Overview*

The most notable risk factors for diabetes are listed in Table 2, and may be broadly categorized as modifiable or non-modifiable. However, it should be noted that a single factor could have both modifiable and non-modifiable components. In addition, these factors may interact to jointly influence disease risk. Particularly in the case of complex diseases such as diabetes that do not have a single cause, disease risk is likely caused by a combination of characteristics of the individuals, behaviors, and the environment in which the behaviors are performed. Several studies have used risk scores to predict incident diabetes,[83-88] but due to the multifaceted nature of diabetes—and given the rising incidence of diabetes and high prevalence of undiagnosed cases—more work is clearly still needed in this area.

Table 2. Common known risk factors for diabetes.

Modifiable	**Non-modifiable**
Obesity/overweight	Age
Smoking	Sex
Low dietary fiber	Genetic factors
High saturated fat in the diet	Family history
Physical inactivity	Prior gestational diabetes

Several risk prediction models have been proposed to identify individuals at high risk for diabetes and/or to screen for undiagnosed cases of diabetes. Listed below are components of some of the most cited published prediction models and/or diabetes risk scores.

Table 3. Components of diabetes risk prediction scores from a review of the literature.

Variable Associated With Increased Risk	**Reference**
Black (compared to White) race	[84]
Chinese (compared to White participants from England and Wales) ancestry	[84]
Current and/or former smoker status (versus never smoker)	[84,87,89]
High or no alcohol consumption	[87]
Higher 2-Hour OGTT finding	[88]
Higher blood pressure	[88,90]
Higher consumption of red meat	[87]
Higher levels of fasting glucose	[88,90]
Higher resting heart rate/pulse	[89]
Higher triglycerides	[88,90]
Larger waist circumference	[87-89,91,92]
Less years of formal education	[93,94]
Low consumption of fruits, berries, or vegetables	[92]
Lower consumption of coffee	[87]
Lower consumption of whole-grain bread	[87]
Lower HDL cholesterol	[88,90]
Lower physical activity levels	[83,87,91,92]
Male sex	[83,84,88]
Non-Caucasian race	[91]
Obesity	[83,84,88,90,92]

Variable Associated With Increased Risk	Reference
Older age	[83,84,87-89,91,92]
Positive family history of diabetes	[83,84,88,89,91,95]
Positive history of antihypertensive drug treatment	[92]
Positive history of drug treatment for high blood glucose	[92]
Positive history of gestational diabetes	[91]
Positive history of hypertension	[83,87-89,91]
Shorter height/stature	[89,91]

3.8.2 *Non-Modifiable Risk Factors*

Non-modifiable risk factors for diabetes include age, sex, family history for diabetes, prior gestational diabetes, and genetic predisposition. These factors are considered unchangeable and thus comprise the underlying risk for disease for a particular individual. As mentioned previously, sex differences in the burden of diabetes have been observed in some (but not all) population-based studies.[78] Additionally, the incidence of diabetes tends to increase with increasing age and women with history of gestational diabetes are significantly higher risk for type 2 diabetes.[1,74,78,79]

In an analysis conducted in the ARIC cohort that observed 1,302 incident diabetes cases during a median follow-up of 9 years, a higher CHD family risk score was predictive of incident type 2 diabetes among participants with positive family history of diabetes.[95] While characteristics such as age, gestational diabetes history, and family history of diabetes are clearly trait characteristics, this does not imply that these factors themselves are direct causes of disease development. This can be most clearly seen in the genetics literature, where effects of genetic factors present at birth are modified by behaviors and the individual's environment.

Genetic factors have been implicated in predisposing individuals to developing diabetes, either directly, through intermediate factors (e.g., weight gain and obesity), or via interactions between multiple genes as well as between genes and environmental/behavioral factors.[96-99] In one of the dissertation manuscripts, the influence of three genes (the adiponectin gene, the fat mass and obesity associated gene, and the peroxisome proliferator-activated receptor gamma gene) on diabetes risk are explored and thus are described in further detail in the subsections that follow.

3.8.2.1 Adiponectin (*ADIPOQ*) Gene

The *ADIPOQ* gene consists of three exons and two introns that span a 17-kb region and has been located on chromosome 3q27.[100,101] *ADIPOQ* encodes the circulating protein adiponectin; low adiponectin levels are associated with central adiposity, overall obesity, insulin resistance, and type 2 diabetes.[102-104] For example, in a Japanese population of 480 participants without diabetes and 384 participants with diabetes, participants with G/G genotype at position 45 or the G/G genotype at position 276 had a significantly increased risk of type 2 diabetes (odds ratio 1.70 [94% CI: 1.09 - 2.65] and 2.17 [1.22 - 3.95], respectively).[102] Expressed exclusively in the adipose tissue, *ADIPOQ* functions in the homeostatic control of glucose, lipid, and energy metabolism.[105] To our knowledge, there are no published data examining interactions between *ADIPOQ* SNPs and cardiorespiratory fitness (the evaluation of which was conducted as part of this dissertation).

3.8.2.2 Fat Mass and Obesity Associated (*FTO*) Gene

The *FTO* gene is a protein-coding gene located on chromosome 16q12.2. Variants of the *FTO* gene have been associated with obesity-related traits that include: body mass

index, body weight, waist circumference, and body fat percentage.[106-109] In a recent study conducted by Rampersaud and colleagues in a community-based population of men and women of European origin, the association between *FTO* genotype on body composition — most notably BMI — was attenuated and not statistically significant in participants with high levels of physical activity (measured by accelerometry).[106] Findings such as these support the hypothesis that physiological phenomena are associated with increased levels of physical activity or fitness might act upon variants of the *FTO* gene.

Additional studies have investigated the *FTO* gene in direct relation to diabetes.[110-113] One study of 17,037 White Caucasian men and women reported that each copy of the *FTO* rs9939609 A allele was associated with both higher fasting insulin (0.039 SD [94% CI: 0.013 - 0.062]; $P = 0.003$) and glucose (0.024 SD [0.001 - 0.048]; $P = 0.044$).[111] Additionally, a study in 6,719 Asians (Chinese and Koreans) observed that the A allele of rs8050136 of *FTO* was associated with risk of type 2 diabetes (odds ratio = 1.13 [95% CI: 1.02-1.25]; $P = 0.016$).[113] Finally, the A allele of the *FTO* rs9939609 polymorphism was associated with risk of both obesity and type 2 diabetes in the French MONICA Study.[112] In this study, the AA genotype of rs9939609 was associated with a 1.45 times higher odds of type 2 diabetes compared to participants with the TT genotype (CI: 1.05-1.99, $p = .02$).[112]

Peroxisome Proliferator-Activated Receptor Gamma (*PPARG*) Gene

PPARG is a member of the proliferator-activated receptor (PPAR) subfamily of nuclear receptors and is located on chromosome 3p25. *PPARG* mRNA is primarily expressed in white adipocytes, the placenta, and macrophages.[96] The protein encoded by *PPARG* has been associated with obesity, diabetes, and atherosclerosis.[114] *PPARG* is

believed to regulate adipocyte differentiation and insulin sensitivity by promoting transcription of numerous target genes.[115]

A number of *PPARG* variants have been studied in relation to diabetes and diabetes risk factors, but most epidemiologic studies have focused on the Pro12Ala variant of rs1801282 (particularly in the context of gene-environment interactions).[96,116-118] For example, the alanine variant of Pro12Ala has been associated with enhanced insulin sensitivity that functions to protect against type 2 diabetes.[97] In Caucasian populations, the high-risk proline allele of the Pro12Ala polymorphism is present in approximately 75% of the population.[10] In a meta-analysis with an additive genetic model, the type 2 diabetes odds ratio (OR) for the Pro12Ala polymorphism across 66 studies was 0.85 using a fixed-effects model and 0.86 using a random-effects model (comparing alanine to proline).[116] These ORs were statistically significant, but among the 66 studies included in this meta-analysis, a moderate degree of inconsistency was observed.[116]

The influence of *PPARG* on diabetes might in part be due to interactions with health behaviors or environmental factors such as exercise.[118-121] One study of 262 healthy Japanese men found that the Pro12Ala polymorphism was a reliable indicator of whether a three-month exercise intervention improved insulin resistance.[119] Participants with the Ala allele improved more in fasting insulin levels compared to those without it (mean ± SD was -2.83 ± 1.47 μU/ml versus 0.55 ± 3.49 μU/ml, $p = 0.02$).[119] Additionally, homeostasis model of insulin resistance (HOMA-IR) levels improved more in the Ala allele group (-0.61 ± 0.32 versus 0.09 ± 0.86, $p = 0.05$).[119] However, the frequency of the Pro12Ala allele in the exercise group was only 0.024 (a total of only 6

participants).[119] A second study including 139 Caucasian type 2 diabetic patients (age 54.4 ± 7.2 years) examined the effects of a three-month supervised exercise program on changes in blood glucose. The frequency of the Ala allele in this population was higher, at 8.3% of the study population. The change in fasting plasma glucose was significantly different between *PPARG* genotypes (-1.66 mmol/l versus -0.54 mmol/l, for Ala carriers and Pro/Pro homozygotes, respectively; p=0.0034 unadjusted and p=0.089 adjusted for baseline glucose).[118]

3.8.3 *Modifiable Risk Factors*

The most cited modifiable risk factors for diabetes include smoking, obesity, dietary factors, and physical inactivity. Other modifiable factors also likely influence diabetes risk and include psychosocial characteristics and environmental factors (e.g., accessibility to health services), though these factors have received receive less attention in the literature.

3.8.3.1 Smoking

A meta-analysis of 25 prospective cohort studies containing 1.2 million participants and 48,844 incident diabetes cases examined the effect of active smoking on risk of type 2 diabetes.[122] Active smoking was associated with an increased risk of type 2 diabetes in all subgroups of the populations included. The relative risk of diabetes among heavy smokers (≥ 20 cigarettes per day) compared to never smokers was 1.61 (95% CI: 1.43-1.80); for light smokers was 1.29 (95% CI: 1.13-1.48); and for former smokers was 1.23 (95% CI: 1.14-1.33).[122]

3.8.3.2 Obesity, Diet, and Physical Activity/Physical Fitness

Diet and physical activity are the two factors most often discussed with regards to diabetes prevention.[21] Each has independent effects on disease development, and jointly these two factors contribute to obesity. Weight control is indisputably linked both to diabetes risk and successful diabetes management efforts.[24,48] Energy intake is the dietary factor most consistently associated with risk of diabetes, but types of carbohydrates, fats, and fiber have also been implicated. Interventions aimed at reducing dietary intake (sometimes combined with increasing physical activity) show favorable effects on metabolic parameters in both non-diabetics and diabetics.[21,123,124]

Dietary carbohydrates increase blood glucose concentrations, particularly in the period immediately following a meal.[23] In observational studies, low glycemic-index and high fiber foods appear to lower the risk of incident diabetes, though further research is needed to clarify the effect of specific foods (as well as the net effect of combinations of foods consumed concurrently) on blood glucose metabolism.[23,125] One study in the Coronary Artery Risk Development in Young Adults (CARDIA) cohort observed that, for example, dairy consumption was inversely associated with the incidence of insulin resistance syndrome among individuals who were overweight (body mass index ≥ 25 kg/m^2).[126]

The literature provides strong evidence of an association between physical activity and diabetes risk, independent of dietary or weight loss changes in people with pre-diabetes.[26] Regular physical activity both directly and indirectly increases muscle glucose uptake and insulin sensitivity, and therefore plays an important role in the

prevention and treatment of metabolic disorders.[29,127] While most intervention studies focus on aerobic exercise, other forms of physical activity are also likely important.

In this document, physical fitness will be considered as defined by Sigal et al. [29] Broadly, "physical fitness" includes cardiorespiratory fitness, muscular fitness, and flexibility. "Physical activity" will refer to physical bodily movement that is produced by the contraction of skeletal muscle that requires energy expenditure in excess of resting energy expenditure. "Exercise" will refer to a subset of physical activity, defined as a planned, structured, and repetitive bodily movement specifically performed to improve or maintain one or more components of physical fitness. The primary characteristic of interest for manuscript #3 is cardiorespiratory fitness, referring to the degree to which the circulatory and respiratory system can supply oxygen during sustained activity. The gold standard for measurement of cardiorespiratory fitness is a test of maximal oxygen uptake (VO_2 max), typically using indirect calorimetry to measure oxygen consumption and carbon dioxide release. Genetically influenced and modifiable with physical training, cardiorespiratory fitness plays an important role in glucoregulation and is associated with risk for type 2 diabetes.[28,29,128,129]

Exercise and dietary modification have been considered key components of diabetes prevention and management (particularly among pre-diabetics), both due to direct effects on blood glucose control and indirectly by facilitating weight management.[20,26,29,123,129-131] Of primary interest in this project are modifiable risk factors that can be changed, ideally though lifestyle change rather than with the use of cost-prohibitive medications. The strongest modifiable diabetes risk factors that have been identified to date include obesity, smoking, diet, and physical activity. However, the

specific attributes of these factors, as well as the exact mechanisms through which each affects the pathogenesis of diabetes, are not fully understood. Of particular interest is to better understand modifiable risk factors, and to improve our ability to identify specific aspects of those factors that have the strongest relative impact on disease risk.

4. OVERVIEW OF DISSERTATION AIMS

While a large body of work exists regarding the causes and consequences of diabetes, several areas—particularly in the domain of early identification of high risk individuals and modifiable risk factors—still require further investigation. The completed manuscripts address literature gaps in defining high-risk in diverse populations and identifying modifiable factors that might help to prevent or delay diabetes onset. The Multi-Ethnic Study of Atherosclerosis (MESA) is an ethnically diverse cohort examined in the first manuscript to study racial/ethnic differences in prediction of risk for diabetes, using fasting glucose and HbA1c measures. A second study, the Coronary Artery Disease in Young Adults (CARDIA) study, enrolled healthy young adults free of cardiovascular disease. The study provides 20 years of follow-up information on demographics, health behaviors, and other cardiovascular risk factors and health-related variables. The CARDIA study allowed for examination of predictors of incident diabetes in a high-risk population (individuals with pre-diabetes). Finally, cardiorespiratory fitness level data—measured in the core CARDIA study at baseline and follow-up exam year 7, as well as in the ancillary CARDIA Fitness Study (CFS) at year 20—allowed for examination of the role of fitness in observed associations between single nucleotide polymorphisms of three candidate genes and diabetes-related outcomes.

5. STUDY DESIGN AND OBJECTIVES

5.1 *Multi-Ethnic Study of Atherosclerosis (MESA) Study*

5.1.1 *MESA: Overview, Study Design, and Study Population*

MESA is a multi-center, prospective cohort study funded by the National Heart, Lung, and Blood Institute. The study aims to further our understanding of the pathophysiology of subclinical cardiovascular disease development and progression as well as the role of subclinical disease in clinical cardiovascular disease. The primary objectives of MESA are to identify characteristics related to subclinical cardiovascular disease and progression to clinical cardiovascular disease across race/ethnic populations.

The study enrolled 6,814 asymptomatic men and women aged 45 to 84 years between July of 2000 and July of 2002 from six field centers across the United States (Baltimore City and Baltimore County, MD; Chicago, IL; Forsyth County, NC; Los Angeles County, CA; Northern Manhattan and the Bronx, NY; and St. Paul, MN). The source population varied in size and ethnic composition by field center. Participants were recruited from four self-reported race/ethnic groups: African Americans, Caucasians, Chinese Americans, and Hispanics. Three follow-up examinations have been conducted to date: exam 2 occurred two years after baseline, exam 3 began in January of 2004 and continued for 18 months, and exam 4 started in July 2005 and was complete two years later. Periodic follow-up via phone and mail occurred every 6-12 months to maintain contact, update contact information, and ascertain information about medical events. The research conducted for this project included 5,006 MESA participants (73% of the

original cohort) who were free of diabetes at exam 1 and who completed exam 4 (approximately 5-7 years after baseline).

At baseline, exclusion criteria were: age outside of the target range, physician-diagnosed myocardial infarction, angina, nitroglycerin use, stroke, transient ischemic attack, or heart failure; atrial fibrillation at time of recruitment; history of coronary artery bypass graft, angioplasty, valve replacement, pacemaker or defibrillator, or any surgery on the heart or arteries; active treatment for cancer; current pregnancy; weight over 300 pounds; serious medical condition(s) that would inhibit long-term participation; cognitive inability to participate; living in or on the waiting list to live in a nursing home; self-reported plans to leave the community within five years; language barrier; and chest CT scan in the past year. Participants were recruited via telephone or in person; recruitment strategies varied by field center based on available resources, past experience, site-specific logistics, and the characteristics of the community. Approximately 390 study participants were obtained through referrals of other consented individuals. Among those households contacted, study staff explained the study to 29.0%; 39.8% of these individuals agreed to participate. The overall participation rate was 59.8% among those screened and who met the eligibility criteria.

5.1.2 *MESA: Summary of Relevant Study Components*

5.1.2.1 Fasting Glucose

Fasting glucose was measured at each study exam. Blood for this assay was collected in a serum tube, allowed to clot for at least forty minutes, centrifuged, and stored for later processing by rate reflectance spectrophotometry on the Vitros analyzer (Johnson & Johnson Clinical Diagnostics, Inc., Rochester, NY); the laboratory coefficient

of variation is 1.1%. Fasting glucose measures are currently available for all four exams. However, the September 2006 MESA Steering Committee report observed unexpected trends in fasting glucose levels and classification of diabetes status across the first three exams that suggested a drift in assay results over time.

After an observed drift in the assay over time, MESA investigators re-assayed 800 fasting glucose samples (200 from each of the four exams). Subsequently, observed levels from Exam 2, Exam 3, and Exam 4 were recalibrated to the Exam 1 scale [for Exam 2 glucose values, estimated value = -1.261 + (original*0.964); for exam 3 glucose values, estimated value = 1.003 + (original*0.963); and for Exam 4 glucose values, estimated value = -5.163 + (original*1.009)]. From the estimated values, calibrated values were obtained using the following formula: calibrated value = 6.453 + estimated value * 1.005. A summary of the effect of calibration on diabetes status for all exams is illustrated in Tables 1 and 2 (unpublished data obtained from the MESA Study). Recalibrated values were used for all analyses.

Table 1. Frequency and percentage of diabetes status by MESA exam (uncalibrated).

	Exam 1		**Exam 2**		**Exam 3**		**Exam 4***	
	Frequency	Percent	Frequency	Percent	Frequency	Percent	Frequency	Percent
Normal	2,890	59.9	2,766	57.3	3,069	63.6	2,530	52.4
Impaired Fasting Glucose	1,323	27.4	1,386	28.7	1,065	22.1	1,552	32.1
Untreated Diabetes	196	4.1	207	4.3	148	3.1	200	4.1
Treated Diabetes	419	8.7	469	9.7	546	11.3	546	11.3
Total	4,828	100.0	4,828	100.0	4,828	100.0	4,828	100.0

*The medication use information for Exam 4 is from Exam 3 since the information is not yet available for Exam 4.

Table 2. Frequency and percentage of diabetes status by MESA exam (recalibrated).

	Exam 1		Exam 2		Exam 3		Exam 4*	
	Frequency	Percent	Frequency	Percent	Frequency	Percent	Frequency	Percent
Normal	2,890	59.9	2,428	50.3	2,520	52.2	1,986	41.1
Impaired Fasting Glucose	1,323	27.4	1,713	35.5	1,568	32.5	2,044	42.3
Untreated Diabetes	196	4.1	218	4.5	194	4.0	252	5.2
Treated Diabetes	419	8.7	469	9.7	546	11.3	546	11.3
Total	4,828	100.0	4,828	100.0	4,828	100.0	4,828	100.0

***The medication use information for Exam 4 is from Exam 3 since the information is not yet available for Exam 4.

5.1.2.2 Hemoglobin A1c

Hemoglobin A1c was measured at Exam 2 using a sample preparation vial containing an aqueous solution of EDTA and potassium cyanide (HbA1c Sample Preparation Kit, Bio-Rad, Hercules, CA 84547). Specimens were analyzed on the Tosoh A1c 2.2 Plus Glucohemoglobin Analyzer (Tosoh Medics, Inc., San Francisco, CA 94080) using an automated high performance liquid chromatography method. The reference range was 4.3 - 6.0% with a laboratory CV range of 1.4 - 1.9%. Measurements were made at the Collaborative Studies Clinical Laboratory at Fairview-University Medical Center (Minneapolis, MN).

5.1.2.3 Ascertainment of Diabetes Status

Diabetes status was documented either as treated or untreated. Treated cases were those where participants reported medications for diabetes. Untreated cases were those first detected through the study examination based on measured fasting glucose values. Fasting glucose values of >125 mg/dL with no previous reported diagnosis were coded as diabetes (untreated). Both untreated and treated cases were pooled for analyses;

participants who identified with incident diabetes at any of the follow-up examinations were considered a case regardless of diabetes status at the subsequent examinations.

5.1.2.4 Questionnaires

Standardized questionnaires were administered to participants at baseline and at follow-up examinations. Information was collected regarding demographics, socioeconomic measures, psychosocial status, medical and family history, medication use, dietary and alcohol intake, smoking, and physical activity. Information was self-reported by study participants. Data collection details for these variables are described elsewhere.[132]

5.1.2.5 Directly-Measured Cardiovascular Disease Risk Factors

Resting blood pressure was measured at the clinic visit (in the right arm after five minutes in the seated position) using a Dinamap automated oscillometric device. Height and weight were measured to the nearest 0.1 cm and 0.5 kg, respectively. Additionally, girths (waist at the umbilicus and hips at the maximum circumference of the buttocks) were measured to the nearest 0.1 cm using a steel measuring tape by trained and certified technicians.

Lipids were measured in plasma and were assayed at a central location, the Laboratory for Clinical Biochemistry Research (University of Vermont, Burlington, Vermont), using standardized methods and reagents. HDL-cholesterol (HDL-c) was measured using the cholesterol oxidase cholesterol method (Roche Diagnostics, Indianapolis, IN) with a laboratory CV of 1.6%. Triglycerides were measured using the Triglyceride GB reagent (Roche Diagnostics, Indianapolis, IN) on the Roche COBAS FARA centrifugal analyzer. The laboratory CV was reported as 4.0%. LDL-cholesterol

(LDL-c) was calculated using the Friedewald equation among those with fasting triglycerides (TG) ≤ 400 mg/dL.

5.1.3 *MESA: Human Subjects*

All participants provided written informed consent prior to baseline assessments. The informed consent process involved communicating information about the study aims, procedures, expectations, confidentiality, risks, and benefits associated with participating in the research study. The risks to participants in this study were minimal and included lab assessments that could cause slight discomfort (e.g., blood draw) as well as questions about behaviors and feelings that could cause psychological distress. Benefits included receiving results from study assessments. The Institutional Review Board (IRB) at each study site, laboratory, and reading center reviewed and approved the study. All personal identifiers were separated from the data and stored in locked file cabinets.

5.2 *Coronary Artery Risk Development in Young Adults (CARDIA) Study and the CARDIA Fitness Study (CFS)*

5.2.1 *CARDIA and CFS: Overview, Study Design, and Study Population*

CARDIA (Coronary Artery Risk Development in Young Adults) is a population-based longitudinal observational study of 5,115 men and women aged 18-30 years enrolled between 1985-1986 at four geographic sites (Birmingham, AL; Chicago, IL; Minneapolis, MN; and Oakland, CA). Investigators recruited the sample to obtain balanced subgroups by gender, race (African American or Black, and Caucasian), education (high school or less and more than high school), and age (18-24 years and 25-30 years). At baseline, the study sample was 45.5% male and 51.5% Black. Six follow-up

examinations are currently complete, providing twenty years of observational data for 72% of the surviving original cohort (69.4% of the total original cohort). The primary aim of CARDIA is to examine how heart disease risk factors develop in adults.

Details regarding methodology for recruitment and enrollment are described elsewhere. [133] Briefly, two of the four study sites (Chicago and Minneapolis) sampled participants using census track information. The Oakland center selected participants from the total Kaiser-Permanente health plan membership plan who resided in Oakland, Berkeley, and adjacent communities. Finally, Birmingham used telephone exchanges to sample from the city population. All study sites used telephone as the primary means of recruitment, though some efforts were supplemented by in-person household visits. The study population was limited to healthy, free-living individuals residing in the four study areas. Chronically ill, disabled, or institutionalized individuals were excluded from the study population.

Fitness assessment was part of the core CARDIA study at the baseline examination and at exam year 7. The CARDIA Fitness Study (CFS) was funded as an ancillary study for the year 20 examination, with the primary aim to study and characterize 20-year longitudinal changes in fitness and body composition in relation to cardiovascular disease risk. Additionally, the study aimed to examine genetic factors and their effects on the development of obesity, the metabolic syndrome, and subclinical cardiovascular disease.

5.2.2 *CARDIA and CFS: Summary of Relevant Study Components*

5.2.2.1 Anthropometrics

Certified technicians measured height and weight at each exam using a standardized protocol. [133] Body weight was measured to the nearest 0.2 kg using a calibrated balance beam scale, and height was measured to the nearest 0.5 cm using a vertical ruler. Participants were instructed to remove shoes and to wear light clothing for the measurements (light clothing was provided if participants arrived without appropriate attire). Body mass index (BMI) was calculated as weight in kilograms divided by squared height in meters. Waist circumference was measured to the nearest 0.5 cm at the midpoint between the iliac crest and the lowest lateral portion of the rib cage, and anteriorly midway between the xiphoid process of the sternum and umbilicus.

5.2.2.2 Fasting glucose

Fasting glucose was measured in serum at baseline and at years 7, 10, 15, and 20. Blood samples were assayed using the hexokinase ultraviolet method on a Cobas Mira Plus chemistry analyzer. Individuals were excluded from measurement of fasting glucose if they were diabetic and taking insulin and/or oral diabetes medications, were using steroids (not including inhalers), or were pregnant. Length of fasting was noted and fasting glucose values were excluded if participants did not fast for at least eight hours on the exam day. Based on reassays of glucose in 2006 and 2007 in about 200 samples per examination drawn at years 7, 10, 15, and 20, glucose values were recalibrated to account for an observed drift in the assay. Recalibrated glucose values were: 6.98 + 0.94 * year 7 glucose concentration, 7.15 + 0.96 * year 10 glucose concentration, 6.99 + 1.01 * year 15 glucose concentration, and 4.06 + 0.97 * year 20 glucose concentration.

