CARDIOVASCULAR DISEASE IN TYPE 1 DIABETES: QUANTIFYING RISK AND ADDRESSING LIMITATIONS IN THE ANALYSIS OF LONGITUDINAL COHORT STUDIES

by

Rachel G. Miller

BS in Microbiology, University of Pittsburgh, 2003

MS in Biostatistics, University of Pittsburgh, 2005

Submitted to the Graduate Faculty of

the Graduate School of Public Health in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2016

UNIVERSITY OF PITTSBURGH

GRADUATE SCHOOL OF PUBLIC HEALTH

This dissertation was presented

by

Rachel G. Miller

It was defended on

July 20, 2016

and approved by

Stewart J. Anderson, PhD, Professor, Department of Biostatistics,
Graduate School of Public Health, University of Pittsburgh

Akira Sekikawa, MD, MPH, PhD, Associate Professor, Department of Epidemiology,
Graduate School of Public Health, University of Pittsburgh

Tina Costacou, PhD, Assistant Professor, Department of Epidemiology,
Graduate School of Public Health, University of Pittsburgh

Dissertation Advisor: Trevor J. Orchard, MD, MMedSci, Distinguished Professor,
Department of Epidemiology, Graduate School of Public Health, University of Pittsburgh

Trevor J. Orchard, MD, MMedSci

CARDIOVASCULAR DISEASE IN TYPE 1 DIABETES: QUANTIFYING RISK AND ADDRESSING LIMITATIONS IN THE ANALYSIS OF LONGITUDINAL COHORT STUDIES

Rachel G. Miller, PhD

University of Pittsburgh, 2016

ABSTRACT

Cardiovascular disease (CVD) has historically been increased in type 1 diabetes compared to the general population, but no contemporary estimates of risk are available in the United States. Additionally, the reasons for this increased risk are not fully understood, as the hyperglycemia that characterizes type 1 diabetes is itself an inconsistent predictor of CVD incidence. Thus, the objective of this dissertation is to quantify the contemporary incidence and excess risk of CVD in young adults <45 years old with type 1 diabetes and to utilize novel statistical methods to address limitations in the analyses of longitudinal cohort studies, in an effort to better understand the risk factor patterns that lead to CVD in this population.

Data are from the Pittsburgh Epidemiology of Diabetes Complications study, a prospective cohort study of childhood-onset type 1 diabetes diagnosed at Children's Hospital of Pittsburgh between 1950 and 1980. CVD data from the background Allegheny County, Pennsylvania population were used to calculate age- and sex-matched standardized mortality (SMR) and incidence rate ratios (IRR). Using tree-structured survival analysis (TSSA), formal subgroup analysis was performed to identify groups at varying levels of risk for CVD, based on threshold effects of continuous risk factors. Joint models were used to simultaneously model the longitudinal trajectory of HbA1c and time to CVD incidence.

CVD risk was shown remain significantly increased in this type 1 diabetes cohort. TSSA identified a range of risk groups, which were defined by combinations of diabetes duration, non-HDL cholesterol, albumin excretion rate, and white blood cell count. The longitudinal trajectory of HbA1c was associated with CVD risk, similarly across all manifestations of CVD, including coronary artery disease, stroke, and lower extremity arterial disease, which is a new finding in this cohort. This work has important impacts on public health, as it confirms that individuals with type 1 diabetes continue to be at increased risk for CVD and demonstrates that novel statistical methods should be utilized as a complement to traditional methods to increase understanding of disease etiology.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

When I began working as a staff member on the Epidemiology of Diabetes Complications study 11 years ago, I could not have imagined that I would complete a PhD in Epidemiology. As both my thesis advisor and job supervisor, Dr. Trevor Orchard has encouraged and inspired me to go beyond my job description and to develop a path toward independent research. Over the past 30 years, Dr. Orchard's EDC study has contributed a wealth of knowledge to the epidemiology of type 1 diabetes. Being able to work on this landmark study has been both humbling and an exceptional learning experience. I am ever grateful for Dr. Orchard's constant support and guidance.

I also thank the other members of my committee: Dr. Tina Costacou, for her expertise on the EDC cohort and her guidance and friendship over the years, Dr. Stewart Anderson, for his invaluable statistical advise and mentorship, and Dr. Akira Sekikawa for lending his expertise on cardiovascular disease.

Finally, this work could not have been completed without the support of my family. I thank my husband, Patrick, for his love, encouragement, and unceasing hard work for our family, our daughter, Stella, for her love and patience, and our soon-to-be-born son, Daniel. I am grateful for the support of my parents, who always lovingly encouraged me in my academic endeavors, especially when I was tempted to take an easier path. I also thank my siblings, friends, and my husband's family for their continual love, support, and encouragement.

1.0 INTRODUCTION

1.1 TYPE 1 DIABETES

1.1.1 Introduction and Epidemiology

Type 1 diabetes is an autoimmune disease in which the insulin-secreting beta cells in the pancreas are destroyed, so that the body is unable to produce insulin, a hormone that is required to convert sugar into energy. The pathogenesis of type 1 diabetes is complex and not well understood, with both genetic and environmental factors likely interacting to trigger the autoimmune response that targets the pancreatic beta cells. The clinical presentation of type 1 diabetes includes weight loss, frequent urination, extreme thirst, and fatigue (1). In the United States, it is estimated that between 740,000 and 970,000 individuals have type 1 diabetes (2) and the incidence has been increasing significantly in recent decades (3–6) by greater than 3% per year (5,7) with a corresponding increase in prevalence of 23% between 2001-2009 reported by the SEARCH for Diabetes in Youth Study (8). Type 1 diabetes most often manifests in childhood or adolescence, reflecting the initiation of islet autoimmunity early in life (9). In high incidence countries, males are at a greater risk of developing type 1 diabetes than females, and this difference does not seem to be explained by sex differences in environmental exposures, while in lower incidence countries this sex difference is attenuated (10). There are also

geographic differences in type 1 diabetes rates, with the highest rates in Finland and Sardinia (7) and the lowest rates in Asia (11). Generally, incidence is higher in the northern hemisphere compared to the southern (11).

1.1.2 Pathophysiology and Risk Factors

The pathophysiology of type 1 diabetes is dominated by an autoimmune response of CD4+ T-lymphocytes against the pancreatic beta cells, however the underlying cause for this process is not known. Various prospective cohort studies have shown that the clinical presentation of type 1 diabetes is preceded by the presence of islet autoantibodies (9,12–14). Additionally, as the progression from autoimmunity to clinical diabetes has been shown to proceed in a linear fashion, the age when these autoantibodies are first detected seems to relate to an increased risk of progression to type 1 diabetes (15) and determine the age of type 1 diabetes onset (16). The Finnish DIPP study reported a median of 2.5 years from autoantibody seroconversion until the onset of type 1 diabetes (17), and a similar rate of progression has also been observed in other cohorts (18). Islet autoantibodies have been found to be present in over 90% of individuals at type 1 diabetes diagnosis and though these autoantibodies do not directly cause type 1 diabetes, they can be used clinically to distinguish type 1 and type 2 diabetes (19).

While the cause of beta cell autoimmunity is not fully understood, there is evidence for contributions from genetic and environmental factors, including dietary, lifestyle, and infectious factors. The most well established genetic factor is the human leukocyte antigen (HLA) locus, located on chromosome 6p21. Individuals with the HLA-DR3-DQ2/DR4-DQ8 genotype are at the highest risk for developing type 1 diabetes, with an approximate 20-fold increased risk

compared to the general population (20). The risk in these individuals is even higher if they also have first-degree relatives with type 1 diabetes (21). Additionally, other HLA genes confer intermediate risk (22) and more recently the protein tyrosine phosphatase nonreceptor type 22 (PTPN22) has also been shown to be a strong predictor of type 1 diabetes (23–26), as well as other autoimmune diseases (25).

The evidence for potential environmental contributors to type 1 diabetes include outbreaks (27) and the increasing incidence of type 1 diabetes among the lower-risk HLA genotypes that has been observed in recent years (28–30). Dietary factors have been explored and in both the DAISY (31) and BABYDIAB (32) cohorts, the introduction of cereals before 3 months of age was associated with an increased risk of islet autoimmunity, particularly in the high risk HLA genotype. Early exposure to cow's milk has also been implicated as a predictor of islet autoimmunity (33–36) and the Ig-G antibodies that bind bovine insulin show cross-reactivity with human insulin (33,37,38). Insulin resistance and increasing rates of childhood obesity have also been hypothesized to be risk factors for type 1 diabetes and underlie the "accelerator hypothesis" (39). In support of this hypothesis, studies have found that insulin resistance is associated an increased risk of progression to type 1 diabetes in individuals with islet autoimmunity (40–44). Finally, enteroviruses have been proposed as another potential trigger of islet autoimmunity (45–47) and their effect may interact with other environmental factors, such as diet, to increase the risk of progression to type 1 diabetes (16,48).

1.2 MORTALITY TRENDS IN TYPE 1 DIABETES

Type 1 diabetes (T1DM) is consistently associated with increased mortality, though the excess risk seems to have decreased over recent years (49). Large differences in T1DM mortality are seen across countries, with smaller excess mortality generally observed in European countries and larger excess observed in the United States (49). There is also a sex difference, with women with T1DM having a greater excess mortality compared to men (50). More recent reports from the Scottish Registry Linkage Study and the Swedish National Diabetes Register both show that T1DM continues to be associated with an increased mortality compared to the general population across all age groups, with younger women at particularly high risk (51,52). A recent report from Australia further suggests that younger patients (<40 years) with T1DM are not experiencing a decline in "diabetes" mortality, in contrast to other causes and age groups (53). Contemporary data from the United States focused on young adults with long-duration T1DM are lacking, though recent findings from the Diabetes Control and Complications Trial/Epidemiology of Diabetes Interventions and Complications (DCCT/EDIC) study show lower all-cause mortality with intensive insulin therapy compared to conventional therapy (54).

1.3 CARDIOVASCULAR DISEASE IN TYPE 1 DIABETES

1.3.1 Introduction and Epidemiology

Cardiovascular disease (CVD), including coronary artery disease (CAD), cerebrovascular disease, and peripheral vascular disease, is a major complication of type 1 diabetes. As in the general population, CVD is a major contributor to morbidity and mortality in individuals with type 1 diabetes, however the risk of CVD is greatly increased in type 1 diabetes (55). The reasons for this increased risk are not fully understood, as the hyperglycemia that characterizes type 1 diabetes is itself an inconsistent predictor of CVD incidence (56–64). The cumulative incidence for CAD by age 55 years has been estimated to be as high as 35% in type 1 diabetes, compared to <10% in the nondiabetic population (55). The risk of stroke is also elevated in type 1 diabetes and has been estimated to be approximately four times higher in type 1 diabetes compared to the nondiabetic population (65,66). Likewise, the risk of peripheral vascular disease (PVD) is also higher in diabetes, with the incidence estimated to be approximately 5 times greater in individuals with diabetes compared to those without in the Framingham Study (67). Notably, the gender difference in CVD incidence seen in the general population seems to be narrowed in diabetes, so that women have nearly the same risk as men (56,68).

During the 1980s, CVD risk was much greater in T1DM compared to the general population in two United States cohorts (55,69), including CVD mortality (70,71), though data from the DCCT/EDIC study has shown a reduction in CVD incidence with intensive insulin therapy (72). The Scottish Registry Linkage Study concludes that while the relative risk for CVD mortality associated with T1DM has declined, there remains a significantly elevated risk compared to the general population (51). The new American College of Cardiology/American

Heart Association (ACC/AHA) cholesterol guidelines have proposed a ten-year CVD risk threshold of ≥7.5% for statin therapy for those aged 40-75 years with an LDL-c >70 mg/dl (73), but there are no clear guidelines for patients with long duration childhood-onset T1DM aged <40 years. There are also no recent reports from the United States quantifying contemporary CVD risk in T1DM.

1.3.2 Risk Factors

As stated above, glycemia has not consistently been associated with CVD incidence in type 1 diabetes. In the DCCT/EDIC study, however, the intensive diabetes therapy arm had significantly lower incidence of CAD compared to the conventional therapy arm (74). In contrast, in observational studies, HbA1c has not been a consistent predictor of CVD events (58,60,61,66,75–77). One potential explanation for these seemingly discordant results is that, in the DCCT, intensive therapy was begun early in the course of diabetes, raising the suggestion that glycemia may play a role in the initiation of atherogenesis (78). Despite the level of HbA1c not predicting CAD incidence, in the EDC study, decreasing HbA1c over time has been associated with a lower risk of CAD (62), as shown in Figure 1.

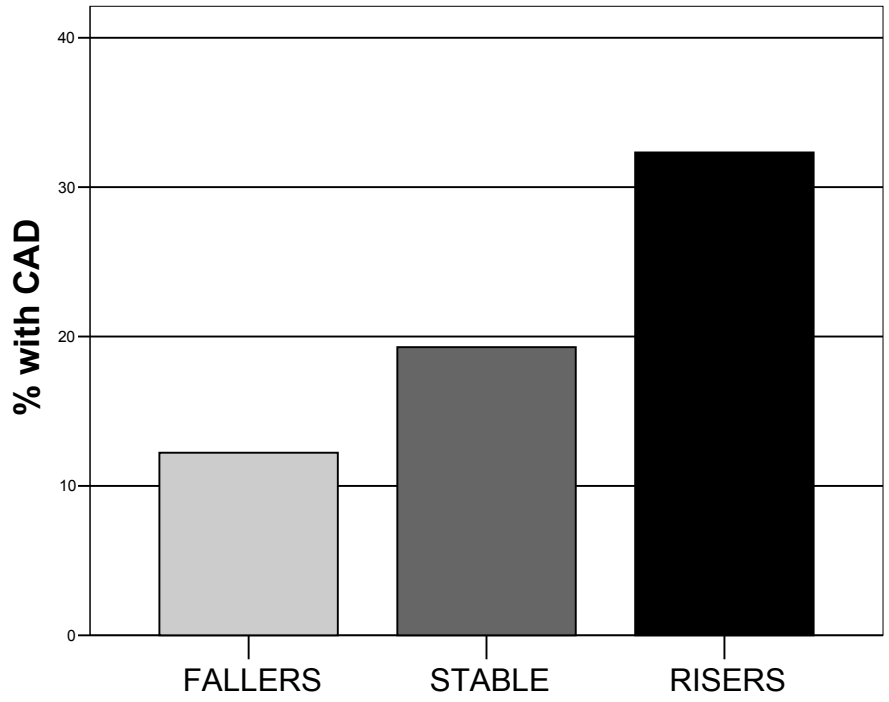

Figure 1. Coronary Artery Disease (CAD) incidence by HbA1c change category in the EDC study. Fallers: a fall of 1 standard deviation or greater; Risers: a rise of 1 standard deviation or greater, Stable: change within 1 standard deviation. Adapted from Prince et al. Diabetologia. 2007; 50(11): 2280-8.

Additionally, HbA1c variability, but not mean HbA1c over time, was predictive of CVD events in the FinnDiane study (77). In contrast to this inconsistent relationship between HbA1c and CAD, glycemia has been more consistently associated with an increased risk of PAD (58,79,80). This difference may indicate that hyperglycemia is more strongly associated with stable atherosclerosis, which characterizes PAD, rather than plaque rupture, which characterizes the acute coronary events (78). In the general population, the majority of acute coronary events are precipitated by plaque rupture and a small proportion is due to plaque erosion (81,82). In type 1 diabetes, these proportions are reversed with the majority of events due to plaque erosion (81,82), adding more evidence that hyperglycemia is associated with stable atherosclerosis.

Other diabetes complications, specifically nephropathy and neuropathy are strong predictors of CVD incidence. While nephropathy is considered a major risk factor for CVD in type 1 diabetes (83–86), this relationship is not well understood. In some cases, adjusting for the traditional cardiovascular risk factors that are potential confounders of the renal-CAD link does attenuate the relationship (60). Also, in the EDC study, nephropathy seems to affect the presentation of CAD, such that individuals with ON are more likely to present with a "harder" or more severe event, like a myocardial infarction (63). Additionally, nephropathy was found to be a stronger risk factor for CAD in men than in women (56), while in the Eurodiab study nephropathy predicted CAD incidence in both sexes (61). Nephropathy is also a strong predictor of stroke (66). Neuropathy has also been associated with an increased risk of CAD (75,87–91), but the nature of this association is not well understood.

In addition to the diabetes-specific factors already discussed, standard CVD risk factors are important predictors of CVD incidence in type 1 diabetes. CVD risk factors appear at younger ages in type 1 diabetes (92). Generally, lipid levels are not particularly high in type 1 diabetes, but it has been reported by the SEARCH study investigators that in youth with type 1 diabetes and poor glycemic control, lipids are often elevated (93). The major standard risk factors also predict CVD incidence in type 1 diabetes, including blood pressure, lipids, and waist-hip ratio (56,75,94,95). Inflammatory factors are also associated with CVD (96–98) and have recently been linked to the progression of coronary artery calcification in type 1 diabetes (99). In general, type 1 diabetes is characterized by high levels of inflammatory cytokines (100–103), which may contribute to the elevated risk of CVD in type 1 diabetes. Although the risk of developing CVD is approximately equivalent by gender in type 1 diabetes, there is evidence that the risk factors for CVD differ in men and women (56). Besides the aforementioned sex

difference in the relationship between CAD and nephropathy, smoking is a stronger risk factor in men, perhaps due to the fact that men are more likely to smoke or smoke more than women, while depressive symptomatology and physical activity appear to predict CAD more strongly in women (56).

There also seems to be some variation in the relationship between specific risk factors and each manifestation of CVD, though few studies of type 1 diabetes have performed the direct comparisons needed to fully address this issue. In a comparison of risk factors for CAD versus LEAD in the Pittsburgh EDC study, after 6 years of follow-up, hypertension, depressive symptomatology, and nephropathy were independently predictive of CAD only, while HbA1c and smoking were independently predictive of LEAD only (58). Diabetes duration and serum lipids (HDL-cholesterol for CAD and LDL-cholesterol for LEAD) were predictive of both manifestations of CVD. A report from the Wisconsin Epidemiologic Study of Diabetic Retinopathy comparing risk factors for CAD and stroke showed that smoking was an independent predictor of CAD only, while HbA1c was predictive of stroke only (75).

Genetic markers have also been identified as potentially important risk factors. The haptoglobin gene seems to be strongly associated with CAD events (104) and coronary artery calcification (105) in type 1 diabetes, consistent with the relationship that had previously been reported in type 2 diabetes (106–108). The endothelial nitric oxide synthase (NOS3) gene and the receptor for advanced glycation end-products (RAGE) gene have also been linked to CVD in type 1 diabetes (109,110).

Many questions remain regarding the prediction of CVD in type 1 diabetes. One important issue is the need to develop a better understanding of the role of glycemia in the pathophysiology of CVD. As discussed above, findings on the association between glycemia

and CVD have been mixed. Another problem that remains is lack of type 1 diabetes-specific target levels for risk factors, as no trials have been undertaken in type 1 diabetes populations. For example, the results of recent trials in type 2 diabetes show that intensive blood pressure control does not seem to decrease the risk of mortality or CVD (111–113), but it is unclear if this relationship is the same in type 1 diabetes, where individuals have longer disease duration at younger ages due to the earlier onset of type 1 diabetes. Finally, more research needs to be done to determine if there are additional high risk subgroups of patients within type 1 diabetes.

1.4 LIMITATIONS OF COMMONLY USED STATISTICAL METHODS IN LONGITUDINAL COHORT STUDIES

Typically, in longitudinal cohort studies, the association between risk factors and the occurrence of the outcome of interest is assessed using Cox proportional hazards regression. The dependent variable in Cox regression is the hazard function of the outcome of interest, while the independent variable(s) are generally entered into the model in their continuous form to obtain a hazard ratio, an estimate of the relative risk (114). When modeled in this manner, the association between risk factor and log cumulative hazard is assumed to be linear (114). Thus, the possibility of a nonlinear relationship or a threshold which stratifies a risk factor into important subgroups is not considered. It is also difficult to model and interpret higher order interactions using traditional regression methods. Additionally, using these types of analyses, the full potential of the longitudinal trajectory of the risk factor data is not often utilized. Instead, only

either baseline variables or the average values of risk factors are often used as predictors. We will expand on these limitations in sections 1.4.1-1.4.3.

1.4.1 Limitation 1: Inadequate assessment of subgroups and higher order interactions

Formal identification of important subgroups is not often attempted in observational cohort studies. Even if identification of such subgroups is attempted, interactions between various groups are not assessed because it is very cumbersome to do so with traditional modeling techniques. Studies where these types of analyses have been undertaken demonstrate the potential importance of such an approach. For example, decision tree analyses of the Western Collaborative Group Study data identified various groups of men, based on interactions between obesity, age, and hypertension status, who were at very high risk of coronary heart disease mortality (115). In the EDC study, these methods were used to identify subgroups of participants based on biologic risk factors who were at low risk for complications (116). Genetic markers were then examined within the low risk groups and, using this approach, associations between the genes and complications emerged that were not observed in the overall study population (116).

1.4.2 Limitation 2: Failure to assess nonlinearity

As stated above, the commonly used regression techniques and tests for trend assume an underlying linear relationship between the exposure and the outcome. If the true relationship is nonlinear, then inferences will be invalid, potentially leading to important clinical consequences

(117–119). An example of this can been seen in the relationship between blood pressure and CVD (120). Blood pressure has been shown to have a J-shaped relationship with CVD. However, in most studies only a linear relationship is examined. Hence, there is the potential for misinformed clinical decisions to be made, if the increased risk at lower blood pressure is not recognized. Determining an accurate representation of the shape of the relationship is critical for prevention efforts (121). In cases where the relationship between a risk factor and outcome is U-shaped, testing for a linear association may not detect a relationship at all (117,122). Another drawback of assuming an underlying linear relationship is that the potential for threshold effects is not assessed.

A common way to assess nonlinearity is to categorize or dichotomize the risk factor. While this is a straightforward way to examine the shape of the association, it is usually based on arbitrary or sample-based groupings, which may not adequately describe the association or identify the optimal cut-points (123). In many cases, different choices of groupings may also lead to differing conclusions (124–126). Categorization also leads to less statistical efficiency, compared to using continuous variables (127) and results in a step-function, which is not biologically plausible (125).

1.4.3 Limitation 3: Neglecting the full potential of longitudinal data

Analyses of data from longitudinal cohort studies also rarely make full use of the information contained in the serially measured risk factors. Commonly, baseline variables are used as predictors, but this approach, while generally providing valid results, does not account for changes in the variables over time that may affect risk (128). For example, this approach

may underestimate the effect of risk factors that become important closer to the time of the event. Stability and change of predictors is relevant for determining the likelihood of an individual having an event, thus incorporating within-subject patterns is important for obtaining a more complete understanding of the relationship between risk factor and outcome (128–131). In some cases, failing to utilize all available data may lead to invalid results (128).

One common approach for accommodating change of risk factors in statistical analyses is to model change (132). However, this approach does not account for the actual pattern of change and assumes a constant, linear increase or decrease over the follow-up and may also underestimate the variance of the factor (133). Another common approach for utilizing longitudinal data is to model the average (or updated mean) risk factor over the follow-up, however this has similar drawbacks to modeling change, leading to a loss of important information (134).

1.5 STATISTICAL APPROACHES TO ADDRESS LIMITATIONS OF TRADITIONAL METHODS

In an attempt to overcome the aforementioned issues, the objective of this dissertation is to utilize statistical methods which can explore higher order interactions and important subgroups, accommodate potential nonlinearity, and make use of serially measured risk factors to better understand the association between the 20-year incidence of the various first manifestations of CVD and risk factors in the Pittsburgh Epidemiology of Diabetes Complications Study. The details of these methods will be described in sections 1.5.1-1.5.3.

1.5.1 Identification of high risk subgroups and interactions using tree structured survival analysis

The aims in using this approach are: 1) to determine whether there are meaningful subgroups at high risk for CAD, stroke, or PVD derived from the many variables measured in the EDC study and 2) to identify interactions between the groups. These goals will be accomplished through the identification of particularly important categorical variables as well as identifying threshold effects or cut-points for continuous variables that divide the sample into high-risk subgroups. Tree-structured survival analysis (TSSA), a decision tree approach that is able to take time to CVD event into account, will be used to accomplish this aim. This approach will also facilitate the identification of important interactions between risk factors. It is important to note that analyses using classification tree methods are exploratory in nature (135), however, they may be useful in identifying previously unknown high-risk subgroups and interactions between risk factors.

TSSA falls under the general machine learning method of decision trees, which were initially proposed by Morgan and Sonquist in the 1960s (136) and refined by Quinlan in the 1970s (137,138). In 1984, Breiman *et al.* developed Classification and Regression Trees (CART), an algorithm which generated decision trees with continuous and categorical outcomes (139). The general idea underlying decision tree construction is that attributes are split into subsets (nodes) and then these are again split by another attribute. The process is repeated recursively until all instances within a node have the same outcome classification (140). Traditionally, decision trees are used to derive subgroups for the prediction of a binary outcome. However, in the 1980s and 90s, the decision tree framework was extended to accommodate time-to-event survival data with censored observations (141–143). Cross-

validation is the most commonly used method for training and testing to find the best tree. Rather than simply building a single tree and holding out a proportion of the data for testing, cross-validation is an iterative process which repeats the process several times on random samples from the data (140). Then error rates from the various iterations are used to identify the best fitting and most parsimonious tree (144).

To accomplish the proposed aims, the rpart (recursive partitioning) routine in the statistical package R version 3.0.2 (145) will be utilized. The rpart routine has been described in detail by Therneau and Atkinson (144). The routine uses the basic method of recursive partitioning described above. Subgroup splits are chosen based on an "impurity function", such as the information index and the Gini index, where the value that minimizes these functions is chosen as the best cut-point. In order to model data where the outcome is a time to event with censored observations, the time values are scaled so that log(survival) is linear under an exponential model, which results in a tree that is equivalent to the local full likelihood decision tree for relative risks (143). To avoid over-fitting, the resulting tree is pruned by selecting a tree size with the minimum cross-validated error, as measured by a complexity parameter (144). Importantly, rpart is able to accommodate missing data and as long as each subject has data for at least one independent variable, they are included.

TSSA has previously been utilized to analyze 10-year follow-up data from the EDC study (116). In those analyses, separate decision trees were constructed to identify the subgroup at the lowest risk of coronary artery disease, confirmed distal symmetric polyneuropathy, and overt nephropathy, based on biologic risk factors only. Within the lowest risk groups, the relationship between genetic markers and the outcome were examined. The current proposed TSSA analyses differ from the prior analyses in the following ways: 1) the follow-up period now reaches 24

years, so there are now 192 cases of CAD, compared to only 99 in the original analysis, and 2) the objective of the current analysis is to identify subgroups at particular risk for both total CVD and the different first manifestations of CVD (CAD, stroke, and PAD). Data from EDC participants who were free of CVD at baseline will be utilized to construct a decision tree with time to CVD event as the outcome with the rpart routine in R. Separate trees will also be constructed for each of the three first manifestations of CVD, in an attempt to identify subgroups and potentially important interactions between risk factors which lead to particularly high or low risk for developing each specific manifestation.

Decision trees have many advantages in epidemiologic research, most notably that the results are presented in a manner that is similar to the thinking process used to approach medical questions. Thus, the resulting trees are well-suited for clinical application (135,146). Another important advantage is that trees provide a clear representation of interactions between attribute subgroups. Demonstration of higher-order interactions using statistical models is complicated and cumbersome, but they can be intuitively presented using decision trees (115,135,146). Decision trees also have the advantage of avoiding various statistical assumptions, such as linearity or proportional hazards (146). As mentioned above, the major limitation of decision trees is that the method is considered to be exploratory and does not assess the relative importance of risk factors (135). Therefore, the resulting subgroups and interactions may not be definitive and should be replicated in other populations. Finally, it is important to consider the biologic and clinical plausibility of the results, and to perform sensitivity analyses to avoid misleading inferences (135,147).

1.5.2 Assessment of nonlinear associations between risk factors and incidence of CVD outcomes using restricted cubic splines

The aim of using this approach is to establish the shape of the relationship between risk factors and risk of CVD using restricted cubic splines. Allowing for potential nonlinear associations between risk factors and CVD incidence may lead to the detection of threshold effects or protective effects at certain levels of a factor (i.e. U- or J-shaped relationships, as described above). Utilizing splines to assess nonlinear relationships helps to avoid pitfalls of categorization or dichotomization, such as loss of power and bias due to choice of cut-points (117,124,126) and also allow the general shape of the relationship to be inferred (117). These issues are detailed in the 2005 paper by May and Bigelow, in which various methods of assessing nonlinear, dose response relationships were compared (117). They concluded that, for detection of a nonlinear, dose-response relationship, restricted splines were best at representing the shape of the association, compared to linear, categorization, and fractional polynomial approaches (117). Spline regression is very flexible and can represent almost any possible shape of an association (119). While underutilized in epidemiologic studies, spline regression has revealed interesting nonlinear relationships between risk factors and CVD outcomes. In a recent report from the Tehran Lipid and Glucose study, fasting plasma glucose was found to have a curvilinear relationship with CVD events and both all-cause and CVD mortality (148). Blood pressure has been shown to have a J- or U-shaped relationship with CVD events in many studies (149–154). Examples where this issue was formally examined using spline regression include the Cardiovascular Health Study (155), where the relationship was determined to be linear, and the Treating to New Targets (TNT) trial (120), where the relationship was J-shaped.

Briefly, spline regression is characterized by piecewise functions where each piece is a low-order polynomial defined over a certain interval of the continuous variable (126). The point between any two intervals is called a "knot", denoted k. Usually, continuity constraints of the curve itself and perhaps the corresponding first or second derivatives are imposed at the knots that connect the piecewise polynomials giving the spline a smooth appearance. Generally, five knots are used (or fewer for smaller data sets) and these are located at fixed percentiles (119). Placing the knots at fixed percentiles helps to avoid subjective bias and makes results comparable across variables and studies (156). Commonly, for k=5, knots are placed at the 5th, 25th, 50th, 75th, and 95th percentiles (156). Spline functions tend to be quite flexible and knot placement does not significantly affect results (157). An important characteristic of spline functions is that the slope of each interval is constrained to be the same as the slope of the next interval at their common knot to avoid a sudden change in risk between intervals (i.e. smoothed) (126,158). Cubic spline functions, which are sums of third degree polynomials, are commonly chosen and produce very smooth curves (126). "Restricted" spline functions are constrained to be linear before the first and after the last knot. This restriction helps to avoid poor model fit that often occurs at extreme ends of the distributions (117,126,158).

Restricted cubic splines are created by 1) dividing the continuous variable into intervals using knots at fixed percentiles, 2) create a new variable for each knot that is zero below the knot and a third order polynomial above, 3) smooth the slopes of the piecewise functions at the knots, as described above, and 4) restrict the first and last intervals to be linear (119). To determine statistical significance, first, the null hypothesis that the slopes of all of the piecewise functions are equal to zero is tested. If this initial test is not statistically significant, then the factor and

outcome are not associated, but if the initial test is significant, then a second test is performed with the null hypothesis that the nonlinear coefficients are zero.