5.2.2.3 Oral Glucose Tolerance Test (OGTT)

Two-hour post-load glucose was measured in non-diabetic and non-pregnant participants at exam years 10 and 20. Participants were provided with 75 grams of a glucose solution (Glucola) for ingestion within a period not to exceed five minutes; study staff checked to make certain that all of the solution was consumed. Glucose was then measured in the same manner as the fasting glucose, two hours following consumption of the glucose solution. A red top Vacutainer tube was used to collect 5 ml of blood; this second glucose measurement was not taken in participants that vomited after consumption of the solution.

5.2.2.4 Fasting Insulin

Fasting insulin was measured at baseline and years 7, 10, 15, and 20 by radioimmunoassay. Participants were required to fast for at least eight hours. As with fasting glucose, individuals were excluded from measurement of fasting insulin if they were diabetic and taking insulin and/or oral diabetes medications, were using steroids (not including inhalers), or were pregnant. Fasting insulin values were excluded from analyses if participants reported not fasting for at least eight hours on the exam day. Samples for measurement is fasting insulin were collected in a Vacutainer tube containing no preservative. The sample was centrifuged within 60 minutes, stored in cryovials, and frozen at -70°C within 90 minutes. Based on reassays of insulin in 2006 and 2007 in 100 samples stored since year 15, insulin values at year 20 were recalibrated using the following formula: -0.36 + 0.93 * year 20 insulin concentration.

5.2.2.5 Hemoglobin A1c

Hemoglobin A1c was measured at exam year 20 using an HbA1c kit (HbA1c Sample Preparation Kit, Bio-Rad, Hercules, CA 84547) and was collected in a capillary tube. The sample was centrifuged and stored at -70 C until processed. Measurements were conducted at the Collaborative Studies Clinical Laboratory at Fairview-University Medical Center (Minneapolis, MN).

5.2.2.6 Blood Pressure

Systolic and diastolic blood pressures were measured at each exam in the seated position after a five-minute rest period in a quiet room. Measurements were obtained before any physical examination, blood draw, or potentially stressful interview. Three blood pressure measurements were obtained using a random zero sphygmomanometer, and were acquired a minimum of 30 seconds apart. The second and third readings were retained and averaged to produce mean resting systolic and diastolic blood pressure values. At exam year 20, an Omron automated device was used for blood pressure measurement. A comparability study was conducted that showed no systematic difference between the two devices.

5.2.2.7 Lipids

A plasma lipid profile was obtained from a blood sample drawn from an anticubital vein. Blood was collected into an EDTA Vacutainer tube, centrifuged for 20 minutes at 2000 RPM at 4°C, and frozen to -75° to -65°C until the sample was assayed at Northwest Lipid Laboratories in Seattle, WA. Details regarding the specific protocols used are described elsewhere.[133] Total cholesterol was determined using enzymatic methods, HDL-cholesterol was measured after dextran sulfate precipitation, and LDL-

cholesterol was calculated using the Friedewald equation among those with fasting TG ≤ 400 mg/dL.

5.2.2.8 Physical Activity History Questionnaire

Self-reported physical activity was measured using the CARDIA Physical Activity History, an interviewer-administered modified version of the Minnesota Leisure Time Physical Activity Questionnaire.[134,135] The instrument assesses thirteen categories of activity covering various intensity levels. For each activity, a trained and certified interviewer recorded whether the activity was performed at any time in the prior year, the number of months during which the activity was performed, and the number of months that the activity was performed frequently. The total score represented total amount of moderate to vigorous physical activity expressed in exercise units. Further details of the scoring system have been described previously.[136]

5.2.2.9 Other Questionnaires

Self-administered questionnaires were used to obtain information about demographics, socioeconomic status, access to health services, health behaviors, and psychosocial characteristics. Variables of interest in this dissertation project include: educational attainment, smoking and alcohol use, family history of diabetes and cardiovascular disease, access to and ability to pay for needed health services, food frequency and diet history information, depressive symptoms, anxiety, social support, anger expression, and discrimination.

5.2.2.10 Exercise Treadmill Test

Participants participated in a symptom-limited maximal graded exercise treadmill testing (GXT) at the baseline examination and at years 7 and 20, where participants were

asked to exercise to maximal exertion. A modified Balke protocol was used with 2-minute stages.[134] Pulse rate, blood pressure, and 12-lead electrocardiograms were obtained on each subject at test. Every 2 minutes for up to 18 minutes (9 states), the speed or grade of the treadmill was increased to a maximum speed of 5.6 miles per hour and a 25% grade. Heart rate, blood pressure, and three-lead electrocardiogram data were obtained at the end of each stage, at maximum exercise, and at every minute for 3 minutes post-exercise. Self-reported ratings of perceived exertion (RPEs) were collected near the end of each test stage to assess symptoms.

Participants were eligible to participate in the GXT if they had no history of ischemic or congenital heart disease, were not using cardiovascular medications other than antihypertensive medications, had blood pressure <160/90 mmHg, and no concurrent febrile illness. Exercise test data were determined valid if participants achieved 85% or greater of their age predicted maximum heart rate determined using the Tanaka formula (208 - 0.7*age).[137] Cardiorespiratory fitness was defined as the duration that participants were able to walk or run on the treadmill (possible range: 0-1080 seconds). Data from the Minneapolis field center at year 7 are excluded from all analyses due to a protocol violation.

5.2.2.11 Genotyping

SNPs were genotyped using a multiplex matrix-assisted laser desorption/ionization time-of-flight mass spectrometry method (Sequenom, San Diego, CA). Targeted sequences were amplified by polymerase chain reaction using specific forward and reverse primers for each polymorphism, followed by extension reactions utilizing specific oligonucleotide primers that annealed immediately upstream of each

polymorphic site and extended the amplification product by a single base pair in the forward direction. The mass of the extension reaction products was determined and translated into genotype using the MassARRAY RT software (Sequenom Inc., San Diego, CA). Quality control (QC) activities employed by the CARDIA study included barcode identification of samples, robotic sample handling, and blind replicate genotype assessment on 5% of the total sample. Genotype concordance between QC pairs was ≥ 99% for all SNPs.

5.2.3 *CARDIA AND CFS: Human Subjects*

The CARDIA study investigators obtained written informed consent from all participants prior to assessments. The informed consent process involved communicating information about the study aims, procedures, expectations, confidentiality, risks, and benefits associated with participating in the research study. The risks to participants in this study were minimal and included lab assessments that could cause slight discomfort (e.g., blood draw and exercise treadmill test) as well as questions about behaviors and feelings that could cause psychological discomfort. Participants were informed of their right to refuse storage and use of their blood samples for genotyping. Benefits included receiving results from study assessments. An IRB at each study site approved all study procedures. Personal identifiers were removed from all data and stored in a secure location at the coordinating center.

6. MANUSCRIPT 1: Ethnic Differences in the Predictive Validity of Impaired Fasting Glucose and Hemoglobin A1c Levels in the Multi-Ethnic Study of Atherosclerosis

6.1 *Synopsis*

Objectives

This study evaluated racial/ethnic differences in the prediction of incident diabetes from fasting glucose and hemoglobin A1c levels in the Multi-Ethnic Study of Atherosclerosis (MESA).

Research Design and Methods

MESA is a community-based study that enrolled 6,814 Caucasian Americans, Chinese Americans, African Americans, and Hispanic Americans aged 45-84 years and free of clinical CVD. Logistic regression and receiver-operator characteristic curves were utilized to examine associations between fasting glucose and HbA1c levels with cumulative incidence of diabetes, and to evaluate potential effect modification by race/ethnicity.

Results

Over five years, 692 incident diabetes cases occurred. The association between fasting glucose at baseline and cumulative incidence of diabetes did not differ by ethnic group ($\chi^2_{3df} = 1.87$, $p = 0.60$). However, the association between HbA1c and diabetes incidence differed significantly by ethnicity ($\chi^2_{3df} = 15.47$, $p < 0.01$). In adjusted models, the c-statistics ranged from 0.75 in African Americans to 0.87 in Caucasians. In further

exploration, the significant overall interaction was primarily driven by the African American subgroup; the c-statistic for this group differed significantly from each of the others.

Conclusions

Improving diabetes prevention in population subgroups that are disproportionately affected is dependent upon effective screening tools. This study illustrates that measures used to identify high-risk populations may not perform equally well across populations. While the predictive validity of HbA1c was comparable to fasting glucose, the strength of association between HbA1c and diabetes incidence varied significantly by race/ethnicity. Therefore, consideration of multiple measures of hyperglycemia may be preferable for risk stratification in diverse populations.

6.2 *Introduction*

Type 2 diabetes accounts for an estimated ninety to ninety-five percent of all diabetes cases among U.S. adults; individuals with impaired fasting glucose (IFG, defined as 100-125 mg/dL after a minimum of eight hours of fasting) are at high risk for developing type 2 diabetes.[13,69,90] One notable advantage of identifying individuals with IFG is that non-pharmacological prevention strategies are efficacious in reversing or delaying the disease process in this population.[49] Improving our ability to identify those at highest risk can facilitate more efficient focusing of resources on population subgroups that would benefit most from prevention efforts.

While generally not recognized as a distinct clinical entity, IFG is regarded as a statistical risk factor by association. The current lower cut point of 100 mg/dL used to

define IFG is not population-specific, though the prevalence of both impaired fasting glucose and diabetes differs markedly across population subgroups.[138] For example, NHANES (NCHS) data show that the physician-diagnosed prevalence of type 2 diabetes among adults aged twenty years and older is almost twice that for Hispanic or Latino Americans compared to non-Hispanic White Americans. Additionally, the prevalence of physician-diagnosed diabetes mellitus is almost twice as high in non-Hispanic Black Americans compared to non-Hispanic White Americans.[139] Given significant observed ethnic differences in the population burden of diabetes, the point on the continuum of blood glucose levels—that confers comparable risk—might also differ by ethnicity.

The Multi-Ethnic Study of Atherosclerosis (MESA) comprises a population-based sample of men and women self-identifying with one of four ethnic groups: Caucasian, Chinese American, African American, or Hispanic. This diverse cohort allowed for the systematic evaluation of potential ethnic differences in the prediction of incident diabetes from two measures of hyperglycemia, fasting glucose and hemoglobin A1c (HbA1c). Thus, the aim of the current study was to determine whether ethnic-specific cut points for IFG are warranted to more precisely identify high-risk individuals. Secondary analyses considered HbA1c values as a predictor of future risk for diabetes in the MESA cohort, since the literature suggests that measurement of HbA1c — either in place of or in conjunction with fasting glucose — may be a preferable screening tool for identifying high-risk individuals who would benefit from early intervention and prevention efforts.[69] Previous studies in non-diabetic populations have consistently observed ethnic differences in HbA1c levels; ethnic differences in the prediction of incident diabetes from HbA1c levels have yet to be fully evaluated.[43,140,141]

6.3 *Methods*

Study Population

MESA is a multi-center, prospective cohort study funded by the National Heart, Lung, and Blood Institute. The primary objectives of MESA are to identify characteristics related to subclinical cardiovascular disease and progression to clinical cardiovascular disease. The study enrolled 6,814 men and women aged 45-84 years in six U.S. communities. All participants were free of cardiovascular disease and self-identified with one of four ethnic groups (Caucasian, Chinese American, African American or non-Hispanic Black, and Hispanic); ethnicity was verified at the time of enrollment. Four study exams were completed between 2000 and 2007. Participants with fasting glucose measured at the baseline examination and with complete diabetes status information through follow-up were included in the analyses; individuals with prevalent diabetes at baseline will be excluded. Using these criteria, the sample size for the primary analysis included 5,006 participants. For secondary analyses examining HbA1c levels as the predictor of incident diabetes, individuals with prevalent diabetes or missing HbA1c data at Exam 2 were excluded; Exam 2 was the baseline for this analysis because this was the sole time point where HbA1c was measured. For the HbA1c analyses, 4,812 participants were included in the analytic sample.

Data Collection Methods

Fasting Glucose

Fasting glucose was measured at each study exam by rate reflectance spectrophotometry using thin film adaptation of the glucose oxidase method on the Vitros

analyzer (Johnson & Johnson Clinical Diagnostics, Inc., Rochester, NY). Samples were processed at the Collaborative Studies Clinical Laboratory at Fairview-University Medical Center (Minneapolis, MN); the reference range for this assay is 60 - 115 mg/dL and the laboratory coefficient of variation (CV) is 1.1%. After unexpected trends in fasting glucose levels and classification of diabetes status across the first three examinations were observed, MESA investigators re-assayed 800 fasting glucose samples (200 from each of the four exams). Subsequently, observed levels from Exam 2, Exam 3, and Exam 4 were recalibrated to the Exam 1 scale [for Exam 2 glucose values, estimated value = -1.261 + (original*0.964); for exam 3 glucose values, estimated value = 1.003 + (original*0.963); and for Exam 4 glucose values, estimated value = -5.163 + (original4*1.009)]. From the estimated values, calibrated values were obtained using the following formula: calibrated value = 6.453 + estimated value * 1.005. The recalibrated values were used for all analyses.

Hemoglobin A1c

Hemoglobin A1c was measured at Exam 2 using a sample preparation vial containing an aqueous solution of EDTA and potassium cyanide (HbA1c Sample Preparation Kit, Bio-Rad, Hercules, CA 84547). Specimens were analyzed on the Tosoh A1c 2.2 Plus Glucohemoglobin Analyzer (Tosoh Medics, Inc., San Francisco, CA 94080) using an automated high performance liquid chromatography method. The reference range was 4.3 - 6.0% with a laboratory CV range of 1.4 - 1.9%. Measurements were made at the Collaborative Studies Clinical Laboratory at Fairview-University Medical Center (Minneapolis, MN).

Ascertainment of Diabetes Status

Diabetes status was documented either as treated or untreated. Treated cases were those where participants reported medication use for diabetes. Untreated cases were those detected at the clinic visit based on measured fasting glucose. Fasting glucose values of 126 mg/dL and greater and with no previous reported diagnosis were coded as diabetes (untreated). For this study, both untreated and treated cases were pooled for the primary analyses. Undiagnosed cases were included since the primary objective was to evaluate prediction of progression to frank diabetes, regardless of treatment status. Additionally, this approach minimized potential confounding by unmeasured factors that may be associated with ethnic differences in the detection/treatment of diabetes. Sensitivity analyses were conducted by examining the subset of individuals with physician diagnosis or documentation of medications only (i.e., excluding the undiagnosed cases); the purpose of this analysis was to examine the potential effect of factors that may be associated with ethnic differences in diabetes diagnosis and treatment.

Questionnaires

Standardized questionnaires were administered at baseline and at follow-up examinations. Information was collected regarding demographics, socioeconomic characteristics, medical and family history, medication use, dietary and alcohol intake, smoking, and physical activity. Information was self-reported by study participants and data collection procedures are described in detail elsewhere. [132]

Directly-Measured Clinical Characteristics

Resting blood pressure was measured at each clinic visit (in the right arm after five minutes in the seated position) using a Dinamap automated oscillometric device.

Three measurements were obtained, with the mean of the last two averaged for use in analyses. Height and weight were measured to the nearest 0.1 cm and 0.5 kg, respectively. Body mass index (BMI) was calculated as weight in kilograms divided by height in meters squared for use as a measure of overall adiposity.

Fasting blood samples were collected to measure lipids. HDL-cholesterol (HDL-c) was measured using the cholesterol oxidase cholesterol method (Roche Diagnostics, Indianapolis, IN) with a laboratory CV of 1.6%. Triglycerides were measured using the Triglyceride GB reagent (Roche Diagnostics, Indianapolis, IN) on the Roche COBAS FARA centrifugal analyzer. The laboratory CV was reported as 4.0%. LDL-cholesterol (LDL-c) was calculated using the Friedewald equation among those with fasting triglycerides (TG) $\leq$ 400 mg/dL.

Statistical Analyses

Baseline participant characteristics, stratified by self-reported ethnic group, were examined among individuals without diabetes. Analysis of variance was used to compare study groups for continuous measures. Chi-square tests were used to test for differences between study groups for categorical characteristics.

The primary analysis excluded individuals with prevalent diabetes at baseline and used logistic regression to predict incident diabetes from measured values of fasting glucose at baseline. Sensitivity (the true positive rate) and specificity (the false positive or true negative rate) of current established cut points (fasting glucose level of 100-125 mg/dL) for IFG by race/ethnic group were examined in addition to the area under the receiver-operating characteristic (ROC) curve (AUC, or c-statistic). Logistic regression

was also used to evaluate the strength of association between IFG status (yes/no) and cumulative incidence of diabetes. Effect modification by self-reported ethnic group was also considered.

Exploratory analyses using various cut points along the continuous range of glucose values evaluated the appropriateness of ethnic-specific cut points relative to 100-125 mg/dL. Sensitivity, specificity, sensitivity + specificity, the c-statistic, and the Bayesian information criterion (BIC, defined as -2*ln(likelihood) + ln(N)*k, where k represents the model degrees of freedom and N signifies number of observations) were used to develop recommendations regarding optimum cut points for identifying high-risk individuals in each race/ethnic group.

ROC curves and logistic regression estimates were also used to examine the prediction of incident diabetes from HbA1c values measured at Exam 2 and potential ethnic differences in the prediction of incident diabetes. Correlated *U* statistics were used to compare the use of fasting glucose versus HbA1c for the prediction of incident diabetes using methods described by Delong et al.[142] All analyses were conducted using the SAS System, Version 9.2 (Cary, NC).

6.4 *Results*

Participant Characteristics

Participant characteristics at baseline are summarized in Table 1, stratified by self-reported race/ethnic group. On average, the Caucasian participants were older and had a higher percentage of individuals with a high school diploma or greater compared to the other study groups. Mean fasting glucose and insulin levels were greater in the

Chinese, in addition to the prevalence of IFG. Mean BMI, waist circumference, systolic and diastolic blood pressure, HDL-cholesterol, and HbA1c were greatest in the African American group. In this group, the prevalence of antihypertensive medication and cholesterol-lowering medication was also highest. Prevalence of smoking was highest in the Hispanic subgroup; mean total cholesterol, LDL-cholesterol, and triglycerides were greatest in this group.

At baseline there were 941 prevalent cases of diabetes (680 treated and 261 untreated cases) that were excluded from analyses (Table 2). Over a mean follow-up period of five years, 692 incident diabetes cases occurred. We observed 203 incident diabetes cases among Caucasians, 68 cases among Chinese Americans, 227 cases among African Americans, and 194 cases among Hispanic participants. Mean levels of fasting glucose at each study exam and HbA1c at exam 2 are summarized in Table 3, stratified by self-reported ethnic group; these same data are described by sex in Table 4.

Fasting Glucose and Cumulative Incidence of Diabetes

Figure 1 plots the crude ROC curves predicting five-year cumulative incidence of diabetes across the range of measured fasting glucose values, stratified by self-reported ethnic group. In models containing fasting glucose, age, sex, BMI, and clinic site, the AUCs ranged from 0.83 in the Chinese subgroup to 0.88 in the Caucasian subgroup (Table 5 and Figure 2). After adjustment, the association between fasting glucose at baseline and cumulative incidence of diabetes did not differ by self-reported ethnic group, $\chi^2_{3df} = 1.87$, $p = 0.60$). However, in the subgroup analysis including only prevalent untreated diabetes cases at baseline (n=261) and excluding the 438 untreated incident

diabetes cases identified during follow-up, effect modification by ethnic group was observed ($\chi^2_{3df} = 14.02$, $p < 0.01$).

Table 6 summarizes the results of exploratory analysis to identify the optimum lower cut point for IFG for the full study population and for each ethnic group. For all participants pooled, the current IFG cut points of 100-125 mg/dL was associated with 62% sensitivity and 88% specificity for predicting cumulative incidence of diabetes in the cohort; the c-statistic was 0.75 and individuals in this range for fasting glucose were over twelve times more likely to develop diabetes than individuals with fasting glucose values less than 100 mg/dL at baseline (OR = 12.42, 95% CI = 9.85-15.66).

For all ethnic groups, with the exception of the Caucasian subgroup, the optimal lower cut point of 100 mg/dL was associated with the most favorable measures of model fit. For the Caucasian subgroup, lowering the cut point to 95 mg/dL significantly improved sensitivity (75% versus 58% for 100 mg/dL). Additionally, the sum of sensitivity and specificity increased from 1.44 to 1.50 and the c-statistic increased from 0.72 to 0.75; this improvement was statistically significant ($U = 4.74$, $p < 0.05$).

HbA1c and Cumulative Incidence of Diabetes

Figure 2 illustrates the ROC curves predicting three-year cumulative incidence of diabetes from measured HbA1c values. In a model including HbA1c in addition to age, sex, BMI, and site, the AUCs ranged from 0.75 in African Americans to 0.87 in Caucasians (Table 7 and Figure 4). The association between HbA1c and cumulative incidence of diabetes differed significantly across ethnic groups ($\chi^2_{3df} = 15.47$, $p < 0.01$). In further exploration, the significant overall interaction was primarily driven by the African American subgroup; the c-statistic for this group was significantly different from

each of the other three study groups. Limiting the definition of diabetes to only include treated cases did not materially alter the observed effect modification by ethnic group ($\chi^2_{3df} = 12.37$, $p < 0.01$).

Fasting Glucose Versus HbA1c in the Prediction of Diabetes

In order to compare the prediction of diabetes from fasting glucose compared to HbA1c, measured fasting glucose values at the second study exam were considered. Table 8 compared the c-statistics predicting diabetes from fasting glucose alone to HbA1c alone. In general, fasting glucose values predicted diabetes slightly better than HbA1c, but this difference was only statistically significant ($p < 0.05$) in the Chinese subgroup. Overall (after controlling for age, sex, BMI, clinic site, and race/ethnicity), a 1-SD increment in fasting glucose at the second exam (9.6 mg/dL) was associated with 2.45 times higher odds of diabetes (95% CI: 2.08-2.87), while the OR for HbA1c for a 1-SD increment (0.42%) was 3.10 (95% CI: 2.67-3.61). However, this comparison should be interpreted with caution, as we excluded prevalent diabetes cases at this exam based on fasting glucose values at the highest end of the range.

After controlling for age, sex, BMI, clinic site, and fasting glucose, HbA1c remained an independent predictor of incident diabetes ($\chi^2_{1df} = 57.33$, $p < 0.0001$); the odds ration for incident diabetes per SD increment in HbA1c (0.42% based on the full study population SD) was 3.14 for Caucasians (95%CI: 1.49-3.06), 2.03 for the Chinese group (95% CI: 1.98-3.80), 1.67 for the African American group (95% CI: 1.39-2.17), and 2.10 for the Hispanic group (95% CI: 1.52-2.91). We also considered the utility of both fasting glucose used in conjunction with HbA1c for predicting diabetes — since each of these measures is independently predictive (with all participants pooled) — and

found that fasting glucose and HbA1c together predicted significantly better than either measure alone ($p < 0.05$). After stratifying by ethnic group, we observed significant improvement in the prediction with the use of both fasting glucose and HbA1c compared to HbA1c alone in both the Caucasian and Chinese subgroups.

6.5 *Discussion*

In this population-based cohort study, the strength of association between fasting glucose and cumulative incidence of diabetes did not differ by self-reported race/ethnic group in the primary analysis. However, measures of model fit suggested that decreasing the lower cut point for IFG from 100 mg/dL to 95 mg/dL in Caucasians improved the sensitivity of the measure in identifying individuals most likely to progress to frank diabetes without significantly increasing false positives. Significant effect modification by race/ethnic group was observed for the association between HbA1c levels and cumulative incidence of diabetes, which may be a reflection of differences in the degree of risk conferred at a given level of HbA1c in different populations over a set period of time. Alternately, there may be factors that differentially affect measured levels of HbA1c and the degree of risk associated with a given HbA1c value. The association was weakest in the African American subgroup and the multivariable-adjusted odds ratio was significantly lower than in the Caucasian, Chinese, and Hispanic subgroups. When diabetes cases were limited to those with self-reported diabetes, ethnic differences emerged for fasting glucose and were even more pronounced for HbA1c. Most likely, these differences were driven by unmeasured factors associated with race/ethnic differences in diagnostic/treatment practices and/or access to screening services.

The low prevalence of diabetes is in the Chinese subgroup limited statistical power for subgroup analyses in this population. Treated diabetes status was also based on self-reported information, which could have led to misclassification of disease status for some participants. Further, HbA1c was only measured at exam 2 and no oral glucose tolerance test was performed. Therefore, only fasting glucose could be used to identify undiagnosed incident cases. Despite these limitations, our study was strengthened by the fact that the MESA study was comprised of a large, ethnically diverse population-based cohort with standardized data collection protocols and assays. Many other studies that have examined similar prediction models have been restricted to one or two ethnic groups. The number of events was also large, with almost 700 available incident diabetes cases. Further, the MESA study measured HbA1c at Exam 2, providing an opportunity to further our understanding of the role of this measure as an adjunct tool in predicting diabetes risk in pre-diabetics. Finally, the study employed stringent quality control measured, used standardized measures and assays to evaluate fasting glucose and HbA1c.

Because the MESA study systematically measured a wide variety of participant characteristics, we explored possible attributes that might help to explain the observed effect modification by ethnic group for the association between HbA1c and diabetes, but were unable to statistically account for the ethnic differences. These attributes included demographic factors such as age, sex, and educational level; clinical factors including lipids and blood pressure; and health behaviors such as smoking and alcohol use. Therefore, while HbA1c is clearly predictive of future diabetes risk, underlying reasons for the observed differences by population subgroup certainly need additional exploration. Further examination of this issue in a study with larger sample size in each

of the ethnic groups (or more follow-up time to increase the number of incident events) might help to shed light upon the underlying cause of the observed race/ethnic differences. Based on our data, consideration of HbA1c in addition to fasting glucose would likely provide improved risk prediction in non-diabetic, ethnically diverse populations.

Recent recommendations for the diagnosis of diabetes and for screening for high-risk persons have incorporated HbA1c, since this measure offers several practical advantages over fasting glucose.[14] HbA1c can be measured in the non-fasting state, reflects longer-term exposure to blood glucose, and the assay has a lower intraindividual variability compared to fasting glucose. HbA1c in general provides a more integrated measure of chronic hyperglycemia, though our understanding of the meaning of HbA1c levels in non-diabetic populations is still limited.

Improving diabetes prevention in population subgroups that are disproportionately affected is dependent upon effective screening tools. This study illustrates that the tools used to identify high-risk populations may not perform equally well across populations and do not capture all individuals who will progress to diabetes. Therefore, consideration of multiple measures of hyperglycemia might allow for more comparable identification of individuals at increased risk across populations at a minimal additional cost.

Because fasting glucose and HbA1c are measured on a continuum, the choice of distinct criteria to define "high risk" will always be somewhat arbitrary. However, the degree of risk conferred by a given level of these measures may not be equivalent across populations. Thus, using a single measure for risk stratification may not be appropriate. Refinement of our ability to better identify those most likely to progress to frank diabetes

in diverse populations is paramount for developing more effective clinical guidelines for risk factor management as well as identifying target populations who will most benefit from prevention efforts. In this ethnically diverse, population-based cohort, the use of fasting glucose in conjunction with HbA1c might improve our ability to correctly identify individuals at highest risk for future development of type 2 diabetes.

6.6 *Tables*

Table 1. MESA participant characteristics at baseline.

	Caucasian (n = 2145)		**Chinese American (n = 565)**		**African American (n = 1275)**		**Hispanic American (n = 1021)**		***P***
Age, mean(SD), years	61.85	(10)	60.47	(9.9)	60.95	(9.9)	60.21	(10.1)	<.0001
Female, n(%)	1129	(52.6)	290	(51.3)	714	(56)	544	(53.3)	0.176
Field center									<.0001
Columbia, n(%)	195	(9.1)	2	(0.4)	266	(20.9)	377	(36.9)	
JHU, n(%)	396	(18.5)	0	(0)	341	(26.7)	0	(0)	
NWU, n(%)	480	(22.4)	243	(43)	226	(17.7)	0	(0)	
UCLA, n(%)	116	(5.4)	320	(56.6)	107	(8.4)	334	(32.7)	
UMN, n(%)	494	(23)	0	(0)	0	(0)	307	(30.1)	
WFU, n(%)	464	(21.6)	0	(0)	335	(26.3)	3	(0.3)	
Education less than high school, n(%)	84	(3.9)	112	(19.8)	115	(9.1)	405	(39.7)	<.0001
Income									<.0001
<$25,000, n(%)	299	(14.2)	250	(44.5)	304	(25.7)	439	(43.7)	
$25,000-$49,999, n(%)	554	(26.4)	130	(23.1)	388	(32.8)	350	(34.8)	
$50,000-$99,999, n(%)	697	(33.2)	112	(19.9)	378	(32)	185	(18.4)	
≥$100,000, n(%)	552	(26.3)	70	(12.5)	113	(9.6)	31	(3.1)	
Fasting glucose, mean(SD), mg/dL	87.6	(9.9)	91.4	(9.9)	90.2	(10.8)	90.9	(10.8)	<.0001
Fasting insulin, mean(SD), mg/dL	5.73	(4)	5.9	(3.7)	6.8	(4.6)	7.7	(4.9)	<.0001
Body mass index, mean(SD), kg/m^2	27.5	(4.9)	24	(3.3)	29.9	(5.7)	29.2	(5)	<.0001
Waist circumference, mean(SD), cm	97.1	(14.2)	86.62	(10)	99.8	(14.1)	99.7	(13)	<.0001
Systolic blood pressure, mean(SD), mmHg	122.3	(19.9)	122.2	(20.4)	130.2	(21)	125.0	(21)	<.0001
Diastolic blood pressure, mean(SD), mmHg	70.2	(9.9)	72.02	(10.4)	74.7	(10)	71.8	(10.1)	<.0001
Total cholesterol, mean(SD), mg/dL	196.5	(34.3)	193.1	(30.8)	190.5	(35.3)	199.2	(36.1)	<.0001
LDL cholesterol, mean(SD), mg/dL	117.8	(29.9)	115.6	(27.8)	117.5	(32.4)	121.1	(31.5)	<.01

	Caucasian (n = 2145)		Chinese American (n = 565)		African American (n = 1275)		Hispanic American (n = 1021)		*P*
HDL cholesterol, mean(SD), mg/dL	52.9	(15.7)	50.0	(13)	53.29	(15.5)	48.5	(13.3)	<.0001
Triglycerides, mean(SD), mg/dL	128.9	(74)	138.8	(76.1)	98.6	(53.8)	150.1	(94.1)	<.0001
Impaired fasting glucose, n(%)*	235	(11)	105	(18.7)	230	(18.1)	189	(18.5)	<.0001
Antihypertensive medication, n(%)	633	(29.5)	136	(24.1)	564	(44.3)	281	(27.5)	<.0001
Cholesterol-lowing medication, n(%)	348	(16.2)	69	(12.2)	181	(14.2)	136	(13.3)	<.05
Smoking status									<.0001
Never Smoker, n(%)	962	(44.9)	431	(76.3)	577	(45.5)	561	(54.9)	
Former Smoker, n(%)	952	(44.5)	104	(18.4)	463	(36.5)	327	(32)	
Current Smoker, n(%)	227	(10.6)	30	(5.3)	228	(18)	133	(13)	
HbA1c (exam 2) , mean(SD), %†	5.35	(0.44)	5.54	(0.51)	5.61	(0.59)	5.52	(0.50)	<.0001

* Impaired fasting glucose defined by American Diabetes Association criteria (fasting glucose 100-125 mg/dL).
† n=2097 for European Americans, 549 for Chinese Americans, 1217 for African Americans, and 987 for Hispanic Americans.
Note: P-value from ANOVA or chi-square as appropriate.

Table 2. Incidence and prevalence of IFG and diabetes by ethnic group.

	Caucasian (n=2614)	Chinese American (n=801)	African American (n=1881)	Hispanic (n=1494)	All Participants (n=6790)
Exam 1					
Normal Fasting Glucose: Prevalence, n (%)	2163 (83)	559 (70)	1271 (68)	999 (67)	**4992 (74)**
IFG: Prevalence, n (%)	293 (11)	137 (17)	278 (15)	231 (15)	**939 (14)**
Untreated Diabetes: Prevalence, n (%)	46 (2)	23 (3)	63 (3)	47 (3)	**179 (3)**
Treated Diabetes: Prevalence, n (%)	112 (4)	82 (10)	269 (14)	217 (15)	**680 (10)**
Exams 2-4*					
Untreated Diabetes: Cumulative Incidence, n (%)	79 (3)	28 (4)	75 (5)	72 (6)	**254 (4)**
Treated Diabetes: Cumulative Incidence, n (%)	124 (5)	40 (6)	152 (10)	122 (10)	**438 (7)**
Total Diabetes: Cumulative Incidence, n (%)	203 (8)	68 (10)	227 (15)	194 (16)	**692 (12)**

*Denominators for percentages exclude diabetes cases at baseline.

Table 3. Unadjusted mean levels of fasting glucose and HbA1c by exam and ethnic group.*

	Caucasian		Chinese American		African American		Hispanic American		
Fasting Glucose	**n**	**Mean mg/dL (SD)**	**n**	**Mean mg/dL (SD)**	**n**	**Mean mg/dL (SD)**	**n**	**Mean mg/dL (SD)**	**p-value**
Exam 1	2138	87.6 (10.0)	563	91.4 (9.4)	1270	90.2 (9.36)	1019	90.9 (10.8)	<.0001
Exam 2	2098	92.2 (12.2)	550	94.4 (13.0)	1237	94.0 (14.1)	991	94.6 (15.0)	<.0001
Exam 3	2102	90.9 (13.8)	551	94.0 (12.7)	1227	93.9 (14.8)	973	95.4 (16.5)	<.0001
Exam 4	2143	94.8 (17.7)	563	95.8 (14.4)	1271	95.7 (14.4)	1021	99.3 (21.0)	<.0001
HbA1c	**n**	**Mean % (SD)**	**n**	**Mean % (SD)**	**n**	**Mean % (SD)**	**n**	**Mean % (SD)**	**p-value**
Exam 2	2097	5.35 (0.44)	549	5.54 (0.51)	1217	5.61 (0.59)	987	5.52 (0.50)	<.0001

* Excludes individuals with prevalent diabetes at Exam 1 or with missing fasting glucose data (n=1824).

Table 4. Unadjusted mean levels of fasting glucose and HbA1c by exam and sex.

	Female		Male		p-value
Fasting Glucose	**n**	**Mean mg/dL (SD)**	**n**	**Mean mg/dL (SD)**	
Exam 1	2668	87.7 (10.3)	2332	91.3 (10.3)	<.0001
Exam 2	2597	91.9 (12.7)	2279	95.1 (14.1)	<.0001
Exam 3	2593	91.3 (13.5)	2260	94.8 (15.6)	<.0001
Exam 4	2670	94.6 (17.0)	2328	97.7 (17.7)	<.0001
HbA1c	**n**	**Mean % (SD)**	**n**	**Mean mg/dL (SD)**	**p-value**
Exam 2	2575	5.49 (0.51)	2265	5.46 (0.51)	<.05

Table 5. Goodness-of-fit statistics by ethnic group for fasting glucose.

Ethnic Group	Model 1*: C Statistic	Model 2†: C Statistic
All Groups (pooled)	0.83	0.86
Caucasian	0.85	0.88
Chinese American	0.83	0.86
African American	0.82	0.83
Hispanic American	0.81	0.84

*Crude model.
†Model containing fasting glucose, age, sex, clinic site, and BMI. Model for all groups pooled additionally included ethnic group.
Note: P-value for glucose x ethnic group interaction after adjusting for age, sex, clinic site, and BMI = 0.60.

Table 6. Evaluation of adjusted cut points for IFG (all participants and by ethnic group).

All Participants

IFG Cut Point	n	No. Within IFG Limits	No. With Diabetes at Exam 4	Sensitivity	Specificity	1-Specificity	Sensitivity + Specificity	C-Statistic	BIC	OR for Incident Diabetes
90-125	5006	2200	357	0.857	0.593	0.407	1.450	0.725	2301	8.73 (6.45-11.81)
95-125	5006	1328	357	0.742	0.771	0.229	1.514	0.757	2204	9.72 (7.59-12.44)
100-125*	5006	759	357	0.619	0.884	0.116	1.503	0.752	2136	12.42 (9.85-15.66)
105-125	5006	434	357	0.496	0.945	0.055	1.441	0.720	2120	16.80 (13.18-21.42)
110-125	5006	243	357	0.356	0.975	0.025	1.331	0.665	2195	21.58 (16.23-28.68)
115-125	5006	139	357	0.238	0.988	0.012	1.227	0.613	2299	26.59 (18.51-38.20)
120-125	5006	72	357	0.140	0.995	0.005	1.135	0.568	2403	34.25 (20.48-57.30)

Caucasian Americans

IFG Cut Point	n	No. Within IFG Limits	No. With Diabetes at Exam 4	Sensitivity	Specificity	1-Specificity	Sensitivity + Specificity	C-Statistic	BIC	OR for Incident Diabetes
90-125	2145	787	110	0.864	0.66	0.34	1.524	0.762	760	12.29 (7.08-21.35)
95-125	2145	442	110	0.755	0.824	0.176	1.578	0.789	719	14.35 (9.16-22.48)
100-125*	2145	235	110	0.582	0.916	0.084	1.498	0.749	724	15.17 (10.07-22.85)
105-125	2145	140	110	0.482	0.957	0.043	1.439	0.720	719	20.82 (13.53-32.05)
110-125	2145	74	110	0.346	0.982	0.018	1.328	0.664	743	29.31 (17.55-48.95)
115-125	2145	35	110	0.2	0.994	0.006	1.194	0.597	792	38.88 (18.96-79.73)
120-125	2145	19	110	0.136	0.998	0.002	1.134	0.567	810	80.17 (26.11-246.19)

Chinese Americans

IFG Cut Point	n	No. Within IFG Limits	No. With Diabetes at Exam 4	Sensitivity	Specificity	1-Specificity	Sensitivity + Specificity	C-Statistic	BIC	OR for Incident Diabetes
90-125	565	288	36	0.917	0.518	0.482	1.4347	0.717	251	11.82 (3.58-39.01)
95-125	565	168	36	0.778	0.735	0.265	1.5131	0.757	242	9.73 (4.33-21.84)
100-125*	565	105	36	0.667	0.847	0.153	1.5136	0.757	237	11.06 (5.31-23.01)
105-125	565	57	36	0.5	0.926	0.074	1.4263	0.713	239	12.56 (6.05-26.08)
110-125	565	33	36	0.361	0.962	0.038	1.3233	0.662	246	14.38 (6.38-32.46)
115-125	565	18	36	0.278	0.985	0.015	1.2627	0.631	246	25.05 (9.13-68.75)
120-125	565	11	36	0.167	0.991	0.010	1.1572	0.579	261	20.96 (6.05-72.61)

African Americans

IFG Cut Point	n	No. Within IFG Limits	No. With Diabetes at Exam 4	Sensitivity	Specificity	1-Specificity	Sensitivity + Specificity	C-Statistic	BIC	OR for Incident Diabetes
90-125	1275	602	118	0.848	0.566	0.434	1.4136	0.702	637	7.25 (4.33-12.13)
95-125	1275	392	118	0.754	0.738	0.262	1.4923	0.750	602	8.65 (5.58-13.42)
100-125*	1275	230	118	0.661	0.867	0.131	1.5296	0.768	565	12.89 (8.49-19.58)
105-125	1275	129	118	0.5	0.940	0.061	1.4395	0.726	573	15.53 (10.06-23.97)
110-125	1275	71	118	0.331	0.972	0.028	1.3028	0.658	608	17.36 (10.32-29.20)
115-125	1275	42	118	0.237	0.988	0.012	1.2252	0.615	625	25.40 (12.91-49.96)
120-125	1275	22	118	0.136	0.995	0.005	1.1304	0.572	651	30.09 (11.52-78.58)

Hispanic Americans

IFG Cut Point	n	No. Within IFG Limits	No. With Diabetes at Exam 4	Sensitivity	Specificity	1-Specificity	Sensitivity + Specificity	C-Statistic	BIC	OR for Incident Diabetes
90-125	1021	523	93	0.839	0.51	0.480	1.359	0.680	589	5.64 (3.20-9.95)
95-125	1021	326	93	0.699	0.719	0.281	1.418	0.709	574	5.93 (3.72-9.45)
100-125*	1021	189	93	0.591	0.856	0.144	1.447	0.724	550	8.58 (5.46-13.48)
105-125	1021	108	93	0.505	0.934	0.066	1.440	0.720	526	14.52 (8.97-23.52)
110-125	1021	65	93	0.398	0.970	0.030	1.368	0.684	529	21.24 (12.13-37.19)
115-125	1021	44	93	0.2689	0.980	0.021	1.248	0.624	567	17.59 (9.22-33.54)
120-125	1021	20	93	0.140	0.993	0.008	1.132	0.566	597	21.38 (8.30-55.11)

*This category represents the current American Diabetes cut points for IFG.[14]

Table 7. Goodness-of-fit statistics by ethnic group for HbA1c as a predictor of incident diabetes.

Ethnic Group	**Model 1*: C Statistic**	**Model 2†: C Statistic**
All Groups (pooled)	0.79	0.82
Caucasian	0.80	0.87
Chinese American	0.81	0.82
African American	0.72	0.75
Hispanic American	0.82	0.84

*Crude model.
†Model containing fasting glucose, age, sex, clinic site, and BMI. Model for all groups pooled additionally included ethnic group.
Note: P-value for HbA1c x ethnic group interaction after adjusting for age, sex, clinic site, and BMI = 0.002.

Table 8. Prediction of cumulative incidence of diabetes at Exam 4 from fasting glucose and HbA1c at Exam 2.

Group	C-Statistic (Fasting Glucose)	C-Statistic (HbA1C)	C-Statistic (Fasting Glucose + HbA1C)
All Participants	0.82	0.79	0.84†‡
Caucasian	0.85	0.80	0.86†
Chinese	0.90	0.80*	0.90†
African American	0.75	0.73	0.78
Hispanic	0.83	0.81	0.86

* $p < .05$ compared to fasting glucose.
† $p < .05$ compared to HbA1c alone.
‡ $p < .05$ compared to fasting glucose alone.

6.6. *Figures*

Figure 1. Receiver-operating characteristic curves predicting cumulative incidence of diabetes at Exam 4 from fasting glucose at Exam 1.

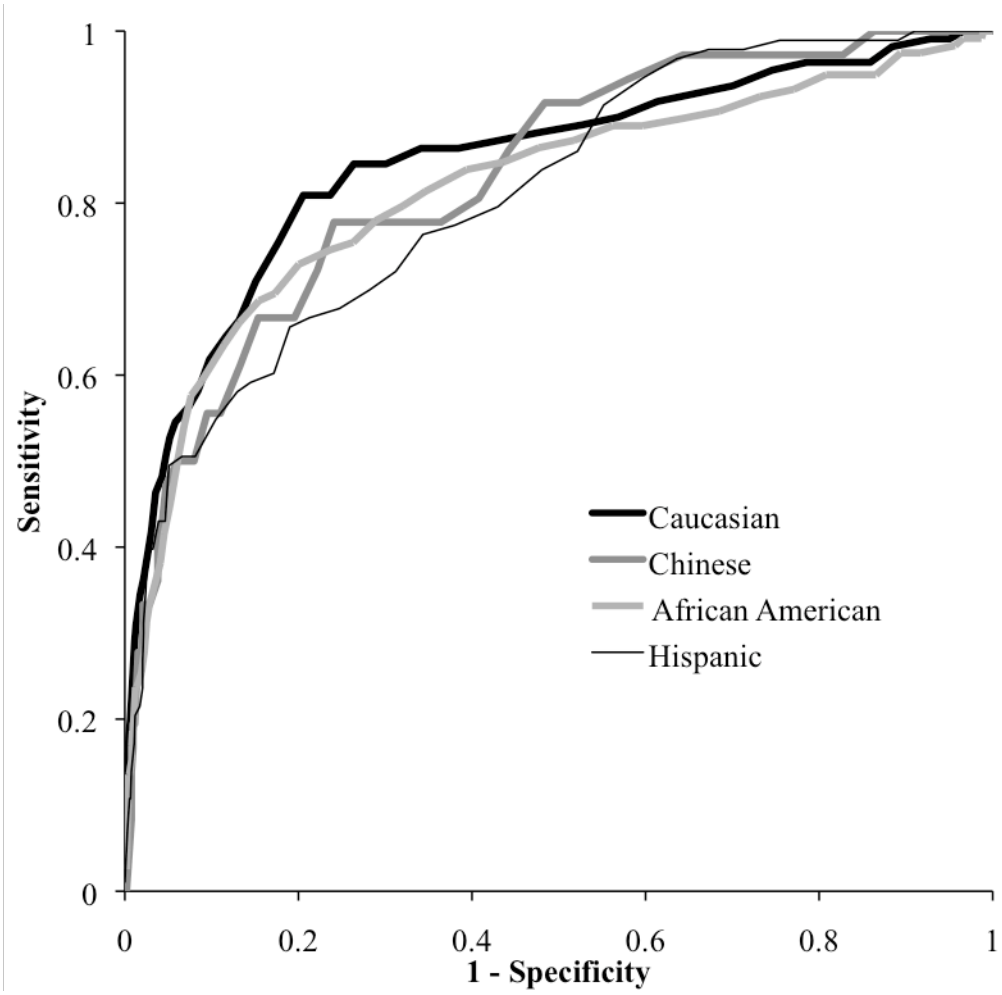
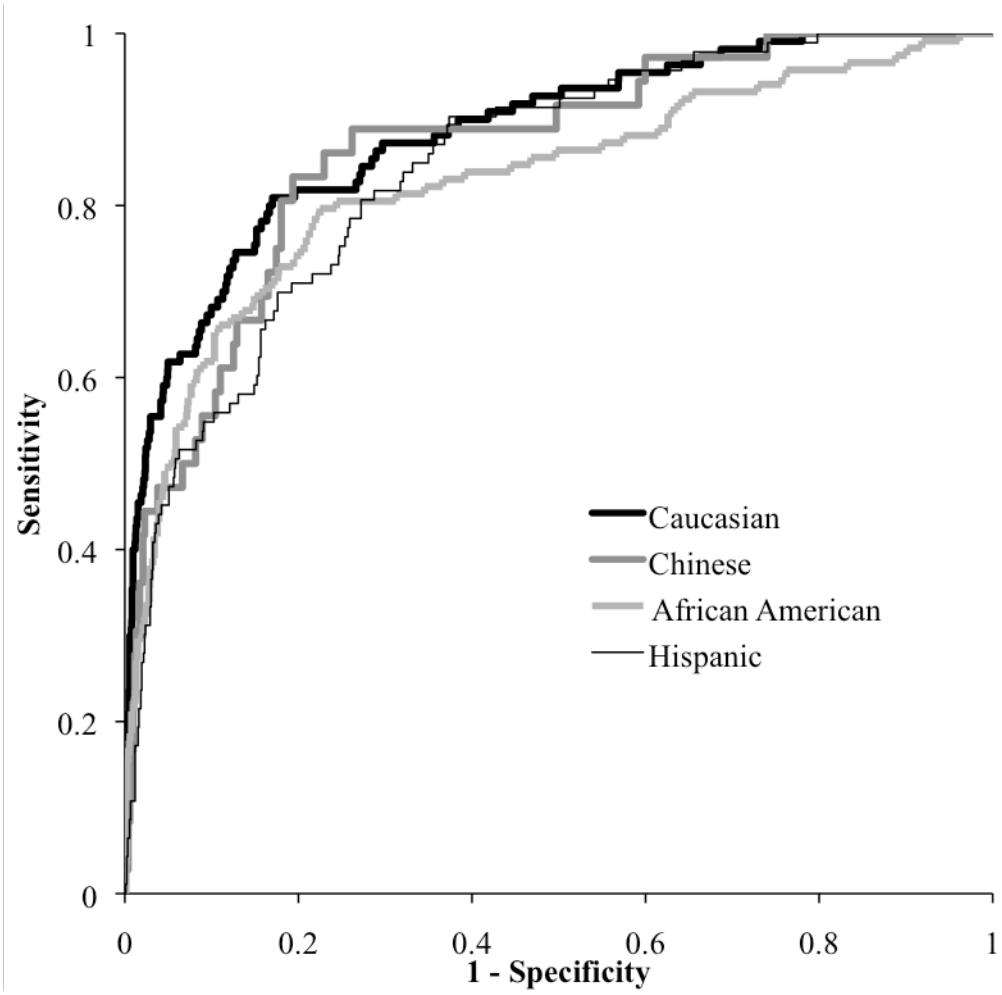


Figure 2. Model 2* receiver-operating characteristic curves predicting cumulative incidence of diabetes at Exam 4 from fasting glucose at Exam 1.



*Model 2 contained fasting glucose, age, sex, BMI, and clinic site.

Figure 3. Receiver-operating characteristic curves predicting cumulative incidence of diabetes at Exam 4 from HbA1c at Exam 2.

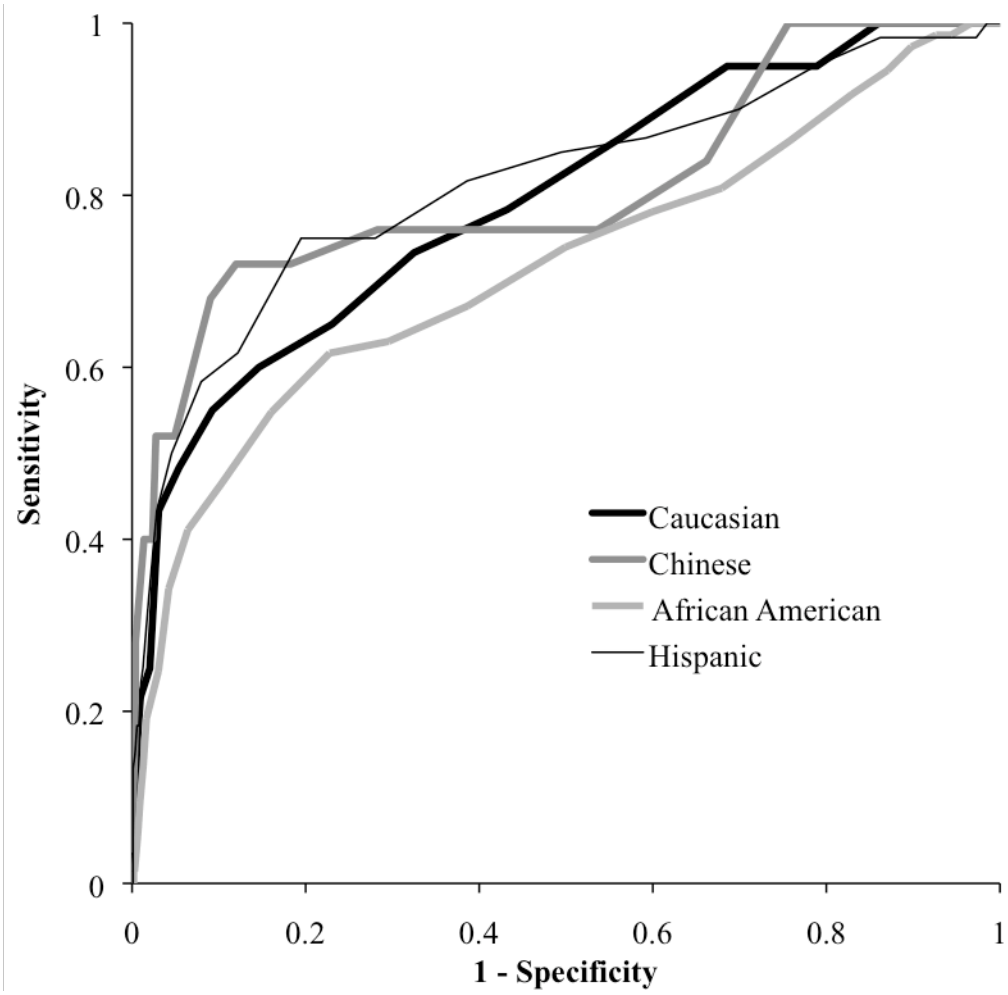
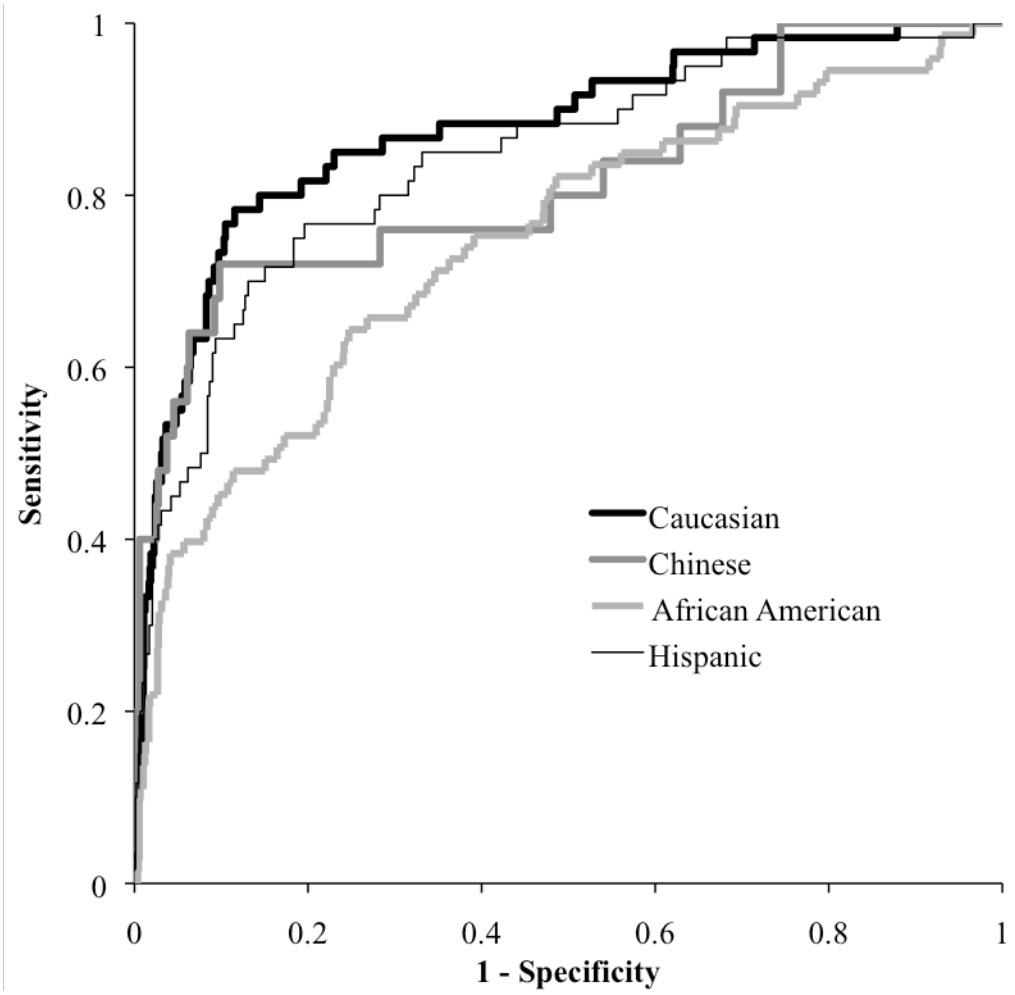


Figure 4. Model 2* receiver-operating characteristic curves predicting cumulative incidence of diabetes at Exam 4 from HbA1c at Exam 2.