The restricted cubic spline method will be utilized with the EDC data to determine whether risk factors are related to CVD incidence in a nonlinear fashion and whether the shape of the relationships differs by the first manifestation of CVD. No systematic assessment of nonlinear relationships between risk factors and CVD incidence has been previously undertaken using EDC data. Prevalent cases of CVD will be excluded and risk factor variables will be modeled using restricted cubic splines in Cox proportional hazards models for risk of incident CVD. All analyses will be performed using the RCS macro for Cox regression with restricted cubic splines (156) in SAS 9.3 (SAS Institute Cary, NC). The RCS macro also provides a graphical representation of the estimated spline function, which facilitates interpretation of the results (156).

The major advantage of using restricted cubic splines is that the shape of the relationship between risk factors and outcomes can be inferred (117). Failing to consider possible nonlinearity could lead to false null results or misrepresentation of the true relationship as being linear when it is not (117). Splines also provide more power to detect a relationship compared to other methods of assessing nonlinearity, such as categorization (159). The main limitation of cubic splines is that they can be sensitive to the choice of knot location (126), so it is important to test different locations and perform sensitivity analyses.

The results of these analyses examining the dose-response relationship between risk factors and cardiovascular disease in the EDC cohort are presented in Appendix A.

1.5.3 Utilizing the serial measured risk factor data in the EDC study to predict the incidence of cardiovascular disease using joint modeling of longitudinal and time-to-event data

The aim of this approach is to assess the predictive capability of serially measured risk factor data on the incidence of total cardiovascular disease, as well as its various manifestations, using joint longitudinal and time-to-event models. Utilizing the longitudinal data available in the EDC study in this manner will accommodate information on both the rate of change of covariates, as well as their relationship with the time to CVD events. Joint modeling is not commonly used in epidemiologic studies, though the method has been applied in varied settings. The HIV literature has several examples of joint modeling being used to determine whether longitudinal CD4+ counts improve the prediction of survival (128,160–162). Additional recent examples of joint modeling in epidemiologic studies include analyses to determine whether longitudinal measures of allograft valve function were predictive of patient survival (163) and to examine the relationship between longitudinal measures of red blood cell distribution width and survival in heart failure patients (164).

Joint modeling allows both the longitudinal repeated covariates, usually modeled using random effects, and the time-to-event data to be modeled together. Briefly, a joint model consists of two sub-models. A mixed-effects model defines the change over time of the longitudinal covariate. The modeled random effects are simultaneously included in a survival model, so that the relationship between the pattern of change of covariates and time to the event of interest can be estimated (131). These random effects, or subject-specific deviations from the average effect, are assumed to be the underlying link between the longitudinal repeated measures and the subject-specific hazard for the time-to-event outcome (128).

Joint modeling can be performed using the joineR routine in R. Before starting, the data must be stored in three "data frames": 1) the repeated measures data must be stored in a frame with the format of one row per observation per subject (i.e. multiple rows per subject), with a column for id and another for the corresponding follow-up time, 2) the time-to-event survival data must be stored in a frame with one row per subject, and 3) the baseline covariate data for adjustment must also be stored in a frame with one row per subject (165). Beginning with exploratory analyses, the individual profiles of repeated measures data can be plotted using joineR for subgroups of the subjects, to examine any potential important patterns or outlying subjects. Then longitudinal and time-to-event hazard sub-models can be fit, specifying a random intercept and slope, which corresponds to the previously described method detailed by Wulfsohn and Tsiatis (160,165).

Joint modeling will be utilized with the EDC data to determine whether longitudinal patterns of covariate change predict CVD incidence and whether these relationships differ by the first manifestation of CVD. While survival analyses and mixed models analyses have been performed extensively on EDC data, jointly modeling both processes has not previously been carried out. The 24-year follow-up data for CVD incidence will be used in the current analyses. Clinical risk factors were measured every two years from baseline (1986-1988) until the 10-year follow-up (1996-1998) and then again at 18-years. Prevalent cases of CVD will be excluded and repeatedly measured risk factor variables will be modeled as predictors in joint models where the outcome is time to first CVD event. Models will be repeated separately with time to CAD, stroke, and PVD as the outcome. All analyses will be performed using the joineR routine in R (145,165).

The major advantage of joint modeling is that it allows the relationship between longitudinal repeated measures and a time-to-event outcome to be examined. Using simpler methods, such as modeling the two processes separately, may lead to bias if some subjects survive far beyond the most recent follow-up. Our proposed joint model is able to account for this differential follow-up time by including the subject-specific random effects within the survival model (128,166). As emphasized by Lim *et al.,* joint modeling may be particularly useful when the goal is to use longitudinal measures of biomarkers to improve prediction of an event (128). The main limitation of joint modeling is the potential for bias if the censoring process is related to the longitudinal repeated measures data (i.e. informative censoring) (160). Another limitation is that joint modeling is not included in any of the widely used commercial software packages, so model fitting is more difficult than when using more established methods.

2.0 ESTIMATING CONTEMPORARY TOTAL MORTALITY AND CARDIOVASCULAR DISEASE RISK IN YOUNG ADULTS WITH TYPE 1 DIABETES

2.1 ABSTRACT

The degree to which mortality and cardiovascular disease (CVD) incidence remains elevated in young US adults with type 1 diabetes (T1DM) is unclear. We determined contemporary rates for adults <45 years old with long-standing, childhood-onset T1DM. Members of the Epidemiology of Diabetes Complications (EDC) study cohort <45 years old during the 1996-2012 follow-up period (n=502) were studied. Mortality and CVD rates were calculated for ages 30-39 and 40-44 years. Data from the background Allegheny County, Pennsylvania population were used to calculate age- and sex-matched standardized mortality (SMR) and incidence rate ratios (IRR). In both age groups, the SMR for total mortality was approximately 5. CVD mortality SMRs ranged from 20-33. Hospitalized CVD IRR was approximately 8; revascularization procedures account for much of the increased risk. For all outcomes the relative risk was larger in women. Participants aged 30-39 years had 6.3% absolute 10-year CVD risk, approaching the American College of Cardiology/American Heart Association recommended cut-point of 7.5% for initiation of statin therapy in older adults. Total and CVD mortality and hospitalized CVD are all significantly increased in this contemporary US

cohort of young adults with long-standing T1DM. These findings support more aggressive risk factor management in T1DM, especially among women.

2.2 INTRODUCTION

Type 1 diabetes (T1DM) is consistently associated with increased mortality, though the excess risk seems to have decreased over recent years (49). Large differences in T1DM mortality are seen across countries, with smaller excess mortality generally observed in European countries and larger excess observed in the United States (49). There is also a sex difference, with women with T1DM having a greater excess mortality compared to men (50). More recent reports from the Scottish Registry Linkage Study and the Swedish National Diabetes Register both show that T1DM continues to be associated with an increased mortality compared to the general population across all age groups, with younger women at particularly high risk (51,52). A recent report from Australia further suggests that younger patients (<40 years) with T1DM are not experiencing a decline in "diabetes" mortality, in contrast to other causes and age groups (53). Contemporary data from the United States focused on young adults with long-duration T1DM are lacking, though recent findings from the Diabetes Control and Complications Trial/Epidemiology of Diabetes Interventions and Complications (DCCT/EDIC) study show lower all-cause mortality with intensive insulin therapy compared to conventional therapy (54).

Atherosclerotic cardiovascular disease (CVD), including coronary artery disease (CAD) is a major complication of T1DM (78). During the 1980s, CVD risk was much greater in T1DM

compared to the general population in two United States cohorts (55,69), including CVD mortality (70,71), though data from the DCCT/EDIC study has shown a reduction in CVD incidence with intensive insulin therapy (72). The Scottish Registry Linkage Study concludes that while the relative risk for CVD mortality associated with T1DM has declined, there remains a significantly elevated risk compared to the general population (51). The new American College of Cardiology/American Heart Association (ACC/AHA) cholesterol guidelines have proposed a ten-year CVD risk threshold of ≥7.5% for statin therapy for those aged 40-75 years with an LDL-c >70mg/dl (73), but there are no clear guidelines for patients with long duration childhood-onset T1DM aged <40 years.

As there are no reports from the United States quantifying contemporary mortality and CVD risk in young adults with long duration T1DM, our objective was to determine such rates for individuals <45 years old with long-standing, childhood-onset T1DM for the years 1996-2012 and to compare these data from the Pittsburgh EDC Study to the age-matched background population in Allegheny County, Pennsylvania.

2.3 METHODS

2.3.1 Study Design and Population

The Pittsburgh EDC Study is a prospective cohort study of childhood-onset (<17 years old) T1DM. All participants were diagnosed, or seen within one year of diagnosis, at Children's Hospital of Pittsburgh between 1950 and 1980. The cohort has been described in detail

elsewhere (167,168). In brief, participants have been followed since 1986-1988, initially with biennial examinations and questionnaires for ten years and thereafter with biennial questionnaires and further examinations at 18 and 25 years post-baseline. The current analyses were restricted to EDC participants who were <45 years old during the January 1, 1996 to December 31, 2012 follow-up period (n=502). Research protocols were approved by the University of Pittsburgh institutional review board and all participants provided written informed consent.

2.3.2 Ascertainment of Mortality and Cardiovascular Outcomes

In the EDC study, mortality was ascertained using medical records, death certificates, autopsy reports, and/or interview with next of kin and classified according to the Diabetes Epidemiology Research International (DERI) system (169) by a committee of physicians. CVD mortality was defined as fatal CAD, myocardial infarction, or stroke. Incidence of nonfatal CVD events (i.e. myocardial infarction, hospitalized coronary artery bypass graft, hospitalized angioplasty, and stroke) was self-reported by the study participant and confirmed by medical records.

Mortality data for the age-matched background Allegheny County, Pennsylvania population between 1996-2012 were obtained through the Allegheny County Health Department and provided by the Division of Health Informatics, Pennsylvania Department of Health (the Department of Health specifically disclaims responsibility for any analyses, interpretations, or conclusions). The data were abstracted using the Pennsylvania Department of Health Epidemiologic Query and Mapping System (EpiQMS) by selecting total deaths, and separately,

coronary heart disease deaths (ICD10/ICD9: I11, I20-I25, I516/402, 410-414, 4292) and stroke deaths (ICD10/ICD9: I60-69/430-438). CVD hospitalization data for the age-matched background Allegheny County, Pennsylvania population were only available for the years 2004-2010 and were abstracted from Pennsylvania Health Care Cost Containment Council (PAHC4) data and obtained through the Allegheny County Health Department.

2.3.3 Risk Factor Assessment

Fasting blood samples were taken to measure HbA1, lipids, and lipoproteins. HbA1 values were converted to DCCT-aligned values using a regression equation derived from duplicate assays (DCCT HbA1c=0.14 + 0.83[EDC HbA1]) (62). Serum total cholesterol and triglycerides were determined enzymatically (170,171) and HDL cholesterol was determined using a modified precipitation technique (172) based on the lipid research clinics method (173). LDL-cholesterol levels were calculated from the measurements of total cholesterol, triglycerides and HDL-cholesterol using the Friedwald equation (174). Use of lipid-lowering therapy was self-reported. Blood pressure readings were taken according to the Hypertension Detection and Follow-Up protocol (175) and hypertension was defined as having either blood pressure >140/90 or self-reported use of blood pressure lowering therapy. Urinary albumin was measured by immunonephelometry (176). Albumin excretion rate (AER) was calculated for each of three timed urine samples (24-hour, overnight, and 4-hour collections obtained over a two week period); the median of the three AERs was used in analyses. Current smoking status was obtained by self-report.

2.3.4 Statistical Analyses

Analyses to quantify total mortality, CVD mortality, and CVD event rates included EDC data from the complete follow-up period between 1996 and 2012. Crude rates were calculated as the total number of events per 100,000 person years with 95% confidence intervals for the age groups 30-39 years and 40-44 years. Outcomes were defined as total mortality, CVD mortality (i.e. fatal CAD or stroke), total CVD (i.e. Fatal CAD or stroke, myocardial infarction, nonfatal stroke, or hospitalized coronary artery bypass graft or angioplasty), ACC/AHA CVD (i.e. Fatal CAD or stroke, nonfatal myocardial infarction, nonfatal stroke), and total CAD (i.e. Total CVD excluding fatal/nonfatal stroke).

To quantify the excess risk of total mortality, CVD mortality, CVD, and CAD incidence in the EDC cohort, the expected number of events for each age group was calculated by multiplying the person-years at risk in the EDC cohort by the mortality and CVD and CAD incidence rates in the background Allegheny County population. Standardized mortality (SMR) and incidence rate ratios (IRR) were calculated, as appropriate, as the ratio of the observed number of events in EDC to the expected number of events. Due to the hospitalization data for the background Allegheny County population being restricted to 2004 to 2010, comparisons to the general population were carried out by restricting the EDC data to participants who were <45 years old between 2004-2010 and performing a comparison of their CVD rates to the 2004-2010 Allegheny County data. Finally, to address the possibility of survivor bias in the EDC data, a sensitivity analysis was performed in which all mortality and CVD events occurring between age 30 and 44 ascertained during the complete 1986-2012 EDC study follow-up were included.

Statistical analyses were performed using SAS version 9.3 (SAS Institute Inc. Cary, NC) and the Open Source Epidemiologic Statistics for Public Health (OpenEpi) calculators (177).

2.4 RESULTS

Of the 391 participants who contributed follow-up between ages 30-39 during 1996-2012, 5.6% (22) had at least one CVD event and 4.6% died during the 1996-2012 follow-up period. Of the 474 participants contributing follow-up between ages 40-44, 8.9% (n=42) had at least one CVD event and 5.1% (n=24) died. Participant characteristics by age group at the 1996-1998 baseline examination are shown in Table 1. In both age groups, those who had CVD events and who died were significantly more likely to have hypertension, be using blood pressure-lowering therapy, and have higher albumin excretion rates. Current smoking was significantly associated with CVD in those aged 30-39 and mortality in those aged 40-44. LDL-cholesterol was associated with CVD in those aged 30-39, but not in those 40-44 years old. HbA1c did not differ significantly in either age group by mortality or CVD status.

Table 1. Baseline (1996-1998) Characteristics by 1996-2012 Total Mortality and Cardiovascular Disease (CVD) Event Status

Age Group[1]	Characteristic	Overall	Died during follow-up	Alive through follow-up	p-value	CVD Event	No CVD	p-value
30-39	Total n (n with clinical data)	391 (260)	18 (10)	373 (252)		22 (14)	369 (246)	
	Diabetes Duration (yrs)	25.1 (5.0)	24.8 (5.2)	25.1 (5.0)	0.77	25.3 (5.5)	25.1 (5.0)	0.81
	Female Sex[2]	49.4% (193)	66.7% (12)	48.5% (181)	0.13	50.0% (11)	49.3% (182)	0.95
	HbA1c (%)	8.4 (1.4)	8.5 (0.8)	8.4 (1.5)	0.93	9.0 (1.5)	8.4 (1.4)	0.13
	LDL-cholesterol (mg/dl)	116.4 (31.2)	122.0 (35.4)	116.2 (31.1)	0.59	132.9 (37.1)	115.3 (30.6)	0.03
	HDL-cholesterol (mg/dl)	53.3 (13.6)	61.0 (22.6)	53.0 (13.1)	0.29	50.9 (14.5)	53.4 (13.6)	0.48
	Triglycerides (mg/dl)[3]	94 (66-144)	94 (73-180)	94 (66-143)	0.76	139 (82-180)	93 (66-132)	0.04
	Lipid-Lowering Therapy[2]	5.1% (19)	13.3% (2)	4.8% (17)	0.18	14.3%(3)	4.6% (16)	0.08
	Hypertension[2]	25.8% (76)	50.0% (9)	24.6% (69)	0.05	52.6% (11)	23.9% (66)	0.006
	Blood Pressure-Lowering Therapy[2]	18.6% (68)	39.0% (7)	17.7% (62)	0.04	47.6% (10)	16.9% (58)	<.0001
	Albumin Excretion Rate (μg/min)[3]	11.5 (5.8-77.4)	310.6 (209-1717)	10.7 (5.6-56.8)	0.0002	296.7 (18.9-1369)	10.5 (5.6-50.2)	0.0004
	Current Smoker[2]	19.8% (73)	25.0% (4)	19.6% (69)	0.53	38.1% (8)	18.7% (65)	0.04
40-44	Total n (n with clinical data)	474 (318)	24 (9)	450 (310)		42 (25)	432 (293)	
	Diabetes Duration (yrs)	27.5 (6.2)	29.0 (5.7)	27.4 (6.2)	0.21	28.3 (5.9)	27.4 (6.3)	0.35
	Female Sex[2]	49.4% (234)	41.7% (10)	49.8% (224)	0.44	40.5% (17)	50.2% (217)	0.23
	HbA1c (%)	8.4 (1.5)	8.7 (1.6)	8.4 (1.5)	0.39	8.6 (2.1)	8.3 (1.4)	0.06
	LDL-cholesterol (mg/dl)	118.2 (31.2)	109.5 (28.6)	118.4 (31.3)	0.45	122.9 (31.0)	117.8 (31.2)	0.47
	HDL-cholesterol (mg/dl)	53.8 (14.0)	61.1 (25.5)	53.6 (13.5)	0.41	55.5 (17.3)	53.7 (13.7)	0.54
	Triglycerides (mg/dl)[3]	90 (66-129)	114 (73-125)	89 (66-129.5)	0.36	99.5 (76-154)	89 (66-126)	0.22
	Lipid-Lowering Therapy[2]	5.8% (26)	15.8% (3)	5.3% (23)	0.09	10.5% (4)	5.3% (22)	0.26
	Hypertension[2]	31.7% (115)	62.5% (10)	30.3% (105)	0.007	50.0% (17)	29.8% (98)	0.02
	Blood Pressure-Lowering Therapy[2]	23.0% (102)	47.4% (9)	21.9% (93)	0.02	42.1% (16)	21.2% (86)	0.003
	Albumin Excretion Rate (μg/min)[3]	12.5 (5.7-57.2)	81.8 (34.8-527.8)	11.0 (5.6-51.6)	0.003	42.0 (14.4-201.3)	10.7 (5.6-50.9)	0.006
	Current Smoker[2]	17.3% (77)	42.1% (8)	16.2% (69)	0.009	25.6% (10)	16.5% (67)	0.15

Values are mean (SD) unless noted. [1]Participant included in an age group if contributed any follow-up within the age range, [2] %(n), [3] median (interquartile range)

2.4.1 Total and CVD Mortality

Age-matched comparisons of total mortality and CVD mortality between the EDC and the background Allegheny County population are presented in Table 2. Total mortality 4.6 times higher than background in the 30-39 year age group (SMR 4.6, 95% CI 2.8, 7.2) 5.4 times higher in the 40-44 year age group (SMR 5.4, 95% CI 3.6, 8.0). In both age groups, relative increase in mortality was greater in women than in men (Table 2). CVD mortality showed greater relative increases than total mortality in both age groups (age 30-39: SMR 95% CI 16.9, 59.4, age 40-44: SMR 20.6, 95% CI 11.9, 33.1) and women again had relative increases than men.

Table 2. Total and Cardiovascular (CVD) Mortality Rates by Age and Sex in the in Epidemiology of Diabetes Complications (EDC) Study Cohort Compared to the Background Allegheny County Population (1996-2012)

Event	Age and Sex	EDC Study Cohort (1996-2012)			Allegheny County (1996-2012)			SMR (95% CI)
		Events	Person Years	Rate per 100,000 (95% CI)	Events	Person Years	Rate per 100,000 (95% CI)	
Total Mortality	30-39	17	2,705	628 (379, 984)	3,817	2,765,085	138 (134, 142)	4.6 (2.8, 7.2)
	Men		1,352	444 (180, 920)	2,537	1,354,618	187 (180, 195)	2.4 (0.97, 5.0)
	Women	11	1,353	813 (428, 1409)	1,280	1,410,467	91 (86, 96)	9.2 (4.8, 15.9)
	40-44	24	1,910	1257 (826, 1835)	3,590	1,551,328	231 (224, 239)	5.4 (3.6, 8.0)
	Men	14	998	1403 (800, 2287)	2,266	750,841	302 (289, 314)	4.7 (2.7, 7.6)
	Women	10	912	1096 (558, 1946)	1,324	800,487	165 (157, 174)	6.7 (3.4, 11.9)
CVD Mortality[1]	30-39	10	2,705	370 (188, 658)	312	2,765,085	11 (10, 13)	33.3 (16.9, 59.4)
	Men	4	1,352	296 (94, 712)	213	1,354,618	16 (14, 18)	18.5 (5.9, 44.7)
	Women	6	1,353	443 (180, 920)	99	1,410,467	7 (6, 8)	63.4 (25.7, 131.8)
	40-44	15	1,910	785 (457, 1263)	583	1,551,328	38 (35, 41)	20.6 (11.9, 33.1)
	Men	9	998	902 (441, 1648)	412	750,841	55 (50, 60)	16.4 (8.0, 30.0)
	Women	6	912	658 (267, 1363)	171	800,487	21 (18, 25)	31.6 (12.8, 65.7)

[1]Fatal coronary artery disease or stroke

2.4.2 CVD Incidence Rates

CVD incidences rates in young adults with T1DM are shown in Table 3. Using the CVD definition, which includes hospitalized revascularization procedures, young adults in aged 30-39 had a CVD rate of 961 events per 100,000 person years (95% CI 643, 1385) those aged 40-44 had a rate of 2,356 events per 100,000 (95% CI 1744, 3113). If the ACC/AHA definition of CVD, which does not include revascularization procedures, is used, the decrease to 628 events per 100,000 person years (95% CI 379, 984) in those aged 30-1,518 events per 100,000 person years (95% CI 1038, 2145) in those aged 40-44. For CAD (which includes hospitalized revascularizations) those aged 30-39 had a rate of 850 per 100,000 person years (95% CI 553, 1253), while the rate for those aged 40-44 was 2,042 per 100,000 (95% CI 1476, 2753).

Age-matched IRRs of CVD rates between young adults in the EDC cohort and background Allegheny County population during 2004-2010 are presented in Figure absolute CVD rates are presented in Table 4). EDC participants aged 30-39 had a total CVD of 599 per 100,000 person years (95% CI 190, 1438), which is 8 times higher than the Allegheny County (95% CI 2·5, 19·3). Those aged 40-44 had a rate of 1,821 per 100,000 years (95% CI 1040, 2964), which is also approximately 8 times higher than the background (Rate Ratio (RR) 7.8, 95% CI 4.4, 12.7). If revascularization procedures are excluded, and ACC/AHA CVD definition is used, the rate for EDC participants aged 30-39 decreases per 100,000 person years (95% CI 50, 986), and is no longer significantly increased background (RR 5.0, 95% CI 0.8, 16.5).

Table 3. Cardiovascular Disease (CVD) Incidence Rates by Age and Sex in the Epidemiology of Diabetes Complications Study 1996-2012

Event	Age and Sex	Events	Person Years	Rate per 100,000 (95% CI)
Total CVD	30-39	26	2,705	961 (643, 1385)
	Men	14	1,352	1036 (591, 1690)
	Women	12	1,353	887 (481, 1503)
	40-44	45	1,910	2356 (1744, 3113)
	Men	27	998	2705 (1827, 3858)
	Women	18	912	1974 (1211, 3041)
ACC/AHA CVD	30-39	17	2,705	628 (379, 984)
	Men	8	1,352	592 (275, 1121)
	Women	9	1,353	665 (325, 1217)
	40-44	29	1,910	1518 (1038, 2145)
	Men	19	998	1904 (1184, 2903)
	Women	10	912	1096 (558, 1946)
CAD	30-39	23	2,705	850 (553, 1253)
	Men	14	1,352	1036 (591, 1690)
	Women	9	1,353	665 (325, 1217)
	40-44	39	1,910	2042 (1476, 2753)
	Men	24	998	2405 (1583, 3503)
	Women	15	912	1645 (959, 2638)

For those aged 40-44, using the ACC/AHA CVD definition reduced the rate to 650 per 100,000 person years (95% CI 238, 1435), which is 4 times higher than background (95% CI 1.4, 8.7). For CAD alone, EDC participants aged 30-39 had 12-fold higher risk than background (RR 11.8, 95% CI 3.7, 28.4), while those aged 40-44 had 8-fold higher risk (RR 8.3, 95% CI 4.4, 14.5). Again, for all definitions and both age groups, women in the EDC cohort had higher relative increases in risk over background compared to men, despite men having higher absolute risk.

2.4.3 Sensitivity Analysis

When the estimates for total mortality and total CVD events were repeated including the entire 1986-2012 follow-up, EDC participants aged 30-39 had 10-fold greater total mortality (95% CI 7.2, 13.5) and 22-fold greater CVD incidence (95% CI 17.1, 28.2) than background. In those aged 40-44, EDC participants had 9-fold higher mortality (95% CI 6.3, 11.7) and 13-fold greater CVD incidence (95% CI 10.0, 16.1) than background.

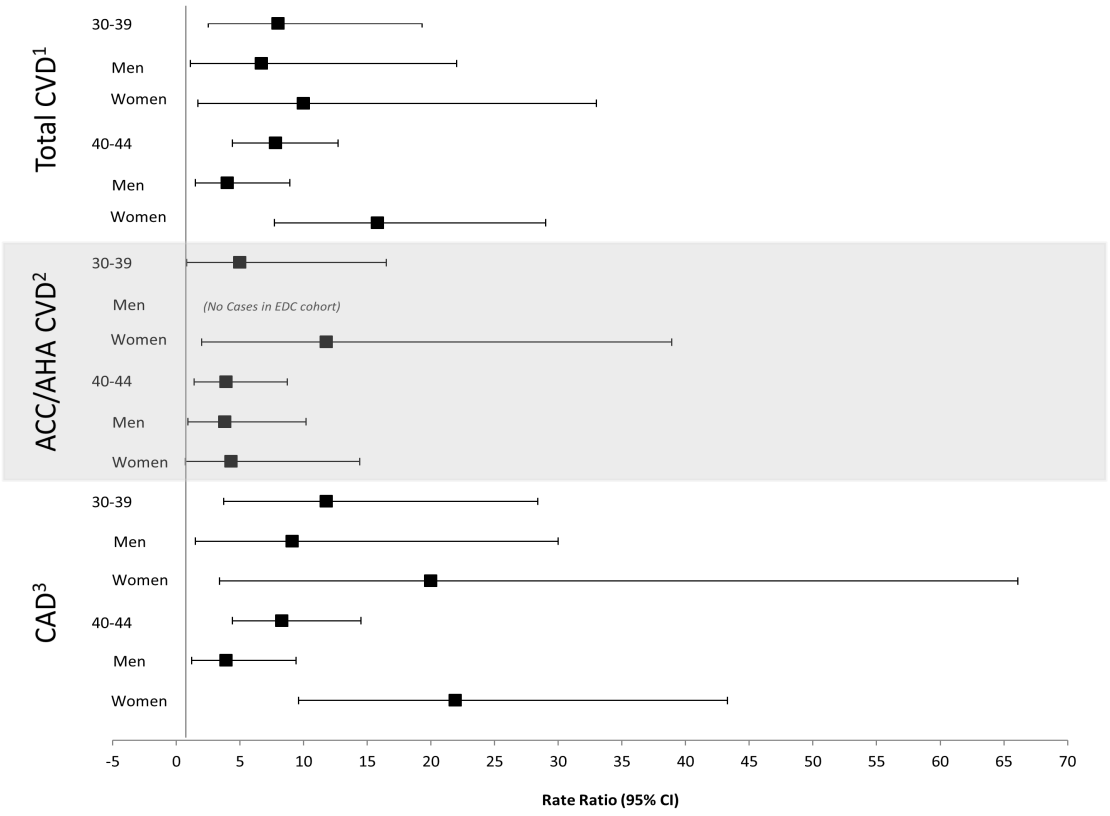

Figure 2. Inicidence rate ratios for the Epidemiology of Diabetes Complications Study cohort compared to the background Allegheny County population, 2004-2010

2.5 DISCUSSION

These results demonstrate that for a contemporary cohort of young adults <45 years old with long duration T1DM, total mortality, CVD mortality, and hospitalized CVD events are still significantly increased compared to the age-matched background population. Total mortality was increased approximately 5-fold, while CVD mortality was increased 20- to 30-fold. Hospitalized CVD events were increased 8-fold, with revascularization procedures accounting for much of the increased risk. If revascularization procedures are excluded, CVD risk was increased 4- to 5-fold. For CAD alone, the increase in risk was even larger than for CVD, at 8 to 11 times that of background. For all outcomes, the relative risk in women was larger than for men.

2.5.1 Relative Risk of Mortality

The excess risk of mortality associated with T1DM in the current analyses is consistent with prior reports. A recent systematic review of 23 reports on population-based T1DM cohorts concluded that T1DM is associated with excess mortality worldwide (49). The magnitude of excess risk varied widely, from no excess in a small cohort from Iceland (178) to an estimated 8.5-fold excess in a cohort from Cuba (179). In a 2010 report from Allegheny County, Pennsylvania, a 6-fold greater mortality than the age-matched general population was observed by 20 years of diabetes duration (180).

More recently, the Scottish Registry Linkage Study reported age-specific mortality and CVD rates (51). Total mortality for men aged 30-39 years with T1DM was 3.4 times that of the general population, while it was 5.2 times that of the general population for women in the same

age group. The SMR for total mortality in women 30-39 years old was markedly higher here in the EDC cohort at 9.2. In the Scottish report, the T1DM population aged 40-49 years had 4 and 4.7 times higher mortality than the general population for men and women, respectively. These increases were again somewhat lower than seen here in the EDC cohort, but as our data examined those 40-44 years old, these results are not directly comparable. It is also important to note that at any given age, the EDC cohort represents a longer duration of diabetes than the Scottish Registry Linkage Study. The Scottish registry includes adult-onset cases of T1DM, and thus the median diabetes duration of 17.5 years at baseline is significantly shorter than that of the EDC cohort (26.6 years), which is restricted to exclusively childhood-onset (<17 years old) T1DM. This difference in the length of exposure to T1DM may explain the lower excess mortality seen in the Scottish study.

Another recent report, from the Swedish National Diabetes Register, showed that even for patients with T1DM who had on-target glycemic control for the past 8 years, both total and cardiovascular mortality were increased compared to the general population (52). When examined by age and sex, regardless of glycemic control, the study reported that men between 18-49 years old had 3- to 4-fold greater total mortality compared to the general population, while women in the same age range had more than 4-fold greater total mortality. Relative increases were larger for deaths due to cardiovascular causes in both sexes, but especially for women, where CVD mortality was more than 7 times that of the general population. These relative increases are again lower than those reported here for young adults in the EDC cohort, particularly for CVD mortality. As in the Scottish Registry Linkage Study, however, the Swedish National Diabetes Register includes adult-onset cases of T1DM, so again, the average diabetes duration of 20.4 years is shorter than that of the EDC cohort, and may at least partially explain

the lower excess mortality. In a separate report on data from the Swedish National Diabetes Register, life expectancy between the time periods of 2002-2006 and 2007-2011 was found to have increased by approximately two years in men but was unchanged in women (181).