*Model 2 contained HbA1c, age, sex, BMI, and clinic site at Exam 2.

7. MANUSCRIPT 2: Predictors of Incident Diabetes Among Individuals with Pre-Diabetes in the Coronary Artery Risk Development in Young Adults (CARDIA) Study

7.1 *Synopsis*

Objectives

This study examined predictors of progression from pre-diabetes to diabetes in the Coronary Artery Risk Development in Young Adults (CARDIA) study.

Research Design and Methods

CARDIA is a population-based observational research study that enrolled 5,115 Black and White men and women aged 18-30 years at enrollment. Participants were free of clinical cardiovascular disease (CVD) at baseline and were followed for 20 years. Unconditional logistic regression was used to evaluate predictors of incident diabetes; proportional hazards regression was used to evaluate the association or these predictors with time to progression to diabetes among pre-diabetes participants with impaired fasting glucose (IFG).

Results

In the CARDIA cohort, body mass index (BMI) was the strongest predictor of incidence odds of diabetes and time to progression to diabetes among individuals with pre-diabetes; other predictors included race/ethnicity, diastolic blood pressure, HDL-cholesterol, and alcohol consumption. After adjustment, the odds of diabetes were 1.87 times higher for per 7 kg/m^2 increment in BMI and the hazard ratio for BMI for the same increment was 1.52 among those initially with IFG.

Conclusions

Early attention to CVD and metabolic syndrome risk factors should be paid to pre-diabetic populations. Risk factor modification in this high-risk population subgroup can have important impact both on preventing the eventual development of diabetes or for delaying the length of time that passes before frank diabetes and associated complication are apparent.

7.2 *Introduction*

Individuals with pre-diabetes are at high risk for developing type 2 diabetes; lifestyle behavior changes including weight loss and physical activity are advised to prevent or delay diabetes onset among these high-risk individuals.[4,46,49,60,143,144] However, general advice to decrease body mass and increase physical activity is rarely successful in practice. Epidemiologic evidence suggests that the prevalence of diagnosed diabetes has increased over time while the prevalence of pre-diabetes has remained relatively stable and the prevalence of normoglycemia has fallen.[75] Thus, current measures to prevent progression from pre-diabetes to diabetes are insufficient. Identifying and understanding the independent and joint contributions of factors associated with progression to frank diabetes may inform efforts to reduce diabetes incidence rates, particularly in this high-risk pre-diabetes group.

Specific guidelines for risk reduction activities include decreasing caloric intake; following general recommendations for a diet lower in fat, higher in dietary fiber, lower in sodium intake, and void of excess alcohol consumption; engaging in moderate intensity physical activity for 30-60 minutes at least five days per week; and managing

CVD risk factors (e.g., hyperlipidemia and hypertension).[48,143] This advice is in line with general health promotion recommendations, but in high-risk populations, a more immediate improvement in the risk profile is often needed. Due to a potentially limited time window of opportunity to alter pre-diabetes progression, quantification of the magnitude of impact of individual risk factors on pre-diabetes progression can help to inform decisions regarding where resources should be focused. A few studies have assessed pharmacologic treatments in individuals with pre-diabetes and have found that this approach may be efficacious.[143,145] However, beginning pharmacologic treatment for pre-diabetes may not be universally accepted and may be cost-prohibitive for many individuals; potential associated side effects and safety concerns are also introduced along with adoption of anti-diabetic medications.[27,143,145,146]

Several risk scores have been developed for targeted detection of undiagnosed diabetes or for identification of individuals at high risk for subsequent diabetes in the general population.[83,84,87,89,92,93] In this study, we sought to develop a prediction model that identifies patient characteristics and behaviors associated with incidence of diabetes in a population of non-diabetic participants who meet the criteria for pre-diabetes in the Coronary Artery Risk Development in Young Adults (CARDIA) study. Thus, the primary goal of this investigation was to identify a set of factors with the most significant associations with progression to diabetes, and to examine the joint impact of these factors on progression to diabetes. For these analyses, we aimed to evaluate the strength of association between traditional risk factors and risk of progression to diabetes in a high-risk population [individuals with pre-diabetes, defined by the presence of either impaired fasting glucose (IFG) or impaired glucose tolerance (IGT)], as well as to potentially

identify new modifiable risk factors. Secondary analyses identified participant characteristics most strongly associated with time to progression to frank diabetes in individuals with IFG. Characteristics evaluated for both aims included: measures of adiposity, physical activity, dietary behaviors, smoking behaviors, alcohol consumption, blood pressure, lipids, insulin levels, access to health services, medication use, and psychosocial factors. Non-modifiable factors — such as race, age, and sex — were also considered as covariates.

7.3 *Methods*

Study Population

The CARDIA Study is a population-based longitudinal observational study initially designed to examine how risk factors for heart disease develop in 5,115 men and women aged 18-30 years at the time of enrollment between 1985-1986 at four geographically distributed U.S. sites (Birmingham, AL; Chicago, IL; Minneapolis, MN; and Oakland, CA). The sample was recruited to obtain balanced subgroups by sex, race (African American/non-Hispanic Black or White), education, and age. In brief, CARDIA collected information at baseline and at follow-up examinations 2, 5, 7, 10, 15, and 20 years later; at the year 20 exam, 72% of the surviving original cohort remained. This investigation utilized data from exam years 0, 7, 10, 15, and 20 (i.e., all exams where fasting glucose was measured). Information about diabetes status at Year 20 was required for inclusion in the analytic sample. Details regarding methodology for recruitment and enrollment are described elsewhere.[133]

Data Collection Methods

Diabetes status was determined by self-report or measured values. Treated cases were those where participants reported medication use for diabetes. Untreated diabetes cases were those detected at the clinic visit with fasting serum glucose values of 126 mg/dL and greater. Documentation of type of diabetes was not collected, but the analytic sample only included incident diabetes cases in non-pregnant adults and thus we expect that most diabetes cases included were of type 2 diabetes. Pre-diabetes cases were defined by the presence of either measured serum fasting glucose values of 100-125 mg/dL (IFG), or serum post-load glucose levels of 140-199 mg/dL (IGT).

Blood samples for fasting glucose were assayed using the hexokinase ultraviolet method on a Cobas Mira Plus chemistry analyzer. Individuals were excluded from measurement of fasting glucose if they did not fast for at least eight hours on the exam day, were diabetic and taking insulin and/or oral diabetes medications, were using steroids (not including inhalers), or were pregnant. Based on reassays of glucose in 2006 and 2007 in about 200 samples per examination drawn at years 7, 10, 15, and 20, glucose values were recalibrated to account for an observed drift in the assay. Recalibrated glucose values used in analyses were: 6.98 + 0.94 * year 7 glucose concentration, 7.15 + 0.96 * year 10 glucose concentration, 6.99 + 1.01 * year 15 glucose concentration, and 4.06 + 0.97 * year 20 glucose concentration.

Two-hour post-load serum glucose was measured in non-pregnant and non-diabetic participants at exam years 10 and 20 with an oral glucose tolerance test (OGTT). Participants were provided with 75 grams of a glucose solution (Glucola) and instructed to ingest the solution completely within a period not to exceed five minutes; study staff

checked to make certain that all of the solution was consumed. Glucose was then measured in the same manner as fasting glucose, two hours after consumption of the glucose solution. A red top Vacutainer tube was used to collect 5 ml of blood; this second glucose measurement was not obtained in participants that vomited after consumption of the solution.

Fasting serum insulin was measured at baseline and years 7, 10, 15, and 20 by radioimmunoassay. Participants were required to fast for a minimum of eight hours prior to measurement. As with fasting glucose, individuals were excluded from measurement of fasting insulin if they did not fast for at least eight hours on the exam day, were diabetic and taking insulin and/or oral diabetes medications, were steroid users (not including inhalers), or were pregnant. Samples for measurement of fasting insulin were collected in a Vacutainer tube containing no preservative. The sample was centrifuged within 60 minutes, stored in cryovials, and frozen at -70°C within 90 minutes until laboratory analysis. Based on reassays of specimens in 2006 and 2007 in 100 samples stored since year 15, insulin values at year 20 were recalibrated using the following formula: -0.36 + 0.93 * year 20 insulin concentration.

Systolic and diastolic blood pressures (SBP and DBP) were measured at each clinic exam in the seated position following a five-minute rest period in a quiet room. Three blood pressure measurements were obtained using a random zero sphygmomanometer, and were acquired a minimum of thirty seconds apart. The second and third blood pressure readings were averaged to produce mean resting SBP and DBP values.

A plasma lipid profile was obtained from a blood sample drawn from an anticubital vein (at all exams excluding year 2). Blood samples were collected in an EDTA Vacutainer tube, centrifuged for 20 minutes at 2000 RPM at 4°C, and frozen to -75° to -65°C until the sample was assayed at Northwest Lipid Laboratories in Seattle, WA. Details regarding the specific protocols used are described elsewhere.[133] Total cholesterol was determined using enzymatic methods, HDL-cholesterol was measured after dextran sulfate precipitation, and LDL-cholesterol was calculated using the Friedewald equation among those with fasting triglycerides ≤ 400 mg/dL.

Self-reported physical activity was measured using the CARDIA Physical Activity History,[136] an interviewer-administered modified version of the Minnesota Leisure Time Physical Activity Questionnaire.[147] The instrument assesses thirteen categories of activity covering various intensity levels. For each activity, a trained and certified interviewer recorded whether the activity was performed at any time in the prior year, the number of months that the activity was performed, and the number of months that the activity was performed. The total score represented total amount of moderate to vigorous physical activity expressed in exercise units. Further details of the scoring system have been described previously.[136]

Self-administered questionnaires were used to obtain additional information about demographics, access to health services, health behaviors, and psychosocial characteristics. Variables of interest included: educational attainment, smoking and alcohol use, family history of diabetes, access to and ability to pay for needed health services, dietary intake, and potential sources of psychological stress (depressive

symptoms, anxiety, anger expression, and job strain). Details regarding methodology and measurement tools used are described elsewhere.[133]

For use in analyses, the alcohol consumption variable was categorized into three levels: abstainers, those who consumed alcohol but on average less than two drinks per day, and those who consumed two or more drinks per day. Participants were asked, "Did you drink any alcoholic beverages in the past year?" and three follow-up questions ascertained how many drinks of wine, beer, and liquor were typically consumed per week. Assuming that one drink of beer, wine, or liquor contains 16.7 ml, 17.0 ml, or 19.1 ml of ethanol, respectively (per the CARDIA protocol), total ethanol consumption per day in milliliters of ethanol was averaged and divided by 17.24 ml of ethanol per average drink, to estimate the usual number of drinks per day that each participant reported.

Statistical Analysis

Participant characteristics at baseline were summarized by means and standard deviations for continuous variables, or by frequencies for categorical characteristics. These characteristics included age, sex, educational attainment, cardiovascular disease risk factors, and health behaviors. Participants with prevalent diabetes at baseline were excluded.

For the primary analysis, exam year 10 served as the baseline since both fasting glucose and post-load glucose data were available at this examination to more completely identify individuals with pre-diabetes. Therefore, individuals with prevalent diabetes or normoglycemia at this examination were excluded. At the year 10 exam, the cohort was old enough (aged 28-40) to accumulate a significant number of participants with pre-

diabetes (n = 400). Additionally, this time point was early enough to provide sufficient follow-up time to observe an adequate number of participants progressing to frank diabetes (10 years of follow-up). Participant characteristics were summarized in this subgroup, stratified by year 20 diabetes status. Student's t tests were used to compare study groups for continuous measures. Chi-square tests were used to test for differences between study groups for categorical characteristics.

Unconditional logistic regression was utilized to determine the associations between the predictors and cumulative incidence of diabetes at year 20 among individuals defined with pre-diabetes at year 10. For this analysis, factors considered included: adiposity, physical activity, dietary intake (food patterns and macronutrient intake), tobacco use, alcohol consumption, blood pressure, lipids, psychosocial factors (depression, anxiety, anger, hostility, and job strain), and access to health services (ability to access and pay for needed health services). Variables were selected based on previously published reports of independent associations with incident diabetes as well as those factors included in other prediction models and/or diabetes risk scores; a listing of these variables considered appears in Table 1.

An improvement-of-fit modeling approach was used based on predictive rather than etiologic associations. A backward stepwise method was employed within groups of related variables, where variables were sequentially removed from a fuller model. The aforementioned groups of variables evaluated demographics, biochemical and health history items, health behaviors, psychosocial variables, and finally, access to health services items. Variables with significant independent associations with incident diabetes status were retained; variables with no evident independent association with diabetes risk

were eliminated in sequence based on the *P* values. The goal of this analysis was to identify a set of factors comprising those variables with significant independent contributions, thus identifying the factors that most strongly predicted odds of progression to frank diabetes by year 20. The area under the receiver operating characteristic curve (AUC, or c statistic) was used as a measure of overall goodness-of-fit. The likelihood ratio chi-square was also used to assess improvement-of-fit during model building. Each variable considered was examined both for its independent association with incident diabetes as well as its contribution to the overall model fit.

A secondary analysis was conducted using data from all study examinations. Because fasting glucose was measured at five of the clinic visits, we were able to examine the association between diabetes risk factors and time to development of frank diabetes. This analysis included all individuals with incident IFG at any exam year prior to year 20. Proportional hazards regression was used to examine the effects of the predictors identified in the primary analysis on time to incident diabetes among individuals with IFG. The baseline exam was estimated as the first exam where an individual who had no history of diabetes met the criteria for IFG based on measured values (or year 0 for those with prevalent IFG at enrollment); participant characteristics at this exam were modeled as time-independent predictors for the proportional hazards regression model. Prior to modeling, the proportional hazards assumption was tested and confirmed by examining interactions between follow-up time and each predictor ($p > 0.05$ for all interactions). Hazard ratios and 95% confidence intervals were calculated; this analysis served to provide information about whether certain modifiable factors predict the time to conversion from IFG to frank diabetes. Because of the limitation of the

dataset — which does not contain exact amount of time from IFG to diabetes diagnosis (but rather only the clinic exams when IFG and diabetes were first documented) — these analyses were considered exploratory. All analyses were conducted using the SAS System, Version 9.2 (Cary, NC).

7.4 *Results*

Four hundred participants met the definition of pre-diabetes at Exam 10 (by meeting the definition for either IFG or IGT); 22% of these participants developed diabetes over a mean follow-up of 10 years (n = 86). This subgroup of 400 served as the analytic sample for identifying predictors of progression from pre-diabetes to diabetes. Compared to those who did not progress to frank diabetes over the 10-year follow-up period, the group that remained non-diabetic had a lower prevalence of Black participants; lower mean BMI and waist circumference, lower mean values for fasting glucose, blood pressure, and LDL cholesterol; and higher mean values for HDL cholesterol (Table 2).

In examination of univariate associations, the following characteristics (summarized in Table 3) were predictive of odds of incident diabetes: Black race/ethnicity, BMI, SBP, HDL-cholesterol, and alcohol consumption. For alcohol consumption, participants who reported drinking any alcohol were at decreased risk. Participants with pre-diabetes who were Black were at 1.66 times higher odds of developing diabetes compared to Caucasian participants (95% CI = 1.03-2.69). Additionally, those who reported consuming alcohol demonstrated lower odds for developing diabetes. The odds of diabetes increased almost twofold (OR = 1.92, 95% CI

= 1.50-2.42) for each standard deviation increment in BMI (7 kg/m^2), the odds ratio (OR) for diabetes for each 10 mmHg increment in DBP was 1.67 (95% CI = 1.30-2.13), and odds of diabetes were almost 40% lower per 13.5 mg/dL increment in HDL-cholesterol (OR = 0.66, 95% CI = 0.59-0.88). None of the other variables listed in Table 1 showed any association with diabetes incidence among pre-diabetic participants. Among the other groups of variables summarized in Table 1 that did not remain in the final model, the strongest psychosocial factor was anxiety (measured with the Spielberger State-Trait Anxiety Inventory[148]) and the strongest access to health services item was self-reported barriers to paying for health services; associations with these variables were not statistically significant ($p > 0.05$ for both).

Table 4 presents multivariable adjusted ORs. Of the demographic characteristics, the odds of diabetes remained higher among Black participants compared to Caucasian participants after adjustment for age, clinic site, and sex (Model 1, OR = 1.65, 95% CI = 1.00-2.73). The second model additionally included the clinical factors that showed significant univariate associations. The odds of diabetes were 67% higher for each 10 mmHg increment in DBP (OR = 1.61, 95% CI = 1.19-2.19), 30% lower per 13.5 mg/dL increment in HDL-cholesterol, and 1.87 per 7 kg/m^2 increment in BMI (95% CI = 1.37-2.56). The AUC for the prediction of diabetes from age, sex, clinic site, race, BMI, DBP, and HDL-cholesterol was 0.76. Adjusting for clinical factors attenuated the association between race and diabetes odds. After adjustment for demographic and clinical factors, alcohol use was not independently associated with diabetes nor did this variable contribute to the overall fit of the model (based on the c-statistic and 2df LR chi-square test) and therefore this variable was eliminated from the subsequent analyses.

Finally, we used proportional hazards regression to examine the association between the race, BMI, DBP, and HDL-cholesterol with time to development of diabetes among participants with IFG. For these analyses, there were data for fasting glucose available at exams years 0, 7, 10, 15, and 20. Because the OGTT was only conducted at two of the CARDIA exams, this analysis only focused on participants with pre-diabetes who had IFG and those with isolated IGT were not included. We excluded prevalent cases of diabetes at exams 0 and 7; the length of follow-up time ranged from 3-20 years.

In this analysis, 984 participants with IFG were identified; by exam year 20, 15% of these cases progressed to frank diabetes (n=151). The mean time period between first identification of IFG to first identification of incident diabetes was 7.2 years. After adjustment for age, sex, and clinic site, Black race, higher BMI, higher DBP, and lower HDL-cholesterol were each independently associated with increased rate of incident diabetes. The rate of diabetes was 67% higher among Black participants compared to Caucasian participants, 45% higher per 7 kg/m^2 increase in BMI, 23% higher per 10 mmHg increment in DBP, and 45% lower per 13.5 mg/dL increment in HDL-cholesterol (Table 5). In follow-up analyses, a statistically significant quadratic effect of BMI was present ($p = 0.01$), suggesting that there may be a non-linear component to the association between BMI and incident diabetes. In a model including BMI as a quadratic term, the associations with diabetes for race, DBP, and HDL-cholesterol were slightly attenuated but still remained statistically significant (HRs became 1.59, 1.21, and 0.70, respectively). Change in BMI during the follow-up time was not significantly associated with rate of incident diabetes in this sample of individuals with IFG ($p > 0.05$). Excluding

prevalent IFG at year 0 only slightly attenuated the associations; for this subgroup analyses there were 896 participants and 136 incident diabetes cases.

For all observed associations, additional analyses were conducted adjusting for use of hypertension medication, use of lipid-lowering medication, fasting glucose values, and post-load glucose values (for those analyses including participants with IGT). None of the additional covariates materially altered the observed associations and therefore they were removed from the final models. Additionally, interactions of predictors with race, age, and sex were considered; there was no evidence that the association between the factors identified and progression to diabetes differed by these key participant characteristics.

7.5 *Discussion*

As anticipated, in this population-based sample of young adults with pre-diabetes, BMI was the factor most strongly related to risk for development of diabetes. While other measures of adiposity were tested (including waist circumference and waist-to-hip ratio), additional measures did not significantly improve diabetes prediction. However, independent of BMI, DBP and HDL-cholesterol were also associated with diabetes risk. This evidence highlights the need for early attention to CVD and metabolic syndrome risk factors in pre-diabetic populations, a group where risk for CVD is already notably elevated.[149]

The final prediction model for the primary analysis had an AUC of 0.76. This statistic indicates a fairly good (but not perfect) prediction of incident diabetes events based on a small number of key risk factors, particularly in light of the fact that the

analytic sample was comprised in high-risk individuals with similar risk factors already present. Factors included in previously published diabetes risk scores that were not associated with increased risk of diabetes in our sample included age, family history of diabetes, smoking, sex, and physical activity.[83,89,92,150] A potential reason for this may be that our sample focused on a high-risk population that had higher prevalence of known diabetes risk factors at baseline compared to individuals with normoglycemia (e.g., physical inactivity). Additionally, while factors such as family history of diabetes and genetic predisposition are known to significantly contribute to one's risk for diabetes, they likely only have a small impact on whether or not a person who already has evidence of pre-diabetes progresses to diabetes.

The clinical factors identified in this study as predictors had strong associations with risk for diabetes and because of this, weaker factors that contributed to diabetes risk in univariate models did not contribute independently to diabetes prediction in multivariable adjusted models (the full list of variables considered appears in Table 1). For example, numerous studies show that psychological stress can exacerbate and/or accelerate many chronic disease processes (either directly through mechanisms including hypothalamic-pituitary-adrenal axis dysregulation or through behavioral factors such as poor compliance with lifestyle behavior changes).[151] However, none of the psychosocial factors (depression, anxiety, hostility, job strain, or anger) significantly contributed to the prediction model. A larger sample size with a wider distribution of potential risk factors would allow for the reevaluation of this issue, but given our findings no strong effect is likely; these risk factors may act primarily through their role in biological risk factors that are on the causal pathway to diabetes.

This study has several important strengths that should be noted, primarily with respect to the unique study population. First, the sample size of the CARDIA is large and provides twenty years of risk factor data for the time-to-event analyses (ten years of data for the primary analyses). Enrollment of young and disease-free adults at baseline allows for observation of the natural history of diabetes development from young adulthood until clinical disease onset and prior to concerns about confounding by medication use or pre-existing conditions. This aspect of the study design provides important information about participant characteristics prior to incident diabetes that strengthens inferences about temporality.

Both the fact that CARDIA study participants were enrolled as young adults and followed for twenty years (and that the study enrolled a large number of Black and White men and women) provides much needed data for predictors of diabetes in high-risk, diverse population. Further, while previous population-based cohort studies have looked at the prediction of incident diabetes in non-diabetic populations, this study focused on a defined high-risk group (individuals with pre-diabetes). Our results demonstrated that information regarding a few key participant characteristics could facilitate risk stratification to identify individuals within the population who are most likely to progress to diabetes. The identification of the strongest predictors demonstrates that these differences in risk factor levels, even in a group a markedly elevated risk compared to the general population, is predictive of likelihood of progressing from pre-diabetes to diabetes. Finally, an OGTT was conducted at the year 10 examination, allowing for inclusion of individuals with either IFG or IGT; the study, however, was underpowered to compare predictors between individuals with isolated IFG to those with isolated IGT.

Limitations of this study include the fact that the proportional hazards regression analysis was only conducted based on IFG criteria because the OGTT was not conducted at each study examination. Therefore, the results for this analysis may not be generalizable to all persons with pre-diabetes. In addition, hemoglobin A1c (a measure recently incorporated in the diagnostic and screening guidelines) was only measured at one study examination (year 20). Therefore, we were unable to incorporate this measure into our case definitions. Further, the health behavior information came from self-reported data and therefore misclassification of risk factors is always a concern. For example, dietary intake was assessed using a self-administered food frequency questionnaire (FFQ). A final limitation is that while participants were queried about duration of fasting prior to the blood draw for fasting glucose, some of these measures may not be true fasting values (and thus some participants might be misclassified in terms of IFG/diabetes status).

Although this study did not identify any novel risk factors for diabetes, the findings contribute important information about the magnitude of impact of factors associated with progression to diabetes in a high-risk group. Additionally, the longitudinal nature of the study allowed for the evaluation of contributors to the length of time from first identification with IFG to progression to diabetes. Risk factor modification — even in those individuals in whom we might not prevent the eventual development of diabetes — can have important impacts delaying the length of time that passes before frank diabetes and associated complications are apparent.

7.6 *Tables*

Table 1. List of participant characteristic considered as predictors for primary analysis.