In a population-based study from Australia during two time periods, 1997-2003 and 2004-2010, T1DM was also consistently associated with increased total mortality across both sexes and all age groups, except in children <10 years old (182). For men, the SMR was 4.1 in both time periods for those 30-39 years and 4.7 and 4.4, respectively, for those aged 40-49 years in each time period. For women, SMRs were 4.6 and 5 in those aged 30-39 years and 4.7 and 4.2 in those aged 40-49. Total and CVD mortality in the source population declined significantly between 2000 and 2011 (53), with similar trends by sex. The SMRs for total mortality that we report here for the EDC cohort are comparable to the Australian report for men, but are markedly higher in women. The higher relative mortality observed in women was also shown in the 2010 report from the Allegheny County T1DM Registry (180). As noted in that report by Secrest *et al*., it may not be appropriate to directly compare differences in SMR due to variation in methodology or characteristics across the populations. The higher relative mortality observed in women in the EDC cohort and the Allegheny County T1DM Registry compared to the other populations discussed may reflect differences in access to health care across countries, but this hypothesis is difficult to assess. The higher relative risk for women was also noted in a recent meta-analysis of 26 T1DM studies published between 1966 and 2014, where the excess risk of all-cause mortality associated with T1DM was 37% higher in women than in men when pooled across studies (50). Interestingly, in a DCCT-eligible subgroup of EDC participants, the mortality rate of 363 per 100,000 (95% CI 202, 604) is very similar to the rate of 346 per

100,000 (95% CI 272, 440) observed in the DCCT conventional treatment arm (54), lending support to the generalizability of the DCCT data in the United States.

2.5.2 Relative Risk of Cardiovascular Disease

Our finding that young adults with T1DM remain at increased risk of CVD is also consistent with other recent reports. In the Scottish Registry Linkage Study, men and women with T1DM aged 20-39 years had CVD risk 4.8 and 5.5 times that of the general population, respectively (51). The sex difference was larger for those aged 40-49 years, where men with T1DM had 3.1 times the risk of the general population, while women with T1DM had greater than 5 times the risk. The sex difference we observed in the EDC cohort supports the finding that women with T1DM have greater excess CVD risk than men, but the excess was larger than seen in the Scottish study (51). The aforementioned meta-analysis by Huxley et al. reported that the sex difference was more extreme for CAD incidence than it was for mortality (50), which is consistent with our findings. In the pooled results from the meta-analysis, women with T1DM were >13 times more likely to have a CAD event than women without diabetes, while men with T1DM were approximately 6 times more likely to have a CAD event (50). It has been suggested that the effect of hyperglycemia and diabetes on vascular risk is greater in women then in men, possibly due to poorer glycemic control in women (50,183). However, Colhoun et al. also reported that the greater excess risk in women with T1DM was not explained by a worse CVD risk factor profile for women compared to men (184). In an earlier report from the EDC study, depressive symptomatology was found to be associated with CAD incidence in women, but not

in men, suggesting that nontraditional risk factors may explain the excess CVD risk in women and merit further exploration (56).

A comparison of the relative risk estimates reported across countries can be found in Appendix B.

2.5.3 Absolute Risk of Cardiovascular Disease

Despite this consistent evidence that young adults with T1DM are at increased risk of CVD mortality and morbidity, treatment guidelines are based on extrapolations from data on type 2 diabetes and the general population due to a lack of trial data in T1DM. The validity of using these extrapolations has been questioned (185). The 2013 ACC/AHA guidelines recommend high-intensity statin therapy in individuals between ages 40-75 years with diabetes and a ≥7.5% estimated 10-year risk of CVD (73). In these analyses of the EDC cohort, young adults 40-44 years old had a 10-year ACC/AHA definition CVD risk of 15.2%. The guidelines state that, in individuals with diabetes who are younger than 40 years, the decision to begin statin therapy should be individualized (73). Here, in the EDC cohort, young adults 30-39 years old had a 6.3% 10-year CVD risk (ACC/AHA definition), which approaches, but is slightly below, the recommended cut-point of 7.5%. However, three additional CVD events occurred in the 30-39 year old age group prior to the 1996 baseline used in these analyses and approximately 5% of 30-39 year olds in the EDC were already on lipid-lowering medication at baseline (Table 1), with the proportion rising to >20% by 2012 (data not shown). Thus, it is likely that the 10-year CVD risk would exceed 7.5% if it were possible to include the prevalent cases and account for existing lipid-lowering therapy. Additionally, if hospitalized revascularization procedures are included in the CVD definition, the 10-year risk in 30-39 year olds increases to 9.6%. These results

suggest that CVD risk in young adults with T1DM duration >20 years who are between 30-39 years old is most likely high enough to merit the use of statin therapy, given current ACC/AHA thresholds. Furthermore, a recent report has suggested that an individualized approach to statin therapy may be more appropriate, which would further justify statin use in this group of T1DM patients (186). The importance of glycemic control in preventing CVD events remains a critical question. Here, while CVD cases did have higher levels of HbA1c at baseline, the difference did not reach statistical significance. Interestingly, in an examination of CVD risk factors from the DCCT/EDIC study, updated weighted-mean HbA1c was found to be a strong predictor of CVD incidence, second only to age, while LDL-cholesterol was a significant, but weaker predictor (187). These results stress the need for intensive glycemic control, in addition to lipid-lowering therapy, to prevent CVD events.

A comparison of the absolute risk estimates reported across countries can be found in Appendix B.

2.5.4 Strengths and Limitations

Our study has noted strengths. The Pittsburgh EDC is a well-characterized cohort with nearly thirty years of prospective follow-up to date. Deaths and CVD events were verified using death certificates, interviews with next of kin, autopsy reports, and medical records and classified by a committee of physicians using all available information. The limitations include the potential for "survivor bias" due to deaths occurring prior to the 1996-2012 follow-up period of interest here. This issue may lead to an underestimation of the CVD risk of the EDC cohort, as the highest risk participants may have died prior to the beginning of follow-up. We have attempted to assess this possibility by performing a sensitivity analysis, which includes all deaths

and CVD events that occurred in young adults <45 years old during the entire 1986-2012 follow-up of the EDC cohort. These estimates provide an upper limit to the excess mortality and CVD risk in these young adults. Another limitation of these analyses is the restricted availability of CVD hospitalization data from the background population to 2004-2010. Also, using the available data set from Allegheny County, it is not possible to identify which individuals in the background population data have T1DM which, in addition to the EDC participants who are already using lipid-lowering therapy, leads to a potential underestimation of the relative increased risk of the EDC cohort. Finally, these data are from a cohort of people from Southwestern Pennsylvania and may not reflect the experience of all young adults with T1DM in the United States. Also, the EDC is a hospital-based cohort, which may restrict the generalizability of these results. However, it has previously been shown that the epidemiologic characteristics (167) and the mortality experience (188) in the hospital-based EDC cohort closely reflects that of the population-based Allegheny County T1DM cohort for participants who were diagnosed during the same period.

In conclusion, total mortality, CVD mortality, and hospitalized CVD events remain significantly increased in this US contemporary cohort of young adults <45 years old with long-duration T1DM, compared to the background population. As has been consistently demonstrated across international studies, the relative increase in risk is greater for women than for men. These findings support the need for more aggressive CVD risk factor management particularly in women with T1DM.

Table 4. Comparison of Cardiovascular Disease (CVD) Incidence Rates by Age and Sex between the Epidemiology of Diabetes Complications (EDC) Study and Background Allegheny County Population (2004-2010)

Event	Age and Sex	EDC Study Cohort (2004-2010)			Allegheny County (2004-2010)			Rate Ratio (95% CI)
		Events	Person Years	Rate per 100,000 (95% CI)	Events	Person Years	Rate per 100,000 (95% CI)	
Total CVD	30-39	4	668	599 (190, 1438)	797	1,032,759	77 (72, 83)	8.0 (2.5, 19.3)
	Men	2	307	651 (109, 2135)	501	505,822	99 (90, 108)	6.7 (1.1, 22.0)
	Women	2	361	554 (93, 1818)	296	526,937	56 (50,63)	10.0 (1.7, 33.0)
	40-44	14	769	1,821 (1040, 2964)	1,411	604,102	234 (221, 246)	7.8 (4.4, 12.7)
	Men	5	390	1,282 (471, 2818)	934	291,303	321 (300, 341)	4.0 (1.5, 8.9)
	Women	9	379	2,375 (1164, 4313)	477	312,799	152 (139, 166)	15.8 (7.7, 29.0)
ACC/AHA CVD	30-39	2	668	229 (50, 986)	635	1,032,759	61 (57, 66)	5.0 (0.8, 16.5)
	Men	0	307	-	383	505,822	76 (68, 83)	-
	Women	2	361	554 (93, 1818)	252	526,937	48 (42, 54)	11.8 (2.0, 38.9)
	40-44	5	769	650 (238, 1435)	1,001	604,102	166 (155, 176)	3.9 (1.4, 8.7)
	Men	3	390	769 (196, 2079)	619	291,303	212 (196, 229)	3.8 (0.9, 10.2)
	Women	2	379	528 (88, 1732)	382	312,799	122 (110, 134)	4.3 (0.7, 14.4)
CAD	30-39	4	668	599 (190, 1438)	524	1,032,759	51 (46, 55)	11.8 (3.7, 28.4)
	Men	2	307	651 (109, 2135)	376	505,822	74 (67, 82)	9.1 (1.5, 30.0)
	Women	2	361	554 (93, 1818)	148	526,937	28 (24, 33)	20.0 (3.4, 66.1)
	40-44	11	769	1,430 (755, 2472)	1,036	604,102	171 (161, 182)	8.3 (4.4, 14.5)
	Men	4	390	1,026 (327, 2455)	768	291,303	263 (245, 282)	3.9 (1.2, 9.4)
	Women	7	379	1,847 (811, 3619)	268	312,799	86 (75, 96)	21.9 (9.6, 43.3)

3.0 RISK STRATIFICATION FOR 25-YEAR CARDIOVASCULAR DISEASE INCIDENCE IN TYPE 1 DIABETES: TREE-STRUCTURED SURVIVAL ANALYSIS

3.1 ABSTRACT

The formal identification of high risk subgroups, though clinically useful, is generally undertaken in observational cohort studies of cardiovascular disease (CVD). Tree-strucured survival analysis (TSSA) was utilized to determine whether there are meaningful subgroups varying levels of CVD risk in the Pittsburgh Epidemiology of Diabetes Complications study, a prospective cohort study of childhood-onset (<17 years old) type 1 diabetes. Of the participants free of CVD (defined as the first instance of coronary artery disease, stroke, or extremity arterial disease) at baseline, 263 (46.9%) had an incident CVD event over the 25 follow-up. TSSA revealed a range of risk groups, from 24 to 85%, which were define combinations of diabetes duration, non-HDL-cholesterol, albumin excretion rate (AER), white blood cell count. Interestingly, those with short duration diabetes but elevated cholesterol had similar risk to long duration diabetes, while major sex differences were appa for long diabetes duration participants. Our findings suggest that subgroups with major risk differences exist in this cohort of childhood-onset type 1 diabetes. Further study, inc external validation, is needed to better determine the prognostic utility of TSSA for CVD management.

3.2 INTRODUCTION

As in the general population, cardiovascular disease (CVD) is a major contributor to morbidity and mortality in individuals with type 1 diabetes, however the risk of CVD is greatly increased (55,65–67). The reasons for this increased risk are not fully understood, as the hyperglycemia that characterizes type 1 diabetes has been an inconsistent predictor of CVD incidence across studies (57–64,189). Notably, the gender difference in CVD incidence observed in the general population seems to be narrowed in type 1 diabetes, so that women have nearly the same risk as men (56,68).

There is evidence to support the existence of important subgroups with respect to CVD incidence and risk factors in type 1 diabetes. While nephropathy is considered a major risk factor for CVD in type 1 diabetes (83–86), this relationship is also not well understood. In some cases, adjusting for the traditional cardiovascular risk factors that are potential confounders or mediators of the renal-CAD link does attenuate the relationship (60). Nephropathy may also affect the presentation of CAD, such that individuals with nephropathy are more likely to present with a "harder" or more severe event, like a myocardial infarction (63). Additionally, though nephropathy was found to be a stronger risk factor for CAD in men than in women in the Pittsburgh Epidemiology of Diabetes Complications (EDC) study (56), in the Eurodiab study nephropathy predicted CAD incidence in both sexes (61). Although the risk of developing CVD is approximately equivalent by sex in type 1 diabetes, there is evidence that the risk factors for CVD differ in men and women. Besides the aforementioned sex difference in the relationship between CAD and nephropathy, smoking may be a stronger risk factor for CAD in men, while depressive symptomatology and less physical activity appear to predict CAD more strongly in women (56).

Formal identification of important subgroups is not often undertaken in observational cohort studies of CVD. Utilizing methodology to formally identify subgroups at varying levels of risk may reveal relationships between risk factors and outcomes that are not detected using usual methods, such as comparing means, which assesses aggregate data, or regression, which generally assumes a constant increasing risk of an outcome for each unit increment of a risk factor. Neither of these commonly used approaches accounts for potential heterogeneity within the group being studied, the possibility of complex higher-order interactions between risk factors (115), or a potential non-linearity or threshold effect of a risk factor on the outcome of interest (146). Even when identification of subgroups is attempted, interactions between groups are often not assessed because it is cumbersome to do so with traditional regression techniques (135,146). Studies where decision tree methods have been utilized demonstrate the potential importance of such an approach in identifying CVD risk groups (115,116,190). Decision tree analysis of the Western Collaborative Group Study data identified different groups of men, based on interactions between obesity, age, and hypertension status, who were at very high risk of coronary heart disease mortality (115). Also, in an earlier exploration of the Pittsburgh EDC study, decision tree methods were used to identify subgroups of participants who were at low risk for diabetes complications based on biologic risk factors. Genetic markers were then examined within the lowest risk groups identified by the trees and, using this approach, associations between the genes and complications emerged that were not observed in the overall study population (116). These examples demonstrate that decision trees offer advantages to epidemiologic research and may reveal relationships between risk factors and outcomes that would otherwise remain undetected. Additionally, the resulting trees are potentially well suited for clinical applications, as they are presented in a manner that is often similar to the medical

decision-making process (135,146). Thus, we used tree-structured survival analysis, a hierarchical decision tree method that can accommodate time-to-event outcomes, to determine whether there are meaningful subgroups at varying levels of CVD risk in the Pittsburgh EDC study. In addition to total CVD, specific manifestations of CVD, including CAD, stroke, lower extremity arterial disease, and fatal CAD and stroke events were assessed as separate outcomes to explore whether differences in subgroups and higher-order interactions would emerge.

3.3 METHODS

3.3.1 Study population

The Pittsburgh EDC Study is a prospective cohort study of childhood-onset (<17 years old) type 1 diabetes. All participants were diagnosed, or seen within 1 year of diagnosis, at Children's Hospital of Pittsburgh between 1950 and 1980. The cohort has been described in detail elsewhere (167,191). In brief, participants have been followed since 1986-1988, initially with biennial examinations and questionnaires for 10 years and thereafter with biennial questionnaires and further examinations at 18 and 25 years post-baseline. Research protocols were approved by the University of Pittsburgh institutional review board and all participants provided written informed consent.

3.3.2 Ascertainment of cardiovascular outcomes

Participants were followed for 25 years to assess the incidence of total cardiovascular disease (CVD), defined as the first instance of coronary artery disease (CAD) (i.e. fatal CAD, myocardial infarction, revascularization procedure, blockage ≥50%, ischemic ECG, or angina), stroke, or lower extremity arterial disease (LEAD - Ankle brachial index (ABI) <0.8, lower limb amputation due to vascular cause, or intermittent claudication). Fatal events were ascertained using medical records, death certificates, autopsy reports, and/or interview with next of kin and classified according to the Diabetes Epidemiology Research International (DERI) system (169). Myocardial infarction, revascularization/blockage, and stroke were self-reported by the study participants and confirmed by medical records. Amputation was ascertained by self-report and/or physician examination. ABI was determined using a Doppler blood-flow detector with the subject supine and ratios for each of the right and left tibialis posterior and doralis pedis systolic pressures were calculated with arm systolic pressure as the denominator. Intermittent claudication was assessed using the Rose questionnaire (192). In addition to total CVD, the following alternative outcomes were assessed: total CAD, stroke, LEAD, and CVD mortality (CAD death, fatal MI, or fatal stroke). Participants with prevalent CVD at baseline were excluded from analyses of the corresponding analyses outcome.

3.3.3 Risk Factor Assessment

Fasting blood samples were taken to measure HbA1, lipids, lipoproteins, serum creatinine, and serum albumin. HbA1 values were converted to DCCT-aligned values using a regression equation derived from duplicate assays (DCCT HbA1c=0.14 + 0.83[EDC HbA1])

(62). Serum total cholesterol and triglycerides were determined enzymatically (170,171) and HDL cholesterol was determined using a modified precipitation technique (172) based on the lipid research clinics method (173). Non-HDL cholesterol (non-HDLc) was calculated by subtracting HDL cholesterol from total cholesterol. Three seated blood pressure readings were taken with a random-zero sphygmomanometer, and the mean of the second and third readings was used in analyses, according to the Hypertension Detection and Follow-Up protocol (175). Serum creatinine was measured using an Ectachem 400 Analyzer (Eastman Kodak Co.). Pulse rate (beats/minute) was determined by palpitating the radial pulse for 30 seconds and multiplying by two. Serum and urinary albumin were measured by immunonephelometry (176). Albumin excretion rate (AER) was calculated for each of three timed urine samples (24-hour, overnight, and 4-hour collections obtained over a two week period); the median of the three AERs was used in analyses. Glomerular filtration rate was estimated by the CKD-EPI creatinine equation (193). White blood cell count (WBC), red blood cell count, hemoglobin, and hematocrit were measured using a Coulter Counter S-Plus IV. Fibrinogen was measured using the Biuret method. Participants were weighed in light clothing on a balance beam scale. Height was measured using a wall-mounted stadiometer. Body mass index (BMI) was calculated as the weight in kilograms divided by the square of the height in meters. Waist circumference was measured twice at the midpoint between the highest point of the iliac crest and the lowest part of the costal margin in the midaxillary line. If the two measurements differed by >0.5 cm, a third measurement was performed. The mean of the waist measurements was recorded as waist circumference. Hip girth was measured at the widest point of the glutei, usually at the level of the greater femoral trochanter, and the mean of two measures was recorded as the hip circumference, in the same manner as the waist measurements. The ratio of the waist to hip circumference (WHR) was used

in analyses. Current smoking status was obtained by self-report. Insulin dose was calculated by dividing units of insulin administered per day by body weight in kilograms.

3.3.4 Tree-Structured Survival Analysis

In 1984, Breiman *et al.* developed Classification and Regression Trees (CART), an algorithm which generated decision trees for continuous and categorical outcomes (139). The general concept underlying decision tree construction is that attributes are split into subsets (nodes), each of which is homogeneous with respect to the outcome of interest. Split points are based on values of attributes that maximize a criterion quantifying heterogeneity of outcomes among each of the daughter nodes. The largest heterogeneity occurs in the initial split. Subsequent splits can be made on either the same or other attributes. This hierarchical process is repeated recursively until all instances within a node are as homogenous as possible, with regard to outcome classification (140). Traditionally, decision trees are used to derive subgroups for the prediction of a binary outcome. However, in the 1980s and 90s, the decision tree framework was extended to accommodate time-to-event survival data with censored observations (141–143). Cross-validation is the most commonly used method for training and testing to find the best tree. Rather than simply building a single tree and holding out a proportion of the data for testing, cross-validation is an iterative process which repeats the process several times on random samples that partition the data into n relatively equally sized subsets (n is usually 5 or 10) (140). One of the subsets is dropped and the tree is refit. This process is repeated for each of the other n-1 subsets and then error rates from the n fits are used to identify the best fitting and most parsimonious tree (144).

In our analyses, we used the rpart routine of R version 3.0.2 (R Development Team 2013) to fit decision trees to the EDC data for the time to CVD event outcome. The rpart routine has been described in detail by Therneau and Atkinson (144). Data from EDC participants who were free of CVD at baseline were included in these analyses (n=561). The list of baseline variables made available for tree building is shown in Table 5. A minimum of 15 observations in each child node was required for further splitting to be considered. The complexity parameter corresponding to the minimum cross-validated error was chosen to prune the tree. Separate trees were also constructed for each manifestation of CVD outcome variable described above, in an attempt to identify subgroups and potentially important interactions between risk factors. Additionally, separate trees were constructed for men and women and for the two diabetes diagnosis subcohorts (1950-1964 versus 1965-1980). We previously examined these two subcohorts and found that they have significantly different survival experiences (188).

3.4 RESULTS

Of the 561 participants free of CVD at baseline, there were 263 (46.9%) cases of incident CVD over the 25-year follow-up. Participant characteristics at baseline are shown in Table 5. EDC participants who developed CVD over the 25 years of follow-up were significantly older, had longer diabetes duration, higher BMI, WHR, non-HDL cholesterol, triglycerides, systolic and diastolic blood pressures, pulse rate, WBC, serum creatinine, and AER, and lower serum albumin, eGFR, and insulin dose, and were more likely to be smokers.

Table 5. Baseline Charateristics of the Pittsburgh EDC Cohort by 25-Year Total Cardiovascular Disease (CVD) Incidence Status (n=561 free of prevalent CVD at baseline)

Characteristic	CVD Incidence (n=263)	No CVD (n=298)	p-value
Age (years)	**29.9 (7.4)**	**24.3 (7.0)**	**<.0001**
Diabetes Duration (years)	**21.5 (7.4)**	**16.2 (6.4)**	**<.0001**
Female Sex, % (n)	49% (130)	49% (147)	0.98
Non Caucasian, % (n)	1.9% (5)	2.4% (7)	0.78
BMI (kg/m^2)	**23.8 (3.2)**	**23.3 (3.2)**	**0.04**
Waist-Hip Ratio	**0.84 (0.07)**	**0.81 (0.07)**	**<.0001**
HbA1c (%)	8.78 (1.50)	8.71 (1.53)	0.59
HDL Cholesterol (mg/dl)	53.7 (12.6)	54.9 (12.0)	0.25
Non-HDL Cholesterol (mg/dl)	**146.5 (42.8)**	**122.7 (34.8)**	**<.0001**
Triglycerides (mg/dl), median (IQR)	**88.5 (65-131)**	**75.0 (56-102)**	**0.005**
Systolic Blood Pressure (mmHg)	**116.3 (16.2)**	**109.0 (11.4)**	**<.0001**
Diastolic Blood Pressure (mmHg)	**74.1 (11.6)**	**70.6 (9.6)**	**0.0001**
Blood Pressure Med Use, % (n)	**14% (37)**	**3.2% (9)**	**<.0001**
Pulse Rate (bpm)	**79.5 (9.7)**	**76.8 (9.7)**	**0.001**
Hemoglobin (g/dl)	15.0 (1.7)	15.3 (1.5)	0.06
White Blood Cell Count ($x10^3/\mu l$)	**6.77 (1.97)**	**6.17 (1.56)**	**<.0001**
Fibrinogen (mg/dl)	301.8 (97.8)	285.8 (126.5)	0.10
Serum Albumin (g/dl)	**4.58 (0.80)**	**4.73 (0.71)**	**0.02**
Serum Creatinine (mg/dl)	**1.11 (1.18)**	**0.90 (0.61)**	**0.007**
Albumin Excretion Rate (µg/min), median (IQR)	**27.8 (8.8-406.3)**	**10.9 (6.3-28.7)**	**<.0001**
Estimated Glomerular Filtration Rate ($mL/min/1.73\ m^2$)	**98.4 (33.2)**	**111.7 (28.4)**	**<.0001**
Insulin Dose (units/kg body weight)	**0.77 (0.27)**	**0.81 (0.24)**	**0.08**
Ever smoker, % (n)	**44% (115)**	**30% (89)**	**0.0007**
Current smoker, % (n)	26% (68)	19% (56)	0.05

Values are mean (sd), unless noted

3.4.1 Total Cardiovascular Disease Decision Tree

The best survival tree for total CVD incidence identified five distinct risk groups (Figure 4). The tree first split on diabetes duration at the selected cut-point of 20.9 years. In those with diabetes duration <20.9 years, and thus low risk, the tree further split on non-HDLc at 132.5 mg/dl (CVD incidence 24.2% below (Group A) and 50.4% above (Group B) the non-HDLc cut-point). In those with diabetes duration ≥20.9 years the tree first split on AER at a cut-point of 91.7 μg/min and in those with AER<91.7 μg/min, a second split on WBC at 7.15 $x10^3/\mu l$ is identified (CVD incidence 52.5% below (Group C) and 81.5% above (Group D) the WBC cut-point). Finally, in those defined by duration≥20.9 years and AER≥91.7 μg/min, the highest risk group had a CVD incidence of 85.2% (Group E). Kaplan-Meier curves for the five groups confirmed the pattern of CVD incidence identified by the tree (Figure 5). The baseline characteristics of the five groups identified by TSSA are shown in Table 6.

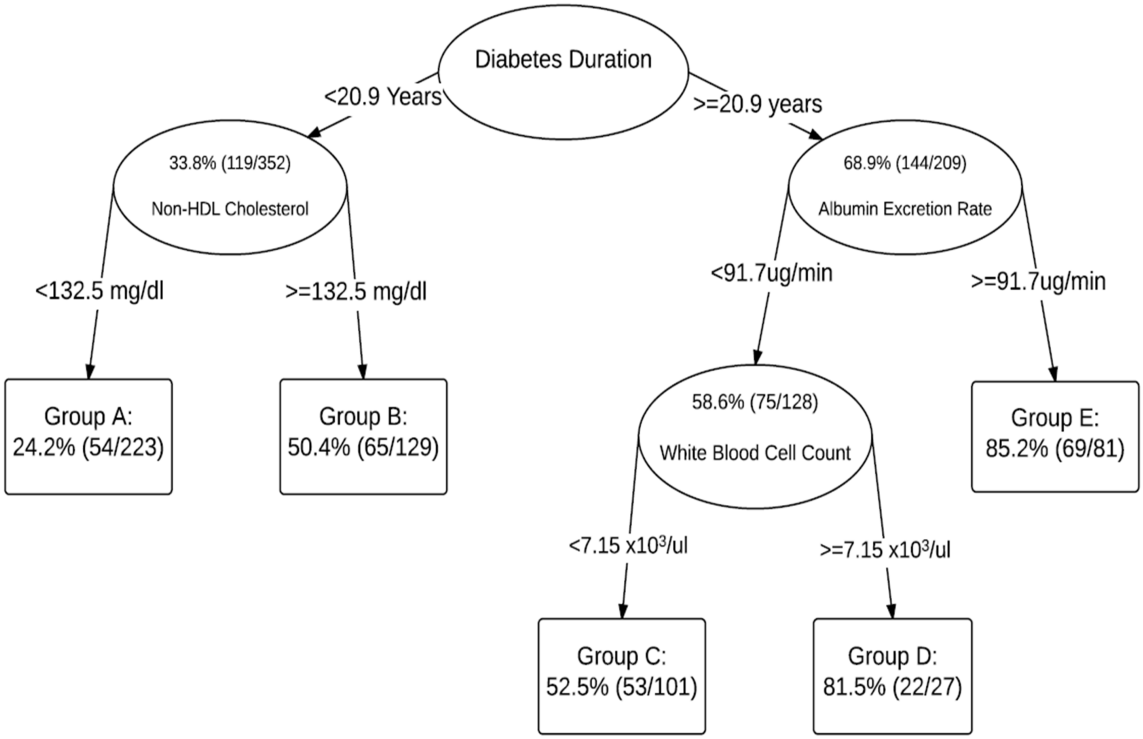

Figure 3. Survival Tree for the 25-Year Incidence of Total Cardiovascular Disease in the Pittsburgh EDC Study Overall Cohort

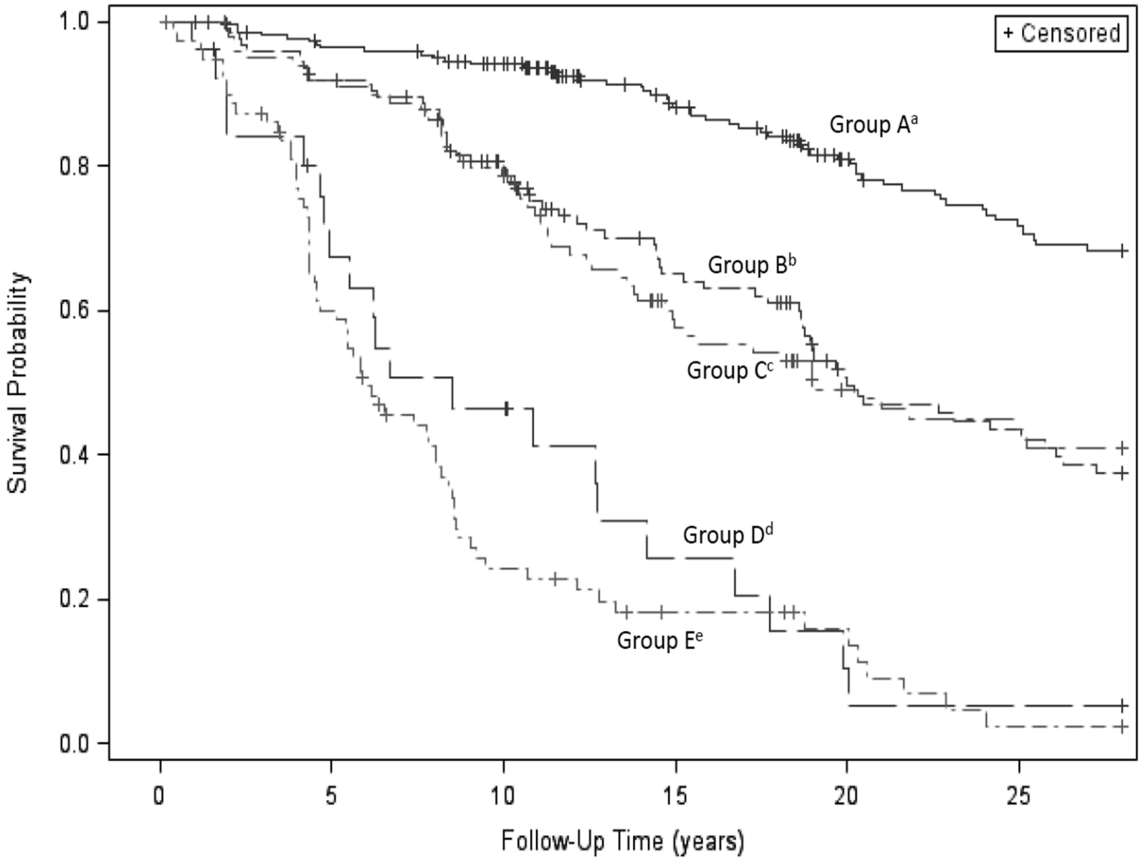

Figure 4. Kaplan Meier Survival Curves for the 25-Year Incidence of Total Cardiovascular Disease by Risk Groups Identified by Tree-Structured Survival Analysis

[a]Group A=Diabetes duration<20.9years, nonHDL-cholesterol<132.5mg/dl

[b]Group B=Diabetes duration<20.9years, nonHDL-cholesterol≥132.5mg/dl

[c]Group C=Diabetes duration≥20.9years, albumin excretion rate<91.7µg/min, white cell count <7.15 x10^3/µl

[d]Group D=Diabetes duration≥20.9years, albumin excretion rate<91.7µg/min, white cell count ≥7.15 x10^3/µl