Demographic Factors
Age
Race/Ethnicity
Sex
Clinic Site
Family history of diabetes
Educational attainment
Clinical/Physiological Factors
BMI
Waist circumference
Height
SBP
DBP
History of antihypertensive drug treatment
Total cholesterol
Triglycerides
HDL-cholesterol
History of lipid-lowering drug treatment
Insulin
HOMA-IR
Parity (in female participants)
Resting heart rate/pulse
Lifestyle/Behavioral Factors
Alcohol consumption
Macronutrient intake
Red meat consumption
Fruits, berries, and vegetable consumption
Whole-grain bread consumption
Coffee consumption
Self-reported moderate to vigorous physical activity
Cigarette smoking
Self-reported binge eating
Psychosocial Factors
Depressive symptoms (Center for Epidemiologic Studies Depression Scale)
Anxiety (Spielberger State-Trait Anxiety Inventory)
Anger (State-Trait Anger Expression Inventory, anger-in subscale)
Job strain (Karasek Job Strain Scale)
Hostility (Cook-Medley Hostility Scale)

Access to Health Services
Barriers to accessing basic health services
Barriers to paying for basic health services

Table 2. CARDIA Year 10 participant characteristics among pre-diabetes cases, stratified by diabetes status at Year 20.

	Mean (standard deviation) or n (%)		***P***
Characteristic	**Non-diabetic at Year 20 (n=314)**	**Diabetic at Year 20 (n = 86)**	
Age, yrs	35.6 (3.5)	35.6 (3.9)	0.99
Female	130 (41.4)	37 (43.0)	0.79
Black	143 (45.5)	50 (58.1)	<0.05
Field Center			0.41
Birmingham, AL	88 (28.0)	29 (33.7)	
Chicago, IL	53 (16.9)	10 (11.6)	
Minneapolis, MN	95 (30.3)	22 (25.6)	
Oakland, CA	78 (24.8)	25 (29.1)	
Education high school diploma or less	106 (33.8)	36 (41.9)	0.16
Body mass index, kg/m^2	29.1 (6.4)	33.9 (7.8)	<0.0001
Waist circumference, cm	91.4 (13.6)	100.5 (12.5)	<0.0001
Fasting glucose, mg/dL	101.4 (7.4)	105.2 (8.5)	<0.0001
Systolic blood pressure, mm Hg	114 (12)	119 (12)	<0.01
Diastolic blood pressure, mm Hg	75 (10)	80 (10)	<0.0001
Hypertensive	43 (13.8)	17 (28.3)	0.16
Total cholesterol, mg/dL	183.1 (37.7)	187.9 (34.2)	0.28
LDL cholesterol, mg/dL	112.9 (33.9)	117.4 (31.7)	0.27
HDL-cholesterol, mg/dL	45.6 (12.9)	41.1 (11.0)	<0.01
Triglycerides, mg/dL	121.9 (84.8)	154.8 (154.9)	<0.01
Hyperlipidemia	46 (14.7)	17 (19.8)	0.51
Ever smoker	135 (74.6)	37 (75.5)	0.89
Any alcohol in past year	253 (80.6)	65 (75.6)	0.31
Total physical activity score	308.9 (236.5)	277.8 (298.7)	0.31

Note: Physical activity was measured in exercise units, where 100 exercise units is approximately equivalent to 2-3 hours per week in vigorous activity for 6 months of the year. [136]

Table 3. Univariate significant predictors of incident diabetes among CARDIA Year 10 pre-diabetes participants (n=400).

Predictor	Odds Ratio	95% Confidence Interval	*P*
Black*	1.66	1.03-2.69	0.04
BMI (per 7 kg/m^2)	1.92	1.50-2.42	<0.0001
DBP (per 10 mm Hg)	1.67	1.30-2.13	<0.0001
HDL-cholesterol (per 13.5 mg/dL)	0.66	0.59-0.88	0.005
Alcohol†			
< 2 drinks per day	0.55	0.44-0.92	0.02
≥ 2 drinks per day	0.27	0.09-0.78	0.01

*Reference group = White participants.
†Reference group = non-drinkers.
Note: For continuous variables, ORs are per standard deviation.

Table 4. Odds ratios* for incident diabetes among CARDIA Year 10 pre-diabetes participants (n=400).

Predictor	Odds Ratio	95% Confidence Interval
Model #1 *(AUC=0.59)*		
Black†	1.65	1.00-2.73
Model #2: Model 1 + Clinical Factors *(AUC=0.76)*		
Black**	1.49	0.81-2.74
BMI (per 7 kg/m^2)	1.87	1.37-2.56
DBP (per 10 mm Hg)	1.61	1.19-2.19
HDL-cholesterol (per 13.5 mg/dL)	0.70	0.49-1.01
Model #3: Model 2 + Alcohol *(AUC=0.75)*		
Black†	1.26	0.70-2.33
BMI (per 7 kg/m^2)	1.74	1.31-2.31
DBP (per 10 mm Hg)	1.59	1.20-2.11
HDL-cholesterol (per 13.5 mg/dL)	0.80	0.56-1.14
Alcohol‡		
$<$ 2 drinks per day	0.67	0.38-1.18
$\geq$ 2 drinks per day	0.37	0.12-1.17

*All models adjusted for age, sex, and clinic site, in addition to other variables listed in the model.
†Reference group = White/Caucasian participants.
‡Reference group = non-drinkers.
Note: For continuous variables, ORs are per standard deviation derived from univariate associations.

Table 5. Adjusted* hazard ratios and 95% confidence intervals for the association of modifiable risk factors and other covariates with time to incident diabetes.

	Hazard Ratio	**95% Confidence Interval**
Black†	1.67	1.15-2.39
BMI (per 7 kg/m^2)	1.45	1.25-1.68
DBP (per 10 mm Hg)	1.23	1.09-1.38
HDL-cholesterol (per 13.5 mg/dL)	0.65	0.52-0.81

*All models adjusted for age, sex, and clinic site, in addition to other variables listed in the model.
†Reference group = White/Caucasian participants.
Note: For continuous variables, hazard ratios are per standard deviation derived from univariate associations.

8. MANUSCRIPT 3: Gene-Fitness Interactions and Type 2 Diabetes-Related Phenotypes: the CARDIA Study

8.1 *Synopsis*

Objectives

The purpose of this study was to evaluate the joint effects of single nucleotide polymorphisms (SNPs) of adiponectin (*ADIPOQ*), fat mass and obesity associated (*FTO*), and peroxisome proliferator-activated receptor gamma (*PPARG*) genes and cardiorespiratory fitness on diabetes-related measures among participants in the Coronary Artery Risk Development in Young Adults (CARDIA) cohort.

Research Design and Methods

CARDIA is a population-based observational research study that enrolled 5,115 Black and White men and women aged 18-30 years at enrollment. Participants were free of clinical cardiovascular disease (CVD) at baseline and were followed for 20 years. Unconditional logistic regression and linear regression were used to examine associations with SNPs and diabetes-related outcomes, as well as to evaluate potential gene-fitness interactions in relation to these outcomes.

Results

Several SNPs were associated with diabetes-related outcomes in the CARDIA cohort. Two SNPs showed significant evidence of interactions where the association between cardiorespiratory fitness and HbA1c levels may differ by genotype (rs8050136 of *FTO* in Caucasians and rs3856806 of *PPARG* in African Americans).

Conclusions

Increased risk of diabetes-related outcomes attributable to low cardiorespiratory fitness may be modified by genetic factors. Replication of these findings in other populations and further research regarding the role of genetic factors on fitness-diabetes associations would contribute to an improved understanding of the role of genetic risk factors and environmental/behavioral factors in impacting diabetes risk.

8.2 *Introduction*

Genetic factors predispose individuals to type 2 diabetes,[97,99,152] yet the pathways through which these attributes impact disease pathogenesis are unclear. Identifying interactions between genetic characteristics, environmental factors, and lifestyle behaviors is a requisite step in elucidating the joint impact of these variables on diabetes development.[96-99,153] Improved characterization of gene-environment interactions in the context of diabetes-related outcomes will both provide information regarding the etiology of diabetes as well as inform disease prevention efforts.

Cardiorespiratory fitness is associated with diabetes risk and refers to the degree to which the circulatory and respiratory system can supply oxygen to the body during sustained activity.[128,154] Low cardiorespiratory fitness is involved in the progression from normal glucose metabolism to diabetes; the gold standard for measurement of cardiorespiratory fitness is a test of maximal oxygen uptake (VO_2 max), typically using indirect calorimetry to assess oxygen consumption and carbon dioxide release.[155,156] Influenced by inherited factors and modifiable with physical training, cardiorespiratory

fitness plays an important role in glucoregulation and may alter observed associations between single nucleotide polymorphisms (SNPs) and diabetes-related characteristics.[28,29,97,128,129,156] Increased physical activity is considered a key component of diabetes prevention and management, both due to direct effects on blood glucose control and metabolism as well as indirectly by facilitating weight management.[20,26,29,123,129-131]

The adiponectin (*ADIPOQ*), fat mass and obesity associated (*FTO*), and peroxisome proliferator-activated receptor gamma (*PPARG*) genes have been linked to diabetes risk though hypothesized effects on insulin sensitivity and/or glucose metabolism. These genes are also associated with adiposity, cardiorespiratory fitness, and/or adaption of physical training.[98,103,106,108,115,117,118,157-161] The *ADIPOQ* gene (located at 3q27 and consisting of three exons and two introns that span a 17-kb region) encodes the circulating protein adiponectin; low adiponectin levels are associated with central adiposity, overall obesity, insulin sensitivity, and type 2 diabetes.[102-104] Variants of the *FTO* gene (located at 16q12.2) are most strongly associated with obesity-related traits that include body mass index, body weight, waist circumference, and body fat percentage.[106-109,162] However, recent investigations have also suggested links with type 2 diabetes risk and other metabolic abnormalities (though these observed associations are less consistent).[110,112] The protein encoded by *PPARG* has been associated with obesity, diabetes, and atherosclerosis;[114] *PPARG* (located at 3p25) is believed to regulate adipocyte differentiation and insulin sensitivity by promoting transcription of numerous target genes.[115]

This study evaluated the joint effects of SNPs of *ADIPOQ*, *FTO*, and *PPARG* and cardiorespiratory fitness on diabetes-related measures among African American and Caucasian participants in the Coronary Artery Risk Development in Young Adults (CARDIA) cohort. In this study, participants were enrolled between the ages of 18-30 years and were followed for twenty years. Associations between SNPs and incident diabetes, fasting glucose levels, fasting insulin levels, and hemoglobin A1c (HbA1c) levels measured after twenty years of follow-up were examined. Subsequent analyses evaluated interactions between SNPs and cardiorespiratory fitness level, with the use of standardized and objectively-measured cardiorespiratory fitness data (via a graded treadmill test).

8.3 *Methods*

Study Population

CARDIA is a population-based observational cohort study of 5,115 men and women aged 18-30 years enrolled between 1985-1986 at four geographic sites (Birmingham, AL; Chicago, IL; Minneapolis, MN; and Oakland, CA). Investigators aimed to examine how heart disease risk factors develop from young adulthood and recruited the sample to obtain balanced subgroups by gender, race/ethnicity (African American and Caucasian), education (high school or less and more than high school), and age (18-24 years and 25-30 years). At baseline, the study sample composition was 45.5% male and 51.5% African American or Black. Six follow-up examinations — that occurred 2, 5, 7, 10, 15, and 20 years after baseline — are currently complete, providing twenty years of observational data for 72% of the surviving original cohort. Details regarding

methodology for recruitment and enrollment are described elsewhere.[133] This study includes participants who attended the most recent examination and who had complete treadmill test data at baseline, blood glucose data at follow-up, and genotype information. The analytic sample included 3,648 CARDIA participants (1,669 African Americans and 1,979 Caucasians). An institutional review board at each site approved all study procedures; written informed consent was obtained from study participants prior to assessments.

SNP Selection and Genotyping

SNPs spanning the *ADIPOQ* and *PPARG* genes were selected based on the estimated pairwise linkage disequilibrium (LD), r^2, between all common SNPs (i.e., those with minor allele frequencies greater than 10%) as described by Carlson et al.[163] One SNP of the *FTO* gene was selected by the investigative team for genotyping based on observed associations with obesity-related traits in the literature. Our analysis included 10 *ADIPOQ* SNPs (rs1063539, rs17300539, rs17366743, rs182052, rs2036373, rs2241766, rs3774261, rs822393, rs822395, and rs9882205), one *FTO* SNP (rs8050136), and 15 *PPARG* SNPs (rs10865710, rs1797912, rs1801282, rs1822825, rs2881654, rs2938395, rs2959272, rs3856806, rs4135247, rs4135263, rs4135275, rs4135304, rs4135317, rs709151, and rs709156).

SNPs were genotyped using a multiplex matrix-assisted laser desorption/ionization time-of-flight mass spectrometry method (Sequenom, San Diego, CA). Targeted sequences were amplified by polymerase chain reaction using specific forward and reverse primers for each polymorphism, followed by extension reactions

utilizing specific oligonucleotide primers that annealed immediately upstream of each polymorphic site and extended the amplification product by a single base pair in the forward direction. The mass of the extension reaction products was determined and translated into genotype using the MassARRAY RT software (Sequenom Inc., San Diego, CA). Quality control (QC) activities employed by the CARDIA study included barcode identification of samples, robotic sample handling, and blind replicate genotype assessment on 5% of the total sample. Genotype concordance between QC pairs was ≥ 99% for all SNPs.

Data Collection Methods

Participants completed a symptom-limited maximal graded exercise treadmill test (GXT) at the baseline examination and at exam year 20, in which individuals were asked to exercise to maximal exertion. A modified Balke protocol was used with 2-minute stages.[134,164] Pulse rate, blood pressure, and 12-lead electrocardiograms were obtained on each subject. Every 2 minutes for up to 18 minutes (9 states), the speed or grade of the treadmill was increased to a maximum speed of 5.6 miles per hour and a 25% grade. Heart rate, blood pressure, and three-lead electrocardiogram data were obtained at the end of each stage, at maximum exercise, and at every minute for 3 minutes post-exercise. Self-reported ratings of perceived exertion (RPEs) were collected near the end of each test stage to assess symptoms.

Participants were eligible to participate in the GXT if they had no history of ischemic or congenital heart disease, were not using cardiovascular medications other than antihypertensive medications, had measured blood pressure values <160/90 mmHg,

and had no concurrent febrile illness. Exercise test data were considered valid if participants achieved 85% or greater of their age predicted maximum heart rate determined using the Tanaka formula (208 - 0.7*age).[137] Cardiorespiratory fitness was defined as the duration in seconds (up to 1,080 seconds, or 18 minutes) that participants were able to walk or run on the treadmill. The current study considered both baseline fitness and twenty-year change in fitness level.

Diabetes status was documented either as treated or untreated. Treated cases were those where participants self-reported medication use for diabetes. Untreated cases were those detected at the clinic visit based on measured fasting glucose. Fasting glucose values of 126 mg/dL and greater and with no previous reported diagnosis were coded as diabetes (untreated). For this study, both untreated and treated cases were pooled for analyses with diabetes as the outcome.

Blood samples for fasting glucose were assayed using the hexokinase ultraviolet method on a Cobas Mira Plus chemistry analyzer. Individuals were excluded from analysis of fasting glucose if they reported not fasting for at least eight hours on the exam day, were diabetic and taking insulin and/or oral diabetes medications, were using steroids (not including inhalers), or were pregnant. Based on reassays of glucose in 2006 and 2007 in about 200 samples per examination drawn at years 7, 10, 15, and 20, glucose values were recalibrated to account for an observed drift in the assay. Recalibrated glucose values at year 20 used for these analyses were 4.06 + 0.97 * year 20 glucose concentration.

As with fasting glucose, individuals were excluded from measurement of fasting insulin if they were diabetic and taking insulin and/or oral diabetes medications, were

using steroids (not including inhalers), or were pregnant. Individuals who reported not fasting for at least eight hours on the exam day were excluded from analyses. Samples for measurement is fasting insulin were collected in a Vacutainer tube containing no preservative. The sample was centrifuged within 60 minutes, stored in cryovials, and frozen at -70°C within 90 minutes until processing. Serum insulin was then measured by radioimmunoassay (Linco Research Inc., St. Charles, MO). Based on reassays of specimens in 2006 and 2007 in 100 samples stored since year 15, insulin values at year 20 were recalibrated using the following formula: -0.36 + 0.93 * year 20 insulin concentration.

For measurement of two-hour post-load glucose, participants were provided with 75 grams of a glucose solution (Glucola) and instructed to ingest the solution completely within a period not to exceed five minutes; study staff checked to make certain that all of the solution was consumed. Glucose was then measured in the same manner as the fasting glucose, two hours after consumption of the glucose solution. A red top Vacutainer tube was used to collect 5 ml of serum; this second glucose measurement was not taken in participants who vomited after consumption of the solution. Hemoglobin A1c was measured in whole blood at exam year 20 using a HbA1c sample preparation kit (HbA1c Sample Preparation Kit, Bio-Rad, Hercules, CA) and was collected in a capillary tube. The sample was centrifuged and stored at -70 C until processed.

Systolic and diastolic blood pressures were measured in the seated position after a five-minute rest period in a quiet room. Measurements were obtained before any physical examination, blood draw, or potentially stressful interview. Three blood pressure measurements were obtained using a random zero sphygmomanometer, and

were acquired a minimum of 30 seconds apart. The second and third readings were retained and averaged to produce mean resting systolic and diastolic blood pressure values. At exam year 20, a Omron automated device was used for blood pressure measurement. A comparability study was performed to ensure that the blood pressure values from the two devices used were consistent.

A plasma lipid profile was obtained from a blood sample drawn from an anticubital vein. Blood was collected into an EDTA Vacutainer tube, centrifuged for 20 minutes at 2000 RPM at 4°C, and frozen to -70°C until the sample was assayed at Northwest Lipid Laboratories in Seattle, WA. Details regarding the specific protocols used are described elsewhere. [133] Total cholesterol was determined using enzymatic methods, HDL-cholesterol was measured after dextran sulfate precipitation, and LDL-cholesterol was calculated using the Friedewald equation among those with fasting triglycerides $\leq$ 400 mg/dL.

Certified technicians measured height, weight and waist circumference using a standardized protocol previously described. [133] Body weight was measured to the nearest 0.2 kg, without shoes and with light clothing. Height and waist circumference were measured to the nearest 0.5 cm. Waist circumference was measured at the midpoint between the iliac crest and the lowest lateral portion of the rib cage, and anteriorly midway between the xiphoid process of the sternum and umbilicus. Body mass index (BMI) was calculated as weight in kilograms divided by height in meters squared. Information regarding sex, race/ethnicity, health history, health behaviors, and demographic information was self-reported either by questionnaire or via interview.

Statistical Analysis

Means and frequencies of year 20 participant characteristics were used to describe the study population and were stratified by self-reported race/ethnic group. These characteristics included age, sex, educational attainment, cardiovascular disease risk factors, diabetes prevalence, impaired fasting glucose (IFG) prevalence, and health behaviors. Participant characteristics were compared across race/ethnic groups using a t-test or χ^2 test as appropriate.

Deviations from Hardy-Weinberg equilibrium (HWE) — within self-reported race/ethnic group — were evaluated using Fisher's exact test for each SNP. Calculations were performed using the HaploView 4.2 software.[165] Since deviations can be due to various factors (e.g., population stratification or selection) and the specific cause is difficult to pinpoint[166], SNPs with deviations from HWE at the significance level of $\propto = 10^{-3}$ were explored further for potential coding errors. If no errors were found, these SNPs were discarded from the analytic sample to eliminate concerns about the validity of the genotyping.

Diabetes-related measures served as the dependent variables. Prevalent diabetes status at exam year 20 was modeled as a dichotomous outcome, and the following measures were considered as continuous outcomes: fasting glucose, fasting insulin, two-hour post-load glucose, and HbA1c. For those SNPs with significant associations with fasting insulin, associations with HOMA-IR were also examined but no significant findings were observed and therefore these data do not appear in the results section. All analyses were stratified by self-reported race/ethnic group to account for population stratification. Unconditional logistic regression was used to assess associations between

SNPs and Year 20 diabetes status. Linear regression was used for analyses with continuous outcomes; continuous outcomes were log-transformed to normalize their distributions using the natural logarithm and therefore all means presented are geometric means. Main effect associations between SNPs and diabetes-related variables were initially evaluated by 2df tests; models were adjusted for field center (Birmingham, Chicago, Minneapolis, or Oakland), age, and sex. The purpose of this step of the investigation was to identify the SNPs most strongly associated with diabetes-related measures, and thus the strongest candidates for gene-fitness interaction analyses. SNPs with *P* values for any association where $\propto < 0.10$ were retained for consideration in the interaction analyses.

For SNPs that satisfied the specified a priori significance level (0.10), the role of cardiorespiratory fitness (at baseline and change from baseline to year 20) as a potential effect modifier was evaluated. For genes where multiple SNPs met the criteria for further analysis, the Type I error rate was adjusted for multiple comparisons using a Bonferroni correction. For example, if three SNPs of a gene met the criteria for further analysis for the fasting glucose outcome, the Type I error rate used for evaluating interactions was 0.05/3=0.017. The analyses evaluating gene-fitness interactions assume an additive genetic model (1df), where each increasing risk allele is assumed to confer a proportional increase in disease risk; fitness was modeled as continuous variable. Interaction models were adjusted for age, sex, and clinic site; secondary models also adjusted for body mass index (BMI), given the modest correlation between cardiorespiratory fitness and adiposity. To further examine the effect of cardiorespiratory fitness and the polymorphisms on diabetes-related outcomes for those interactions that were significant,

mean levels at various points on the continuum of fitness were observed by genotype. All analyses were conducted using the SAS System, Version 9.2 (Cary, NC).

8.4 *Results*

Participant characteristics are described in Table 1 and are stratified by self-reported race/ethnicity. Compared to the African American participants, Caucasian participants were on average older and had higher mean total cholesterol levels. Additionally, the Caucasian group had a higher proportion of former smokers, female participants, and individuals with hypercholesterolemia. The proportion of participants with less then a high school education was higher in African Americans; this group also had higher mean BMI, fasting glucose, post-load glucose, fasting insulin, and HbA1c values. Blood pressure was higher in the African American group and HDL-cholesterol was lower. There was a higher frequency of hypertension in the African American group, as well as a higher prevalence of current smokers.

All SNPs were in HWE ($p>.001$) with the exception of rs1797912 in African Americans (Table 2). Therefore, this SNP was eliminated from all further analysis in African Americans. Figures 1 and 2 show the LD plots for the *ADIPOQ* SNPs by race/ethnic group. One pair of *ADIPOQ* SNPs had an r^2 of 0.88 in (rs2241766 and rs1063539 in both race/ethnic groups). Additionally, several pairs of *PPARG* SNPs had r^2 values greater than 0.80 (Figures 3 and 4).

Table 3 presents the p-values for the associations between each SNP and each outcome, for both race/ethnic groups. This analysis served as a screening step to identify those SNPs most strongly associated with several diabetes-related outcomes. Seven

ADIPOQ SNPs, eight *PPARG* SNPs, and the *FTO* SNP had at least one association test with $p<0.10$. Four *ADIPOQ* SNPs had associations with $p < 0.05$. The strongest association with incident diabetes was for rs709156 of *PPARG* in Caucasians ($p = 0.04$), for fasting glucose was rs1063539 of *FTO* in Caucasians ($p = 0.003$), for 2-hour post-load glucose was rs10801282 of *PPARG* in both African Americans and Caucasians ($p = 0.02$ and $p = 0.003$, respectively), for fasting insulin was rs2881654 of *PPARG* in African Americans and Caucasians ($p = 0.003$ and $p = 0.03$, respectively), and for HbA1c was rs1063539 of *FTO* in Caucasians ($p = 0.003$).

For all SNPs that showed associations at $\alpha = 0.10$, interactions between baseline cardiorespiratory fitness as well as 20-year change in fitness were evaluated (adjusted for age, sex, and clinic site). Four SNPs met the Bonferonni threshold of significance for interactions with baseline cardiorespiratory fitness and one SNP and 20-year change in cardiorespiratory fitness met the threshold of significance.

No genotype-fitness interactions were observed for the diabetes outcome. Only one interaction with baseline fitness in Caucasians was observed with *ADIPOQ*; this interaction was seen with rs2036373 and fasting insulin, but only after adjustment for BMI (Table 4). Additionally, the association between *ADIPOQ* SNP rs822393 and fasting insulin was significantly modified by 20-year change in fitness level in African Americans (Table 5, $p = 0.007$ without adjustment for BMI and $p = .003$ after adjustment for BMI); there was no interaction for the SNP and baseline fitness.

The rs8050136 of *FTO* and baseline fitness interaction was strong for the HbA1c outcome ($p = 0.003$ without adjustment for BMI and $p = 0.005$ after adjustment for BMI). The interaction was modestly significant for the fasting glucose outcome without

adjustment for BMI ($p = 0.03$), but became statistically non-significant after adjustment for BMI ($p = 0.06$). In Caucasians, a significant gene-baseline fitness interaction was observed for rs1801282 of *PPARG* both with and without adjustment for BMI. However, in follow-up analyses this significant interaction appeared to be driven by the GG genotype that contained only 25 participants. In African Americans, there was a significant interaction between rs3856806 and baseline fitness for HbA1c ($p = 0.001$ with and without adjustment for BMI). Excluding cases of treated diabetes did not materially alter any of the observed interactions (data not shown).

The two SNPs that showed the strongest and most consistent interactions were rs8050136 of *FTO* and baseline fitness with HbA1c in Caucasians, and rs3856806 of *PPARG* with baseline fitness and HbA1c in African Americans. To evaluate these SNPs further, we first examined the main effect associations between these SNPs and HbA1c levels (Table 6). The unadjusted and adjusted geometric means for the outcome variable were examined by genotype. Individuals with the AC or AA genotype for rs8050136 of *FTO* had higher mean levels of HbA1c compared to those with the CC genotype. There was a monotonic association with HbA1c across the genotypes, though for this SNP the mean differences between the heterozygote group and the reference group were not statistically significant. This pattern was consistent in the African American group, though there were no significant differences observed across genotypes. Individuals with the CT or TT genotype for rs3856806 of *PPARG* had higher mean levels of HbA1c compared to those with the CC genotype in African Americans. Again, this same pattern was observed in Caucasians but the difference in means was not statistically significant.