[e]Group E=Diabetes duration≥20.9years, albumin excretion rate≥91.7µg/min

Table 6. Baseline Characteristics of the Five Pittsburgh EDC 25-Year Total Cardiovascular Disease (CVD) Risk Groups Identified by Tree-Structured Survival Analysis

Characteristic	Group A: 24.2% CVD[a]	Group B: 50.4% CVD[b]	Group C: 52.5% CVD[c]	Group D: 81.5% CVD[d]	Group E: 85.2% CVD[e]
Age (years)	22.3 (5.6)	23.9 (5.1)	33.3 (5.7)	37.0 (4.9)	33.3 (5.4)
Diabetes Duration (years)	13.5 (3.9)	14.7 (3.7)	26.5 (3.8)	28.5 (3.7)	26.6 (4.2)
Female Sex, % (n)	47% (106)	53% (69)	51% (51)	63% (17)	42% (34)
Non Caucasian, % (n)	1.8% (4)	3.1% (4)	2.0% (2)	3.7% (1)	1.2% (1)
BMI (kg/m^2)	23.0 (3.3)	23.8 (3.2)	23.9 (2.9)	23.8 (3.3)	24.0 (3.3)
Waist-Hip Ratio	0.81 (0.06)	0.83 (0.06)	0.82 (0.07)	0.85 (0.09)	0.85 (0.08)
HbA1c (%)	8.55 (1.26)	9.50 (1.77)	8.16 (1.30)	9.02 (1.33)	8.77 (1.64)
Insulin Dose (units/kg)	0.84 (0.23)	0.83 (0.25)	0.69 (0.19)	0.72 (0.27)	0.76 (0.31)
HDL Cholesterol (mg/dl)	54.9 (11.3)	52.4 (11.4)	57.3 (13.7)	57.9 (12.5)	50.8 (13.2)
Non-HDL Cholesterol (mg/dl)	105.7 (16.6)	165.5 (37.7)	129.8 (35.0)	151.2 (32.9)	160.6 (43.4)
Triglycerides (mg/dl), median (IQR)	67.5 (55-89)	112 (79-153)	71 (52-97)	94.5 (68.5-136)	115 (80-173)
Systolic Blood Pressure (mmHg)	107.5 (10.8)	111.3 (12.4)	112.4 (11.7)	115.6 (13.6)	126.6 (18.7)
Diastolic Blood Pressure (mmHg)	69.3 (10.0)	73.0 (9.2)	71.1 (8.5)	72.3 (9.4)	80.6 (13.0)
Blood Pressure Med Use, % (n)	0.95% (2)	5.6% (7)	6.1% (6)	15% (4)	34% (27)
Pulse Rate (bpm)	76.2 (9.8)	79.7 (8.8)	77.2 (8.3)	80.4 (11.1)	81.0 (11.0)
Hemoglobin (g/dl)	15.3 (1.5)	15.4 (1.6)	15.0 (1.5)	14.7 (1.9)	14.8 (1.9)
White Blood Cell Count ($x10^3/\mu l$)	6.12 (1.58)	6.83 (1.70)	5.42 (0.92)	8.52 (1.50)	7.35 (2.25)
Fibrinogen (mg/dl), median (IQR)	250 (210-300)	280 (240-360)	250 (210-310)	300 (240-330)	320 (275-382.5)
Serum Albumin (g/dl)	4.71 (0.71)	4.64 (0.81)	4.82 (0.61)	4.50 (0.59)	4.43 (0.92)
Serum Creatinine (mg/dl)	0.84 (0.29)	0.99 (1.12)	0.88 (0.27)	1.30 (1.75)	1.50 (1.51)
Albumin Excretion Rate (μg/min), median (IQR)	9.75 (6.08-20.2)	20.22 (8.81-168.6)	11.07 (6.51-23.6)	12.36 (6.94-39.4)	830.6 (291.1-1961.4)
Estimated Glomerular Filtration Rate ($mL/min/1.73\ m^2$)	114.8 (25.5)	112.0 (32.6)	102.6 (24.8)	91.6 (34.5)	77.5 (32.6)
Ever smoker, % (n)	21% (48)	42% (54)	41% (41)	63% (17)	56% (45)
Current smoker, % (n)	15% (34)	29% (37)	9.9% (10)	44% (12)	26% (32)

[a]Group A (n=223): Diabetes duration<20.9 yrs, nonHDL-cholesterol<132.5 mg/dl

[b]Group B (n=129): Diabetes duration<20.9 yrs, nonHDL-cholesterol≥132.5 mg/dl

[c]Group C (n=101): Diabetes duration≥20.9 years, albumin excretion rate<91.7μg/min, white cell count $<7.15\ x10^3/\mu l$

[d]Group D (n=27): Diabetes duration≥20.9 years, albumin excretion rate<91.7μg/min, white cell count $\geq 7.15\ x10^3/\mu l$

[e]Group E (n=81): Diabetes duration≥20.9 years, albumin excretion rate≥91.7μg/min

Separate decision trees were also derived for total CVD incidence by sex and diabetes diagnosis subcohort (1950-1964 versus 1965-1980) (Figure 6). In women (panel A), the 25-year

incidence of CVD was 46.9% (130 cases) and the tree first split on diabetes duration at 17.6 years. In lower risk women with diabetes duration <17.6 years, non-HDLc further discriminated CVD risk, with a chosen cut-point of 125.3 mg/dl (CVD incidence 13% below and 46% above cut-point). In women with diabetes duration ≥17.6 years, no further discriminating factors were identified and the 25-year risk of CVD was 66%. In men (panel B), there was a similar 46.8% (133 cases) 25-year incidence of CVD, but the tree first split on AER at 104.3 μg/min. In men with AER<104.3 μg/min, the tree split on non-HDLc. Those with non-HDLc <133.2 mg/dl had a 24.1% 25-year risk of CVD, while in those with non-HDLc ≥133.2 mg/dl waist-hip ratio further discriminated risk (CVD incidence 0% below and 64.3% above cut-point). Men with AER ≥104.3 μg/min had the highest risk and the tree further split on diabetes duration at 26 years (CVD incidence 65.2% below and 96.2% above cut-point).

In the 1950-1964 subcohort (panel C), the 25-year incidence of CVD was 70.6% (120 cases) and the tree split only on AER at 91.7 μg/min (CVD incidence 59.8% below and 88.9% above cut-point). The 1965-1980 subcohort (panel D) had a 36.6% (143 cases) 25-year CVD incidence. The tree first split on non-HDLc at 133.2 mg/dl. In those with non-HDLc <133.2 mg/dl, the tree split on AER: those with AER <458 μg/min had 23.4% risk, while those with AER ≥458 μg/min had 71.4% risk. In those with non-HDLc≥133.2 mg/dl, diabetes duration further discriminated risk, such that those with duration <18.2 years had 44.3% risk and those with duration ≥18.2 years had 73.5% risk.

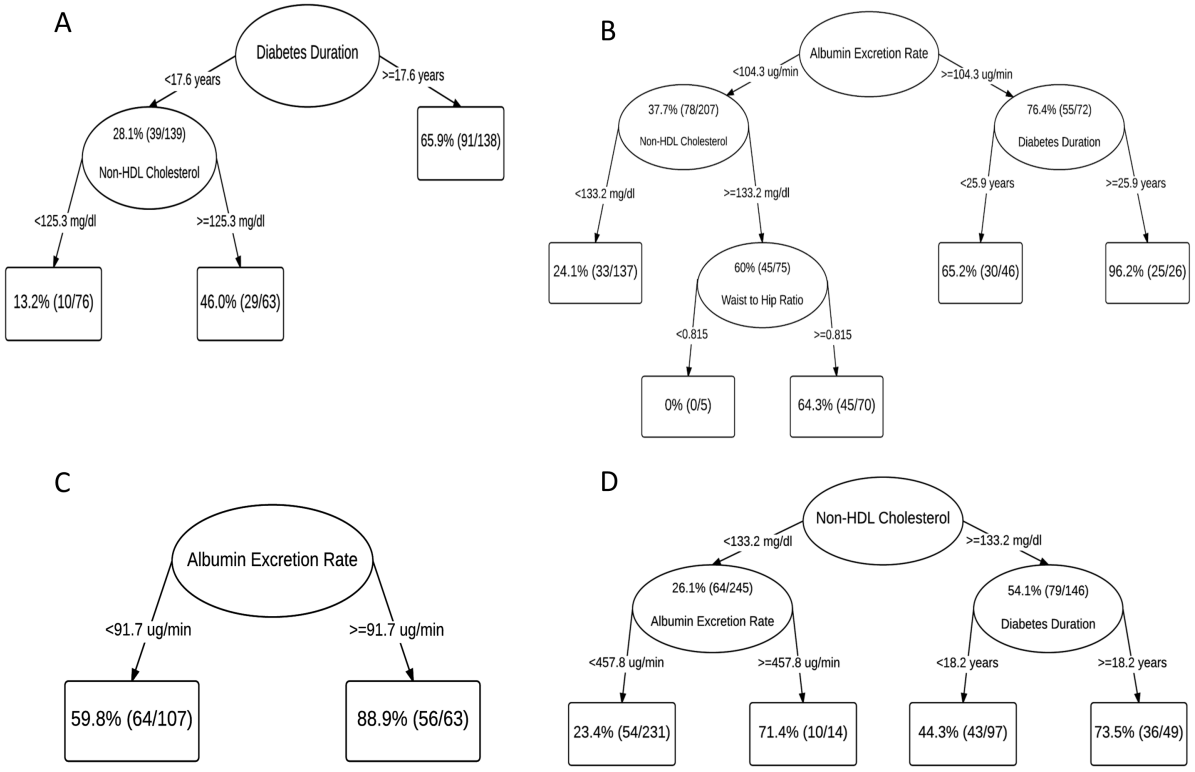

Figure 5. Survival Trees for the 25-Year Incidence of Total Cardiovascular Disease By Sex and Type 1 Diabetes Diagnosis Cohort
A)Women, B) Men, C) Type 1 Diabetes Diagnosed 1950-1964, D) Type 1 Diabetes Diagnosed 1965-1980

3.4.2 Specific Manifestations of Cardiovascular Disease

The best fitting survival trees for the specific manifestations of CVD are shown in Figure 7. The 25-year incidence of CVD death was 17.0% (112 deaths). The tree for CVD mortality (panel A) first split on AER: in those with AER<26.93 μg/min, diabetes duration further discriminated risk of CVD mortality at a cut-point of 23.09 years. Diabetes duration was also

selected as the discriminating factor in those with AER ≥26.93 μg/min, but at a slightly lower cut-point of 21.95 years (Figure 7A).

For CAD, the 25-year incidence was 34.7% (210 cases out of 605 free of CAD at baseline), with the first split on diabetes duration at the selected cut-point of 22.8 years (panel B). In those with diabetes duration <22.8 years the tree further split on non-HDLc at 123.2 mg/dl, while in those with diabetes duration ≥22.8 years the tree further split on AER at 131.9 μg/min.

The 25-year incidence of stroke was 6.7% (44 cases out of 654 free of stroke at baseline). The tree for stroke (panel C) first split on non-HDLc at 115.6 mg/dl. In those with non-HDLc<115.6 mg/dl, no further splits were selected, but in those with non-HDLc ≥115.6 mg/dl, stroke risk split on AER at 54.3 μg/min, with the lower AER group further splitting on WBC at 9.2 $x10^3/\mu l$.

The 25-year incidence of LEAD was 26.1% (157 cases out of 601 free of LEAD at baseline). For LEAD (panel D), diabetes duration was chosen for the first split, such that with diabetes duration <21 years, the tree split on non-HDL cholesterol at 132.5 mg/dl. Likewise, in those with diabetes duration ≥21 years, the tree also split on non-HDLc, but at a higher cut-point of 164.6 mg/dl.

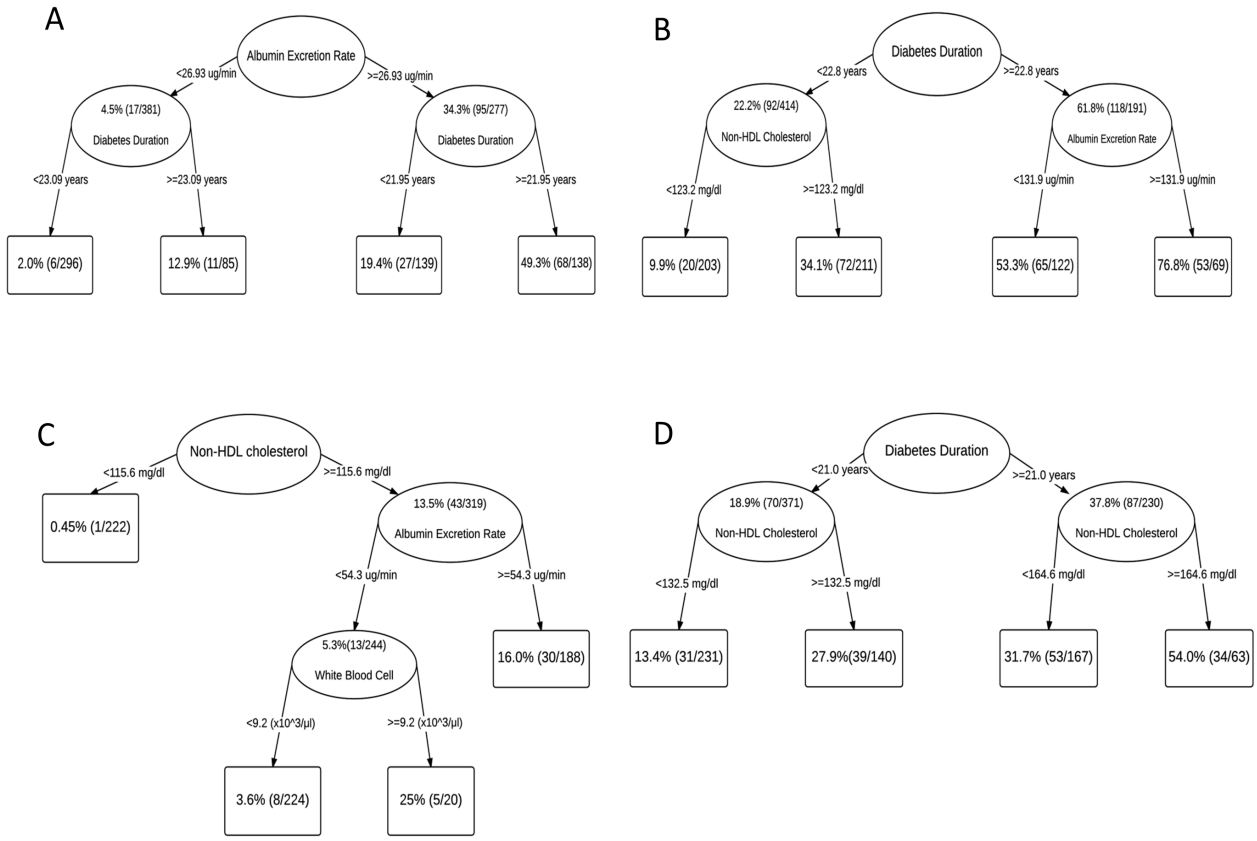

Figure 6. Survival Trees for the 25-Year Incidence of the Specific Manifestations of Cardiovascular Disease

A) Cardiovascular Mortality, B) Coronary Artery Disease, C) Stroke, D) Lower Extremity Arterial Disease

3.5 DISCUSSION

Tree-structured survival analysis of the EDC study data suggest that subgroups with major CVD risk differences exist in this cohort of childhood-onset type 1 diabetes over 25 years of follow-up and that these subgroups are identifiable by threshold effects of continuously measured risk factors measured only once at baseline. A wide range of CVD risk was observed across the TSSA-identified subgroups, ranging from 24 to 85% over the 25-year follow-up. Though various combinations of diabetes duration, non-HDLc and/or AER generally discriminated risk for both total CVD and the specific manifestations of CVD, the identified optimal cut-points of the risk factors varied and WBC and waist-hip ratio also discriminated in some subgroups.

For total CVD, five subgroups were identified through varying combinations of risk factors and they corresponded to three levels of CVD risk. The lowest risk (24.2%) group was characterized by shorter diabetes duration/younger age and had a relatively good risk factor profile (Group A). Two moderate risk groups were identified: one with shorter diabetes duration/younger age and non-HDLc ≥132.5 mg/dl (Group B, CVD risk 50.4%) (Group B) and the other with longer diabetes duration/older age, but a relatively good risk factor profile (Group C, CVD risk 52.5%). The nearly equivalent 25-year CVD risk in these two groups suggests that moderate dyslipidemia can "accelerate" CVD risk in shorter duration type 1 diabetes to the equivalent of older, longer duration individuals. Further examination of the risk factor profile of Group B showed that the mean non-HDLc was 165.5 (SD 37.7) and HbA1c was 9.5 (SD 1.8), both of which were the highest of all of the five groups identified by TSSA (Table 6). These results provide considerable support for more intensive lipid and HbA1c control in short duration type 1 diabetes. Finally, two high risk groups were identified: both had longer diabetes

duration/older age, but one mostly female, characterized by higher WBC, and smoking rates (Group D, CVD risk 81.5%) and the other mostly male, characterized primarily by renal disease (Group E, CVD risk 85.2%).

That AER was the primary factor discriminating risk at longer, but not shorter, duration may be related to more severe albuminuria as diabetes duration increases. The chosen AER cut-point of approximately 92 μg/min falls within the microalbuminuria range (20-200 μg/min). However, further investigation (Table 6) of the group with AER >92 μg/min (Group E) showed that the majority had overt nephropathy (AER >200 μg/min), which is a well-established CVD risk factor in type 1 diabetes (83–86). Within the lower AER group, WBC further discriminated total CVD risk, such that those with lower AER, but higher WBC (Group D, 82% CVD risk) were at nearly equivalent risk to those with higher AER (Group E, 85% CVD risk). Further examination of Group D showed that it had the highest rate of smoking of the five groups identified by the TSSA analyses (Table 6), which may contribute to both high WBC levels and CVD rate of this group.

When separate trees were fit by sex, AER only discriminated CVD risk in men. This supports the earlier observation that renal disease may be a more important CVD risk factor in men compared to women (56). Non-HDLc discriminated risk in men only if AER was lower. In women, diabetes duration was the major discriminating factor and non-HDLc further discriminated CVD risk in those with shorter diabetes duration.

In the earlier, 1950-1964 diagnosis cohort, AER was the only discriminating factor identified, which concurs with the total CVD tree, where AER was important at longer diabetes duration. In contrast, the 1965-1980 diagnosis cohort first split on non-HDLc and AER only discriminated risk at lower levels of non-HDLc. Interestingly, the lower-non-HDLc/higher AER

group had nearly the same CVD risk as the higher-non-HDLc/longer duration group (71 and 74%, respectively). The differences in the trees between the two diagnosis cohorts support the evidence that there may be a declining role of renal disease on CVD. Rates of renal disease, particularly end-stage renal disease, have declined with time (168) and albuminuria has been shown to be less likely to precede CVD than previously reported (194).

The tree for CVD mortality first split on AER at a cut-point of 26.9 μg/min. This cut-point is close to the cut-point used to define the lower limit of microalbuminuria (20-200 μg/min), which is strongly associated with higher all-cause mortality (195,196) and CVD rates (84,86,197) in type 1 diabetes. Diabetes duration modified risk at both levels of AER, such that having duration ≥23 years was associated with a 6-fold increase in risk within the lower AER group and having duration of ≥22 years was associated with an approximately 2.5-fold risk within the higher AER group.

The tree for CAD was similar to the total CVD tree, which is not surprising as CAD was the most common manifestation CVD in this cohort. The chosen optimal cut-points for the risk factors differed slightly than those chosen for total CVD, with the non HDLc cut-point in the CAD tree being 9 units lower than in the total CVD tree (123 and 132 mg/dl, respectively). Conversely, the AER cut-point was 40 units higher in the CAD compared to the total CVD tree (132 and 92 μg/min, respectively).

For stroke, non-HDLc <115.6 mg/dl was associated with very low risk of <0.5% over 25 years. The combination of non-HDLc≥115.6 mg/dl and AER≥54.3 μg/min led to a moderate 16% risk of stroke. Interestingly, having higher non-HDLc and WBC≥9.2x10^3/μl, but lower AER, led to the highest level of risk, at 25%. This relationship between very elevated WBC and stroke merits further study. WBC has been reported as a risk factor for stroke in the EDC study

(198), as well as in the general population (199–201), but a threshold effect at such a high WBC level on the long-term risk of stroke, has not been previously reported, to our knowledge. Generally, little is known about the predictors of stoke in type 1 diabetes, as there have been few reports. Serum lipids have been associated with stroke incidence in the Fremantle (202) and EDC studies (198), but not in the WHO Multinational Study of Vascular Disease in Diabetes (WHO MSVDD) (203), while nephropathy was shown to be strongly associated with stroke in both the EDC and the MSVDD (198,203).

Finally, the tree for LEAD split on diabetes duration and then, within both shorter and longer duration, split on non-HDLc at 132.5 mg/dl and 164.6 mg/dl, respectively. The shorter duration/higher non-HDLc group had nearly the same risk of LEAD as the longer duration/lower non-HDLc group (28 and 32%, respectively), again underscoring the importance of lipid control in type 1 diabetes, especially with diabetes duration <25 years. This relatively short duration corresponds to an average age of only 33 years in this cohort of childhood-onset type 1 diabetes. That non-HDLc was the major discriminating factor concurs with the results from an earlier analysis of the EDC study at 6 years of follow-up, which found that serum lipids predicted LEAD, but renal damage did not (58). In the general population, serum lipids have also been strongly associated with the development of LEAD (204–210).

Strengths and limitations

The EDC cohort is a prospective study with 25 years of follow-up to ascertain CVD outcomes. The cohort is well characterized, with data available for many clinical risk factors, and represents a wide range of diabetes diagnosis years. CVD events were verified using death certificates and medical records by reviewers masked to risk factor status. One limitation of the

study is that the cohort is 98% Caucasian, due to the demographics of Allegheny County, Pennsylvania (<15% African American), and lower type 1 diabetes incidence among African Americans during the period when the cohort was diagnosed, so it is not known if these results apply to other groups. There is also a potential for "survivor bias", particularly in the older 1950-1964 diagnosis subcohort, because prevalent cases of CVD were excluded from these analyses of incident events. Only baseline risk factor data were included as predictors in these analyses, thus we cannot account for the impact of patterns of risk factor change over time. However, this situation is very appropriate for clinicians also have only a single data point available. Finally, TSSA is considered to be an exploratory method and these results require external validation in other T1D cohorts before they can be applied in clinical settings. In particular, TSSA is not easily amenable to the construction of models where inferential statements (using, e.g., p-values) can be made about the data.

In conclusion, tree-structured survival analysis of the EDC study data suggests that subgroups with major CVD risk differences are identifiable by threshold effects of continuously measured risk factors. In many cases, similar levels of CVD risk are reached through different combinations of risk factor groupings. For total CVD, a high-risk group with shorter diabetes duration was identified which suggests that elevated non-HDLc increases CVD risk to the level of a longer diabetes duration group. Within longer diabetes duration individuals, two high-risk groups were identified through distinct pathways: one mostly female, with high rates of smoking, and WBC, and the other mostly male, characterized by renal disease. Utilizing decision tree methods to identify subgroups at higher levels of risk in this manner may help to shed light on the various pathways leading to CVD outcomes. Further study, including external validation,

is needed to better determine the prognostic utility of TSSA and the identification of subgroups at varying levels of CVD risk in type 1 diabetes.

4.0 HEMOGLOBIN A1C AND CARDIOVASCULAR DISEASE INCIDENCE IN TYPE 1 DIABETES: AN APPLICATION OF JOINT MODELING OF LONGITUDINAL AND TIME-TO EVENT DATA IN THE PITTSBURGH EPIDEMIOLOGY OF DIABETES COMPLICATIONS (EDC) STUDY

4.1 ABSTRACT

The reasons for the increased risk of cardiovascular disease (CVD) in type 1 diabetes (T1D) are not fully understood. The hyperglycemia, measured by HbA1c, that characterizes T1D has been an inconsistent predictor of CVD incidence. However, only baseline HbA1c or a summary measure, such as the mean over follow-up, is usually analyzed. Joint modeling allows longitudinal repeated covariates, modeled using random effects, and the time-to-event data to be modeled together. Thus our objective was to assess the association between serially measured HbA1c and CVD incidence using joint modeling. Data were from the Pittsburgh Epidemiology of Diabetes Complications (EDC) study, a prospective cohort study of childhood-onset T1D that has repeatedly reported little association between mean baseline or follow-up HbA1c and coronary artery disease incidence. Of 561 participants without CVD at baseline, 263 (42.1%) developed CVD over 25-years. Each one unit increase in HbA1c was associated with a 1.4-fold increase in the risk of CVD (95% CI: 1.17, 1.72), after adjusting for baseline values of other CVD risk factors, and a 1.3-fold increased risk (95% CI: 1.03, 1.42) after adjusting for updated

means of other factors. These findings, which support the need for good glycemic control to prevent CVD in T1D, underscore the importance of utilizing methods that incorporate within-subject variation over time when analyzing and interpreting data from longitudinal cohort studies.

4.2 INTRODUCTION

Cardiovascular disease (CVD), comprising coronary artery disease (CAD), cerebrovascular disease, and peripheral vascular disease, is a major complication of type 1 diabetes. As in the general population, CVD is a significant contributor to morbidity and early mortality in individuals with type 1 diabetes, however the risk of CVD is greatly increased in type 1 diabetes (55). The reasons for this increased risk are not fully understood, as the hyperglycemia, as measured by HbA1c, that characterizes type 1 diabetes is itself an inconsistent predictor of CVD incidence (56–64). In the Diabetes Control and Complications Trial (DCCT)/Epidemiology of Diabetes Interventions and Complications (EDIC) follow-up study, the intensive diabetes therapy arm had significantly lower incidence of CAD compared to the conventional therapy arm (211). Also, in a more recent analysis, updated weighted-mean HbA1c was found to be an important predictor of CVD incidence, second only to age (187). In contrast, observational studies have not found HbA1c to be a consistent predictor of CVD events (58,60,61,75,66,76,77). One potential explanation for these seemingly discordant results is that, in the DCCT, intensive therapy was begun early in the course of diabetes, raising the suggestion that glycemia may play a role in the initiation of atherogenesis (78). Despite the level of HbA1c not predicting CAD incidence, in the Pittsburgh Epidemiology of Diabetes Complications (EDC)

study, a prospective cohort study of childhood-onset type 1 diabetes, decreasing HbA1c over time was associated with a lower risk of CVD (62). Additionally, HbA1c variability, but not mean HbA1c over time, was predictive of CVD events in the FinnDiane study (77). In contrast to this inconsistent relationship between HbA1c and CAD, glycemia has been more consistently associated with an increased risk of peripheral artery disease (PAD) (58,79,80). This difference may indicate that hyperglycemia is more strongly associated with stable atherosclerosis, which characterizes PAD, rather than plaque rupture, which precipitates the acute coronary events (78). Stable atherosclerosis is characterized by plaques with high fibrous and lower macrophage and lipid content and evenly distributed atherosclerosis around the circumference of the vessel lumen (212). Unstable atherosclerosis is characterized by "vulnerable plaques", which have a necrotic core highly prone to produce a thrombus, or clot, and are covered by thin, fibrous caps (212). In the general population, the majority of acute coronary events are precipitated by the rupture of vulnerable plaques, with only a small proportion resulting from erosion of more stable atherosclerotic plaques (81,82). In type 1 diabetes, these proportions are reversed with the majority of events due to plaque erosion (81,82), consistent with the hypothesis that hyperglycemia is associated with more stable atherosclerosis.

In the observational cohort studies used to examine these associations, baseline HbA1c is most often used as a predictor of CVD incidence. This approach, while generally providing valid results, does not account for changes in glycemic control over time that may affect risk (128). For example, this approach may underestimate the effect of HbA1c if it becomes more important closer to the time of the event. Stability and change of predictors are relevant for determining the likelihood of an individual having an event, thus incorporating within-subject patterns is important for obtaining a more complete understanding of the relationship between

risk factor and outcome (128–131). Analyses of data from longitudinal cohort studies rarely make full use of the information contained in serially measured risk factors and, in some cases, failing to utilize all available data may lead to invalid results (128).

One common approach for accommodating change of risk factors in statistical analyses is to model change directly (132). However, this approach does not account for the actual pattern of change and assumes a constant, linear increase or decrease over the follow-up and may also underestimate the variance of the factor (133). Another approach for utilizing longitudinal data is to model the average (or updated mean) risk factor over the follow-up period. However, this method has similar drawbacks to modeling change, leading to a loss of important information regarding variability over time (134).

Joint modeling allows both longitudinal repeated covariates and time-to-event data to be modeled together, as detailed by Wulfsohn and Tsiatis (160). Comprehensive overviews of joint models have also been provided by Tsiatis and Davidian (213) and by Diggle *et al.* (214). Briefly, a joint model consists of two submodels. A mixed-effects model characterizes the trajectory over time of the longitudinal covariate. The modeled random effects are simultaneously included in a survival model, so that the relationship between the pattern of covariate change and time to the event of interest can be estimated (131). These random effects, or subject-specific deviations from the average effect, are assumed to be the underlying link between the longitudinal repeated measures and the subject-specific hazard for the time-to-event outcome (128).

While survival analyses and repeated measures analyses using mixed models have separately been performed using EDC study data, jointly modeling both processes has not previously been carried out. Thus, our objective is to assess the relationship between serially

measured HbA1c and the incidence of CVD using joint modeling. Additionally, the specific manifestations of CVD (i.e. coronary artery disease, LEAD, CVD mortality, etc.) were examined separately to explore whether the relationship with HbA1c differs by type.

4.3 METHODS

4.3.1 Study Population

The Pittsburgh EDC Study is a prospective cohort study of childhood-onset (<17 years old) T1DM. All participants (n=658) were diagnosed, or seen within one year of diagnosis, at Children's Hospital of Pittsburgh between 1950 and 1980. The cohort has been described in detail elsewhere (167,168). In brief, participants have been followed since 1986-1988, initially with biennial examinations and questionnaires for ten years and thereafter with biennial questionnaires and further examinations at 18 and 25 years post-baseline. Research protocols were approved by the University of Pittsburgh institutional review board and all participants provided written informed consent.

4.3.2 Ascertainment of Cardiovascular Outcomes

Participants were followed for 25 years to assess the incidence of total CVD, defined as the first instance of CAD (fatal CAD, myocardial infarction, revascularization procedure, blockage ≥50%, ischemic ECG, or angina), stroke, or LEAD (Ankle brachial index (ABI) <0.8,

lower limb amputation due to vascular cause, or intermittent claudication). Fatal events were ascertained using medical records, death certificates, autopsy reports, and/or interview with next of kin and classified according to the Diabetes Epidemiology Research International (DERI) system (169). Myocardial infarction, revascularization/blockage, and stroke were self-reported by the study participants and confirmed by medical records. Amputation was ascertained by self-report and/or physician examination. ABI was determined using a Doppler blood-flow detector with the subject supine and ratios for each of the right and left tibialis posterior and dorsalis pedis systolic pressures were calculated with arm systolic pressure as the denominator. Intermittent claudication was assessed using the Rose questionnaire (192). In addition to total CVD, the following alternative outcomes were assessed: total CAD, stroke, LEAD, and CVD mortality (CAD death, fatal MI, or fatal stroke). Participants with prevalent CVD at baseline were excluded from analyses of the corresponding outcome.