The interaction between rs8050136 of *FTO* and baseline fitness with HbA1c in Caucasians is illustrated in Figure 5. Individuals with the AA genotype tended to have higher HbA1c levels at low levels of fitness compared to the AC and CC groups; HbA1c levels were lower in individuals with the AA genotype at higher levels of fitness. For example, the predicted geometric mean was 6.4% for an individual with the AA genotype who was 25 years of age and whose treadmill duration was 2 minutes (using the reference level for all other covariates in the model), while the predicted value was 5.7% for an individual with the AC genotype and 5.6% for an individual with the CC genotype. However, the predicted geometric mean was 4.6% for an individual with the AA genotype who was 25 years of age and whose treadmill duration was 18 minutes, while the predicted value was 4.9% for an individual with either the AC or CC genotype. Further adjustment for BMI did not alter the significant findings. Additionally, we examined main effect associations between this SNP and baseline fitness and found preliminary evidence of an association between rs8050136 and baseline fitness after adjustment for age, sex, clinic site, and BMI ($p = 0.02$). Finally, we reexamined the interactions in African Americans and for 20-year change in fitness but none of these results were statistically significant (though a significant interaction by race was observed, $p < 0.05$).

The interaction between rs3856806 of *PPARG* and baseline fitness with HbA1c in African Americans is depicted in Figure 5. Individuals with the CC genotype had higher HbA1c levels at low levels of fitness compared to those with the CT or TT genotype. However, HbA1c levels decreased for those with the CC genotype as fitness levels increased. The predicted geometric mean was 6.7% for an individual with the CC

genotype who was 25 years of age and whose treadmill duration was 2 minutes (using the reference level for all other covariates in the model), while the predicted value was 5.9% for an individual with the CT or TT genotype. For an individual with the same characteristics with a treadmill duration of 18 minutes, the predicted mean HbA1c level was 4.4% for the CC genotype compared to 4.9% for the CT or TT genotype. No statistically significant associations were observed in Caucasians (though a significant interaction by race was observed, $p < 0.05$), a similar but statistically non-significant interaction was seen with 20-year change in fitness, and no main effect associations were observed between rs3856806 and baseline fitness level.

8.5 *Discussion*

In this study, we identified two SNPs with robust interactions where cardiorespiratory fitness may modify the association between genotype and HbA1c levels (rs8050136 of *FTO* in Caucasians and rs3856806 of *PPARG* in African Americans). For rs8050136, Caucasian participants with the AA genotype had significantly higher HbA1c levels in those with lower cardiorespiratory fitness compared to those with the AC or CC genotypes; the presence of high cardiorespiratory fitness appeared to remove the effect of the risk alleles. In African Americans for rs3856806 of *PPARG*, the effect of higher levels of cardiorespiratory fitness on HbA1c levels was stronger for those with the CC genotype compared to those with the CT or TT genotype. We examined similar directions of associations in the non-significant interactions of these SNPs with the 20-year change in fitness and with the other race group (African Americans for rs8050136 and Caucasians for rs3856806), which supports the consistency of our findings.

Additionally, rs3856806 of *FTO* was associated with cardiorespiratory fitness in a main effect model, suggesting that further evaluation of the role of fitness in the gene-diabetes association and replication in other populations is needed to confirm our findings; it is possible that cardiorespiratory fitness is actually on the causal pathway.

Two previous manuscripts published using CARDIA data found an association between objectively measured cardiorespiratory fitness and risk of diabetes, which persisted after adjustment for self-reported physical activity level and two measures of adiposity (body mass index and waist circumference).[128,129] These findings, in addition to previous literature, suggest that high aerobic fitness may be a trait with independent effects on the development of diabetes beyond contributing to caloric balance and weight management.[26,29,167] Because evidence suggests that cardiorespiratory fitness may play a unique role in the etiology of diabetes, better understanding how genetic factors and cardiorespiratory fitness might jointly influence one's likelihood of developing diabetes can help to better quantify the impact of modifiable and non-modifiable risk factors.

Previous studies have examined the influence of variants of *ADIPOQ*, *FTO*, and/or *PPARG* on risk for type 2 diabetes, but the mechanisms through which these variants affect diabetes risk are still unclear.[97,102-104,110-112,118,161] Therefore, despite the fact that we observed interactions between SNPs and cardiorespiratory fitness, we can only hypothesize about the potential modes of action. For example, the *ADIPOQ* gene encodes the circulating protein adiponectin that influences (among other functions) energy metabolism;[105] physiological characteristics associated with increased cardiorespiratory fitness levels may alter the action of this gene. Aside from the effect of improved fitness levels on caloric balance and body composition, higher physical fitness

levels also contribute directly to systemic vascular resistance, cardiac output, and insulin sensitivity.[128,130,155,168,169] The degree to which fitness level affects these phenotypes might be partially dependent of variants of specific genes. Thus, the presence of a significant SNP-fitness interaction could reflect the effect of fitness on the gene itself or gene variants influence fitness levels, which in turn impacts diabetes risk (preliminary evidence of this was seen for rs8050136 of *FTO* in Caucasians).

This study has important strengths that should be noted. First, CARDIA is a large, population-based bi-racial cohort, with large numbers of both Caucasian and African American participants who were recruited in young adulthood. Additionally, objectively measured cardiorespiratory fitness is quite rare for such a large, observational study. Most studies rely on self-reported physical activity that results in misclassification and biased estimates of fitness, with only high levels of physical activity meaningfully altering fitness categories. With a standardized and supervised protocol, the treadmill test offers a consistent measurement of performance that can be more easily compared between individuals.

The fact that the most consistent interactions that were observed were for the HbA1c outcome is worth noting. Levels of HbA1c represent the average level of blood glucose to which a red blood cell has been exposed during its lifetime, which typically lasts approximately 120 days.[6,33] Therefore, HbA1c levels reflect longer-term exposure to elevated glucose levels in the blood and this measure may be a more reliable indicator of chronic hyperglycemia (compared to measures such as fasting glucose levels that tend to demonstrate greater day-to-day fluctuations). Similar patterns of association were observed in the interaction analyses for rs8050136 and rs3856806 for the glucose and

diabetes outcomes, even though those interactions did not reach the level of statistical significance. Thus, future studies examining genetic markers of risk for diabetes-related outcomes may want to include assessment of associations with HbA1c levels.

This study is limited by the fact that some individuals might have been unable or unwilling to participate in the treadmill test. Additionally, the test was not a direct $V0_2$ max test (i.e., maximal oxygen consumption during exercise was not directly measured). Due to safety concerns, the assessment ended on self-reported symptoms rather on physiological markers. For this reason, there may be some bias in the fitness assessment results (which would most likely bias findings towards the null, since it is more likely that actual fitness level would be underestimated). The subjective nature of this component of the test should be acknowledged, but the protocol has been validated against the gold standard of the VO_2 maximum test and is a widely accepted alternative. We did not observe strong associations between any of the SNPs and 20-year change in fitness, which may be due to the fact that fitness levels in general declined for most participants over time in this population; intervention studies aimed at altering fitness levels might be better suited to address to question about whether changes in fitness levels alter the association between genetic factors and diabetes-related outcomes.

Our results suggest that increased risk of diabetes-related phenotypes attributable to genetic susceptibility may be exacerbated by the presence of low cardiorespiratory fitness. Further, improving fitness levels might decrease the likelihood of development of diabetes, particularly in genetically susceptible individuals. This study also more generally highlights the importance of considering gene-environment interactions in candidate gene studies for type 2 diabetes and other metabolic diseases. Identifying

individuals who would most benefit from interventions aimed at improving cardiorespiratory fitness can facilitate public health efforts that target high-risk individuals.

8.6 *Tables*

Table 1. CARDIA Year 20 participant characteristics by race/ethnicity [mean (standard deviation) or n (%)].

	African American (n=1669)	**Caucasian (n=1979)**	***P* value***
Age, yrs	44.7 (3.8)	45.8 (3.4)	<.0001
Female, %	1029 (61.7)	1067 (53.9)	<.0001
Field Center, %			<.0001
Birmingham, AL	469 (28.1)	408 (20.6)	
Chicago, IL	362 (21.7)	465 (23.5)	
Minneapolis, MN	314 (18.8)	615 (31.1)	
Oakland, CA	524 (31.4)	491 (24.8)	
Impaired fasting glucose, %	374 (22.4)	455 (23.0)	.68
Education less than high school, %	85 (5.0)	25 (1.5)	<.0001
Body mass index, kg/m^2	31.4 (7.6)	27.9 (6.5)	<.0001
Fasting glucose, mg/dL	101.5 (30.6)	97.9 (21.1)	<.0001
OGTT glucose, mg/dL	113.2 (42.3)	104.1 (34.8)	<.0001
Fasting insulin, mg/dL	16.7 (11.3)	13.7 (9.6)	<.0001
HbA1c, mean %	5.7 (1.0)	5.4 (0.7)	<.0001
Systolic blood pressure, mmHg	121 (17)	113 (13)	<.0001
Diastolic blood pressure, mmHg	77 (12)	70 (11)	<.0001
Hypertension (self-reported), %	557 (33.4)	324 (16.4)	<.0001
Total cholesterol, mg/dL	183.8 (35.8)	187.4 (34.3)	<.01
LDL cholesterol, mg/dL	109.9 (33.4)	109.9 (31.1)	.99
HDL-cholesterol, mg/dL	54.4 (16.3)	54.3 (17.2)	.93
Triglycerides, mg/dL	99.3 (68.8)	118.4 (86.4)	<.0001
Hypercholesterolemia (self-reported), %	330 (19.8)	508 (25.7)	<.0001
Current Smoker, %	406 (24.3)	296 (15.0)	<.0001
Former Smoker, %	226 (13.5)	468 (23.6)	<.0001

*P-values from t-test for continuous characteristics or χ^2 for categorical characteristics.

Table 2. Hardy-Weinberg equilibrium tests for single nucleotide polymorphisms (SNPs) by self-reported race/ethnic group.

			African American Participants			Caucasian Participants		
SNP	**Position (base pairs)**	**Location**	**Minor allele**	**Minor allele frequency**	**HWE p-value***	**Minor allele**	**Minor allele frequency**	**HWE p-value***
ADIPOQ								
rs17300539	186559460	5' UTR	A	.015	1	A	.089	.12
rs182052	186560782	intron	A	.376	.33	A	.351	.77
rs822393	186566326	intron	T	.355	.86	T	.256	.04
rs822395	186566807	intron	C	.409	.31	C	.345	.21
rs2036373	186570191	intron	G	.059	.11	G	.068	.99
rs9882205	186570398	intron	A	.196	.003	A	.285	.80
rs2241766	186570892	exon	G	.042	1	G	.114	.79
rs3774261	186571559	intron	G	.462	.18	A	.397	.04
rs17366743	186572089	exon	C	.009	1	C	.027	1
rs1063539	186575392	3' UTR	C	.044	.07	C	.12	.88
FTO								
rs8050136	53816275	intron	A	.437	.39	A	.392	.36
PPARG								
rs10865710	12353198	intron	G	.224	.16	G	.274	.08
rs1801282	12393125	exon	G	.028	.06	G	.132	.43
rs4135304	12394601	intron	A	.408	.08	A	.006	.08
rs4135247	12396588	intron	G	.233	.43	G	.399	.77
rs2881654	12396955	intron	G	.429	.002	A	.145	.11
rs4135317	12408950	intron	T	.062	.48	T	0	1
rs4135263	12423266	intron	C	.117	.28	C	.148	.35
rs2938395	12429468	intron	C	.204	.66	C	.379	.48
rs2959272	12442833	intron	A	.405	.63	C	.473	.09

			African American Participants			Caucasian Participants		
SNP	**Position (base pairs)**	**Location**	**Minor allele**	**Minor allele frequency**	**HWE p-value***	**Minor allele**	**Minor allele frequency**	**HWE p-value***
rs4135275	12443844	intron	G	.076	.57	G	.18	.13
rs1822825	12449963	intron	C	.211	.52	C	.466	.06
rs709151	12454999	intron	T	.151	.46	T	.346	.84
rs709156	12461615	intron	A	.134	.12	A	.311	.88
rs1797912	12470239	intron	C	.152	**5.62×10^{-11}**	C	.366	1
rs3856806	12475557	exon	T	.06	.54	T	.128	.77

*Hardy-Weinberg Equilibrium p-values derived from exact test.

Table 3. P-values for associations of SNPs with Year 20 diabetes outcomes, stratified by self-reported race/ethnic group.

	Prevalent Diabetes	Fasting Glucose	2-hour Post-Load Glucose	Fasting Insulin	HbA1c
ADIPOQ					
rs1063539					
African American*	.93	.47	.83	1.0	.46
Caucasian	.98	.15	.18	.46	.44
rs17300539					
African American*	.38	.46	.30	.98	**.06**
Caucasian	.31	.78	.29	.90	.97
rs17366743					
African American*	.13	.27	**.09**	.21	.14
Caucasian*	.87	.86	.86	.36	.12
rs182052					
African American	.18	**.09**	.48	.33	.72
Caucasian	.27	.83	.47	.67	.60
rs2036373					
African American	.82	1.0	.66	**.03**	.49
Caucasian*	**.08**	.11	**.06**	**.05**	**.01**
rs2241766					
African American*	.99	.23	.50	.47	.35
Caucasian	.86	.23	.38	.30	.68
rs3774261					
African American	.14	.18	**.03**	.29	.46
Caucasian	.61	.61	**.009**	.21	**.09**
rs822393					
African American	.98	.14	.22	**.06**	.56

	Prevalent Diabetes	Fasting Glucose	2-hour Post-Load Glucose	Fasting Insulin	HbA1c
Caucasian	.89	.88	**.003**	.79	.35
rs822395					
African American	.67	.17	**.02**	**.08**	.37
Caucasian	.20	.35	**.02**	.77	.93
rs9882205					
African American	.41	.45	.99	.28	.28
Caucasian	.11	.35	**.05**	.72	**.09**
FTO					
rs8050136					
African American	.20	.12	**.03**	**.06**	.34
Caucasian	**.09**	**.003**	.89	**.03**	**.003**
PPARG					
rs10865710					
African American	.62	.52	.18	**.09**	.98
Caucasian	.12	**.096**	.65	.93	.17
rs1797912					
African American*	--	--	--	--	--
Caucasian	.16	.23	.62	.89	.88
rs1801282					
African American	.89	.35	**.02**	**.02**	.95
Caucasian	.68	.26	**.003**	**.09**	.46
rs1822825					
African American	.37	.83	.81	.47	.88
Caucasian	.26	.59	.24	.17	.83
rs2881654					
African American	.47	.11	**.05**	**.003**	.33
Caucasian	.13	**.02**	**.003**	**.03**	.42

	Prevalent Diabetes	Fasting Glucose	2-hour Post-Load Glucose	Fasting Insulin	HbA1c
rs2938395					
African American	.68	**.04**	**.09**	.86	.31
Caucasian	.26	.79	.96	.31	.17
rs2959272					
African American	.84	.79	.15	.21	.95
Caucasian	.41	.42	.11	.26	.75
rs3856806					
African American*	**.08**	.54	**.06**	.63	**.04**
Caucasian	.67	.52	**.04**	.54	.61
rs4135247					
African American	.65	.57	.65	**.05**	.33
Caucasian	.37	.97	1.0	.93	.57
rs4135263					
African American	.12	.19	.63	**.02**	.29
Caucasian	.17	.13	.36	.65	.83
rs4135275					
African American*	.32	.43	.41	.69	.15
Caucasian	**.04**	.11	.61	.28	.85
rs4135304					
African American	.69	.68	.88	.12	.26
Caucasian*	.24	.20	.14	.36	.43
rs4135317					
African American*	.67	.32	.30	.19	.12
Caucasian	--	--	--	--	--
rs709151					
African American	.61	.21	.11	.67	.60
Caucasian	.16	.34	.99	.32	.86

	Prevalent Diabetes	Fasting Glucose	2-hour Post-Load Glucose	Fasting Insulin	HbA1c
rs709156					
African American	.87	.32	.26	.52	.50
Caucasian	**.04**	**.09**	.98	.74	.38

*1df test (because of genotype with no participants, or because genotypes with less than 10 participants were pooled with the heterozygote group).
Note: Models were adjusted for age, sex, and clinic site.

Table 4. P-values for interactions of SNPs and baseline cardiorespiratory fitness with year 20 diabetes-related outcomes.

	Prevalent Diabetes	**Fasting Glucose**	**2-hour Post-Load Glucose**	**Fasting Insulin**	**HbA1c**
ADIPOQ					
rs17300539					
African American	--	--	--	--	.91 / .71
rs17366743					
African American	--	--	.43 / .52	--	--
rs182052					
African American	--	.08 / .15	--	--	--
rs2036373					
African American	--	--	--	.95 / .62	--
Caucasian	.91 / .53	--	.60 / .84	.17 / **.02**	.68 / .17
rs3774261					
African American	--	--	.68 / .67	--	--
Caucasian	--	--	.65 / .79	--	--
rs822393					
African American	--	--	--	.56 / .33	--
Caucasian	--	--	.12 / .10	--	--
rs822395					
African American	--	--	.15 / .16	.78 / .77	--
Caucasian	--	--	.02 / .02		--
rs9882205					
Caucasian	--	--	.05 / .05	--	.05 / .05
FTO					
rs8050136					
African American	--	--	.83/ .71	.87 / .72	--
Caucasian	.06 / .15	**.03** / .06	--	1.0 / .73	**.003** / **.005**

	Prevalent Diabetes	Fasting Glucose	2-hour Post-Load Glucose	Fasting Insulin	HbA1c
PPARG					
rs10865710					
African American	--	--	--	.86 / .79	--
Caucasian	--	1.0 / .60	--	--	--
rs1801282					
African American	--	--	.71 / .69	.54 / .58	--
Caucasian	--	--	**.003** / **.009**	.08 / .38	--
rs2881654					
African American	--	--	.81 / .83	.79 / .95	--
Caucasian	--	.54 / .12	.02 / .05	.11 / .47	--
rs2938395					
African American	--	.55 / .54	.82 / .80	--	--
rs3856806					
African American	.09 / .07	--	.52 / .54	--	**.001** / **.001**
Caucasian	--	--	.22 / .27	--	--
rs4135263					
African American	--	.07 / .15	--	.12 / .11	--
rs4135275					
Caucasian	.95 / .93	--	--	--	--
rs709156					
Caucasian	.91 / .64	.24 / .07	--	--	--

Note: Interactions were examined only for those SNPs, outcomes, and race groups where associations were observed in Table 1.The first p-value listed is from a model adjusted for age, sex, and clinic site. The second p-value listed is from a model additionally adjusted for BMI.

Table 5. P-values for interactions of SNPs and 20-year change in cardiorespiratory fitness with year 20 diabetes-related outcomes.

	Prevalent Diabetes	Fasting Glucose	2-hour Post-Load Glucose	Fasting Insulin	HbA1c
ADIPOQ					
rs17300539					
African American	--	--	--	--	.34 / .26
rs17366743					
African American	--	--	.35 / .41	--	--
rs182052					
African American	--	.65 / .60	--	--	--
rs2036373					
African American	--	--	--	.71 / .44	--
Caucasian	.58 / .25	--	.55 / .23	.75 / .47	.37 / .25
rs3774261					
African American	--	--	.86 / .82	--	--
Caucasian	--	--	.57 / .67	--	--
rs822393					
African American	--	--	--	.03 / **.01**	--
Caucasian	--	--	.65 / .52	--	--
rs822395					
African American	--	--	.07 / .09	.37 / .51	--
Caucasian	--	--	.46 / .64		--
rs9882205					
Caucasian	--	--	1.0 / .91	--	.53 / .42
FTO					
rs8050136					
African American	--	--	.69/ .56	.86 / .68	--
Caucasian	1.0 / .74	.38 / .61	--	.61 / 1.0	.20 / .42

	Prevalent Diabetes	Fasting Glucose	2-hour Post-Load Glucose	Fasting Insulin	HbA1c
PPARG					
rs10865710					
African American	--	--	--	.87 / .54	--
Caucasian	--	.02 / .02	--	--	--
rs1801282					
African American	--	--	.72 / .71	.55 / .45	--
Caucasian	--	--	.41 / .52	.72 / .91	--
rs2881654					
African American	--	--	.43 / .41	.52 / .34	--
Caucasian	--	.54 / .54	.20 / .26	.74 / .76	--
rs2938395					
African American	--	.63 / .90	.44 / .38	--	--
rs3856806					
African American	.32 / .41	--	.09 / .09	--	.24 / .23
Caucasian	--	--	.70 / .61	--	--
rs4135263					
African American	--	.57 / .52	--	.30 / .32	--
rs4135275					
Caucasian	.66 / .82	--	--	--	--
rs709156					
African American	--	--	--	--	--
Caucasian	.48 / .83	.28 / .65	--	--	--

Note: Interactions were examined only for those SNPs, outcomes, and race groups where associations were observed in Table 1. The first p-value listed is from a model adjusted for age, sex, and clinic site. The second p-value listed is from a model additionally adjusted for BMI.

Table 6. Mean HbA1c values by genotype for SNPs with significant interactions with cardiorespiratory fitness.

HbA1c (%)						
Gene	*SNP*	*Race/ Ethnicity*	*Unadjusted Mean* (95% CI)*	*P*	*Adjusted* Mean (95% CI)*	*P*
FTO	rs8050136	Caucasian				
	CC (n=697)		5.3 (5.3-5.4)	ref	5.3 (5.3-5.3)	ref
	AC (n=925)		5.3 (5.3-5.4)	0.33	5.3 (5.3-5.4)	0.19
	AA (n=282)		5.4 (5.4-5.5)	0.002	5.4 (5.4-5.5)	0.0007
PPARG	rs3856806	African American				
	CC (n=1085)		5.6 (5.6-5.7)	ref	5.6 (5.6-5.7)	ref
	CT/TT (n=159)		5.8 (5.6-5.9)	0.05	5.8 (5.6-5.9)	0.04

*Geometric means. Adjusted means are adjusted for age, sex, and clinic site.

8.7 ***Figures***

Figure 1. Pairwise linkage disequilibrium (r^2) for *ADIPOQ* SNPs among African American participants.

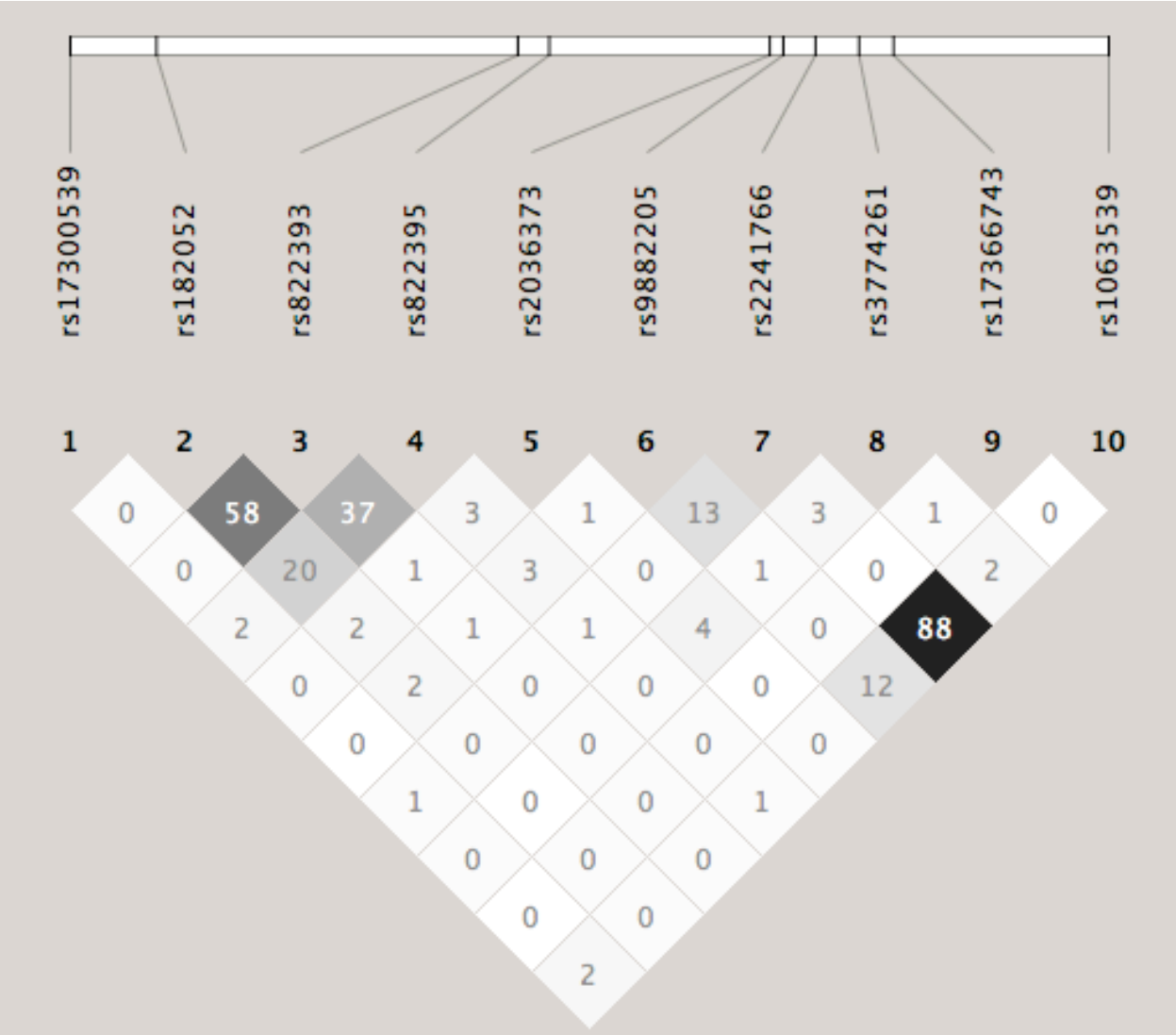


Figure 2. Pairwise linkage disequilibrium (r^2) for *ADIPOQ* SNPs among Caucasian participants.