4.3.3 Risk Factor Assessment

Risk factors were assessed at baseline and repeated at 2-, 4-, 6-, 8-, 10-, and 18-years of follow-up. Fasting blood samples were taken to measure HbA1/HbA1c, lipids, and serum creatinine. From baseline through 10 years of follow-up, HbA1 values were converted to DCCT-aligned HbA1c values using a regression equation derived from duplicate assays (DCCT HbA1c=0.14 + 0.83[EDC HbA1]) (62). At the 18-year examination, HbA1c was measured using the DCA 2000 analyzer (Bayer Healthcare LLC. Elkhart, IN) and converted to DCCT-aligned HbA1c by the equation: DCCT HbA1c=(EDC HbA1c-1.13)/0.81. From baseline through the 10-year examination, serum total cholesterol and triglycerides were determined enzymatically (170,171) and HDL cholesterol was determined using a modified precipitation

technique (172) based on the lipid research clinics method (173). At the 18-year examination, serum lipids were measured using the Cholestech LDX (Cholestech Corp., Hayward, CA). Non-HDL cholesterol (non-HDLc) was calculated by subtracting HDL cholesterol from total cholesterol. White blood cell count (WBC) was determined using a Coulter Counter. Three seated blood pressure readings were taken with a random-zero sphygmomanometer, and the mean of the second and third readings was used in analyses, according to the Hypertension Detection and Follow-Up protocol (175). Hypertension was defined as blood pressure ≥140/90 mmHg or use of blood pressure-lowering medication. Past and current smoking status was obtained by self-report. At study visits, three timed urine samples (24-hour, overnight, and 4-hour collections obtained over a two week period) were collected. Urinary albumin was measured in each sample by immunonephelometry (176) and albumin excretion rate (AER) was calculated; the natural log transformed median of the three AERs was used in analyses. Serum creatinine was measured using an Ectachem 400 Analyzer (Eastman Kodak Co., Rochester, NY). Glomerular filtration rate (eGFR) was estimated by the Chronic Kidney Disease Epidemiology Collaboration (CKD-EPI) creatinine equation (193). A positive renal disease status at baseline was defined as AER≥200μg/min in at least 2 of the 3 timed urine samples, eGFR<60 ml/min/m^2, or a history of renal transplant or dialysis.

4.3.4 Statistical Analyses

Prevalent cases of CVD at baseline were excluded. Baseline characteristics were compared between those who developed CVD during the 25 years of follow-up and those who did not using two-sample t-tests for continuous variables (or Wilcoxon rank-sum, if data were

not normally distributed) or using chi-squared tests for dichotomous variables in SAS v. 9.3 (SAS Institute, Cary, NC). Longitudinal HbA1c and time to first CVD event were the outcomes of interest and were examined simultaneously using joint models (160,165). For the longitudinal sub-model, a subject-specific random effects model with random intercept and slope was used to examine the trajectory of HbA1c over follow-up. HbA1c was censored at the time point immediately prior to the CVD event for cases. For the survival submodel we employed a Cox proportional hazards model. Interactions between longitudinal HbA1c and important pre-defined subgroups, specifically diabetes diagnosis cohort (diabetes diagnosed 1950-64 versus 1965-80), sex, and renal disease at baseline, were tested to determine whether the longitudinal HbA1c trajectory differed by group. Models were repeated separately with time to CAD, stroke, and LEAD as the outcome, respectively. All models were adjusted for other baseline CVD risk factors and final models were selected using backward selection and likelihood ratio tests. The risk factors considered for adjustment were years of diabetes duration, non-HDLc, hypertension, ln(AER), and white blood cell count, except in cases where collinearity was a concern (i.e. models including the diagnosis cohort indicator were not adjusted for diabetes duration and those including the renal disease status indicator were not adjusted for ln(AER)). Alternate models were fit adjusting for the updated mean of each factor (i.e. the mean value of all available measurements prior to event or censoring), instead of baseline values. All joint modeling was performed using the joineR routine in R (165,145).

4.4 RESULTS

Of the 561 participants free of CVD at baseline, 263 (42.1%) went on to develop CVD during the 25-year follow-up. Baseline characteristics by CVD incidence status are shown in Table 5. The participants who developed CVD were significantly older, had longer diabetes duration, higher body mass index, waist-hip ratio, non-HDLc, triglycerides, blood pressures, white blood cell count, and AER and lower eGFR. They were also more likely to be taking blood pressure lowering medication, have hypertension, and report past smoking. The median follow-up time for CVD ascertainment was 15.0 years (range 0.2 to 27.7 years). The survival curve for total CVD is shown in Figure 8 and the observed mean longitudinal HbA1c profile by CVD incidence status is shown in Figure 9.

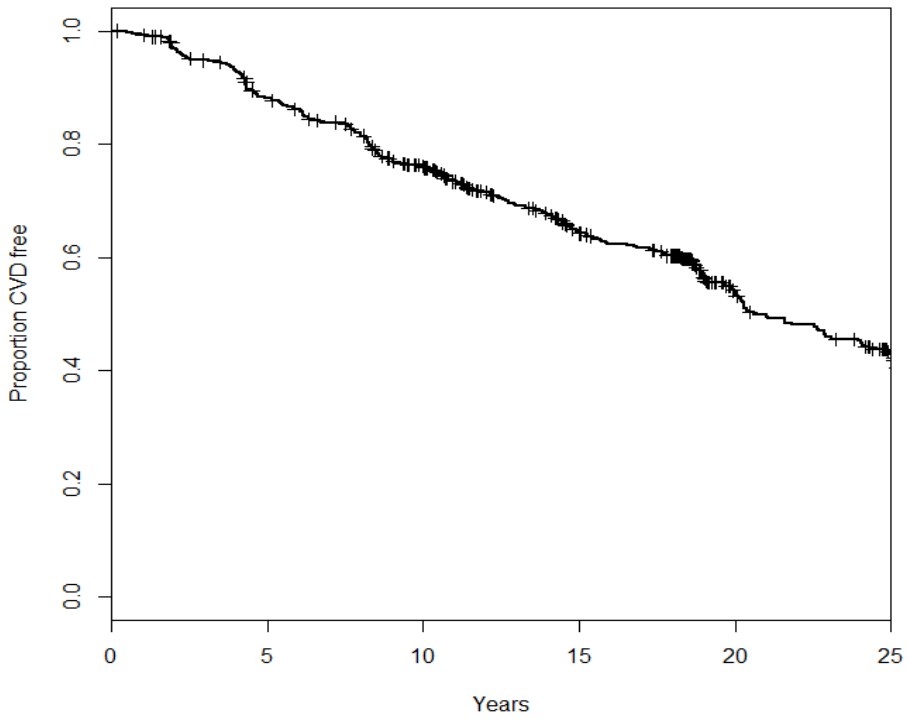

Figure 7. Kaplan Meier Curve for Survival Free of Cardiovascular Disease (CVD) Over 25 Years of Follow-Up

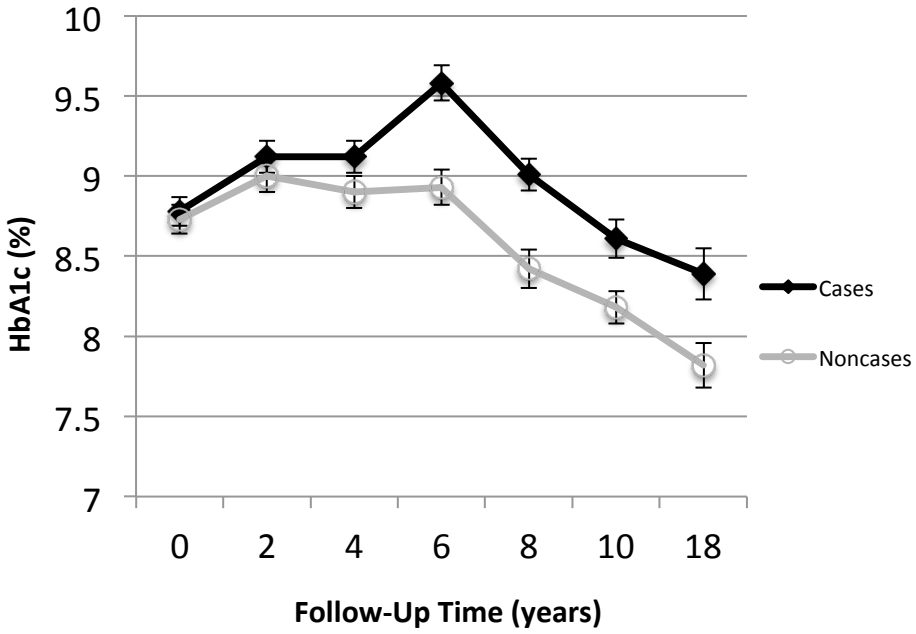

Figure 8. Observed Mean HbA1c by 25-Year Total Cardiovascular Disease Incidence Status
Error bars are ±standard error

Three models were fit to associate HbA1c trajectories (longitudinal component) with time to total CVD (survival component): (1) a joint model where both the longitudinal and survival sub-models were not adjusted for other covariates, (2) a joint model where both the longitudinal and survival sub-models were adjusted for baseline covariates, and (3) a joint model where both the longitudinal and survival sub-models were adjusted for covariates averaged over the follow-up (Table 7). In model 1, there was a statistically significant mean decrease in HbA1c of -0.03 units per year (95% CI: -0.05, -0.02). For the corresponding survival sub-model, the association parameter was 0.43 and was statistically significant (95% CI: 0.26, 0.64), indicating that there is strong evidence of an association between the two sub-models. In other words, HbA1c is associated with CVD risk at any given time t. In model 2, after adjusting for baseline diabetes duration, non-HDLc, hypertension, ln(AER), and WBC, the association parameter of 0.33 was only slightly attenuated and remained statistically significant. Thus, each one unit increase in HbA1c was associated with an exp(0.33)=1.4-fold increase in the risk of CVD (95% CI: 1.17, 1.72), after covariate adjustment. In model 3, adjusted for diabetes duration and updated mean non-HDLc, ln(AER), and WBC, the association parameter was further attenuated, but remained significant at 0.22 (95% CI: 0.03, 0.35), corresponding to a 1.25-fold increased risk of CVD (95% CI: 1.03, 1.42) per unit increase in HbA1c.

Table 7. Estimated Coefficients and 95% Confidence Intervals From a Joint Model with a Random Effects Longitudinal HbA1c Sub-Model and a Cox Survival Sub-Model for Time to Total Cardiovascular Disease (CVD)

Model	Model Adjustment	Sub-Model	Covariates	Estimate	95% CI
1	Unadjusted, Main Effects Only	Longitudinal HbA1c random effects	Intercept	9.03	8.90, 9.12
			Time	-0.03	-0.05, -0.02
		Cox Survival	Association Parameter	0.43	0.26, 0.64
2	Adjusted for Baseline CVD Risk Factors	Longitudinal HbA1c random effects	Intercept	8.29	7.58, 8.53
			Time	-0.035	-0.05, -0.02
			Diabetes Duration	-0.03	-0.04, -0.01
			Non-HDLc	0.009	0.008, 0.01
			Hypertension	-0.36	-0.70, -0.08
		Cox Survival	Diabetes Duration	0.07	0.05, 0.09
			Non-HDLc	0.008	0.004, 0.01
			Hypertension	0.52	0.07, 0.99
			ln(Albumin Excretion Rate)	0.15	0.08, 0.21
			White Blood Cell Count	0.10	0.02, 0.18
			Association Parameter	0.33	0.16, 0.54
3	Adjusted for Updated Mean CVD Risk Factors	Longitudinal HbA1c random effects	Intercept	7.98	7.47, 8.37
			Time	-0.02	-0.04, -0.01
			Diabetes Duration	-0.03	-0.05, -0.02
			Non-HDLc	0.01	0.006, 0.01
			ln(Albumin Excretion Rate)	0.11	0.05, 0.15
		Cox Survival	Diabetes Duration	0.08	0.08, 0.01
			Non-HDLc	0.01	0.005, 0.02
			ln(Albumin Excretion Rate)	0.23	0.17, 0.31
			White Blood Cell Count	0.10	0.03, 0.16
			Association Parameter	0.22	0.03, 0.35

There was a significant interaction between diabetes diagnosis cohort and time with respect to HbA1c trajectory (ß=-0.04, 95% CI: -0.08, -0.01). The observed longitudinal trajectories of HbA1c by diabetes diagnosis cohort are depicted in Figure 10. As seen in the graph, HbA1c was generally lower in the earlier 1950-1964 diagnosis cohort throughout follow-up and this earlier cohort also experienced a more dramatic decline in HbA1c after 6 years of

follow-up compared to the more recent 1965-1980 diagnosis cohort. Despite this interaction between diagnosis cohort and time with respect to HbA1c, the association between HbA1c trajectory and CVD risk was similar in both diagnosis cohorts (1950-1964, HR per 1-unit increase in HbA1c=1.30, 95% CI: 1.05, 1.65; 1965-1980, HR per 1-unit increase in HbA1c=1.21, 95% CI: 1.001, 1.46). There were no significant interactions between sex or renal disease status at baseline and time with respect to HbA1c.

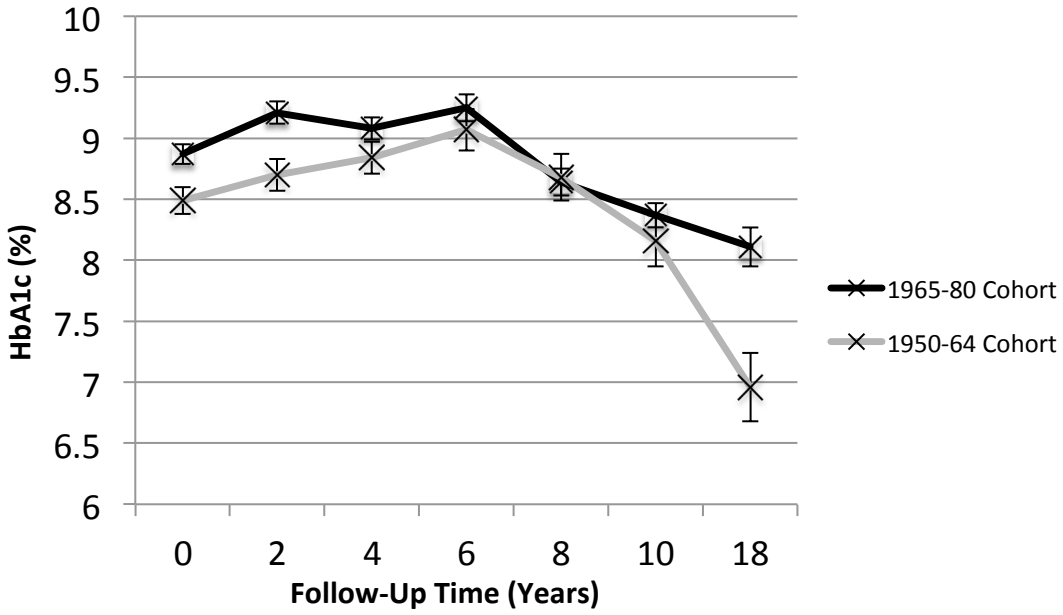

Figure 9. Observed Mean HbA1c by Type 1 Diabetes Diagnosis Cohort
Error bars are ±standard error

The results of joint models for longitudinal HbA1c and the incidence of each specific manifestation of CVD adjusted for baseline CVD risk factors are shown in Table 8. The longitudinal trajectory of HbA1c was significantly associated with each of the manifestations of

CVD, with a one unit increase in HbA1c associated with a 1.4-fold increased risk of both CAD and LEAD (95% CI: 1.05, 1.67 and 1.22, 1.67, respectively) and a 2-fold increased risk of stroke (95% CI: 1.5, 2.5). Adjusting for updated mean CVD risk factors only slightly attenuated these associations and the longitudinal trajectory of HbA1c remained significantly associated with each CVD manifestation (Table 9). In these alternate models, a one unit increase in HbA1c was associated with a 1.2-fold increased risk of CAD (95% CI: 1.03, 1.35), a 1.9-fold increased risk of stroke (95% CI: 1.26, 2.36), and a 1.3-fold increased risk of LEAD (95% CI: 1.11, 1.57).

Table 8. Estimated Coefficients from Joint Models with a Random Effects Longitudinal HbA1c Sub-Model and a Cox Survival Sub-Model for Time to Each Specific Manifestation of Cardiovascular Disease, Adjusted for Baseline Covariates

Outcome	Sub-Model	Covariates	Estimate	95% CI
Coronary Artery Disease	Longitudinal HbA1c random effects	Intercept	8.26	7.88, 8.60
		Time	-0.04	-0.05, -0.03
		Diabetes Duration	-0.03	-0.04, -0.02
		NonHDLc	0.008	0.006, 0.01
		Hypertension	-0.48	-0.81, -0.17
		ln(Albumin Excretion Rate)	0.07	0.01, 0.11
	Cox Survival	Diabetes Duration	0.09	0.07, 0.11
		NonHDLc	0.008	0.004, 0.01
		Hypertension	0.47	0.03, 0.83
		ln(Albumin Excretion Rate)	0.15	0.06, 0.23
		Association Parameter	0.32	0.05, 0.51
Stroke	Longitudinal HbA1c random effects	Intercept	8.33	7.81, 8.62
		Time	-0.04	-0.05, -0.03
		Diabetes Duration	-0.02	-0.03, -0.01
		NonHDLc	0.009	0.007, 0.01
		Hypertension	-0.36	-0.66, -0.07
	Cox Survival	Diabetes Duration	0.08	0.03, 0.11
		NonHDLc	0.01	0.008, 0.02
		ln(Albumin Excretion Rate)	0.27	0.15, 0.39
		Association Parameter	0.71	0.40, 0.92
Lower Extremity Arterial Disease	Longitudinal HbA1c random effects	Intercept	8.39	7.96, 8.71
		Time	-0.03	-0.05, -0.02
		Diabetes Duration	-0.03	-0.05, -0.02
		NonHDLc	0.009	0.007, 0.012
	Cox Survival	Diabetes Duration	0.06	0.04, 0.09
		NonHDLc	0.01	0.008, 0.02
		Association Parameter	0.34	0.20, 0.51

Table 9. Estimated Coefficients from Joint Models with a Random Effects Longitudinal HbA1c Sub-Model and a Cox Survival Sub-Model for Time to Each Specific Manifestation of Cardiovascular Disease, Adjusted for Updated Mean Covariates

Outcome	Sub-Model	Covariates	Estimate	95% CI
Coronary Artery Disease	Longitudinal HbA1c random effects	Intercept	7.92	7.30, 8.33
		Time	-0.03	-0.05, -0.02
		Diabetes Duration	-0.03	-0.05, -0.02
		NonHDLc	0.0009	0.007, 0.013
		ln(Albumin Excretion Rate)	0.11	0.04, 0.17
	Cox Survival	Diabetes Duration	0.10	0.08, 0.12
		NonHDLc	0.008	0.004, 0.01
		ln(Albumin Excretion Rate)	0.24	0.13, 0.31
		White Blood Cell Count	0.10	0.0009, 0.18
		Association Parameter	0.18	0.03, 0.30
Stroke	Longitudinal HbA1c random effects	Intercept	8.04	7.71, 8.41
		Time	-0.04	-0.05, -0.03
		Diabetes Duration	-0.03	-0.05, -0.02
		NonHDLc	0.009	0.006, 0.011
		ln(Albumin Excretion Rate)	0.11	0.06, 0.16
	Cox Survival	Diabetes Duration	0.07	0.04, 0.11
		NonHDLc	0.013	0.004, 0.02
		ln(Albumin Excretion Rate)	0.44	0.28, 0.59
		Association Parameter	0.62	0.23, 0.86
Lower Extremity Arterial Disease	Longitudinal HbA1c random effects	Intercept	8.08	7.61, 8.46
		Time	-0.03	-0.05, -0.02
		Diabetes Duration	-0.04	-0.05, -0.02
		NonHDLc	0.009	0.006, 0.011
		ln(Albumin Excretion Rate)	0.11	0.06, 0.15
	Cox Survival	Diabetes Duration	0.06	0.04, 0.08
		NonHDLc	0.014	0.008, 0.018
		ln(Albumin Excretion Rate)	0.17	0.06, 0.27
		Association Parameter	0.30	0.10, 0.45

4.5 DISCUSSION

Our findings show that the longitudinal trajectory of HbA1c is associated with 25-year CVD incidence in this cohort of long-standing childhood onset type 1 diabetes and that this association does not differ by the specific manifestation of CVD. The results of these joint models show that CVD risk is associated with the combination of a subject's level of Hba1c at baseline and rate of change of HbA1c over follow-up, a longitudinal process represented by a random intercept and slope. A one unit increase in the level of HbA1c is associated with a 40% increased risk of developing CVD, after adjusting for baseline CVD risk factors, and a 25% increased risk of CVD, after adjusting for each participant's mean values of CVD risk factors over the follow-up.

These results provide additional insight into the role of HbA1c on CVD risk, compared to other methods of assessing this relationship, such as modeling baseline HbA1c (62,215), updated mean HbA1c, change between the first and last HbA1c measurements (62), or a cumulative measure of glycemic exposure (58) as predictors. Prior analyses of EDC data using baseline HbA1c, were unable to demonstrate an association between glycemia and CAD incidence (58,215). As evident in the plots of HbA1c trajectories shown in Figure 2, the level of HbA1c at baseline did not differ in those participants who went on to develop CVD versus those who did not. However, as time progressed, the HbA1c trajectories between the two groups grew farther apart. Interestingly, in both groups, mean HbA1c rose until 6 years of follow-up, albeit more so in the CVD cases, which corresponds to calendar years 1992-1994, and began to decline steadily

thereafter, likely reflecting the implementation of intensive insulin therapy into clinical practice following the results of the DCCT (216). In another prior EDC analysis, based on a 16 year follow-up, absolute change in HbA1c was significantly associated with CAD risk (62). In those analyses, CAD cases had a mean increase in HbA1c of +0.62 between their first and last measures, while noncases had a mean decrease of -0.09. The current analyses add additional information by incorporating all available HbA1c measures and within subject-variability, allowing us to quantify the increase in CVD risk associated with each unit increase in HbA1, while also adjusting for the mean level of other risk factors over time.

Attenuation of the association between longitudinal HbA1c and CVD was somewhat greater when adjusting for the updated means of other CVD risk factors, compared to adjusting for the baseline values. Additional analyses to determine which factors lead to the attenuation of the HbA1c effect revealed that mean non-HDLc had the greatest impact (data not shown). Non-HDLc is traditionally an important CVD risk factor in general, regardless of diabetes status and individuals with type 1 diabetes typically do not have higher non-HDLc levels than observed in the general population. A report using data from the National Health and Nutrition Examination Surveys estimated that US adults had a mean non-HDLc of 155 mg/dl (95% CI 153-157) between 1988-1994 (217). This value is markedly higher than the mean non-HDLc observed in the EDC cohort both at the 1986-1988 baseline (Table 5) and over follow-up (mean over all cycles=136.3, SD=36.3), thus non-HDLc level itself does not seem to explain the increased CVD risk associated with type 1 diabetes in this cohort. It is, however, possible that the effect of non-HDLc could be modified by diabetes, so that CVD risk is increased even at lower levels of non-HDLc. Even given the observed attenuation of the effect of the HbA1c trajectory on CVD risk, the HbA1c trajectory remained a significant predictor of CVD independent of non-HDLc,

demonstrating that glycemic control may explain more of the increased CVD risk than has previously been shown in observational studies. These findings are consistent with the recent report from the DCCT/EDIC, where updated weighted-mean HbA1c was found to be an important predictor of CVD incidence independent of lipid levels (187).

From a clinical perspective, these results provoke questions as to how the relationship observed here can be used to determine which patients should be targeted for CVD prevention based on their level of glycemic control. Data from a longitudinal study can be used to examine etiologic pathways that may take many years of follow-up to be revealed. This information is in contrast to clinically available information, where often only a current snapshot of risk factor level is accessible. Potentially, this type of joint modeling could be used to create a "2-step" prediction model. For example, cross-sectional HbA1c and other clinical characteristics are entered to first predict the longitudinal trajectory of HbA1c over time and then to subsequently translate this predicted trajectory into a patient's risk of developing CVD. An important unanswered question is whether some individuals are resistant to the benefits of intensive insulin therapy and would continue to be at high risk for CVD, regardless of HbA1c control. As shown in Figure 9, in the EDC, the pattern of glycemic control was generally worse over time in those who developed CVD compared to those who did not. However, it is unclear whether hyperglycemia in the CVD cases was undertreated or if this group is resistant to the benefits of appropriate treatment. The mean insulin dose between the CVD cases and noncases did not differ at baseline (Table 5) or at any other point during the follow-up (data not shown), suggesting that equivalent treatment for glycemia was less effective in those who went on to develop CVD. It is possible that there are genetic factors that drive both resistance to glycemic control and CVD risk. Identifying such factors is an important area of future research.

Our study has many strengths. The EDC is a well-characterized, prospective cohort study with 25 years of follow-up to ascertain CVD incidence and clinical risk factors. CVD events were verified using death certificates and medical records by reviewers who were masked to risk factor status. Only HbA1c measurements prior to CVD incidence were included in these analyses to minimize the potential for reverse causality. Using joint modeling we were able to incorporate all HbA1c measurements and model the longitudinal trajectory taking the within-subject variability into account. The major advantage of joint modeling is that it allows the relationship between longitudinal repeated measures and a time-to-event outcome to be examined. As emphasized by Lim *et al.,* joint modeling may be particularly useful when the goal is to use longitudinal measures of biomarkers to improve prediction of an event (128). Another advantage of using joint modeling in this type of application is that it is able to reduce the potential for bias associated with event-free survival far beyond the last measurement of a longitudinal covariate by accounting for the subject-specific random effects (131). Using simpler methods, such as modeling the two processes separately, may lead to bias if some subjects survive far beyond the most recent follow-up. Joint modeling is able to account for this differential follow-up time by including the subject-specific random effects within the survival model (128,166).

One limitation of our study is that the cohort is 98% Caucasian. There are two main reasons for this overrepresentation of Caucasians: 1) the population of Allegheny County, Pennsylvania is <15% African American and 2) the incidence of type 1 diabetes was lower among African Americans during the period when the cohort was diagnosed. Consequently, it is unknown whether these results apply to more diverse groups. There is also a potential for "survivor bias", particularly in the older 1950-1964 diagnosis subcohort, because prevalent cases

of CVD at baseline were excluded from these analyses. A limitation associated with joint modeling is that the censoring process is assumed to be independent of the random effects. If this assumption is violated then the parameter estimates may be biased (160).

In conclusion, using joint modeling, we demonstrate that the longitudinal trajectory of HbA1c is associated with 25-year CVD incidence in this cohort of long-standing childhood onset type 1 diabetes and that this association does not differ by the specific manifestation of CVD. These results are in contrast to previous analyses using only baseline or mean HbA1c as predictors, which have not revealed a strong relationship between glycemic control and CVD (62,215). Thus, these findings underscore the importance of utilizing methods that incorporate within-subject variation when analyzing data from longitudinal cohort studies.

5.0 DISCUSSION

5.1 SUMMARY AND CONCLUSIONS

These findings reflect a comprehensive examination of the epidemiology of CVD and its risk factors in type 1 diabetes. Our first objective was to quantify the contemporary incidence and excess risk of CVD in young adults <45 years old with type 1 diabetes, which addresses a critical gap in the knowledge of this population in the United States. We demonstrate that CVD risk continues to be significantly increased in type 1 diabetes, compared to the background population, and that this risk is particularly elevated in women with type 1 diabetes. Our second objective was to utilize novel statistical methods to address limitations in the analyses of longitudinal cohort studies, in an effort to better understand the risk factor patterns that lead to the increased risk of CVD in this population. One limitation of traditional analyses is that associations between risk factors and outcomes are assumed to be linear and continuous, ignoring the possibility of threshold effects or higher-order interactions between factors. Using tree-structured survival analysis, we performed formal subgroup analysis and identified distinct groups at varying levels of risk for CVD, based on threshold effects of continuous risk factors. The factors defining these groups differed by diabetes diagnosis cohort and by sex, suggesting that the relative importance of CVD risk factors have changed over time and are different in men and women. Another limitation in the analysis of cohort studies is the failure to adequately

utilize the repeated measures risk factor data by using only baseline or composite measures of risk factors that underestimate variability over time. We used joint models to simultaneously model the longitudinal trajectory of HbA1c, including all measures obtained during the EDC follow-up, and time to CVD incidence. We found that the longitudinal trajectory of glycemic control is associated with the risk of CVD. Importantly, longitudinal HbA1c was similarly associated with all manifestations of CVD, including coronary artery disease, which is a new finding in this cohort. Interestingly, longitudinal HbA1c was nearly equivalently associated with both CAD and LEAD, while the association was stronger for stroke incidence. These results support the findings from trial data, which show that glycemic control is important for preventing and delaying CVD, in addition to microvascular complications.

5.2 FINDINGS

5.2.1 Contemporary Cardiovascular Disease Burden

A major gap exists regarding the contemporary burden of CVD in type 1 diabetes in the United States. This is particularly true for young adults, who would be expected to have higher excess risk than older ages due to very low rates of CVD in young adults the general population. Additionally, treatment guidelines are unclear for young adults <40 years old, thus it is important to quantify risk in this group to inform clinical practice recommendations. Treatment guidelines are based on extrapolations of data from the general and type 2 diabetes populations, but the usefulness of these recommendations in individuals with long duration type 1 diabetes is

questionable (185). At any given age, the average duration of type 1 diabetes is significantly longer than type 2 diabetes, due to the typically younger age at onset.

By examining data from EDC participants <45 years old between 1996-2012, we have attempted to address the question of whether there remains an excess risk of CVD, as was last demonstrated in the 1980s (55,69). We found that both CVD mortality and hospitalized CVD event rates are still significantly increased compared to the age-matched background population. This increase is especially large for women, an issue discussed in further detail below.

The excess CVD risk is consistent with data from other recent reports from Scotland (51), Sweden (52), and Australia (53), however, in our US cohort, our estimates suggest a greater excess risk than shown in all of these countries. Differences in diabetes duration (i.e. length of exposure to diabetes) may explain some of the higher risk seen in the EDC cohort. The other registries include adult-onset cases of type 1 diabetes, thus there is a shorter average diabetes duration at any given age. There is also the distressing prospect that individuals with type 1 diabetes in the United States are inadequately treated compared to their counterparts in other developed nations. There is historical evidence to support this possibility. A 1991 report from the DERI study found that premature mortality in type 1 diabetes was greater in the US and Japan compared to Finland and Israel (169). Another prior report comparing data across countries showed that the United States has greater excess mortality associated with type 1 diabetes than most of the European countries (with the exception of eastern Europe) and Canada (218). The current findings add more evidence that prevention trials specific to type 1 diabetes are needed to address CVD risk in this high-risk population.

We also quantified the absolute risk of CVD in young adults to address the question of whether their CVD risk is high enough to warrant statin therapy, regardless of cholesterol levels.