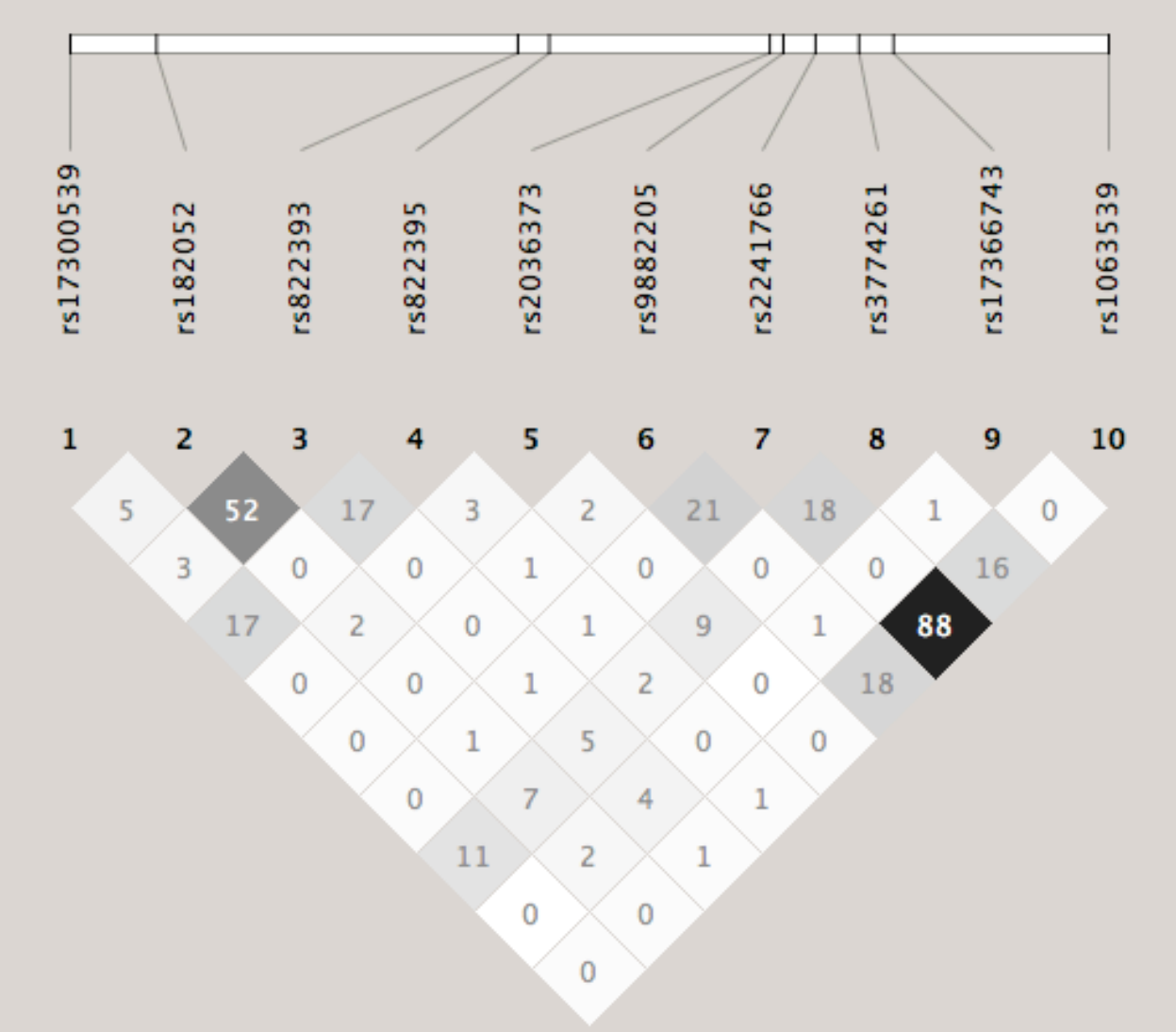


Figure 3. Pairwise linkage disequilibrium (r^2) for *PPARG* SNPs among African American participants.

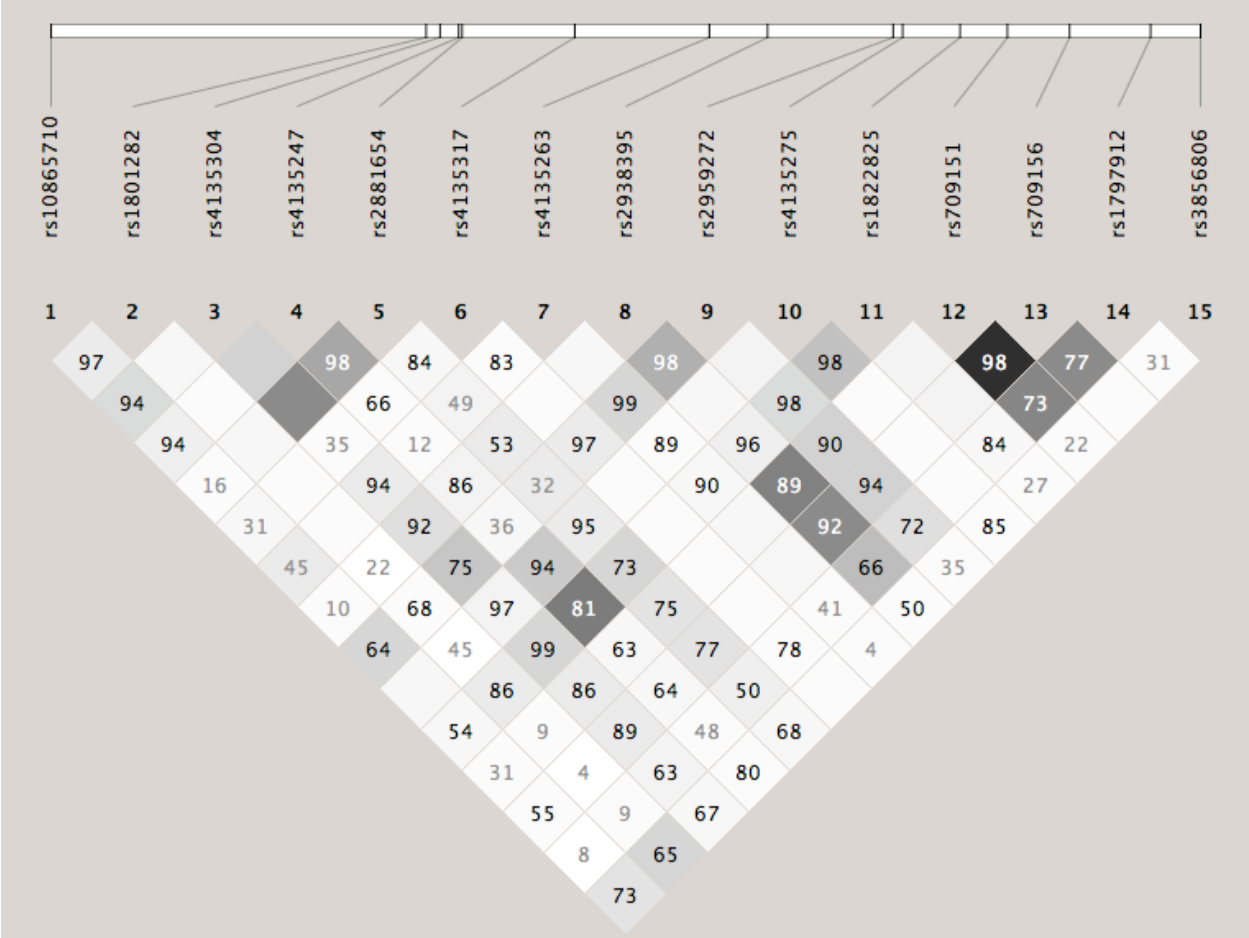


Figure 4. Pairwise linkage disequilibrium (r^2) for *PPARG* SNPs among Caucasian participants.

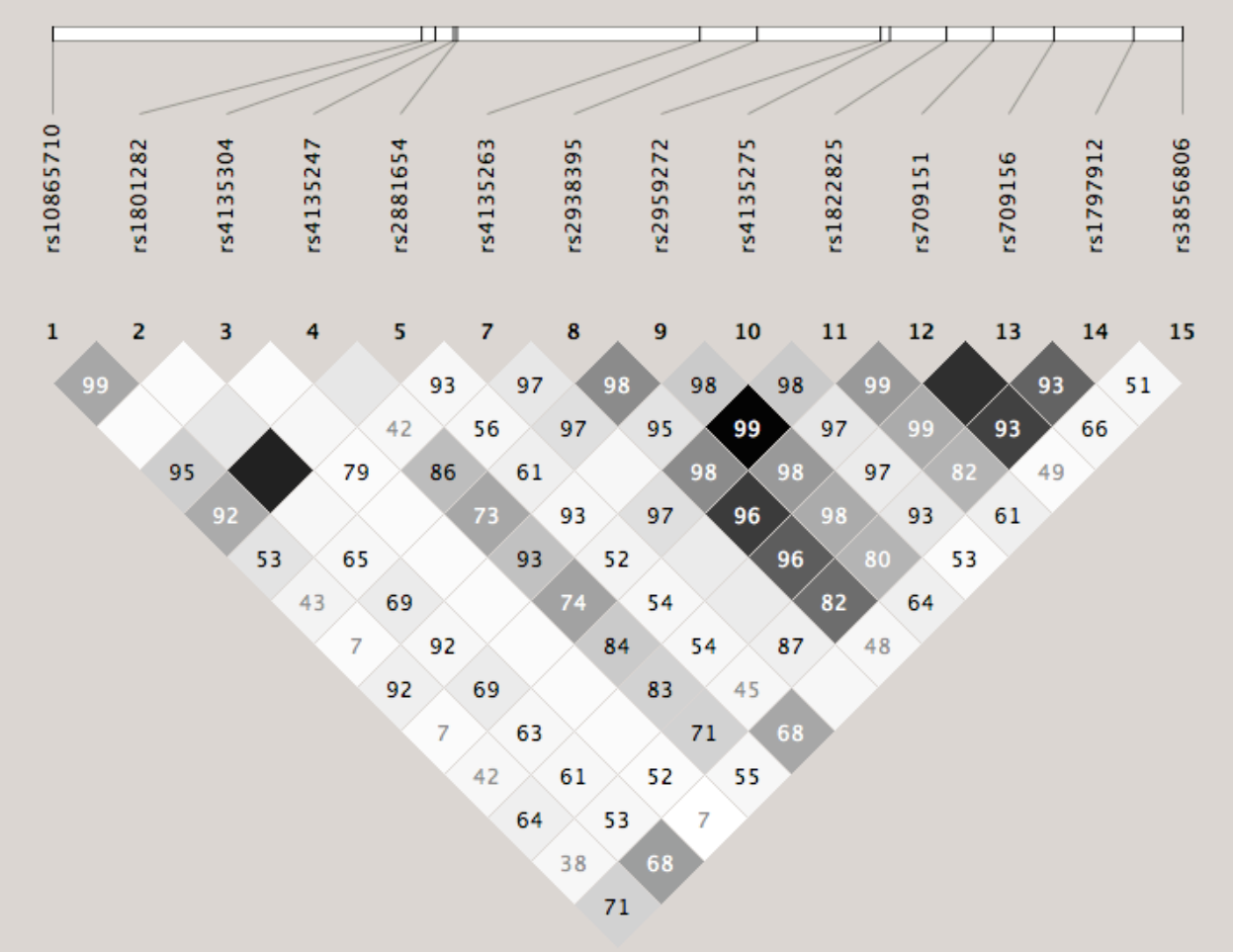
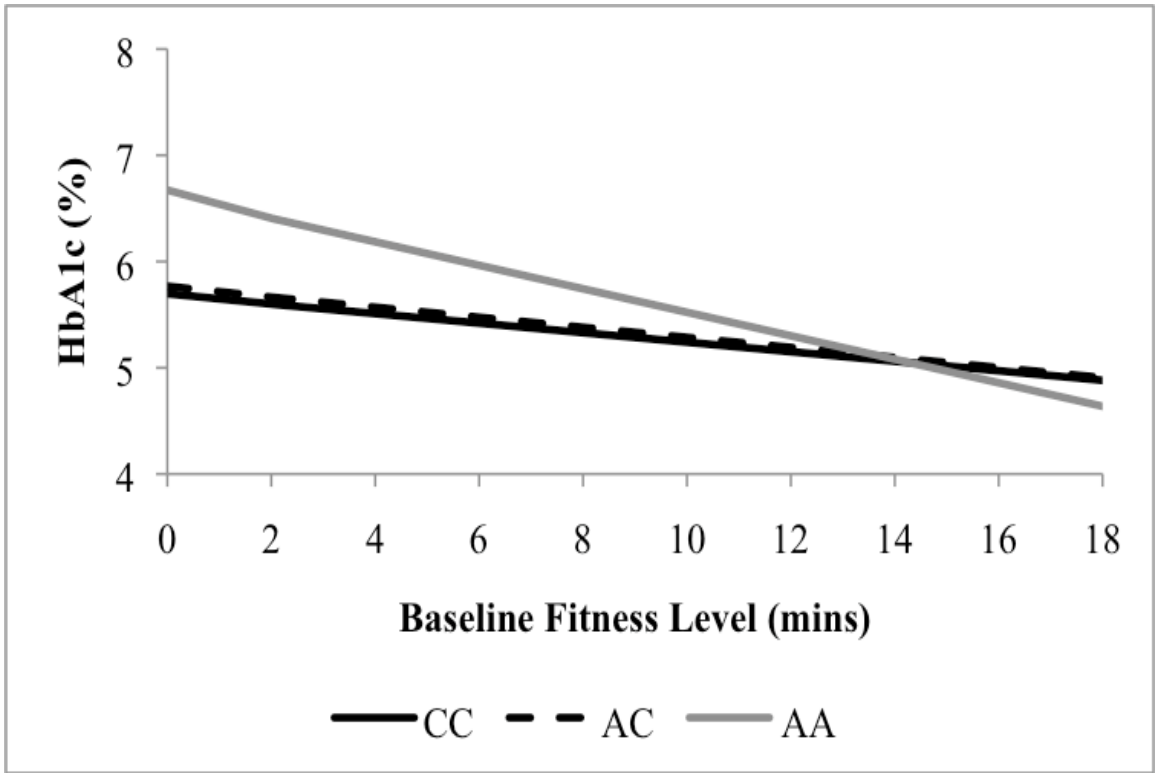
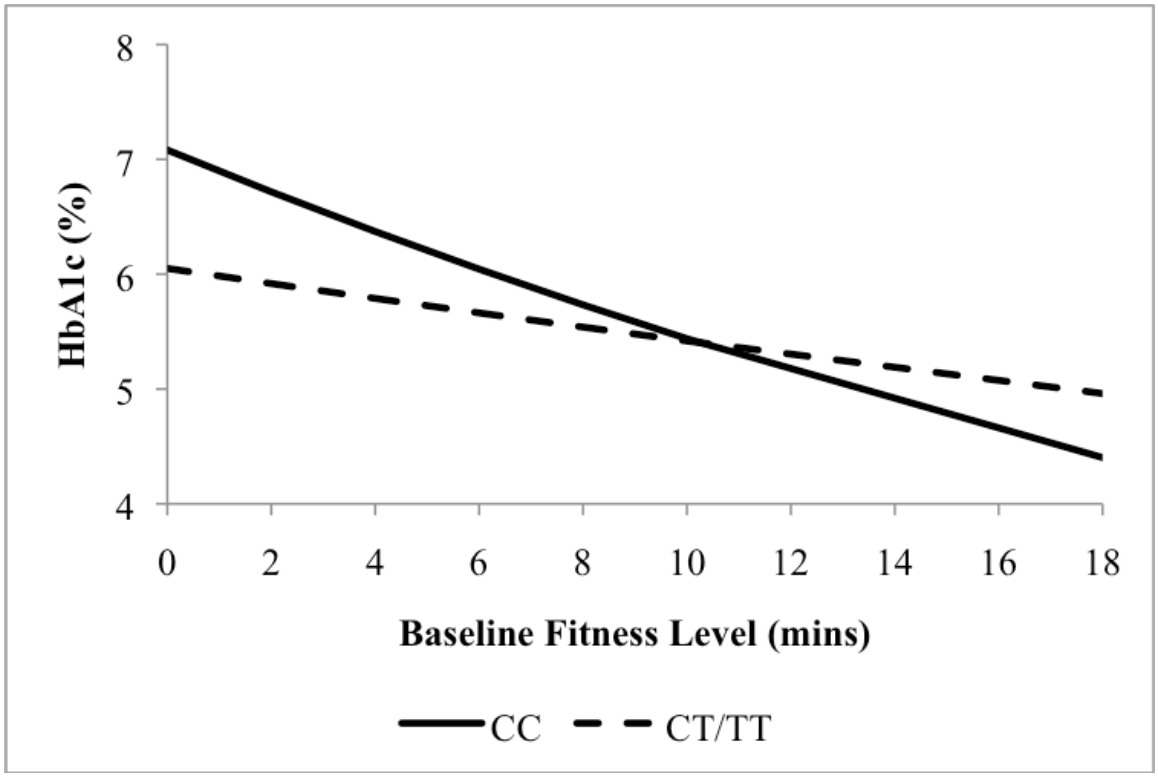


Figure 5. Interaction between the *FTO* rs8050136 genotype and baseline cardiorespiratory fitness level in their association with Year 20 HbA1c in Caucasian participants.



Note: Model adjusted for age, sex, and clinic site. Fitness level ranged from 0-18 minutes. Predicted values are plotted for reference group for sex (female) and clinic site (Oakland, CA), and for an age of 25 years. Mean values are presented as geometric means. *P* for interaction = 0.0003.

Figure 6. Interaction between the *PPARG* rs3856806 genotype and baseline cardiorespiratory fitness level in their association with Year 20 HbA1c among African American participants.



Note: Model adjusted for age, sex, and clinic site. Fitness level ranged from 0-18 minutes. Predicted values are plotted for reference group for sex (female) and clinic site (Oakland, CA), and for an age of 25 years. Mean values are presented as geometric means. *P* for interaction = 0.001.

9. OVERALL CONCLUSIONS

This dissertation served to evaluate the predictive value of current screening tools and markers of risk for type 2 diabetes, and supported previous findings that both non-modifiable and modifiable risk factors significantly contribute to diabetes risk prediction. Given known between-population differences in the burden of diabetes, these papers highlighted the importance of improving diabetes prevention in population subgroups that are disproportionately affected (particularly those with pre-diabetes and in certain race/ethnic groups). Adequate screening and prevention efforts are dependent upon effective screening tools that are validated across populations, as well as a more comprehensive understanding of the role of modifiable and non-modifiable risk factors in influencing one's diabetes risk profile.

10. REFERENCES

1. Lloyd-Jones D, Adams R, Carnethon M, et al. Heart Disease and Stroke Statistics--2009 Update. A Report From the American Heart Association Statistics Committee and Stroke Statistics Subcommittee. *Circulation*. 2008.

2. Zimmet P, Alberti KG, Shaw J. Global and societal implications of the diabetes epidemic. *Nature*. 2001;414(6865):782-787.

3. Selvin E, Coresh J, Golden SH, Brancati FL, Folsom AR, Steffes MW. Glycemic control and coronary heart disease risk in persons with and without diabetes: the Atherosclerosis Risk in Communities study. *Arch Intern Med*. 2005;165(16):1910-1916.

4. National Institute of Diabetes and Digestive and Kidney Diseases. National Diabetes Statistics, 2007 fact sheet. 2008.

5. Anselmino M, Gohlke H, Mellbin L, Ryden L. Cardiovascular prevention in patients with diabetes and prediabetes. *Herz*. 2008;33(3):170-177.

6. American Diabetes Association. Diagnosis and classification of diabetes mellitus. *Diabetes Care*. 2009;32(Supplement 1):S62-67.

7. Ekoé J, Zimmet PZ, Yale J. The clinical syndrome and the biochemical definition. In: Ekoé J, Rewers M, Williams R, Zimmet P, eds. *The Epidemiology of Diabetes Mellitus.* 2nd ed. West Sussex, UK: John Wiley & Sons, Ltd; 2008:5.

8. DeFronzo RA. Pathogenesis of type 2 diabetes mellitus. *Med Clin North Am*. 2004;88(4):787-835.

9. Skyler JS. Diabetes mellitus: pathogenesis and treatment strategies. *J Med Chem*. 2004;47(17):4113-4117.

10. Stumvoll M, Goldstein BJ, van Haeften TW. Type 2 diabetes: principles of pathogenesis and therapy. *Lancet*. 2005;365(9467):1333-1346.

11. Expert Committee on the Diagnosis and Classification of Diabetes Mellitus. Report of the Expert Committee on the Diagnosis and Classification of Diabetes Mellitus. *Diabetes Care*. 2003;26(90001):S5-20.

12. Gable D, Sanderson SC, Humphries SE. Genotypes, obesity and type 2 diabetes--can genetic information motivate weight loss? A review. *Clin Chem Lab Med*. 2007;45(3):301-308.

13. Meigs JB, Cupples LA, Wilson PW. Parental transmission of type 2 diabetes: the Framingham Offspring Study. *Diabetes*. 2000;49(12):2201-2207.

14. American Diabetes Association. Diagnosis and Classification of Diabetes Mellitus. *Diabetes Care*. 2010;33(Supplement 1):S62-S69.

15. Finch CF, Zimmet PZ, Alberti KG. Determining diabetes prevalence: a rational basis for the use of fasting plasma glucose concentrations? *Diabet Med*. 1990;7(7):603-610.

16. Milicevic Z, Raz I, Beattie SD, et al. Natural history of cardiovascular disease in patients with diabetes: role of hyperglycemia. *Diabetes Care*. 2008;31(Suppl 2):S155-60.

17. Skyler JS, Bergenstal R, Bonow RO, et al. Intensive Glycemic Control and the Prevention of Cardiovascular Events: Implications of the ACCORD, ADVANCE, and VA Diabetes Trials: A position statement of the American Diabetes Association and a scientific statement of the American College of Cardiology Foundation and the American Heart Association. *Diabetes Care*. 2009;32(1):187-192.

18. Nathan DM, Buse JB, Davidson MB, et al. Medical management of hyperglycemia in type 2 diabetes: a consensus algorithm for the initiation and adjustment of therapy: a consensus statement of the American Diabetes Association and the European Association for the Study of Diabetes. *Diabetes Care*. 2009;32(1):193-203.

19. Wylie-Rosett J, Albright AA, Apovian C, et al. 2006-2007 American Diabetes Association nutrition recommendations: issues for practice translation. *J Am Diet Assoc*. 2007;107(8):1296-1304.

20. Bassuk SS, Manson JE. Lifestyle and risk of cardiovascular disease and type 2 diabetes in women: a review of the epidemiologic evidence. *American Journal of Lifestyle Medicine*. 2008;2(3):191-213.

21. Cox KL, Burke V, Morton AR, Beilin LJ, Puddey IB. Independent and additive effects of energy restriction and exercise on glucose and insulin concentrations in sedentary overweight men. *Am J Clin Nutr*. 2004;80(2):308-316.

22. Franz MJ, Wylie-Rosett J. The 2006 American Diabetes Association nutrition recommendations and interventions for the prevention and treatment of diabetes. *Diabetes Spectrum*. 2007;20(1):49-52.

23. Riccardi G, Rivellese AA, Giacco R. Role of glycemic index and glycemic load in the healthy state, in prediabetes, and in diabetes. *Am J Clin Nutr*. 2008;87(1):269S-274S.

24. Klein S, Sheard NF, Pi-Sunyer X, et al. Weight management through lifestyle modification for the prevention and management of type 2 diabetes: rationale and strategies. *Diabetes Care*. 2004;27(8):2067-2073.

25. Centers for Disease Control and Prevention (CDC). Self-reported prediabetes and risk-reduction activities--United States, 2006. *MMWR Morb Mortal Wkly Rep*. 2008;57(44):1203-1205.

26. Yates T, Khunti K, Bull F, Gorely T, Davies MJ. The role of physical activity in the management of impaired glucose tolerance: a systematic review. *Diabetologia*. 2007;50(6):1116-1126.

27. Kanaya AM, Narayan KM. Prevention of type 2 diabetes: data from recent trials. *Prim Care*. 2003;30(3):511-526.

28. Wei M, Gibbons LW, Mitchell TL, Kampert JB, Lee CD, Blair SN. The association between cardiorespiratory fitness and impaired fasting glucose and type 2 diabetes mellitus in men. *Ann Intern Med*. 1999;130(2):89-96.

29. Sigal RJ, Kenny GP, Wasserman DH, Castaneda-Sceppa C. Physical activity/exercise and type 2 diabetes. *Diabetes Care*. 2004;27(10):2518-2539.

30. Huisman TH, Martis EA, Dozy A. Chromatography of hemoglobin types on carboxymethylcellulose. *J Lab Clin Med*. 1958;52(2):312-327.

31. Kahn R, Fonseca V. Translating the A1C assay. *Diabetes Care*. 2008;31(8):1704-1707.

32. Sacks DB. Translating hemoglobin A1c into average blood glucose: implications for clinical chemistry. *Clin Chem*. 2008;54(11):1756-1758.

33. Genuth S, Alberti KG, Bennett P, et al. Follow-up report on the diagnosis of diabetes mellitus. *Diabetes Care*. 2003;26(11):3160-3167.

34. Goldstein DE, Little RR, Lorenz RA, et al. Tests of glycemia in diabetes. *Diabetes Care*. 2004;27(7):1761-1773.

35. Nguyen QM, Srinivasan SR, Xu J, Chen W, Berenson GS. Distribution and cardiovascular risk correlates of hemoglobin A1c in nondiabetic younger adults: the Bogalusa Heart Study. *Metabolism*. 2008;57(11):1487-1492.

36. Selvin E, Wattanakit K, Steffes MW, Coresh J, Sharrett AR. HbA1c and peripheral arterial disease in diabetes: the Atherosclerosis Risk in Communities study. *Diabetes Care*. 2006;29(4):877-882.

37. Cefalu WT, Watson K. Intensive glycemic control and cardiovascular disease observations from the ACCORD study: now what can a clinician possibly think? *Diabetes*. 2008;57(5):1163-1165.

38. Traynor K. Diabetes specialists keep ACCORD data in perspective. *American Journal of Health-System Pharmacy*. 2008;65(6):488,490.

39. Ginde AA, Cagliero E, Nathan DM, Camargo CA,Jr. Value of risk stratification to increase the predictive validity of HbA1c in screening for undiagnosed diabetes in the US population. *J Gen Intern Med*. 2008;23(9):1346-1353.

40. Saudek CD, Herman WH, Sacks DB, Bergenstal RM, Edelman D, Davidson MB. A new look at screening and diagnosing diabetes mellitus. *J Clin Endocrinol Metab*. 2008;93(7):2447-2453.

41. Bruns DE, Boyd JC. Few point-of-care hemoglobin A1c assay methods meet clinical needs. *Clin Chem*. 2010;56(1):4-6.

42. Kirk JK, Passmore LV, Bell RA, et al. Disparities in A1C levels between Hispanic and non-Hispanic white adults with diabetes: a meta-analysis. *Diabetes Care*. 2008;31(2):240-246.

43. Kirk JK, D'Agostino RB, Jr., Bell RA, et al. Disparities in HbA1c levels between African-American and Non-Hispanic White adults with diabetes: a meta-analysis. *Diabetes Care*. 2006;29(9):2130-2136.

44. Saudek CD, Derr RL, Kalyani RR. Assessing Glycemia in Diabetes Using Self-monitoring Blood Glucose and Hemoglobin A1c. *JAMA*. 2006;295(14):1688-1697.

45. Pignone M. Value of self-monitoring of blood glucose in non-insulin-using patients with type 2 diabetes. *Clinical Diabetes*. 2009;27(1):17-18.

46. Eldin WS, Emara M, Shoker A. Prediabetes: a must to recognise disease state. *Int J Clin Pract*. 2008;62(4):642-648.

47. Nathan DM, Davidson MB, DeFronzo RA, et al. Impaired fasting glucose and impaired glucose tolerance: implications for care. *Diabetes Care*. 2007;30(3):753-759.

48. Bloomgarden ZT. Approaches to treatment of pre-diabetes and obesity and promising new approaches to type 2 diabetes. *Diabetes Care*. 2008;31(7):1461-1466.

49. Phillips LS, Weintraub WS, Ziemer DC, et al. All pre-diabetes is not the same: Metabolic and vascular risks of impaired fasting glucose at 100 versus 110 mg/dl: The Screening for Impaired Glucose Tolerance Study 1 (SIGT 1). *Diabetes Care*. 2006;29(6):1405-1407.

50. Nichols GA, Hillier TA, Brown JB. Progression from newly acquired impaired fasting glucose to type 2 diabetes: Response to Dinneen and Rizza. *Diabetes Care*. 2007;30(7):e80.

51. Harris MI. Undiagnosed NIDDM: clinical and public health issues. *Diabetes Care*. 1993;16(4):642-652.

52. Harris MI, Klein R, Welborn TA, Knuiman MW. Onset of NIDDM occurs at least 4-7 yr before clinical diagnosis. *Diabetes Care*. 1992;15(7):815-819.

53. Biuso TJ, Butterworth S, Linden A. A conceptual framework for targeting prediabetes with lifestyle, clinical, and behavioral management interventions. *Dis Manage*. 2007;10(1):6-15.

54. Hu FB, Stampfer MJ, Haffner SM, Solomon CG, Willett WC, Manson JE. Elevated risk of cardiovascular disease prior to clinical diagnosis of type 2 diabetes. *Diabetes Care*. 2002;25(7):1129-1134.

55. Nguyen TT, Wang JJ, Wong TY. Retinal vascular changes in pre-diabetes and prehypertension: New findings and their research and clinical implications. *Diabetes Care*. 2007;30(10):2708-2715.

56. Sumner CJ, Sheth S, Griffin JW, Cornblath DR, Polydefkis M. The spectrum of neuropathy in diabetes and impaired glucose tolerance. *Neurology*. 2003;60(1):108-111.

57. Abdul-Ghani MA, Lyssenko V, Tuomi T, DeFronzo RA, Groop L. Fasting versus postload plasma glucose concentration and the risk for future type 2 diabetes: Results from the Botnia Study. *Diabetes Care; Diabetes Care*. 2009;32(2):281-286.

58. Balion CM, Raina PS, Gerstein HC, et al. Reproducibility of impaired glucose tolerance (IGT) and impaired fasting glucose (IFG) classification: a systematic review.[erratum appears in Clin Chem Lab Med. 2007;45(11):1570]. *Clin Chem Lab Med*. 2007;45(9):1180-1185.

59. Meigs JB, Muller DC, Nathan DM, Blake DR, Andres R. The natural history of progression from normal glucose tolerance to type 2 diabetes in the Baltimore Longitudinal Study of Aging. *Diabetes*. 2003;52(6):1475-1484.

60. Ahmad LA, Crandall JP. Type 2 diabetes prevention: A review. *Clinical Diabetes*. 2010;28(2):53-59.

61. Haffner SM, Lehto S, Ronnemaa T, Pyorala K, Laakso M. Mortality from coronary heart disease in subjects with type 2 diabetes and in nondiabetic subjects with and without prior myocardial infarction. *N Engl J Med*. 1998;339(4):229-234.

62. Kannel WB, McGee DL. Diabetes and cardiovascular disease. The Framingham study. *JAMA*. 1979;241(19):2035-2038.

63. Manson JE, Colditz GA, Stampfer MJ, et al. A prospective study of maturity-onset diabetes mellitus and risk of coronary heart disease and stroke in women. *Arch Intern Med*. 1991;151(6):1141-1147.

64. Stamler J, Vaccaro O, Neaton JD, Wentworth D. Diabetes, other risk factors, and 12-yr cardiovascular mortality for men screened in the Multiple Risk Factor Intervention Trial. *Diabetes Care*. 1993;16(2):434-444.

65. Seaquist ER. Microvascular complications of diabetes. Strategies for managing retinopathy, nephropathy, and neuropathy. *Postgrad Med*. 1998;103(1):61-63.

66. Congdon NG, Friedman DS, Lietman T. Important causes of visual impairment in the world today. *JAMA*. 2003;290(15):2057-2060.

67. The Eye Diseases Prevalence Research Group,. The prevalence of diabetic retinopathy among adults in the United States. *Arch Ophthalmol*. 2004;122(4):552-563.

68. Nishikawa T, Edelstein D, Du XL, et al. Normalizing mitochondrial superoxide production blocks three pathways of hyperglycaemic damage. *Nature*. 2000;404(6779):787-790.

69. Deedwania PC, Fonseca VA. Diabetes, prediabetes, and cardiovascular risk: shifting the paradigm.[see comment]. *Am J Med*. 2005;118(9):939-947.

70. Fox CS, Sullivan L, D'Agostino R, Wilson PWF. The significant effect of diabetes duration on coronary heart disease mortality: The Framingham Heart Study. *Diabetes Care*. 2004;27(3):704-708.

71. Goldstein LB, Adams R, Becker K, et al. Primary prevention of ischemic stroke: A statement for healthcare professionals from the Stroke Council of the American Heart Association. *Circulation*. 2001;103(1):163-182.

72. Stout RW. Insulin and atheroma. 20-yr perspective. *Diabetes Care*. 1990;13(6):631-654.

73. Rathmann W, Giani G. Global prevalence of diabetes: Estimates for the year 2000 and projections for 2030: Response to Wild et al. *Diabetes Care*. 2004;27(10):2568-2569.

74. Selvin E, Coresh J, Brancati FL. The burden and treatment of diabetes in elderly individuals in the U.S. *Diabetes Care*. 2006;29(11):2415-2419.

75. Cowie CC, Rust KF, Ford ES, et al. Full accounting of diabetes and pre-diabetes in the U.S. population in 1988-1994 and 2005-2006. *Diabetes Care; Diabetes Care*. 2009;32(2):287-294.

76. World Health Organization. Preventing chronic disease: a vital investment. . 2005.

77. Roglic G, Unwin N, Bennett PH, et al. The burden of mortality attributable to diabetes: Realistic estimates for the year 2000. *Diabetes Care*. 2005;28(9):2130-2135.

78. Wild S, Roglic G, Green A, Sicree R, King H. Global Prevalence of Diabetes: Estimates for the year 2000 and projections for 2030. *Diabetes Care*. 2004;27(5):1047-1053.

79. Cowie CC, Rust KF, Byrd-Holt DD, et al. Prevalence of diabetes and impaired fasting glucose in adults in the U.S. population: National Health and Nutrition Examination Survey 1999-2002. *Diabetes Care*. 2006;29(6):1263-1268.

80. Winker MA. Measuring race and ethnicity: Why and how? *JAMA*. 2004;292(13):1612-1614.

81. Burchard EG, Ziv E, Coyle N, et al. The importance of race and ethnic background in biomedical research and clinical practice. *N Engl J Med ; N Engl J Med*. 2003;348(12):1170-1175.

82. Kaufman JS, Cooper RS. Commentary: Considerations for use of racial/ethnic classification in etiologic research. *Am J Epidemiol ; Am J Epidemiol*. 2001;154(4):291-298.

83. Bang H, Edwards AM, Bomback AS, et al. Development and Validation of a Patient Self-assessment Score for Diabetes Risk. *Ann Intern Med*. 2009;151(11):775.

84. Hippisley-Cox J, Coupland C, Robson J, Sheikh A, Brindle P. Predicting risk of type 2 diabetes in England and Wales: prospective derivation and validation of QDScore. *BMJ; BMJ*. 2009;338(mar17_2):b880.

85. Lindstrom J, Tuomilehto J. The Diabetes Risk Score: A practical tool to predict type 2 diabetes risk. *Diabetes Care*. 2003;26(3):725-731.

86. Schmidt MI, Duncan BB, Bang H, et al. Identifying individuals at high risk for diabetes: The Atherosclerosis Risk in Communities study. *Diabetes Care*. 2005;28(8):2013-2018.

87. Schulze MB, Hoffmann K, Boeing H, et al. An accurate risk score based on anthropometric, dietary, and lifestyle factors to predict the development of type 2 diabetes. *Diabetes Care*. 2007;30(3):510-515.

88. Wilson PWF, Meigs JB, Sullivan L, Fox CS, Nathan DM, D'Agostino RB,Sr. Prediction of incident diabetes mellitus in middle-aged adults: The Framingham Offspring Study. *Arch Intern Med*. 2007;167(10):1068-1074.

89. Kahn HS, Cheng YJ, Thompson TJ, Imperatore G, Gregg EW. Two risk-scoring systems for predicting incident diabetes mellitus in U.S. adults age 45 to 64 years. *Ann Intern Med*. 2009;150(11):741-751.

90. Nichols GA, Hillier TA, Brown JB. Normal fasting plasma glucose and risk of type 2 diabetes diagnosis. *Am J Med*. 2008;121(6):519-524.

91. Heikes KE, Eddy DM, Arondekar B, Schlessinger L. Diabetes Risk Calculator: a simple tool for detecting undiagnosed diabetes and pre-diabetes. *Diabetes Care*. 2008;31(5):1040-1045.

92. Lindstrom J, Tuomilehto J. The diabetes risk score: a practical tool to predict type 2 diabetes risk. *Diabetes Care*. 2003;26(3):725-731.

93. Glumer C, Carstensen B, Sandbaek A, et al. A Danish diabetes risk score for targeted screening: the Inter99 study. *Diabetes Care*. 2004;27(3):727-733.

94. Kumari M, Head J, Marmot M. Prospective study of social and other risk factors for incidence of type 2 diabetes in the Whitehall II Study. *Arch Intern Med*. 2004;164(17):1873-1880.

95. Yeung EH, Pankow JS, Astor BC, Powe NR, Saudek CD, Kao WH. Increased risk of type 2 diabetes from a family history of coronary heart disease and type 2 diabetes. *Diabetes Care*. 2007;30(1):154-156.

96. Franks PW, Mesa JL, Harding AH, Wareham NJ. Gene-lifestyle interaction on risk of type 2 diabetes. *Nutr Metab Cardiovasc Dis*. 2007;17(2):104-124.

97. Lyssenko V, Almgren P, Anevski D, et al. Genetic prediction of future type 2 diabetes. *PLoS Med*. 2005;2(12):e345.

98. Sanghera DK, Ortega L, Han S, et al. Impact of nine common type 2 diabetes risk polymorphisms in Asian Indian Sikhs: PPARG2 (Pro12Ala), IGF2BP2, TCF7L2 and FTO variants confer a significant risk. *BMC Med Genet*. 2008;9:59.

99. Zeggini E, Scott LJ, Saxena R, et al. Meta-analysis of genome-wide association data and large-scale replication identifies additional susceptibility loci for type 2 diabetes. *Nat Genet*. 2008;40(5):638-645.

100. Schaffler A, Barth N, Palitzsch KD, Drobnik W, Scholmerich J, Schmitz G. Mutation analysis of the human adipocyte-specific apM-1 gene. *Eur J Clin Invest*. 2000;30(10):879-887.

101. Takahashi M, Arita Y, Yamagata K, et al. Genomic structure and mutations in adipose-specific gene, adiponectin. *Int J Obes Relat Metab Disord*. 2000;24(7):861-868.

102. Hara K, Boutin P, Mori Y, et al. Genetic variation in the gene encoding adiponectin is associated with an increased risk of type 2 diabetes in the Japanese population. *Diabetes; Diabetes*. 2002;51(2):536-540.

103. Stumvoll M, Tschritter O, Fritsche A, et al. Association of the T-G polymorphism in adiponectin (exon 2) with obesity and insulin sensitivity: Interaction with family history of type 2 diabetes. *Diabetes; Diabetes*. 2002;51(1):37-41.

104. Rasmussen-Torvik LJ, Pankow JS, Jacobs DR,Jr, Steinberger J, Moran A, Sinaiko AR. The association of SNPs in ADIPOQ, ADIPOR1, and ADIPOR2 with insulin sensitivity in a cohort of adolescents and their parents. *Hum Genet*. 2009;125(1):21-28.

105. Vasseur F, Lepretre F, Lacquemant C, Froguel P. The genetics of adiponectin. *Curr Diab Rep*. 2003;3(2):151-158.

106. Rampersaud E, Mitchell BD, Pollin TI, et al. Physical activity and the association of common FTO gene variants with body mass index and obesity. *Arch Intern Med*. 2008;168(16):1791-1797.

107. Frayling TM, Timpson NJ, Weedon MN, et al. A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity. *Science*. 2007;316(5826):889-894.

108. Scuteri A, Sanna S, Chen WM, et al. Genome-wide association scan shows genetic variants in the FTO gene are associated with obesity-related traits. *PLoS Genet*. 2007;3(7):e115.

109. Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature*. 2007;447(7145):661-678.

110. Attaoua R, Ait El Mkadem S, Lautier C, et al. Association of the FTO gene with obesity and the metabolic syndrome is independent of the IRS-2 gene in the female population of Southern France. *Diabetes Metab*. 2009;35(6):476-483.

111. Freathy RM, Timpson NJ, Lawlor DA, et al. Common variation in the FTO gene alters diabetes-related metabolic traits to the extent expected given its effect on BMI. *Diabetes*. 2008;57(5):1419-1426.

112. Legry V, Cottel D, Ferrières J, et al. Effect of an FTO polymorphism on fat mass, obesity, and type 2 diabetes mellitus in the French MONICA Study. *Metab Clin Exp*. 2009;58(7):971-975.

113. Ng MC, Park KS, Oh B, et al. Implication of genetic variants near TCF7L2, SLC30A8, HHEX, CDKAL1, CDKN2A/B, IGF2BP2, and FTO in type 2 diabetes and obesity in 6,719 Asians. *Diabetes*. 2008;57(8):2226-2233.

114. Fornage M, Jacobs DR, Steffes MW, Gross MD, Bray MS, Schreiner PJ. Inverse effects of the PPAR(gamma)2 Pro12Ala polymorphism on measures of adiposity over 15 years in African Americans and whites. The CARDIA study. *Metabolism*. 2005;54(7):910-917.

115. Nelson TL, Fingerlin TE, Moss LK, Barmada MM, Ferrell RE, Norris JM. Association of the peroxisome proliferator–activated receptor γ gene with type 2 diabetes mellitus varies by physical activity among non-Hispanic whites from Colorado. *Metabolism*. 2007;56(3):388-393.

116. Gouda HN, Sagoo GS, Harding AH, Yates J, Sandhu MS, Higgins JP. The association between the peroxisome proliferator-activated receptor-gamma2 (PPARG2) Pro12Ala gene variant and type 2 diabetes mellitus: a HuGE review and meta-analysis. *Am J Epidemiol*. 2010;171(6):645-655.

117. Tonjes A, Stumvoll M. The role of the Pro12Ala polymorphism in peroxisome proliferator-activated receptor gamma in diabetes risk. *Curr Opin Clin Nutr Metab Care*. 2007;10(4):410-414.

118. Adamo KB, Sigal RJ, Williams K, Kenny G, Prud'homme D, Tesson F. Influence of Pro12Ala peroxisome proliferator-activated receptor gamma2 polymorphism on glucose response to exercise training in type 2 diabetes. *Diabetologia*. 2005;48(8):1503-1509.

119. Kahara T, Takamura T, Hayakawa T, et al. PPAR[gamma] gene polymorphism is associated with exercise-mediated changes of insulin resistance in healthy men. *Metabolism*. 2003;52(2):209-212.

120. Kilpelainen TO, Lakka TA, Laaksonen DE, et al. SNPs in PPARG associate with type 2 diabetes and interact with physical activity. *Med Sci Sports Exerc*. 2008;40(1):25-33.

121. Weiss EP, Kulaputana O, Ghiu IA, et al. Endurance training–induced changes in the insulin response to oral glucose are associated with the peroxisome proliferator–activated receptor-γ2 Pro12Ala genotype in men but not in women. *Metabolism,*. 2005;54(1):97-102.

122. Willi C, Bodenmann P, Ghali WA, Faris PD, Cornuz J. Active smoking and the risk of type 2 diabetes: A systematic review and meta-analysis. *JAMA*. 2007;298(22):2654-2664.

123. American Diabetes Association. Nutrition recommendations and interventions for diabetes-2006: A position statement of the American Diabetes Association. *Diabetes Care*. 2006;29(9):2140-2157.

124. Schroder H. Protective mechanisms of the Mediterranean diet in obesity and type 2 diabetes. *J Nutr Biochem*. 2007;18(3):149-160.

125. Wien M. A review of macronutrient considerations for persons with prediabetes. *Topics in Clinical Nutrition*. 2006;21(2):64-75.

126. Pereira MA, Jacobs DR,Jr, Van Horn L, Slattery ML, Kartashov AI, Ludwig DS. Dairy consumption, obesity, and the insulin resistance syndrome in young adults: the CARDIA Study. *JAMA*. 2002;287(16):2081-2089.

127. Young JC. Exercise prescription for individuals with metabolic disorders. Practical considerations. *Sports Med*. 1995;19(1):43-54.

128. Carnethon MR, Sternfeld B, Schreiner PJ, et al. Association of 20 year changes in cardirespiratory fitness with incident type 2 diabetes: The CARDIA Fitness Study. *Diabetes Care*. 2009;32:1284-1288.

129. Carnethon MR, Gidding SS, Nehgme R, Sidney S, Jacobs DR,Jr, Liu K. Cardiorespiratory fitness in young adulthood and the development of cardiovascular disease risk factors. *JAMA*. 2003;290(23):3092-3100.

130. Boule NG, Kenny GP, Haddad E, Wells GA, Sigal RJ. Meta-analysis of the effect of structured exercise training on cardiorespiratory fitness in Type 2 diabetes mellitus. *Diabetologia*. 2003;46(8):1071-1081.

131. Katzmarzyk PT, Craig CL, Gauvin L. Adiposity, physical fitness and incident diabetes: the physical activity longitudinal study. *Diabetologia*. 2007;50(3):538-544.

132. Bild DE, Bluemke DA, Burke GL, et al. Multi-Ethnic Study of Atherosclerosis: Objectives and design. *Am J Epidemiol*. 2002;156(9):871-881.

133. Friedman GD, Cutter GR, Donahue RP, et al. CARDIA: study design, recruitment, and some characteristics of the examined subjects. *J Clin Epidemiol*. 1988;41(11):1105-1116.

134. Balke B. A simple field test for the assessment of physical fitness. Rep 63-6. *Rep Civ Aeromed Res Inst US*. 1963:1-8.

135. Sidney S, Haskell WL, Crow R, et al. Symptom-limited graded treadmill exercise testing in young adults in the CARDIA study. *Clinical Epidemiology*. 1992;24(2).

136. Jacobs DRJ, Hahn LP, Haskell WL, Pirie P, Sidney S. Validity and reliability of Short Physical Activity History: CARDIA and the Minnesota Heart Health Program. *J Cardiopulm Rehabil*. 1989;9(11):448-459.

137. Tanaka H, Monahan KD, Seals DR. Age-predicted maximal heart rate revisited. *J Am Coll Cardiol*. 2001;37(1):153-156.

138. Lloyd-Jones D, Adams RJ, Brown TM, et al. Heart Disease and Stroke Statistics--2010 update. A report from the American Heart Association. *Circulation*. 2009.

139. Centers for Disease Control and Prevention (CDC). Prevalence of diabetes and impaired fasting glucose in adults--United States, 1999-2000. *MMWR Morb Mortal Wkly Rep*. 2003;52(35):833-837.

140. Selvin E, Steffes MW, Zhu H, et al. Glycated hemoglobin, diabetes, and cardiovascular risk in nondiabetic adults. *N Engl J Med*. 2010;362(9):800-811.

141. Selvin E, Zhu H, Brancati FL. Elevated A1C in adults without a history of diabetes in the U.S. *Diabetes Care*. 2009;32(5):828-833.

142. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics*. 1988;44(3):837-845.

143. Pratley RE, Matfin G. Review: Pre-diabetes: clinical relevance and therapeutic approach. *The British Journal of Diabetes & Vascular Disease*. 2007;7(3):120-129.

144. Perreault L, Kahn SE, Christophi CA, Knowler WC, Hamman RF. Regression From pre-diabetes to normal glucose regulation in the Diabetes Prevention Program. *Diabetes Care*. 2009;32(9):1583-1588.

145. Lilly M, Godwin M. Treating prediabetes with metformin: Systematic review and meta-analysis. *Can Fam Physician*. 2009;55(4):363-369.

146. Jermendy G. Can type 2 diabetes mellitus be considered preventable? *Diabetes Res Clin Pract*. 2005;68(Suppl1):S73-81.

147. Sallis JF, Haskell WL, Wood PD, et al. Physical Activity Assessment Methodology in the Five-City Project. *Am J Epidemiol*. 1985;121(1):91-106.

148. Spielberger CD, Gorsuch RL, Lushene RE. STAI manual for the state-trait anxiety inventory. *Palo Alto, CA: Consulting Psychologists*. 1970.

149. Ford ES, Zhao G, Li C. Pre-diabetes and the risk for cardiovascular disease: a systematic review of the evidence. *J Am Coll Cardiol*. 2010;55(13):1310-1317.

150. Edelstein SL, Knowler WC, Bain RP, et al. Predictors of progression from impaired glucose tolerance to NIDDM: an analysis of six prospective studies. *Diabetes*. 1997;46(4):701-710.

151. Cohen S, Janicki-Deverts D, Miller GE. Psychological stress and disease. *JAMA*. 2007;298(14):1685-1687.

152. Sladek R, Rocheleau G, Rung J, et al. A genome-wide association study identifies novel risk loci for type 2 diabetes. *Nature*. 2007;445(7130):881-885.

153. Hunter DJ. Gene-environment interactions in human diseases. *Nat Rev Genet*. 2005;6(4):287-298.

154. Sawada SS, Lee I-, Naito H, et al. Long-term trends in cardiorespiratory fitness and the incidence of type 2 diabetes. *Diabetes Care*. 2010;33(6):1353-1357.

155. LaMonte MJ, Blair SN, Church TS. Physical activity and diabetes prevention. *J Appl Physiol*. 2005;99(3):1205-1213.

156. Sui X, Hooker SP, Lee I, et al. A prospective study of cardiorespiratory fitness and risk of type 2 diabetes in women. *Diabetes Care*. 2008;31(3):550-555.

157. Andreasen CH, Stender-Petersen KL, Mogensen MS, et al. Low physical activity accentuates the effect of the FTO rs9939609 polymorphism on body fat accumulation. *Diabetes*. 2008;57(1):95-101.

158. Hautala AJ, Leon AS, Skinner JS, Rao DC, Bouchard C, Rankinen T. Peroxisome proliferator-activated receptor-{delta} polymorphisms are associated with physical performance and plasma lipids: the HERITAGE Family Study. *Am J Physiol Heart Circ Physiol*. 2007;292(5):H2498-2505.

159. Kraemer RR, Aboudehen KS, Carruth AK, et al. Adiponectin responses to continuous and progressively intense intermittent exercise. *Medicine & Science in Sports & Exercise*. 2003;35(8):1320-1325.

160. Rankinen T, Bray MS, Hagberg JM, et al. The human gene map for performance and health-related fitness phenotypes: the 2005 update. *Med Sci Sports Exerc*. 2006;38(11):1863-1888.

161. Wei Q, Jacobs DR,Jr, Schreiner PJ, Siscovick DS, Steffes MW, Fornage M. Patterns of association between PPARgamma genetic variation and indices of adiposity and insulin action in African-Americans and whites: the CARDIA Study. *J Mol Med*. 2006;84(11):955-965.

162. Rankinen T, Roth SM, Bray MS, et al. Advances in exercise, fitness, and performance genomics. *Medicine & Science in Sports & Exercise*. 2010;42(5):835.

163. Carlson CS, Eberle MA, Rieder MJ, Yi Q, Kruglyak L, Nickerson DA. Selecting a maximally informative set of single-nucleotide polymorphisms for association analyses using linkage disequilibrium. *Am J Hum Genet*. 2004;74(1):106-120.

164. Sidney S, Haskell WL, Crow R, et al. Symptom-limited graded treadmill exercise testing in young adults in the CARDIA study. *Med Sci Sports Exerc*. 1992;24(2):177-183.

165. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics*. 2005;21(2):263-265.

166. Balding DJ. A tutorial on statistical methods for population association studies. *Nature Reviews Genetics*. 2006;7(10):781-791.

167. Eisenmann JC. Aerobic fitness, fatness and the metabolic syndrome in children and adolescents. *Acta Paediat*r. 2007;96(12):1723-1729.

168. Katzmarzyk PT, Church TS, Blair SN. Cardiorespiratory fitness attenuates the effects of the metabolic syndrome on all-cause and cardiovascular disease mortality in men. *Arch Intern Me*d. 2004;164(10):1092-1097.

169. Wing RR, Jakicic J, Neiberg R, et al. Fitness, fatness, and cardiovascular risk factors in type 2 diabetes: look ahead study. *Med Sci Sports Exer*c. 2007;39(12):2107-2116.

11. APPENDIX: ABBREVIATIONS AND ACRONYMS

ADA	American Diabetes Association
ADIPOQ	adiponectin
BMI	body mass index
CARDIA	Coronary Artery Risk Development in Young Adults
CHD	coronary heart disease
CI	confidence interval
CVD	cardiovascular disease
DCCT	Diabetes Control and Complications Trial
DPP	Diabetes Prevention Program
FPG	fasting plasma glucose
FTO	fat mass and obesity
GXT	graded exercise treadmill testing
HbA1c	hemoglobin A1c
HOMA-IR	homeostasis model of insulin resistance
IFG	impaired fasting glucose
IGT	impaired glucose tolerance
MESA	Multi-Ethnic Study of Atherosclerosis
NGSP	National Glycohemoglobin Standardization Program
NHANES	National Health and Nutrition Examination Survey
OGTT	oral glucose tolerance test
PPARG	peroxisome proliferator-activated receptor, gamma
QC	quality control
RPE	rating of perceived exertion
SD	standard deviation
SNP	single nucleotide polymorphism