The American College of Cardiology/American Heart Association guidelines state that high-intensity statin therapy is recommended in individuals between ages 40-75 years with diabetes and a ≥7.5% 10-year risk of CVD, but in individuals with diabetes who are <40 years old, the decision to being statin therapy should be individualized (73). In cases of childhood-onset type 1 diabetes, the duration of diabetes already exceeds 20 years by time an individual reaches age 30, so it is unclear how this recommendation should be applied in clinical practice. In the EDC cohort, we find that young adults 30-39 years old have a 6.3% 10-year CVD risk, which approaches the 7.5% 10-year risk specified by the ACC/AHA guidelines. However, it is important to note that 5% of this age group in the EDC cohort was already on lipid-lowering medication in 1996 and this proportion rose to 20% by 2012. Thus, we can reasonably assume that the 10-year risk would exceed 7.5% if none of the cohort were using lipid-lowering medication. These results support recommending statin therapy to all individuals with type 1 diabetes duration of >20 years who are at least 30 years old and underscore the need for trial data specific to type 1 diabetes.

5.2.2 Cardiovascular Disease Risk Stratification and Interactions Between Risk Factors

The excess CVD risk that persists in type 1 diabetes also points to the need for better identification of individuals who are at high risk for targeted interventions. Recommending statin therapy to a wider group, as discussed above, is an important strategy to consider. It is also likely that assessing the possibility of interactions between multiple risk factors could lead to the identification of subgroups, which may benefit from interventions that target a range of risk factors. Using tree-structured survival analysis (TSSA), we were able to demonstrate the presence of such meaningful subgroups with differing levels of CVD risk in the EDC cohort.

For total CVD, five groups were identified based on threshold effects of continuously measured risk factors. Interestingly, the five groups were based on combinations of different risk factors, but corresponded to 3 levels of CVD risk: a low risk (24%) group characterized by shorter diabetes duration and younger age, with a relatively good risk factor profile, two moderate risk groups, one with shorter diabetes duration and younger age, but high non-HDLc and HbA1c, and the other with longer diabetes duration and older age, but a good risk factor profile (50 and 53% risk, respectively), and two high risk groups, one mostly women with higher WBC and smoking, and the other mostly men with renal disease (82 and 85% risk, respectively). There are several intriguing observations that can be made regarding these results. One observation is that the risk factors profiles of the subgroups differed in ways additional to those characterized by the factors chosen by this computational method as the splitting factors. In other words, the method discriminated subgroups that were distinct from each other in more ways than just the combinations of diabetes duration, non-HDLc, albumin excretion rate, and WBC that defined the groups. Another important observation is that having high non-HDLc cholesterol seems to erase the protection against CVD associated with shorter diabetes duration/younger age. This finding, in combination with our finding that CVD risk remains significantly elevated in young adults with type 1 diabetes, underscores the importance of more aggressive treatment of cholesterol beginning at younger ages in this population. The American Association of Clinical Endocrinologists guidelines state that non-HDLc should be below 130 mg/dl for people with diabetes and below 100 mg/dl in the cases of diabetes with another major risk factor (e.g. advanced age, smoking, family history of CAD) (219). Here, the optimal cut-point for non-HDLc in EDC participants with diabetes duration of <20 years was 132.5 mg/dl, which is close to the recommendation of 130 mg/dl. Prior analyses from the EDC study examining various

LDLc target levels concluded that LDLc>100mg/dl was associated with an increased risk of CAD (220). Generally, a non-HDLc cut-point of 30 mg/dl higher than the LDLc cut-point is recommended (221), so again, the 132.5 mg/dl identified here is consistent with these prior results.

In this TSSA analysis, we used baseline factors, so they account for exposures at only a single time point, which excludes any interpretation regarding cumulative exposures or changes in risk factor patterns over time. However, it is an important finding that baseline factors were able to discriminate between groups so well because this situation is similar to what is encountered in clinical practice, where a care provider may only have access to the most recent measurements. Computational decision tree methods, like TSSA, hold promise for helping to better elucidating the interrelationships between risk factors to increase understanding of the pathophysiology of CVD. Traditional regression-based analysis methods tend to examine whether risk factors are related to outcomes independently, but this approach ignores the complex interactions between risk factors that are certainly occurring (115,146). As more computational tools become available, it will be important to utilize them to continue to increase our understanding of disease processes.

5.2.3 Longitudinal Trajectory of Glycemic Control and Cardiovascular Disease

The inconsistent relationship between glycemic control and cardiovascular outcomes has been a perplexing issue in type 1 diabetes research. As diabetes is characterized by hyperglycemia, it is natural to assume that glycemia should explain the excess CVD risk observed in type 1 diabetes. In the DCCT/EDIC, the intensive insulin therapy arm had significantly lower CVD incidence than the conventional therapy arm (74), which leads to the

logical assumption that lower HbA1c is protective against CVD incidence. However, this relationship has not been consistent across observational studies of type 1 diabetes (56–63). One potential explanation for the discordant results is that the DCCT participants had a shorter diabetes duration at baseline, so the intensive insulin therapy was begun early in the course of disease, raising the possibility that glycemia plays a greater role in the initiation of atherogenesis than in its progression (78). Another possibility is that the DCCT intensive insulin therapy arm, with a mean baseline HbA1c of 9.1%, achieved an average HbA1c of 7.4% after 6 years (74), which is significantly lower than the average HbA1c in the observational studies. For example, in the EDC data shown here, mean HbA1c was 8.75% at the 1986-88 baseline, decreased to 8.4% 10-years later, and was approximately 8.0% at 18 years of follow-up (Figure 9). This difference between the DCCT and observational studies suggests that perhaps the benefit of glycemic control on CVD risk may only be achieved at very low levels of HbA1c. Another major difference is that the observational studies have generally examined only baseline HbA1c as a predictor of CVD incidence. This approach provides valid results but does not account for changes in glycemic control over time that may affect risk (128). The DCCT results encompass six years of intensive insulin therapy, so cumulative improvement in glycemic control over the course of the trial is accounted for, even if it is not directly modeled. Prior analyses of the EDC cohort have attempted to account for the cumulative effect of HbA1c by modeling HbA1c-months, a measure of exposure to excess HbA1c, similar to pack-years of smoking (58,222). This cumulative measure was still not independently associated with coronary artery disease. Another, more recent, analysis of EDC data also confirmed no association between CAD and baseline or mean HbA1c over follow-up, but did detect an inverse relationship with absolute change in HbA1c (62). These results suggest that analyses need to incorporate variability in

HbA1c over time, but still do not provide an estimate of the relative risk of CVD associated with the level of glycemic control at any given time. This point is essential for clinical relevance, as healthcare providers may only have a patient's current HbA1c measure.

To address the limitations discussed above, here we have presented the results of joint models of the longitudinal trajectory of HbA1c and time to CVD event. Joint models incorporate the subject-specific random effects into the time-to-event framework, and thus model the combined effect of a subject's baseline HbA1c level and rate of change over follow-up on the risk of an outcome by utilizing a random intercept and slope. This method allows us to quantify the relative CVD risk associated with each unit increase in HbA1c, averaged over the entire longitudinal trajectory. We find that a one unit increase in the level of HbA1c is significantly associated with a 40% increased risk of developing CVD, after adjusting for baseline CVD risk factors, and a 25% increased risk of CVD, after adjusting for the mean values of CVD risk factors over follow-up. CAD has been the specific CVD outcome with the most inconsistent relationship with HbA1c, but here a one unit increase in HbA1c is associated with a 40% increased risk of CAD, after adjusting for baseline factors, and a 20% increased risk of CAD after adjusting for mean risk factor values over follow-up.

Graphically examining the longitudinal trajectories of HbA1c by CVD incidence status provides further insight into the lack of association that has been previously reported. As seen in Figure 9, baseline HbA1c was nearly identical regardless of subsequent CVD incidence status. With little variability in HbA1c, an association between baseline HbA1c and CVD risk could not possibly be detected. Significant separation in the two trajectories is not observed until 6 years of follow up. These first several years of HbA1c measures would drive much of the mean HbA1c over time. After the first 6 years of follow-up, HbA1c began to steadily decline,

reflecting the implementation of the DCCT results into clinical practice. Even though the rate of decline (i.e. the slope) after 6 years appears to be the same in subsequent CVD cases compared to noncases, the cases had a larger increase in the first 6 years, and their HbA1c was approximately 0.5 units higher than those who did not develop CVD over the remainder of follow-up.

It is intriguing that baseline levels of other CVD risk factors, specifically lipids and blood pressure, are able to predict CVD incidence, while baseline HbA1c does not. Examining the longitudinal trajectories of non-HDLc and systolic (SBP) and diastolic blood pressures (DBP) in EDC reveals that participants who go on to develop CVD have higher levels of these traditional risk factors throughout the entire follow-up, including the baseline examination. The levels of non-HDLc, SBP, and DBP increase over time to a greater degree in the subsequent CVD cases, but there is still obvious separation between the cases and noncases even >20 years prior to the event or censoring time. This observation suggests that lipid and blood pressure disturbances in patients at high risk for CVD may become apparent earlier in the natural history of diabetes. In contrast, HbA1c levels were nearly equivalent at baseline and separation between cases and noncases became greater closer to the time of CVD event. The reasons for this lack of variability in HbA1c at baseline are unclear. Given the wide range of years of diabetes duration (6-36 years) represented at the EDC baseline, it would be expected that glycemic control would also vary widely. One possibility is that while the standard insulin regimen at the time of the EDC baseline examination was generally inadequate, some individuals were more susceptible to the effects of this inadequate regimen than others and their control worsened over time, leading to increased vascular risk. This explanation is supported by the HbA1c trajectories shown in Figure 9, where the participants who subsequently developed CVD had accelerated worsening of

HbA1c under the period of conventional treatment (i.e. through year 6). Beginning in follow-up year 8, after intensive insulin therapy had been initiated in clinical practice, their HbA1c declined at a rate similar to those who did not develop CVD. The traditional CVD risk factors may be more strongly related to lifestyle and genetic factors that are not specific to diabetes, and thus appear earlier in the lifespan. The differences in the trajectories of HbA1c and the traditional CVD risk factors warrants further investigation. Though baseline values of these other traditional factors have consistently been shown to be associated with CVD incidence in type 1 diabetes, using utilizing joint models to examine the relationship between their longitudinal trajectories and CVD incidence should provide more precise estimates of relative risk. This type of analysis may also reveal important differences in longitudinal trajectories between subgroups of participants (e.g. men and women). Performing joint models analysis for other CVD risk factors using EDC data is an important area of future research.

5.2.4 Cardiovascular Disease Risk and Risk Factors in Men and Women

While in the general population, women have lower risk of CVD than men, in diabetes this protection is essentially erased and women are consistently shown to have nearly equivalent absolute CVD risk as men (56,68). This lack of protection in diabetes also means that women with diabetes have a greater excess risk of CVD than men, an effect that has been consistently shown across international studies (50–53,183). In a 2015 meta-analysis, Huxley et al. reported a pooled SMR for CVD mortality of 11.3 in women versus 5.7 in men (50). The difference in excess risk was somewhat greater for CAD events, with a pooled incidence ratio of 13.3 in women and 5.6 in men. The results presented here are consistent, also showing a greater excess risk in women of both CVD mortality (Table 2) and all definitions of CVD and CAD events

(Table 4). The excess risk observed in EDC is even greater than seen in women of comparable ages in other recent reports; some potential reasons for this difference were discussed in section 5.2.1. However, this differential excess risk seen in EDC compared to the Scottish and Swedish registries is even greater in women than in men. This difference is not likely to be explained by the longer average diabetes duration in EDC versus the registries, which should not differ significantly by sex. Rather, the difference in excess risk by sex supports the hypothesis that there is a treatment bias unfavorable to women. A cross-sectional analysis of data from the DCCT/EDIC showed that women were less likely than men to achieve an HbA1c of <7% or <8% (223). Women were also less likely to be treated with statins even if they had elevated LDL cholesterol levels (223). Differences in treatment by sex have also been shown in type 2 diabetes, where women have been shown to be less likely to be using statins (224,225) and aspirin therapy (226,227) and also less likely to achieve target blood pressure and lipid levels than men (225,227). In the EDC, throughout the 1996-2012 follow-up in the young adults <45 years old examined here, women were about half as likely as men to report taking lipid-lowering medications, despite similar levels of LDL cholesterol. This difference in lipid medication use by sex was not statistically significant. Nevertheless, our results suggest the possibility that these treatment differences by sex in type 1 diabetes may be greater in the United States, compared to other developed countries, and warrants further study.

Despite the similar rates of CVD incidence in men and women with type 1 diabetes, there is evidence that the risk factors for CVD differ by sex (56). In the EDC study, nephropathy was found to be a stronger risk factor for CAD in men than in women (56), although in the Eurodiab study nephropathy predicted CAD incidence in both sexes (61). In the EDC, smoking is a stronger risk factor in men, perhaps due to the fact that men are more likely

to smoke or smoke more than women, while depressive symptomatology and physical activity appear to predict CAD more strongly in women (56). Hyperglycemia has also been shown to contribute to the development of hypertension more strongly in men than women in EDC (228). Different distributions of lipid levels by sex can lead to differential relationships with CVD outcomes and may at least partially explain the excess risk seen in women. For example, while HDL cholesterol is generally protective, very high HDL cholesterol (>80mg/dl) appears to increase CAD risk in women with type 1 diabetes (229). In men, an inverse linear association is present in men, but very few have HDL cholesterol levels >80mg/dl.

Here we have also shown that factors for risk stratification differ in men and women. While sex was not chosen as a splitting factor, in the overall CVD decision tree (Figure 4), one of the major differences between the two high risk, long diabetes duration groups (Groups D and E) was a differential sex distribution. Group D was mostly female and characterized by worse glycemic control and a much higher rate of current smoking than Group E. In contrast, Group E was mostly male with greater renal damage, as measured by their very high median AER, and lower mean estimated glomerular filtration rate. When separate trees were fit by sex (Figure 6, Panels a and b), diabetes duration was the primary splitting factor in women. In women with longer diabetes duration at baseline, no other factors discriminated risk – a discouraging finding as diabetes duration is not a modifiable risk factor and the risk in this group is quite high, approaching 70% over 25 years of follow-up. In women with shorter diabetes duration, non-HDL cholesterol was the only discriminating factor. Nearly 50% of the shorter duration women with non-HDL cholesterol >125 mg/dl developed CVD. These women ranged in age from 13 to 30 years old at baseline and thus only 38 to 55 years old at the end of follow-up, so they are at a particularly high CVD risk for such a young group. This finding underscores the importance of

more aggressive treatment of cholesterol, particularly in younger women, as discussed above. Albumin excretion rate only discriminated CVD risk in men. This supports the earlier observation that renal disease may be a more important CVD risk factor in men compared to women (56). Non-HDL cholesterol discriminated risk in men only in the case of lower AER. In men with lower AER, but higher non-HDL cholesterol, waist-hip ratio further discriminated risk, supporting the role of insulin resistance in the development of CVD in men. A measure of estimated insulin sensitivity, estimated glucose disposal rate, has been shown to be inversely related to cardiovascular outcomes in both sexes in EDC, but men do have lower insulin sensitivity, on average, than women (215). So, as in the case of very high HDL cholesterol in women, this observation in men is more likely to be driven by a difference in the distribution of waist-hip ratio by sex, rather than by a true sex-specific effect of insulin resistance on CVD risk.

Finally, when examining the effect of the longitudinal trajectory of HbA1c on CVD risk, there was no difference in the trajectory of HbA1c over time by sex. There was also no evidence of a difference in the association between longitudinal HbA1c and CVD by sex, regardless of the manifestation of CVD. This observation is consistent with the recent analyses from the DCCT/EDIC, where there was also no difference in the relationship between updated mean HbA1c and CVD by sex (187).

5.2.5 Diabetes Diagnosis Cohort Effects

One of the greatest strengths of the EDC study is that it represents a wide range of diabetes diagnosis years. This characteristic means that the study cohort lends itself well to division into meaningful subcohorts that have experienced different natural histories due to the changing nature of diabetes treatment during the follow-up time. Past analyses stratified by

diabetes diagnosis year have helped to quantify the impact of improvements in care, such as the initiation of intensive insulin therapy after the results of the DCCT, in an observational setting. Specifically, these previous analyses have examined changing incidence of complications over the decades of diagnosis years (191), a significant improvement in life expectancy in those diagnosed in 1965-1980 compared to 1950-1964 (188), and changing impacts of risk factors between those diagnosed in the 1960s versus the 1970s (230).

In the analyses presented here, separate decision trees were built by diabetes diagnosis cohort to explore changes in factors for CVD risk stratification. In the earlier, 1950-1964 diagnosis cohort, AER was the only discriminating factor identified, which concurs with the total CVD tree, where AER was important at longer diabetes duration. In contrast, the 1965-1980 diagnosis cohort first split on non-HDL cholesterol. In this more recent subcohort, AER only discriminated risk at lower levels of non-HDL cholesterol. This difference in the decision trees between the two diagnosis subcohorts supports the evidence that there may be a declining role of renal disease on CVD. It has previously been shown that rates of renal disease, particularly end-stage renal disease, have declined with time (168) and albuminuria has been shown to be less likely to precede CVD than reported in earlier studies (194).

When examining the longitudinal trajectory of HbA1c, we observed differences by diagnosis cohort (Figure 10). The 1950-64 diagnosis cohort had a lower mean baseline HbA1c and this difference persisted throughout most of the follow-up. The HbA1c level in both diagnosis cohorts peaked at cycle 6 and then declined thereafter, reflecting the impact of the DCCT results being incorporated into clinical practice. There was, however, a significant interaction between diagnosis cohort and time with respect to HbA1c, driven by the steeper decline in HbA1c observed after cycle 6 in the 1950-64 cohort, compared to the more recent

1965-80 cohort. Despite this interaction with respect to HbA1c trajectory over time, there was no difference in the relationship between longitudinal HbA1c and CVD incidence by diagnosis cohort.

5.3 STRENGTHS AND LIMITATIONS

The work presented here has many strengths. The Pittsburgh EDC is a well-characterized cohort with over 25 years of prospective follow-up to ascertain mortality and complication status, as well as several clinical visits providing repeated measures of clinical characteristics over time. The study includes a wide range of diabetes diagnosis years, which enables analyses of the changing natural history of type 1 diabetes. Deaths and cardiovascular outcomes were verified using death certificates, interviews with next of kin, autopsy reports, and medical records, as appropriate, and were classified by a committee of physicians who were masked to risk factors status. The novel statistical methods applied to the data addressed key limitations of traditional analyses by formally assessing the presence of subgroups and higher-order interactions using tree-structured survival analysis and more complete utilization of the available repeated measures data on HbA1c using joint models.

The major limitation of these data is the potential for "healthy survivor" bias due to deaths and the exclusion of participants with cardiovascular disease prior to baseline. This issue could lead to an underestimation of the CVD risk in this cohort, as the highest risk participants may have died prior to the beginning of follow-up. We have attempted to address this possibility by performing a sensitivity analysis for the estimated CVD rates and also by performing analyses separately between the two diagnosis subcohorts (1950-1964 and 1965-1980), since the bias, if

present, would occur almost exclusively in the earlier subcohort. Another limitation of these analyses is the restricted availability of CVD hospitalization data from the background Allegheny County population. This particularly impairs our ability to make estimates of excess risk that are as precise as those from other countries, where population-level databases are more comprehensive and accessible. Additionally, these data are from a cohort of almost exclusively Caucasian (98%) people from Southwestern Pennsylvania and may not reflect the experience of all individuals with type 1 diabetes in the United States. The EDC study is also a hospital-based cohort, which may further restrict the generalizability of these results. However, it has previously been shown that the epidemiologic characteristics (167) and mortality experience (188) of the EDC closely reflect those of the population-based Allegheny County type 1 diabetes registry for individuals who were diagnosed during the same period. Furthermore, it has also been shown that patterns of complication incidence in the EDC are similar to those of the conventional treatment arm of the DCCT, in a subgroup with similar baseline characteristics (194), lending support to the generalizability of the EDC cohort.

5.4 DIRECTIONS FOR FUTURE RESEARCH

The results presented here highlight several key questions that remain regarding CVD in type 1 diabetes. One major question is: what is the contemporary burden of CVD in individuals with type 1 diabetes who are aged <30 years? The EDC cohort does not have optimal numbers of individuals in this younger age range after 1996 to estimate contemporary CVD risk in this group. Combining EDC data with data from younger individuals also diagnosed at Children's Hospital of Pittsburgh would help to address this issue in a comparable group. A combined

analysis would also allow more exploration into changes in factors for risk stratification, as well as examination of the impact of the HbA1c trajectory in a group who had a greater proportion of their diabetes natural history in the post-DCCT era. These analyses could also be used to explore at what age elevated CVD risk factors begin to translate into clinical CVD events.

Within the EDC cohort, additional analyses that better account for the effect of medications, such as statin therapy, on CVD risk should be undertaken. In these current analyses, the proportion of young adults <45 years old using statins started out very low (<5% in 1996), but rose to approximately 20% over follow-up. Analyses that attempt to incorporate the effects of statins on CVD incidence would give us a better estimate of the untreated/baseline CVD risk and allow us to quantify how this risk is modified by treatment. Performing these analyses by sex is especially important, given the evidence that women are at a greater excess risk for CVD and may be undertreated for CVD risk factors.

As an exploratory method, the decision tree analysis naturally lends itself to many follow-up analyses. One interesting application of these results would be to look within the subgroups identified using clinical risk factors to examine differences in novel markers of risk, similar to the analysis by Costacou *et al*. (116). In that earlier analysis, candidate genes were examined. The EDC study has recently obtained new genotyping data that would now allow a genome wide association study (GWAS) to be performed, opening the possibility of identifying new genetic markers for CVD risk. A key question is: why are some individuals who are expected to be at high risk for CVD according to their clinical characteristics protected and, conversely, why are some individuals expected to be at low risk for CVD susceptible? Many novel risk factors, including genes and cytokines, could be examined in these groups in an attempt to further discriminate risk and better predict which individuals will develop CVD.

Decision tree methods could also be used to identify better cut-points for treatment targets and to inform planning for clinical trials specific to type 1 diabetes. Rather than identifying a single risk factor target level for the entire type 1 diabetes population, this type of machine learning method could be used to identify subgroup-specific target levels based on combinations of risk factors. A key limitation of the decision tree analysis is that because it is an exploratory method, it that requires external validation before the results could be widely translated. Collaborating with other studies of type 1 diabetes cohorts and registries to validate the factors and cut-points for CVD risk stratification should be a priority.

Expanding on the joint models presented here is another area of future study. One potential direction is to utilize this method or another related method to extend the understanding of the progression of glycemia over the natural history of type 1 diabetes and its relationship to the development of macrovascular disease. One problem with translating the results of longitudinal cohort studies like the EDC into clinical practice is the lack of complete information available to the healthcare provider. Data from a longitudinal study can be used to examine etiologic pathways that may take many years of follow-up to be revealed. This is in contrast with clinically available information, where often only a current snapshot of risk factor levels is accessible. Potentially, this type of joint modeling could be used to create a "2-step" prediction model. For example, cross-sectional HbA1c and other clinical characteristics are entered to first predict the longitudinal trajectory of HbA1c over time and then to subsequently translate this predicted trajectory into a patient's risk of developing CVD. Additionally, statisticians are developing and refining methods to jointly model longitudinal risk factors and multiple time-to-event outcomes (231). In these models, the longitudinal trajectory of a risk factor can be used to predict a sequence of time-to-event outcomes, such as a CVD event and subsequent death. Type

1 diabetes is characterized by many different complications, thus applying this type of modeling has great potential to increase understanding of the natural history of complication development in this population. The pathways from renal disease to CVD and CVD to mortality are of particular interest. Finally, the two major approaches described here, TSSA and joint modeling, could be combined. First, TSSA could be used to identify major subgroups and then joint modeling could be applied within the groups to determine whether trajectories of other continuous risk factors differ with respect to time and their relationship to CVD incidence.

5.5 PUBLIC HEALTH IMPLICATIONS

We have demonstrated that young adults with type 1 diabetes remain at a significantly increased risk of CVD mortality and morbidity compared to the age- and sex-matched background population. The excess risk is especially great in younger women. Likewise, risk stratification showed that women with shorter diabetes duration at baseline who also have elevated non-HDL cholesterol had an approximate 50% risk of developing CVD by age 55. We have also shown that young adults with type 1 diabetes who are between 30-39 years old have a 6.3% 10-year risk of developing CVD overall, which approaches the 7.5% 10-year risk threshold for intensive statin therapy recommended for adults >40 years old by the ACC/AHA (73). However, the 10-year risk we estimate for 30-39 years old includes individuals who are already on statin therapy, which rose to 20% of the group by the end of follow-up. Thus, we can reasonably assume that the 10-year risk would exceed 7.5% if none of the cohort were using lipid-lowering medication. These results support recommending statin therapy to all individuals with type 1 diabetes duration of >20 years who are at least 30 years old. These findings also

underscore the need for cholesterol and blood pressure intervention trials specific to type 1 diabetes to address the excess risk that remains in this population, particularly in young adults and women.

The relationship between glycemic control and CVD has historically been inconsistent across trial data and observational studies. Here, by utilizing an analysis method that incorporates the longitudinal trajectory of glycemia, we have been able to quantify the relative CVD risk associated with each unit increase in HbA1c, averaged over the entire longitudinal trajectory. These findings add important information to the understanding of how the relationship between hyperglycemia and CVD develops over time. It supports the need for good glycemic control, in addition to lipid and blood pressure control, to prevent CVD in type 1 diabetes.

From an analytic standpoint, this work demonstrates the need to consider the limitations of traditional statistical methods for analyzing data from longitudinal cohort studies. The relationships between risk factors and disease are complicated and may not be adequately represented by traditional methods with assume characteristics such as linearity or do not account for changes over time. The accessibility of computationally intensive methods is ever increasing due to higher processing capabilities and the availability of open source programs.

This body of work has therefore demonstrated a comprehensive examination of the epidemiology of CVD and its risk factors in type 1 diabetes. We have quantified the contemporary incidence and excess risk of CVD in young adults <45 years old with type 1 diabetes and have utilized novel statistical methods to address limitations in the analyses of longitudinal cohort studies. These novel methods have enabled us to identify subgroups at varying levels of CVD risk and to estimate the association of longitudinal HbA1c with CVD

incidence. This work demonstrates that novel methods should be considered as part of the epidemiologic "toolbox" as a complement to more traditional methods to increase understanding of disease etiology.

APPENDIX A. EXAMINATION OF THE DOSE RESPONSE RELATIONSHIPS BETWEEN RISK FACTORS AND CARDIOVASCULAR DISEASE IN TYPE 1 DIABETES USING RESTRICTED CUBIC SPLINES

Blood Pressure and 25-Year Risk of Coronary Artery Disease (CAD) by Sex in Type 1 Diabetes (T1D): The Pittsburgh Epidemiology of Diabetes Complications (EDC) Study

Diastolic Blood Pressure (DBP) has been shown to have a J-shaped relationship with CAD in many studies. Trial data have suggested that aggressively lowering BP can have detrimental CAD effects, but there is a lack of evidence in T1D. Thus we aim to model the dose-response relationship between Systolic BP (SBP) and DBP with CAD incidence separately by sex to examine whether there is increased risk at low BP levels and if this relationship differs in men and women with T1D.

Participants in the Pittsburgh EDC, a prospective cohort study of childhood onset T1D, were followed for 25 yrs to ascertain CAD incidence (first instance of fatal CAD, myocardial infarction, revascularization, coronary stenosis >=50%, ischemic ECG, or angina). Restricted cubic splines in Cox models were used to model baseline (1986-88) SBP and DBP on risk of

CAD in those who were free of prevalent CAD at baseline (n=304 men, 301 women, mean age=27, T1D duration=19yrs).

Over 25 yrs, 109 (35.9%) men and 101 (33.6%) women developed CAD. SBP had a linear association with CAD in both sexes ($p<.0001$). For DBP, no association was observed in men ($p=0.21$), but a J-shaped association was observed in women (nonlinear $p=0.005$), with a nadir at approximately 75 mmHg (Figure 10). This nonlinear association remained after covariate adjustment ($p=0.03$).

These data suggest that the shape of the relationship between diastolic blood pressure and CAD differs by sex in T1D. More research is needed to determine whether aggressive lowering of blood pressure is associated with increased risk of CAD in women with T1D.

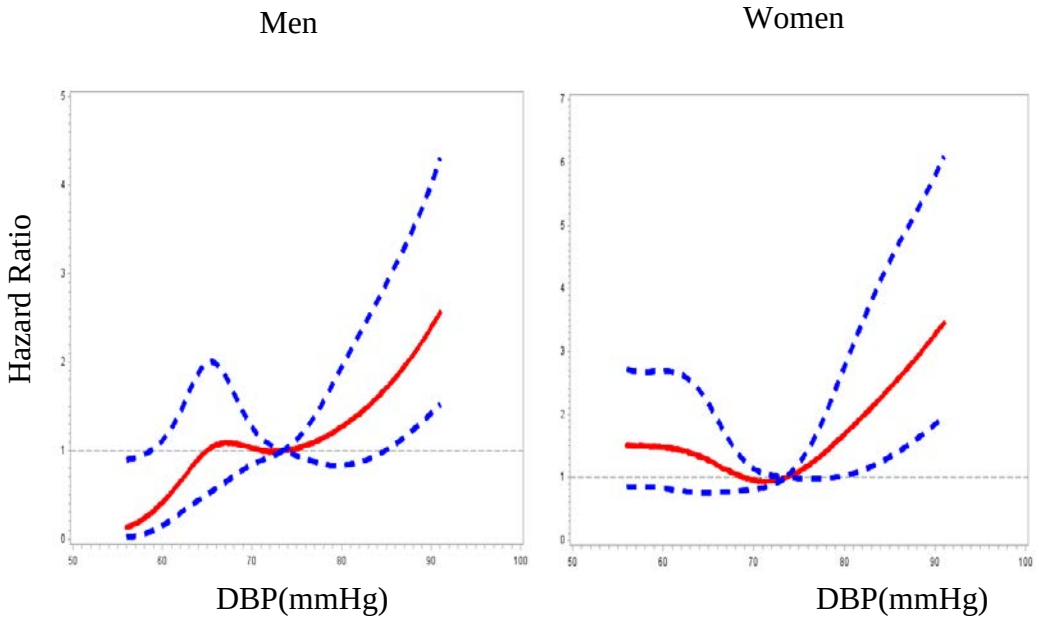

Figure 10. Hazard Rations Over the Range of Diastolic Blood Pressures (DBP) and Coronary Artery Disease by Sex, Restricted Cubic Spline Analysis
Albumin Excretion Rate (AER) Has A Nonlinear Association With Coronary Artery Disease (CAD) Mortality In Type 1 Diabetes (T1D): The Pittsburgh Epidemiology Of Diabetes Complications (EDC) Study

Background: There is evidence that risk for CAD may increase at AER levels below the microalbuminuria threshold (20 μg/min). Accurately determining the shape of this relationship is critical for prevention efforts. We thus modeled the dose-response relationship between AER and CAD.

Methods: EDC study participants diagnosed with T1D from 1965-1980 and without CAD (n=413 mean) were followed for 25 years from baseline (1986-88) when mean T1D duration and age were 15 and 23yrs; 50% male. The association between baseline ln(AER) and CAD incidence (fatal CAD, MI, revascularization procedure/blockage ≥50%, ischemic ECG or angina) and mortality was modeled using Cox proportional hazards with restricted cubic splines (five knots at fixed percentiles - 5^{th}: 3.7μg/min, 25^{th}: 7.2μg/min, 50^{th}: 13.9μg/min, 75^{th}:100.5 μg/min, and 95^{th}:1863 μg/min).

Results: There were 95 (23%) cases of incident CAD. Risk of total CAD increased linearly with increasing AER ($p<.0001$) but adding nonHDL-cholesterol to the model attenuated the relationship ($p=0.19$); there was no evidence for a nonlinear relationship (Figure 11, nonlinear effect $p=0.55$). Risk of fatal CAD (29 cases) increased between ~12-150 μg/min and then remained constant at higher AERs (Figure, adjusted nonlinear effect $p=0.01$).

Conclusions: Risk of fatal CAD seems to increase before the 20 μg/min microalbuminuria threshold is reached and remains constant at AER>150 μg/min.

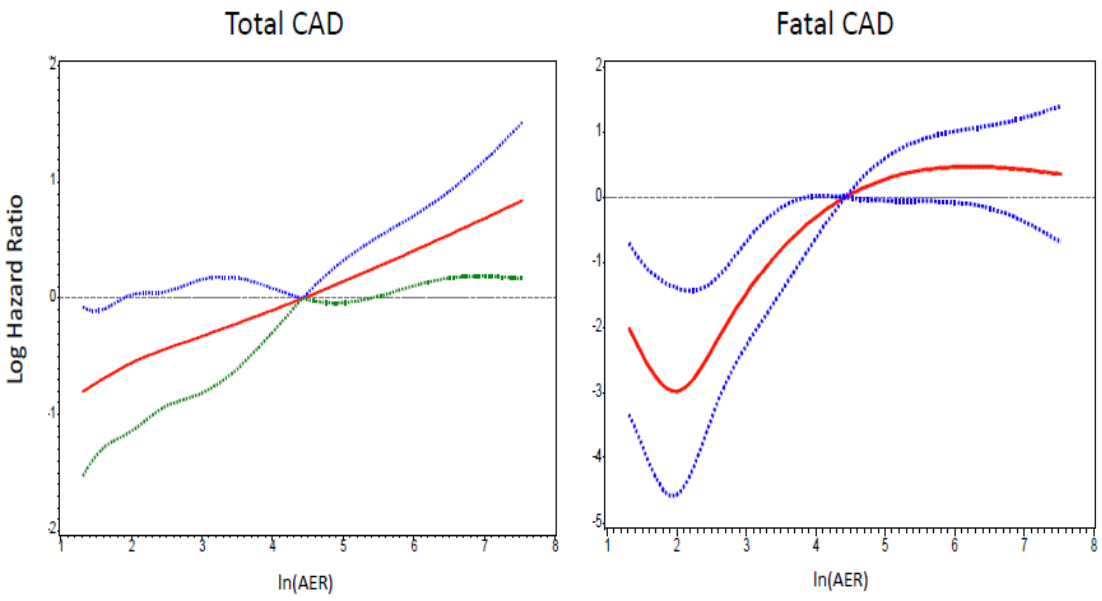

Figure 11. Log Hazard Ratios Over the Range of ln(Albumin Excretion Rate) and Total and Fatal Coronary Artery Disease, Restricted Cubic Spline Analysis

APPENDIX B. COMPARISONS OF RELATIVE AND ABSOLUTE RISKS OF MORTALITY AND CARDIOVASUCLAR DIEASE ACROSS COUNTRIES

Table 10. Comparison of EDC Estimated Excess Mortality and CVD with Recent International Reports

	EDC		**Scottish Registry Linkage Study** (51)		**Swedish National Diabetes Register** (52)		**Australia** (182)	
Follow-Up Period	**1996-2012**		**2005-2007**		**1998-2011**		**1997-2003; 2004-2010***	
	Age Group	**SMR (95% CI)**	**Age Group**	**SMR (95% CI)**	**Age Group**	**HR (95% CI)**	**Age Group**	**SMR (95% CI)**
Total Mortality								
Men	**30-39**	**2.4 (0.97, 5.0)**	**30-39**	**3.4 (2.8, 4.0)**	**18-34**	**2.8 (2.3, 3.5)**	**30-39**	**4.1 (3.6, 4.6); 4.1 (3.6, 4.7)***
	40-44	**4.7 (2.7, 7.6)**	**40-49**	**4.0 (3.8, 4.2)**	**35-49**	**4.3 (3.8, 4.8)**	**40-49**	**4.7 (4.3, 5.1); 4.4 (4.0, 4.8)***
Women	**30-39**	**9.2 (4.8, 15.9)**	**30-39**	**5.2 (3.7, 7.2)**	**18-34**	**4.1 (3.1, 5.4)**	**30-39**	**4.6 (3.9, 5.3); 5.0 (4.2, 5.9)**
	40-44	**6.7 (3.4, 11.9)**	**40-49**	**4.7 (3.6, 6.2)**	**35-49**	**4.4 (3.7, 5.2)**	**40-49**	**4.7 (4.2, 5.3); 4.2 (3.8, 4.7)***
Cardiovascular Mortality								
Men	**30-39**	**18.5 (5.9, 44.7)**	**-**	**-**	**18-34**	**3.5 (1.9, 6.4)**	**-**	**-**
	40-44	**16.4 (8.0, 30.0)**	**-**	**-**	**35-49**	**5.4 (4.3, 6.8)**	**-**	**-**
Women	**30-39**	**63.4 (25.7, 131.8)**	**-**	**-**	**18-34**	**6.9 (2.8, 17.3)**	**-**	**-**
	40-44	**31.6 (12.8, 65.7)**	**-**	**-**	**35-49**	**7.5 (5.2, 10.9)**	**-**	**-**

Table 11. Comparison of EDC Estimated Excess CVD with the Scottish Registry Linkage Study

	EDC		Scottish Registry Linkage Study	
Follow-Up Period	**1996-2012**		**2005-2007**	
Total CVD Incidence	**Age Group**	**Rate Ratio (95% CI)**	**Age Group**	**Rate Ratio (95% CI)**
Men	**30-39**	**6.7 (1.1, 22.0)**	**20-39**	**4.8 (3.7, 6.2)**
	40-44	**4.0 (1.5, 8.9)**	**40-49**	**3.1 (2.7, 3.6)**
Women	**30-39**	**10.0 (1.7, 33.0)**	**20-39**	**5.5 (4.2, 7.2)**
	40-44	**15.8 (7.7, 29.0)**	**40-49**	**5.1 (3.8, 6.8)**

Table 12. Comparison of EDC Absolute and CVD Mortality with Recent International Reports

	EDC		Scottish Registry Linkage Study		Swedish National Diabetes Register		Australia	
Follow-Up Period	**1996-2012**		**2005-2007**		**1998-2011**		**2006**	
	Age Group	**Rate per 100,000 (95% CI)**	**Age Group**	**Rate per 100,000 (95% CI)**	**Age Group**	**Rate per 100,000 (95% CI)**	**Age Group**	**Rate per 100,000**
Total Mortality								
Men	**30-39**	**444 (180, 920)**	**30-39**	**627 (474, 816)**	**≥18**	**997 (959, 1035)**	**0-39**	**269**
	40-44	**1430 800, 2287)**	**40-49**	**1236 (1023, 1481)**			**40-60**	**1198**
Women	**30-39**	**813 428, 1409)**	**30-39**	**418 (281, 599)**			**0-39**	**190**
	40-44	**1096 558, 1946)**	**40-49**	**843 (642, 1088)**			**40-60**	**689**
CVD Mortality								
Men	**30-39**	**296 (94, 712)**	**-**	**-**	**≥18**	**342 (320, 365)**	**0-39**	**49**
	40-44	**902 441, 1648)**	**-**	**-**			**40-60**	**338**
Women	**30-39**	**443 (180, 920)**	**-**	**-**			**0-39**	**22**
	40-44	**658 267, 1363)**	**-**	**-**			**40-60**	**128**

Table 13. Comparisons of Absolute CVD Incidence with the Scottish Registry Linkage Study

	EDC		Scottish Registry Linkage Study	
Follow-Up Period	**1996-2012**		**2005-2007**	
Total CVD Incidence	**Age Group**	**Rate per 100,000 (95% CI)**	**Age Group**	**Rate per 100,000 (95% CI)**
Men	**30-39**	**1036 (591, 1690)**	**20-39**	**273 (197, 369)**
	40-44	**2705 (1827, 3858)**	**40-49**	**1154 (944, 1397)**
Women	**30-39**	**887 (481, 1503)**	**20-39**	**176 (110, 276)**
	40-44	**1974 (1211, 3041)**	**40-49**	**814 (613, 1060)**

BIBLIOGRAPHY

1. American Diabetes Association. Diabetes Basics: Symptoms [Internet]. 2013. Available from: http://www.diabetes.org/diabetes-basics/symptoms/?loc=DropDownDB-symptoms

2. Menke A, Orchard TJ, Imperatore G, Bullard KM, Mayer-Davis E, Cowie CC. The prevalence of type 1 diabetes in the United States. Epidemiology. 2013 Sep;24(5):773–4.

3. Secular trends in incidence of childhood IDDM in 10 countries. Diabetes Epidemiology Research International Group. Diabetes. 1990 Jul;39(7):858–64.

4. Vehik K, Hamman RF, Lezotte D, Norris JM, Klingensmith G, Bloch C, et al. Increasing incidence of type 1 diabetes in 0- to 17-year-old Colorado youth. Diabetes Care. 2007;30(3):503–9.

5. Patterson CC, Dahlquist GG, Gyürüs E, Green A, Soltész G. Incidence trends for childhood type 1 diabetes in Europe during 1989-2003 and predicted new cases 2005-20: a multicentre prospective registration study. Lance. Elsevier Ltd; 2009;373(9680):2027–33.

6. Vehik K, Dabelea D. The changing epidemiology of type 1 diabetes: why is it going through the roof? Diabetesmetabolism Res Rev. 2011;27(1):3–13.

7. Variation and trends in incidence of childhood diabetes in Europe. EURODIAB ACE Study Group. Lancet. 2000 Mar 11;355(9207):873–6.

8. Mayer-Davis E, Dabelea D, Talton J, Hamman R, Divers J, Badaru A, et al. Increase in prevalence of type 1 diabetes from the SEARCH for Diabetes in Youth Study: 2001-2009. Diabetes. 2012;61(Suppl 1):A322.

9. Barker JM, Barriga KJ, Yu L, Miao D, Erlich HA, Norris JM, et al. Prediction of autoantibody positivity and progression to type 1 diabetes: Diabetes Autoimmunity Study in the Young (DAISY). J Clin Endocrinol Metab. 2004;89(8):3896–902.

10. Ostman J, Lönnberg G, Arnqvist HJ, Blohmé G, Bolinder J, Ekbom Schnell A, et al. Gender differences and temporal variation in the incidence of type 1 diabetes: results of 8012 cases in the nationwide Diabetes Incidence Study in Sweden 1983-2002. J Intern Med. 2008 Apr;263(4):386–94.

11. Karvonen M, Tuomilehto J, Libman I, LaPorte R. A review of the recent epidemiological data on the worldwide incidence of type 1 (insulin-dependent) diabetes mellitus. World Health Organization DIAMOND Project Group. Diabetologia. 1993 Oct;36(10):883–92.

12. Ziegler AG, Hummel M, Schenker M, Bonifacio E. Autoantibody appearance and risk for development of childhood diabetes in offspring of parents with type 1 diabetes: the 2-year analysis of the German BABYDIAB Study. Diabetes. 1999;48(3):460–8.

13. Nejentsev S, Sjöroos M, Soukka T, Knip M, Simell O, Lövgren T, et al. Population-based genetic screening for the estimation of Type 1 diabetes mellitus risk in Finland: selective genotyping of markers in the HLA-DQB1, HLA-DQA1 and HLA-DRB1 loci. Diabet Med a J Br Diabet Assoc. 1999;16(12):985–92.

14. Sosenko JM, Krischer JP, Palmer JP, Mahon J, Cowie C, Greenbaum CJ, et al. A risk score for type 1 diabetes derived from autoantibody-positive participants in the diabetes prevention trial-type 1. Diabetes Care. 2008 Mar;31(3):528–33.

15. Hummel M, Bonifacio E, Schmid S, Walter M, Knopff A, Ziegler A-G. Brief communication: early appearance of islet autoantibodies predicts childhood type 1 diabetes in offspring of diabetic parents. Ann Intern Med. 2004 Jun 1;140(11):882–6.

16. Steck AK, Johnson K, Barriga KJ, Miao D, Yu L, Hutton JC, et al. Age of islet autoantibody appearance and mean levels of insulin, but not GAD or IA-2 autoantibodies, predict age of diagnosis of type 1 diabetes: diabetes autoimmunity study in the young. Diabetes Care. 2011 Jun;34(6):1397–9.

17. Parikka V, Näntö-Salonen K, Saarinen M, Simell T, Ilonen J, Hyöty H, et al. Early seroconversion and rapidly increasing autoantibody concentrations predict prepubertal manifestation of type 1 diabetes in children at genetic risk. Diabetologia. 2012 Jul;55(7):1926–36.

18. Ziegler A-G, Bonifacio E. Age-related islet autoantibody incidence in offspring of patients with type 1 diabetes. Diabetologia. 2012 Jul;55(7):1937–43.

19. Wenzlau JM, Juhl K, Yu L, Moua O, Sarkar SA, Gottlieb P, et al. The cation efflux transporter ZnT8 (Slc30A8) is a major autoantigen in human type 1 diabetes. Proc Natl Acad Sci U S A. National Academy of Sciences; 2007;104(43):17040–5.

20. Noble JA, Valdes AM, Cook M, Klitz W, Thomson G, Erlich HA. The role of HLA class II genes in insulin-dependent diabetes mellitus: molecular analysis of 180 Caucasian, multiplex families. Am J Hum Genet. 1996;59(5):1134–48.

21. Hagopian WA, Erlich H, Lernmark A, Rewers M, Ziegler AG, Simell O, et al. The Environmental Determinants of Diabetes in the Young (TEDDY): genetic criteria and international diabetes risk screening of 421 000 infants. Pediatr Diabetes. 2011;12(8):733–43.

22. Valdes AM, Erlich HA, Carlson J, Varney M, Moonsamy P V, Noble JA. Use of class I and class II HLA loci for predicting age at onset of type 1 diabetes in multiple populations. Diabetologia. 2012;

23. Bottini N, Musumeci L, Alonso A, Rahmouni S, Nika K, Rostamkhani M, et al. A functional variant of lymphoid tyrosine phosphatase is associated with type I diabetes. Nat Genet. 2004;36(4):337–8.

24. Steck AK, Zhang W, Bugawan TL, Barriga KJ, Blair A, Erlich HA, et al. Do Non-HLA Genes Influence Development of Persistent Islet Autoimmunity and Type 1 Diabetes in Children With High-Risk HLA-DR,DQ Genotypes? Diabetes. American Diabetes Association; 2009;58(4):1028–33.

25. Cho JH, Gregersen PK. Genomics and the multifactorial nature of human autoimmune disease. N Engl J Med. 2011;365(17):1612–23.

26. Peng H, Zhou M, Xu W-D, Xu K, Zhai Y, Li R, et al. Association of PTPN22 C1858T Polymorphism and Type 1 Diabetes: A Meta-analysis. Immunol Invest. 2012;

27. Rewers M, LaPorte RE, Walczak M, Dmochowski K, Bogaczynska E. Apparent epidemic of insulin-dependent diabetes mellitus in Midwestern Poland. Diabetes. 1987;36(1):106–13.

28. Hermann R, Knip M, Veijola R, Simell O, Laine A-P, Akerblom HK, et al. Temporal changes in the frequencies of HLA genotypes in patients with Type 1 diabetes--indication of an increased environmental pressure? Diabetologia. 2003;46(3):420–5.

29. Vehik K, Hamman RF, Lezotte D, Norris JM, Klingensmith GJ, Rewers M, et al. Trends in High-Risk HLA Susceptibility Genes Among Colorado Youth With Type 1 Diabetes. Diabetes Care. American Diabetes Association; 2008;31(7):1392–6.

30. Fourlanos S, Varney MD, Tait BD, Morahan G, Honeyman MC, Colman PG, et al. The Rising Incidence of Type 1 Diabetes Is Accounted for by Cases With Lower-Risk. Diabetes Care. 2008;31(8):12–5.

31. Norris JM. Infant and childhood diet and type 1 diabetes risk: recent advances and prospects. Curr Diab Rep. 2010;10(5):345–9.

32. Ziegler AG, Schmid S, Huber D, Hummel M, Bonifacio E. Early Infant Feeding and Risk of Developing Type 1 Diabetes-Associated Autoantibodies. Jama. 2003;290(13):1721–8.

33. Vaarala O, Knip M, Paronen J, Hämäläinen AM, Muona P, Väätäinen M, et al. Cow's milk formula feeding induces primary immunization to insulin in infants at genetic risk for type 1 diabetes. Diabetes. 1999.

34. Study design of the Trial to Reduce IDDM in the Genetically at Risk (TRIGR). Pediatr Diabetes. 2007 Jun;8(3):117–37.

35. Knip M, Virtanen SM, Becker D, Dupré J, Krischer JP, Åkerblom HK. Early feeding and risk of type 1 diabetes: experiences from the Trial to Reduce Insulin-dependent diabetes mellitus in the Genetically at Risk (TRIGR). Am J Clin Nutr. 2011 Dec;94(6 Suppl):1814S – 1820S.

36. Vaarala O, Ilonen J, Ruohtula T, Pesola J, Virtanen SM, Härkönen T, et al. Removal of Bovine Insulin From Cow's Milk Formula and Early Initiation of Beta-Cell Autoimmunity in the FINDIA Pilot Study. Arch Pediatr Adolesc Med. 2012 Jul 1;166(7):608–14.

37. Virtanen SM, Läärä E, Hyppönen E, Reijonen H, Räsänen L, Aro A, et al. Cow's milk consumption, HLA-DQB1 genotype, and type 1 diabetes: a nested case-control study of siblings of children with diabetes. Childhood diabetes in Finland study group. Diabetes. 2000 Jun;49(6):912–7.

38. Lempainen J, Vaarala O, Mäkelä M, Veijola R, Simell O, Knip M, et al. Interplay between PTPN22 C1858T polymorphism and cow's milk formula exposure in type 1 diabetes. J Autoimmun. 2009;33(2):155–64.

39. Wilkin TJ. The accelerator hypothesis: weight gain as the missing link between Type I and Type II diabetes. Diabetologia. Springer; 2001;44(7):914–22.

40. Fourlanos S, Narendran P, Byrnes GB, Colman PG, Harrison LC. Insulin resistance is a risk factor for progression to type 1 diabetes. Diabetologia. 2004;47(10):1661–7.

41. Dabelea D, D'Agostino RB, Mayer-Davis EJ, Pettitt DJ, Imperatore G, Dolan LM, et al. Testing the accelerator hypothesis: body size, beta-cell function, and age at onset of type 1 (autoimmune) diabetes. Diabetes Care. 2006;29(2):290–4.

42. Bingley PJ, Mahon JL, Gale EAM. Insulin resistance and progression to type 1 diabetes in the European Nicotinamide Diabetes Intervention Trial (ENDIT). Diabetes Care. 2008;31(1):146–50.

43. Xu P, Cuthbertson D, Greenbaum C, Palmer JP, Krischer JP. Role of insulin resistance in predicting progression to type 1 diabetes. Diabetes Care. 2007 Sep 1;30(9):2314–20.

44. Mrena S, Virtanen SM, Laippala P, Kulmala P, Hannila M-L, Akerblom HK, et al. Models for predicting type 1 diabetes in siblings of affected children. Diabetes Care. 2006 Mar;29(3):662–7.

45. Chapman NM, Kim KS. Persistent coxsackievirus infection: enterovirus persistence in chronic myocarditis and dilated cardiomyopathy. Curr Top Microbiol Immunol. 2008;323:275–92.

46. Viskari H, Knip M, Tauriainen S, Huhtala H, Veijola R, Ilonen J, et al. Maternal enterovirus infection as a risk factor for type 1 diabetes in the exposed offspring. Diabetes Care. 2012 Jun;35(6):1328–32.

47. Hyöty H, Hiltunen M, Knip M, Laakkonen M, Vähäsalo P, Karjalainen J, et al. A prospective study of the role of coxsackie B and other enterovirus infections in the pathogenesis of IDDM. Childhood Diabetes in Finland (DiMe) Study Group. Diabetes. 1995 Jun;44(6):652–7.

48. Stene LC, Barriga K, Hoffman M, Kean J, Klingensmith G, Norris JM, et al. Normal but increasing hemoglobin A1c levels predict progression from islet autoimmunity to overt type 1 diabetes: Diabetes Autoimmunity Study in the Young (DAISY). Pediatr Diabetes. 2006 Oct;7(5):247–53.

49. Morgan E, Cardwell CR, Black CJ, McCance DR, Patterson CC. Excess mortality in Type 1 diabetes diagnosed in childhood and adolescence: a systematic review of population-based cohorts. Acta Diabetol. 2015 Aug;52(4):801–7.

50. Huxley RR, Peters SAE, Mishra GD, Woodward M. Risk of all-cause mortality and vascular events in women versus men with type 1 diabetes: a systematic review and meta-analysis. lancet Diabetes Endocrinol. 2015 Mar;3(3):198–206.

51. Livingstone SJ, Looker HC, Hothersall EJ, Wild SH, Lindsay RS, Chalmers J, et al. Risk of cardiovascular disease and total mortality in adults with type 1 diabetes: Scottish registry linkage study. PLoS Med. 2012;9(10):e1001321.

52. Lind M, Svensson A-M, Kosiborod M, Gudbjörnsdottir S, Pivodic A, Wedel H, et al. Glycemic control and excess mortality in type 1 diabetes. N Engl J Med. 2014 Nov 20;371(21):1972–82.

53. Harding JL, Shaw JE, Peeters A, Davidson S, Magliano DJ. Age-specific trends from 2000-2011 in all-cause and cause-specific mortality in type 1 and type 2 diabetes: a cohort study of more than one million people. Diabetes Care. 2016;Online bef.

54. The Diabetes Control and Complications Trial/Epidemiology of Diabetes Interventions and Complications (DCCT/EDIC) Research Group. Association Between 7 Years of Intensive Insulin Treatment of Type 1 Diabetes and Long-term Mortality. JAMA. 2015;313(1):45–53.

55. Krolewski AS, Kosinski EJ, Warram JH, Leland OS, Busick EJ, Asmal AC, et al. Magnitude and determinants of coronary artery disease in juvenile-onset, insulin-dependent diabetes mellitus. Am J Cardiol. 1987 Apr 1;59(8):750–5.

56. Lloyd CE, Kuller LH, Ellis D, Becker DJ, Wing RR, Orchard TJ. Coronary Artery Disease in IDDM: Gender differences in risk factors, but not risk. ArteriosclerThrombVascBiol. 1996;16(6):720–6.

57. Koivisto VA, Stevens LK, Mattock M, Ebeling P, Muggeo M, Stephenson J, et al. Cardiovascular disease and its risk factors in IDDM in Europe. EURODIAB IDDM Complications Study Group. Diabetes Care. 1996 Jul;19(7):689–97.

58. Forrest KY, Becker DJ, Kuller LH, Wolfson SK, Orchard TJ. Are predictors of coronary heart disease and lower-extremity arterial disease in type 1 diabetes the same? A prospective study. Atherosclerosis. 2000 Jan;148(1):159–69.

59. Weis U, Turner B, Gibney J, Watts GF, Burke V, Shaw KM, et al. Long-term predictors of coronary artery disease and mortality in type 1 diabetes. QJM Mon J Assoc Physicians. 2001;94(11):623–30.

60. Orchard TJ, Olson JC, Erbey JR, Williams K, Forrest KY-Z, Smithline Kinder L, et al. Insulin resistance-related factors, but not glycemia, predict coronary artery disease in type 1 diabetes: 10-year follow-up data from the Pittsburgh Epidemiology of Diabetes Complications Study. Diabetes Care. 2003;26(5):1374–9.

61. Soedamah-Muthu SS, Chaturvedi N, Toeller M, Ferriss B, Reboldi P, Michel G, et al. Risk factors for coronary heart disease in type 1 diabetic patients in Europe: the EURODIAB Prospective Complications Study. Diabetes Care. 2004;24(2):530–7.

62. Prince CT, Becker DJ, Costacou T, Miller RG, Orchard TJ. Changes in glycaemic control and risk of coronary artery disease in type 1 diabetes mellitus: findings from the Pittsburgh Epidemiology of Diabetes Complications Study (EDC). Diabetologia. 2007 Nov;50(11):2280–8.

63. Orchard TJ, Costacou T. When are type 1 diabetic patients at risk for cardiovascular disease? Curr Diab Rep. 2010 Mar;10(1):48–54.

64. Nathan DM, McGee P, Steffes MW, Lachin JM. Relationship of Glycated Albumin to Blood Glucose and Glycated Hemoglobin (HbA1C) Values and to Retinopathy, Nephropathy and Cardiovascular Outcomes in the DCCT/EDIC Study. Diabetes. 2014 Aug 29;63(1):282–90.

65. Soedamah-Muthu SS, Fuller JH, Mulnier HE, Raleigh VS, Lawrenson RA, Colhoun HM. High risk of cardiovascular disease in patients with type 1 diabetes in the U.K.: a cohort study using the general practice research database. Diabetes Care. 2006;29(4):798–804.

66. Secrest AM, Prince CT, Costacou T, Miller RG, Orchard TJ. Predictors of and survival after incident stroke in type 1 diabetes. Diabetes Vasc Dis Res. 2012;10:3–10.

67. Kannel WB, Skinner JJ, Schwartz MJ, Shurtleff D. Intermittent claudication. Incidence in the Framingham Study. Circulation. 1970 May;41(5):875–83.

68. Kalyani RR, Lazo M, Ouyang P, Turkbey E, Chevalier K, Brancati F, et al. Gender Differences in Diabetes and Risk of Incident Coronary Artery Disease in Healthy Young and Middle-Aged Adults. Diabetes Care. 2013 Oct 31;dc13–1755 – .

69. Dorman JS, Laporte RE, Kuller LH, Cruickshanks KJ, Orchard TJ, Wagener DK, et al. The Pittsburgh Insulin-dependent Diabetes Mellitus (IDDM) Morbidity and Mortality Study: Mortality Results. Diabetes. 1984;33(3):271–6.

70. Moss SE, Klein R, Klein BE. Cause-specific mortality in a population-based study of diabetes. Am J Public Health. 1991;81(9):1158–62.

71. Secrest AM, Becker DJ, Kelsey SF, Laporte RE, Orchard TJ. Cause-specific mortality trends in a large population-based cohort with long-standing childhood-onset type 1 diabetes. Diabetes. 2010;59(12):3216–22.

72. The Diabetes Control and Complications Trial/Epidemiology of Diabetes Interventions and Complications (DCCT/EDIC) Research Group. Intensive Diabetes Treatment and Cardiovascular Outcomes in Type 1 Diabetes: The DCCT/EDIC Study 30-Year Follow-Up. Diabetes Care. 2016;(online be.

73. Stone NJ, Robinson JG, Lichtenstein AH, Bairey Merz CN, Blum CB, Eckel RH, et al. 2013 ACC/AHA Guideline on the Treatment of Blood Cholesterol to Reduce Atherosclerotic Cardiovascular Risk in Adults: A Report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. Circulation. 2014 Jun 24;129(25 Suppl 2):S1–45.

74. Nathan DM, Cleary PA, Backlund J-YC, Genuth SM, Lachin JM, Orchard TJ, et al. Intensive diabetes treatment and cardiovascular disease in patients with type 1 diabetes. N Engl J Med. 2005 Dec 22;353(25):2643–53.

75. Klein BEK, Klein R, McBride PE, Cruickshanks KJ, Palta M, Knudtson MD, et al. Cardiovascular disease, mortality, and retinal microvascular characteristics in type 1 diabetes: Wisconsin epidemiologic study of diabetic retinopathy. Arch Intern Med. 2004;164(17):1917–24.

76. Deckert T, Yokoyama H, Mathiesen E, Ronn B, Jensen T, Feldt-Rasmussen B, et al. Cohort study of predictive value of urinary albumin excretion for atherosclerotic vascular disease in patients with insulin dependent diabetes. BMJ. 1996 Apr 6;312(7035):871–4.

77. Wadén J, Forsblom C, Thorn LM, Gordin D, Saraheimo M, Groop P-H. A1C variability predicts incident cardiovascular events, microalbuminuria, and overt diabetic nephropathy in patients with type 1 diabetes. Diabetes. 2009 Nov 1;58(11):2649–55.

78. Orchard TJ, Costacou T, Kretowski A, Nesto RW. Type 1 diabetes and coronary artery disease. Diabetes Care. 2006 Nov 1;29(11):2528–38.

79. Olson JC, Erbey JR, Forrest KYZ, Williams K, Becker DJ, Orchard TJ. Glycemia (or, in women, estimated glucose disposal rate) predict lower extremity arterial disease events in type 1 diabetes. Metabolism. 2002 Feb;51(2):248–54.

80. Sahakyan K, Klein BEK, Lee KE, Myers CE, Klein R. The 25-year cumulative incidence of lower extremity amputations in people with type 1 diabetes. Diabetes Care. 2011 Mar 1;34(3):649–51.

81. Silva JA, Escobar A, Collins TJ, Ramee SR, White CJ. Unstable Angina : A Comparison of Angioscopic Findings Between Diabetic and Nondiabetic Patients. Circulation. 1995 Oct 1;92(7):1731–6.

82. Davies MJ. The composition of coronary-artery plaques. N Engl J Med. 1997 May 1;336(18):1312–4.

83. Krolewski AS, Warram JH, Christlieb AR, Busick EJ, Kahn CR. The changing natural history of nephropathy in type I diabetes. Am J Med. 1985 May;78(5):785–94.

84. Jensen T, Borch-Johnsen K, Kofoed-Enevoldsen A, Deckert T. Coronary heart disease in young type 1 (insulin-dependent) diabetic patients with and without diabetic nephropathy: incidence and risk factors. Diabetologia. 1987;30(3):144–8.

85. Tuomilehto J, Borch-Johnsen K, Molarius A, Forsén T, Rastenyte D, Sarti C, et al. Incidence of cardiovascular disease in Type 1 (insulin-dependent) diabetic subjects with and without diabetic nephropathy in Finland. Diabetologia. 1998 Jul;41(7):784–90.

86. Torffvit O, Lövestam-Adrian M, Agardh E, Agardh C-D. Nephropathy, but not retinopathy, is associated with the development of heart disease in Type 1 diabetes: a 12-year observation study of 462 patients. Diabet Med. 2005 Jun;22(6):723–9.

87. May O, Arildsen H, Damsgaard EM, Mickley H. Cardiovascular autonomic neuropathy in insulin-dependent diabetes mellitus: prevalence and estimated risk of coronary heart disease in the general population. J Intern Med. 2000 Dec;248(6):483–91.

88. Valensi P, Sachs R-N, Harfouche B, Lormeau B, Paries J, Cosson E, et al. Predictive Value of Cardiac Autonomic Neuropathy in Diabetic Patients With or Without Silent Myocardial Ischemia. Diabetes Care. 2001 Feb 1;24(2):339–43.

89. Ewing DJ, Campbell IW, Clarke BF. Mortality in diabetic autonomic neuropathy. Lancet. 1976 Mar 20;1(7960):601–3.

90. Veglio M, Giunti S, Stevens LK, Fuller JH, Perin PC. Prevalence of Q-T interval dispersion in type 1 diabetes and its relation with cardiac ischemia : the EURODIAB IDDM Complications Study Group. Diabetes Care. 2002 Apr;25(4):702–7.

91. Lykke JA, Tarnow L, Parving HH, Hilsted J. A combined abnormality in heart rate variation and QT corrected interval is a strong predictor of cardiovascular death in type 1 diabetes. Scand J Clin Lab Invest. 2008;68(7):654–9.

92. Mayer-Davis EJ, Ma B, Lawson A, D'Agostino RB, Liese AD, Bell RA, et al. Cardiovascular disease risk factors in youth with type 1 and type 2 diabetes: implications of a factor analysis of clustering. Metab Syndr Relat Disord. 2009;7(2):89–95.

93. Guy J, Ogden L, Wadwa RP, Hamman RF, Mayer-Davis EJ, Liese AD, et al. Lipid and lipoprotein profiles in youth with and without type 1 diabetes: the SEARCH for Diabetes in Youth case-control study. Diabetes Care. 2009 Mar 1;32(3):416–20.

94. Soedamah-Muthu SS, Chaturvedi N, Witte DR, Stevens LK, Porta M, Fuller JH. Relationship between risk factors and mortality in type 1 diabetic patients in Europe: the EURODIAB Prospective Complications Study (PCS). Diabetes Care. 2008 Jul;31(7):1360–6.

95. Gordin D, Wadén J, Forsblom C, Thorn L, Rosengård-Bärlund M, Tolonen N, et al. Pulse pressure predicts incident cardiovascular disease but not diabetic nephropathy in patients with type 1 diabetes (The FinnDiane Study). Diabetes Care. 2011;34(4):886–91.

96. Ross R. Atherosclerosis is an inflammatory disease. Am Heart J. 1999;138(5 Pt 2):115–26.

97. Libby P. Inflammation and Atherosclerosis. Circulation. Am Heart Assoc; 2002;105(9):1135–43.

98. Schnabel RB, Yin X, Larson MG, Yamamoto JF, Fontes JD, Kathiresan S, et al. Multiple inflammatory biomarkers in relation to cardiovascular events and mortality in the community. Arterioscler Thromb Vasc Biol. 2013 Jul;33(7):1728–33.

99. Alman AC, Kinney GL, Tracy RP, Maahs DM, Hokanson JE, Rewers MJ, et al. Prospective association between inflammatory markers and progression of coronary artery calcification in adults with and without type 1 diabetes. Diabetes Care. 2013;36(7):1967–73.

100. Gomes MB, Piccirillo LJ, Nogueira VG, Matos HJ. Acute-phase proteins among patients with type 1 diabetes. Diabetes Metab. 2003;29(4 I):405–11.

101. Libby P, Nathan DM, Abraham K, Brunzell JD, Fradkin JE, Haffner SM, et al. Report of the National Heart, Lung, and Blood Institute-National Institute of Diabetes and Digestive and Kidney Diseases Working Group on Cardiovascular Complications of Type 1 Diabetes Mellitus. Circulation. 2005;111:3489–93.

102. Devaraj S, Glaser N, Griffen S, Wang-Polagruto J, Miguelino E, Jialal I. Increased monocytic activity and biomarkers of inflammation in patients with type 1 diabetes. Diabetes. 2006;55(3):774–9.

103. Trinanes J, Salido E, Fernandez J, Rufino M, Gonzalez-Posada JM, Torres A, et al. Type 1 diabetes increases the expression of proinflammatory cytokines and adhesion molecules in the artery wall of candidate patients for kidney transplantation. Diabetes Care. 2012;35:427–33.

104. Costacou T, Ferrell RE, Orchard TJ. Haptoglobin genotype: a determinant of cardiovascular complication risk in type 1 diabetes. Diabetes. 2008;57(6):1702–6.

105. Simpson M, Snell-Bergeon JK, Kinney GL, Lache O, Miller-Lotan R, Anbinder Y, et al. Haptoglobin genotype predicts development of coronary artery calcification in a prospective cohort of patients with type 1 diabetes. Cardiovasc Diabetol. Barbara Davis Center for Childhood Diabetes, Aurora, CO, USA.; 2011;10:99.

106. Asleh R, Marsh S, Shilkrut M, Binah O, Guetta J, Lejbkowicz F, et al. Genetically determined heterogeneity in hemoglobin scavenging and susceptibility to diabetic cardiovascular disease. Circ Res. 2003;92(11):1193–200.

107. Levy AP, Purushothaman KR, Levy NS, Purushothaman M, Strauss M, Asleh R, et al. Downregulation of the hemoglobin scavenger receptor in individuals with diabetes and the Hp 2-2 genotype: implications for the response to intraplaque hemorrhage and plaque vulnerability. Circ Res. 2007;101(1):106–10.

108. Asleh R, Levy AP. In vivo and in vitro studies establishing haptoglobin as a major susceptibility gene for diabetic vascular disease. Vasc Health Risk Manag. 2005;1(1):19–28.

109. Mollsten A, Lajer M, Jorsal A, Tarnow L. The endothelial nitric oxide synthase gene and risk of diabetic nephropathy and development of cardiovascular disease in type 1 diabetes. Mol Genet Metab. 2009;97(1):80–4.

110. Pettersson-Fernholm K, Forsblom C, Hudson BI, Perola M, Grant PJ, Groop P-H. The functional -374 T/A RAGE gene polymorphism is associated with proteinuria and cardiovascular disease in type 1 diabetic patients. Diabetes. 2003 Mar;52(3):891–4.

111. McBrien K, Rabi DM, Campbell N, Barnieh L, Clement F, Hemmelgarn BR, et al. Intensive and standard blood pressure targets in patients with type 2 diabetes mellitus: systematic review and meta-analysis. Arch Intern Med. 2012;172(17):1296–303.

112. Lv J, Neal B, Ehteshami P, Ninomiya T, Woodward M, Rodgers A, et al. Effects of intensive blood pressure lowering on cardiovascular and renal outcomes: a systematic review and meta-analysis. PLoS Med. 2012 Aug;9(8):e1001293.

113. Accord Study Group. Effects of intensive blood pressure control in type 2 diabetes mellitus. N Engl J Med. 2010;362:1575–85.

114. Cox DR. Regression models and life-tables. J R Stat Soc Ser B. 1972;34(2):187–220.

115. Carmelli D, Zhang H, Swan GE. Obesity and 33-year follow-up for coronary heart disease and cancer mortality. Epidemiology. 1997 Jul;8(4):378–83.

116. Costacou T, Chang Y, Ferrell RE, Orchard TJ. Identifying genetic susceptibilities to diabetes-related complications among individuals at low risk of complications: An application of tree-structured survival analysis. Am J Epidemiol. 2006 Nov 1;164(9):862–72.

117. May S, Bigelow C. Modeling nonlinear dose-response relationships in epidemiologic studies: statistical approaches and practical challenges. Dose Response. 2005;3(4):474–90.

118. Maclure M, Greenland S. Test for trend and dose response: misinterpretations and alternatives. Am J Epidemiol. 1992;135:96–104.

119. Marrie RA, Dawson N V, Garland A. Quantile regression and restricted cubic splines are useful for exploring relationships between continuous variables. J Clin Epidemiol. 2009;62:511–7.

120. Bangalore S, Messerli FH, Wun C-C, Zuckerman AL, DeMicco D, Kostis JB, et al. J-curve revisited: An analysis of blood pressure and cardiovascular events in the Treating to New Targets (TNT) Trial. Eur Heart J. 2010 Dec;31(23):2897–908.

121. Law MR, Wald NJ. Risk factor thresholds: their existence under scrutiny. BMJ. 2002;324(7353):1570–6.

122. Pua Y-H, Ong P-H, Chong H-C, Yeo W, Tan CI-C, Lo N-N. Associations of self-report physical function with knee strength and knee range-of-motion in total knee arthroplasty possible nonlinear and threshold effects. J Arthroplasty. 2013 Oct;28(9):1521–7.

123. Turner EL, Dobson JE, Pocock SJ. Categorisation of continuous risk factors in epidemiological publications: a survey of current practice. Epidemiol Perspect Innov. 2010;7:9.

124. Altman DG, Royston P. The cost of dichotomising continuous variables. BMJ. 2006 May 6;332(7549):1080.

125. Royston P, Altman DG, Sauerbrei W. Dichotomizing continuous predictors in multiple regression: a bad idea. Stat Med. 2006;25(1):127–41.

126. Greenland S. Dose-response and trend analysis in epidemiology: alternatives to categorical analysis. Epidemiology. 1995;6:356–65.

127. Zhao LP, Kolonel LN. Efficiency loss from categorizing quantitative exposures into qualitative exposures in case-control studies. Am J Epidemiol. 1992 Aug 15;136(4):464–74.

128. Lim HJ, Mondal P, Skinner S. Joint modeling of longitudinal and event time data: application to HIV study. J Med Stat Informatics. Herbert Publications; 2013 Sep 28;1:1.

129. Edwards LJ. Modern statistical techniques for the analysis of longitudinal data in biomedical research. Pediatr Pulmonol. 2000;30:330–44.

130. Eggleston EP, Laub JH, Sampson RJ. Methodological Sensitivities to Latent Class Analysis of Long-Term Criminal Trajectories. J Quant Criminol. 2004 Mar;20(1):1–26.

131. Tsiatis AA, Davidian M. Joint modeling of Longitudinal and Time-to-Event Data: An Overview. Stat Sin. 2004;14:809–34.

132. Twisk JWR, de Vente W. The analysis of randomised controlled trial data with more than one follow-up measurement. A comparison between different approaches. Eur J Epidemiol. 2008 Jan;23(10):655–60.

133. Locascio JJ, Atri A. An overview of longitudinal data analysis methods for neurological research. Dement Geriatr Cogn Dis Extra. 2011 Jan;1(1):330–57.

134. Gibbons RD, Hedeker D, DuToit S. Advances in analysis of longitudinal data. Annu Rev Clin Psychol. 2010 Jan;6:79–107.

135. Segal MR. Features of tree-structured survival analysis. Epidemiology. 1997 Jul;8(4):344–6.

136. Morgan J, Sonquist J. Problems in the analysis of survey data and a proposal. J Am Stat Assoc. 1963;58:415–34.

137. Quinlan J. Discovering rules by induction from large collections of examples. In: Michie D, editor. Expert systems in the microelectronic age. Endinburgh: Edinburgh University Press; 1979.

138. Quinlan JR. Induction of decision trees. Mach Learn. 1986 Mar;1(1):81–106.

139. Breiman L, Friedman J, Olshen R, Stone C. Classification and regression tree. New York, NY: Wadsworth; 1984.

140. Witten I, Frank E, Hall M. Data Mining: Practical Machine Learning Tools and Techniques. Third. Burlington, MA: Morgan Kaufman Publishers; 2011. 192-203 p.

141. Ciampi A, Lawless F, McKinney M, Singhal K. Regression and recursive partition strategies in the analysis of medical survival data. J Clin Epidemiol. 1988;41(8):737–48.

142. Segal M. Regression trees for censored data. Biometrics. 1988;44(1):35–47.

143. LeBlanc M, Crowley J. Relative risk trees for censored survival data. Biometrics. 1992;48:411–25.

144. Therneau T, Atkinson E. An Introduction to Recursive Partitioning Using the RPART Routines. Mayo Clinic; 2013.

145. R Development Team. 2013.

146. Ciampi A, Negassa A, Lou Z. Tree-structured prediction for censored survival data and the Cox model. J Clin Epidemiol. 1995 May;48(5):675–89.

147. Marshall RJ. The use of classification and regression trees in clinical epidemiology. J Clin Epidemiol. 2001 Jun;54(6):603–9.

148. Yadegari H, Bozorgmanesh M, Hadaegh F, Azizi F. Non-linear contribution of glucose measures to cardiovascular diseases and mortality: reclassifying the Framingham's risk categories: a decade follow-up from the Tehran lipid and glucose study. Int J Cardiol. 2013 Aug 20;167(4):1486–94.

149. Messerli FH, Mancia G, Conti CR, Hewkin AC, Kupfer S, Champion A, et al. Dogma disputed: can aggressively lowering blood pressure in hypertensive patients with coronary artery disease be dangerous? Ann Intern Med. 2006 Jun 20;144(12):884–93.

150. Protogerou AD, Safar ME, Iaria P, Safar H, Le Dudal K, Filipovsky J, et al. Diastolic blood pressure and mortality in the elderly with cardiovascular disease. Hypertension. 2007 Jul;50(1):172–80.

151. Weinberger MH. Do no harm. Antihypertensive therapy and the "J" curve. Arch Intern Med. 1992 Mar;152(3):473–6.

152. Farnett L, Mulrow CD, Linn WD, Lucey CR, Tuley MR. The J-curve phenomenon and the treatment of hypertension. Is there a point beyond which pressure reduction is dangerous? JAMA. 1991;265(4):489–95.

153. Waller PC, Isles CG, Lever AF, Murray GD, McInnes GT. Does therapeutic reduction of diastolic blood pressure cause death from coronary heart disease? J Hum Hypertens. 1988 Jun;2(1):7–10.

154. Hansson L, Zanchetti A, Carruthers SG, Dahlöf B, Elmfeldt D, Julius S, et al. Effects of intensive blood-pressure lowering and low-dose aspirin in patients with hypertension: principal results of the Hypertension Optimal Treatment (HOT) randomised trial. HOT Study Group. Lancet. 1998 Jun 13;351(9118):1755–62.

155. Psaty BM, Furberg CD, Kuller LH, Cushman M, Savage PJ, Levine D, et al. Association between blood pressure level and the risk of myocardial infarction, stroke, and total mortality: the cardiovascular health study. Arch Intern Med. 2001 May 14;161(9):1183–92.

156. Heinzl H, Kaider A. Gaining more flexibility in Cox proportional hazards regression models with cubic spline functions. Comput Methods Programs Biomed. 1997 Nov 1;54(3):201–8.

157. Durrleman S, Simon R. Flexible regression models with cubic splines. Stat Med. 1989 May;8(5):551–61.

158. Desquilbet L, Mariotti F. Dose-response analyses using restricted cubic spline functions in public health research. Stat Med. 2010;29(9):1037–57.

159. Govindarajulu US, Spiegelman D, Thurston SW, Ganguli B, Eisen EA. Comparing smoothing techniques in Cox models for exposure-response relationships. Stat Med. 2007 Sep 10;26(20):3735–52.

160. Wulfsohn MS, Tsiatis AA. A joint model for survival and longitudinal data measured with error. Biometrics. 1997 Mar;53(1):330–9.

161. Guo X, Carlin BP. Separate and Joint Modeling of Longitudinal and Event Time Data Using Standard Computer Packages. Am Stat. Taylor & Francis; 2004 Feb;58(1):16–24.

162. Baghfalaki T, Ganjali M, Berridge D. Robust joint modeling of longitudinal measurements and time to event data using normal/independent distributions: a Bayesian approach. Biom J. 2013 Nov;55(6):844–65.

163. Andrinopoulou E-R, Rizopoulos D, Jin R, Bogers AJJC, Lesaffre E, Takkenberg JJM. An introduction to mixed models and joint modeling: analysis of valve function over time. Ann Thorac Surg. 2012 Jun;93(6):1765–72.

164. Núñez J, Núñez E, Rizopoulos D, Miñana G, Bodí V, Bondanza L, et al. Red blood cell distribution width is longitudinally associated with mortality and anemia in heart failure patients. Circ J. 2014 Jan 24;78(2):410–8.

165. Philipson P, Sousa I, Diggle P. joineR - Joint modeling of repeated measurements and time-to-event data. 2012.

166. Tsiatis AA, Degruttola V, Wulfsohn MS. Modeling the Relationship of Survival to Longitudinal Data Measured with Error. Applications to Survival and CD4 Counts in Patients with AIDS. J Am Stat Assoc. Taylor & Francis Group; 1995 Feb 27;90:27–37.

167. Wagener DK, Sacks JM, LaPorte RE, Macgregor JM. The Pittsburgh study of insulin-dependent diabetes mellitus. Risk for diabetes among relatives of IDDM. Diabetes. 1982 Feb;31(2):136–44.

168. Pambianco G, Costacou T, Ellis D, Becker DJ, Klein R, Orchard TJ. The 30-year natural history of type 1 diabetes complications: the Pittsburgh Epidemiology of Diabetes Complications Study experience. Diabetes. Am Diabetes Assoc; 2006;55(5):1463–9.

169. Diabetes Epidemiology Research International Mortality Study Group. International evaluation of cause-specific mortality and IDDM. Diabetes Epidemiology Research International Mortality Study Group. Diabetes Care. 1991 Jan 1;14(1):55–60.

170. Bucolo G, David H. Quantitative determination of serum triglycerides by the use of enzymes. Clin Chem. 1973;19(5):476–82.

171. Allain CC, Poon LS, Chan CSG, Richmond W, Fu PC. Enzymatic determination of total serum cholesterol. Lipids. Am Assoc Clin Chem; 1974;20(4):470–5.

172. Warnick GR, Albers JJ. Heparin--Mn2+ quantitation of high-density-lipoprotein cholesterol: an ultrafiltration procedure for lipemic samples. Clin Chem. 1978;24(6):900–4.

173. Lipid Research Clinics Program: Manual of laboratory operations, Vol. 1: Lipid and liproprotein analysis. Washington, DC: National Institutes of Health, Department of Health, US Govt. Printing Office; 1974.

174. Friedewald WT, Levy RI, Fredrickson DS. Estimation of the concentration of low-density lipoprotein cholesterol in plasma, without use of the preparative ultracentrifuge. Clin Chem. Am Assoc Clin Chem; 1972;18(6):499–502.

175. The hypertension detection and follow-up program. Hypertension detetection and follow-up program cooperative group. Prev Med (Baltim). 1976;5:207–15.

176. Ellis D, Coonrod BA, Dorman JS, Kelsey SF, Becker DJ, Avner ED, et al. Choice of urine sample predictive of microalbuminuria in patients with insulin-dependent diabetes mellitus. Am J kidney Dis Off J Natl Kidney Found. 1989;13(4):321–8.

177. Dean A, Sullivan K, Soe M. OpenEpi: Open Source Epidemiologic Statistics for Public Health. 2015.

178. Patterson C, Dahlquist G, Harjutsalo V, Joner G, Feltbower R, Svensson J. Early mortality in EURODIAB population-based cohorts of type 1 diabetes diagnosed in childhood since 1989. Diabetologia. 2007;50:2439–42.

179. Collado-Mesa M, Diaz-Diaz O, Melian-Torres R, Surarez-Perez R, Vera-Gonzalez, M, Aldana-Padilla D. Mortality of childhood-onset IDDM patients: a cohort study in Havana City Province, Cuba. Diabetes Care. 1997;20:1237–41.

180. Secrest AM, Becker DJ, Kelsey SF, LaPorte RE, Orchard TJ. All-cause mortality trends in a large population-based cohort with long-standing childhood-onset type 1 diabetes: The Allegheny County type 1 diabetes registry. Diabetes Care. 2010;33(12):2573–9.

181. Petrie D, Lung TWC, Rawshani A, Palmer AJ, Svensson A-M, Eliasson B, et al. Recent trends in life expectancy for people with type 1 diabetes in Sweden. Diabetologia. 2016;(online be.

182. Harding JL, Shaw JE, Peeters A, Guiver T, Davidson S, Magliano DJ. Mortality trends among people with type 1 and type 2 diabetes in Australia: 1997-2010. Diabetes Care. 2014 Sep;37(9):2579–86.

183. Huxley R, Barzi F, Woodward M. Excess risk of fatal coronary heart disease is associated with diabetes in men and women: meta-analysis of 37 prospective cohort studies. BMJ. 2006;332:73–8.

184. Colhoun HM, Rubens MB, Underwood SR, Fuller JH. The effect of type 1 diabetes mellitus on the gender difference in coronary artery calcification. JAmCollCardiol. 2000;36(0735-1097 LA - eng PT - Journal Article):2160–7.

185. The National Collaborating Centre for Chronic Conditions. Type 1 Diabetes in Adults: National Clinical Guidelines for Diagnosis and Management in Primary and Secondary Care. London: Royal College of Physicians; 2004. p. 73–87.

186. Thanassoulis G, Williams K, Kimler Altobelli K, Pencina MJ, Cannon CP, Sniderman AD. Individualized Statin Benefit for Determining Statin Eligibility in the Primary Prevention of Cardiovascular Disease. Circulation. 2016 Mar 4;

187. The Diabetes Control and Complications Trial/Epidemiology of Diabetes Interventions and Complications (DCCT/EDIC) Research Group. Risk Factors for Cardiovascular Disease in Type 1 Diabetes. Diabetes. 2016;65:1370–9.

188. Miller RG, Secrest AM, Sharma RK, Songer TJ, Orchard TJ. Improvements in the life expectancy of type 1 diabetes: the Pittsburgh Epidemiology of Diabetes Complications study cohort. Diabetes. 2012 Nov;61(11):2987–92.

189. Lloyd CE, Becker D, Ellis D, Orchard TJ. Incidence of Complications in Insulin-dependent Diabetes Mellitus: A Survival Analysis. Am J Epidemiol. 1996 Mar 1;143(5):431–41.

190. Carmelli D, Halpern J, Swan GE, Dame A, McElroy M, Gelb AB, et al. 27-Year mortality in the Western Collaborative Group study: construction of risk groups by recursive partitioning. J Clin Epidemiol. 1991;44(12):1341–51.

191. Pambianco G. The 30-Year Natural History of Type 1 Diabetes Complications: The Pittsburgh Epidemiology of Diabetes Complications Study Experience. Diabetes. 2006 May 1;55(5):1463–9.

192. Rose G, Blackburn H. Cardiovascular Survey Methods: WHO Technical Manual No. 56. Generoseva: World Health Organization; 1968. 1-188 p.

193. Levey AS, Stevens LA, Schmid CH, Zhang YL, Castro AF, Feldman HI, et al. A new equation to estimate glomerular filtration rate. Ann Intern Med. 2009;150(9):604–12.

194. Group DC and CT of DI and C (DCCT/EDIC) R, Nathan DM, Zinman B, Cleary PA, Backlund J-Y, Genuth S, et al. Modern-Day Clinical Course of Type 1 Diabetes Mellitus After 30 Years' Duration. Arch Intern Med. 2009;169(14):1307–16.

195. Roy M, Rendas-Baum R, Skurnick J. Mortality in African-Americans with type 1 diabetes: the New Jersey 725. Diabet Med. 2006;23(6):698–706.

196. Orchard TJ, Secrest AM, Miller RG, Costacou T. In the absence of renal disease, 20 year mortality risk in type 1 diabetes is comparable to that of the general population: a report from the Pittsburgh Epidemiology of Diabetes Complications Study. Diabetologia. 2010 Nov;53(11):2312–9.

197. Borch-Johnsen K, Kreiner S. Proteinuria : value as a predictor of cardiovascular mortality in insulin dependent diabetes mellitus. Br Med J (Clin Res Ed). 1987;294:1651–4.

198. Secrest AM, Prince CT, Costacou T, Miller RG, Orchard TJ. Predictors of and survival after incident stroke in type 1 diabetes. Diab Vasc Dis Res. 2013 Jan;10(1):3–10.

199. Folsom AR, Rosamond WD, Shahar E, Cooper LS, Aleksic N, Nieto FJ, et al. Prospective Study of Markers of Hemostatic Function With Risk of Ischemic Stroke. Circulation. 1999 Aug 17;100(7):736–42.

200. Gillum RF, Ingram DD, Makuc DM. White Blood Cell Count and Stroke Incidence and Death: The NHANES I Epidemiologic Follow-up Study. Am J Epidemiol. 1994 May 1;139(9):894–902.

201. Margolis KL, Manson JE, Greenland P, Rodabough RJ, Bray PF, Safford M, et al. Leukocyte count as a predictor of cardiovascular events and mortality in postmenopausal women: the Women's Health Initiative Observational Study. Arch Intern Med. 2005;165(5):500–8.

202. Davis TME, Bruce DG, Davis WA. Predictors of first stroke in Type 1 diabetes: The Fremantle Diabetes Study. Diabet Med. 2005;22(5):551–3.

203. Fuller J, Stevens L, Wang S. Risk factors for cardiovascular mortality and morbidity: The WHO Multinational Study of Vascular Disease in Diabetes. Diabetologia. 2001;44(S2):S54–64.

204. Gerald R. Fowkes F, Housley E, Riemersma RA, Macintyre CCA, Cawood EHH, Prescott RJ, et al. Smoking, Lipids, Glucose Intolerance, and Blood Pressure as Risk Factors for Peripheral Atherosclerosis Compared with Ischemic Heart Disease in the Edinburgh Artery Study. Am J Epidemiol. 1992 Feb 15;135(4):331–40.

205. Newman AB, Siscovick DS, Manolio TA, Polak J, Fried LP, Borhani NO, et al. Ankle-arm index as a marker of atherosclerosis in the Cardiovascular Health Study. Cardiovascular Heart Study (CHS) Collaborative Research Group. Circulation. 1993 Sep 1;88(3):837–45.

206. Bowlin SJ, Medalie JH, Flocke SA, Zyzanski SJ, Goldbourt U. Epidemiology of Intermittent Claudication in Middle-aged Men. Am J Epidemiol. 1994 Sep 1;140(5):418–30.

207. Ingolfsson IO, Sigurdsson G, Sigvaldason H, Thorgeirsson G, Sigfusson N. A marked decline in the prevalence and incidence of intermittent claudication in Icelandic men 1968-1986: a strong relationship to smoking and serum cholesterol--the Reykjavik Study. J Clin Epidemiol. 1994 Nov;47(11):1237–43.

208. Murabito JM, D'Agostino RB, Silbershatz H, Wilson PWF. Intermittent Claudication : A Risk Profile From The Framingham Heart Study. Circulation. 1997 Jul 1;96(1):44–9.

209. Meijer WT, Grobbee DE, Hunink MG, Hofman A, Hoes AW. Determinants of peripheral arterial disease in the elderly: the Rotterdam study. Arch Intern Med. 2000 Oct 23;160(19):2934–8.

210. Joosten MM, Pai JK, Bertoia ML, Rimm EB, Spiegelman D, Mittleman MA, et al. Associations between conventional cardiovascular risk factors and risk of peripheral artery disease in men. JAMA. 2012 Oct 24;308(16):1660–7.

211. Diabetes TDCACTO, Group AC (DCCT/EDIC) SR. Intensive Diabetes Treatment and Cardiovascular Disease in Patients with Type 1 Diabetes. N Engl J Med. 2005;2643–53.

212. Farb A, Burke AP, Tang AL, Liang TY, Mannan P, Smialek J, et al. Coronary plaque erosion without rupture into a lipid core. A frequent cause of coronary thrombosis in sudden coronary death. Circulation. 1996 Apr 1;93(7):1354–63.

213. Anastasios A. Tsiatis MD. Joint modeling of longitudinal and time-to-event data: an overview. Stat Sin. 2004;14:809–34.

214. Diggle P, Sousa I, Chetwynd AG. Joint modelling of repeated measurements and time-to-event outcomes: The fouth Armitage lecture. Stat Med. 2008;27:2981–98.

215. Orchard TJ, Olsen JC, Erbey JR, Williams K, Forrest KY-Z, Smithline Kinder L, et al. Insulin Resistance – Related Factors , but not Glycemia , Predict Coronary Artery Disease in Type 1 Diabetes. Diabetes Care. 2003;26(5):1374–9.

216. The effect of intensive treatment of diabetes on the development and progression of long-term complications in insulin-dependent diabetes mellitus. The Diabetes Control and Complications Trial Research Group. N Engl J Med. 1993 Sep 30;329(14):977–86.

217. Carroll M, Kit B, Lacher D, Shero S, Mussolino M. Trends in lipids and lipoproteins in US adults, 1988-2010. JAMA. 2012;308(15):1545–54.

218. Matsushima M, LaPorte RE, Maruyama M, Shimizu K, Nishimura R, Tajima N. Geographic variation in mortality among individuals with youth-onset diabetes mellitus across the world. DERI Mortality Study Group. Diabetes Epidemiology Research International. Diabetologia. 1997 Feb;40(2):212–6.

219. Jellinger PS, Smith DA, Mehta AE, Ganda O, Handelsman Y, Rodbard HW, et al. American Association of Clinical Endocrinologists' Guidelines for Clinical Management of Dyslipidemia and Prevention of Atherosclerosis. Endocr Pract. 2012;18(Suppl 1):1–78.

220. Orchard TJ, Forrest KY, Kuller LH, Becker DJ, Pittsburgh Epidemiology of Diabetes Complications Study. Lipid and blood pressure treatment goals for type 1 diabetes: 10-year incidence data from the Pittsburgh Epidemiology of Diabetes Complications Study. Diabetes Care. American Diabetes Association; 2001 Jun;24(6):1053–9.

221. National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III). Third Report of the National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III) final report. Circulation. American Heart Association Journals; 2002 Dec 17;106(25):3143–421.

222. Orchard T, Forrest K, Ellis D, Becker D. Cumulative glycemic exposure and microvascular complications in insulindependent diabetes mellitus. Arch Intern Med. 1997;157:1851–6.

223. Larkin ME, Backlund J-Y, Cleary P, Bayless M, Schaefer B, Canady J, et al. Disparity in management of diabetes and coronary heart disease risk factors by sex in DCCT/EDIC. Diabet Med. 2010 Apr;27(4):451–8.

224. Tonstad S, Rosvold E, Furu K, Skurtveit S. Undertreatment and overtreatment with statins: the Oslo Health Study 2000-2001. J Intern Med. 2004;255:494–502.

225. Wexler D, Grant R, Meigs J, Nathan D, Cagliero E. Sex disparities in treatment of cardiac risk factors in patients with type 2 diabetes. Diabetes Care. 2005;28:514–20.

226. Cull C, Neil H, Holman R. Changing aspirin use in patients with type 2 diabetes in the UKPDS. Diabet Med. 2004;21:1368–71.

227. Persell S, Baker D. Aspirin use among adults with diabetes. Arch Intern Med. 2004;164:2492–9.

228. Costacou T, Orchard T. Differential effect of glycemia on the incidence of hypertension by sex: The Epidemiology of Diabetes Complications study. Diabetes Care. 2013;36(1):77–83.

229. Costacou T, Evans R, Orchard T. High density lipoprotein cholesterol in diabetes: Is higher always better? J Clin Lipidol. 2011;5(5):387–94.

230. Miller RG, Secrest AM, Ellis D, Becker DJ, Orchard TJ. Changing impact of modifiable risk factors on the incidence of major outcomes of type 1 diabetes: the Pittsburgh Epidemiology of Diabetes Complications Study. Diabetes Care. 2013 Dec;36(12):3999–4006.

231. Rouanet A, Joly P, Dartigues J-F, Proust-Lima C, Jacqmin-Gadda H. Joint latent class model for longitudinal data and interval-censored semi-competing events: Application to dementia. Biometrics. 2016 Apr 28;e-publish before print.