

# Characterization of direct and/or indirect genetic associations for multiple traits in longitudinal studies of disease progression

Myriam Brossard<sup>1</sup>, Andrew D. Paterson<sup>2,3</sup>, Osvaldo Espin-Garcia<sup>3,4</sup>, Radu V. Craiu<sup>5</sup>, Shelley B. Bull<sup>1,3,\*</sup>

<sup>1</sup>Lunenfeld-Tanenbaum Research Institute, Sinai Health System, Toronto, ON, Canada;

<sup>2</sup>Program in Genetics and Genome Biology, Hospital for Sick Children Research Institute, Toronto, ON, Canada;

<sup>3</sup>Division of Biostatistics, Dalla Lana School of Public Health, University of Toronto, Toronto, ON, Canada;

<sup>4</sup>Department of Biostatistics, Princess Margaret Cancer Centre, Toronto, ON, Canada;

<sup>5</sup>Department of Statistical Sciences, University of Toronto, Toronto, ON, Canada.

**\*Corresponding authors:** Lunenfeld-Tanenbaum Research Institute, 60 Murray Street, Box #18, M5T 3L9, Toronto, ON, Canada. E-mails: [bull@lunenfeld.ca](mailto:bull@lunenfeld.ca), [brossard@lunenfeld.ca](mailto:brossard@lunenfeld.ca).

## Contents

Supplementary Information SI-1. DCCT-based simulation study .....	2
Assumed underlying system model .....	2
Procedure for data generation under the <i>causal</i> genetic scenario specified in Figures 2 and 3 .....	3
DCCT datasets used for the simulation study and specified parameters .....	5
Procedure to simulate SNPs under the <i>global</i> null genetic scenario .....	5
Supplementary Information SI-2. Analysis of the DCCT Genetics Study data .....	6
The Diabetes Control and Complications Trial (DCCT) data .....	6
Selection of the 307 candidate SNPs from the literature .....	7
Analysis of the DCCT data .....	7
Association structures for HbA1c .....	7
Supplementary Information SI-3. Notes on a multi-trait SNP association test for SNP effects estimated under the proposed joint model framework .....	9
Multi-trait SNP association test .....	9
Performances of the multi-trait SNP association test for SNP discovery .....	9
Application of the multi-trait SNP association test in DCCT .....	10
List of Supplementary Tables that are provided in a separate document .....	13
List of Supplementary Figures that are provided in a separate document .....	14
List of the participants of the DCCT/EDIC Research Group (as of January 1, 2021) .....	15
References .....	18

## **Supplementary Information SI-1. DCCT-based simulation study**

The algorithm we developed is designed to mimic the structure of the DCCT data. As described in the paper and presented in **Figure 2**, we simulate  $M=5$  causal SNPs and  $K=2$  non-independent time-to-event traits based on  $L=3$  longitudinal traits:

- two quantitative traits as measured in  $N=667$  DCCT individuals (HbA1C and SBP) to preserve *realistic* dependencies between/within longitudinal traits
- one quantitative trait ( $U$ ) simulated to induce *unexplained dependency* between the time-to-event traits.

We describe in detail the procedure followed to generate such DCCT-based simulated dataset. Example R script is available on GitHub (<https://github.com/brossardMyriam/Joint-model-for-multiple-trait-genetics>).

### **Assumed underlying system model**

#### **(i) Linear mixed models**

##### **Observed HbA1c ( $l=1$ )**

$$Y_{i(1)}(t_{ij}) = \beta_{0(1)} + b_{i0(1)} + (\beta_{1(1)} + b_{i1(1)})t_{ij} + \beta_{g(1)}SNP1_i + \varepsilon_{ij(1)} \\ = Y_{i(1)}^*(t_{ij}) + \varepsilon_{ij(1)}$$

- $b_{i(1)} = (b_{i0(1)}, b_{i1(1)}) \sim N_2(0, D_{(1)})$  is the vector of random effects for HbA1c and  $D_{(1)}$  is the *specified*  $2 \times 2$  positive definite covariance matrix.
- $\beta_{(1)} = (\beta_{0(1)}, \beta_{1(1)}, \beta_{g(1)})$  are *specified* fixed parameter values for the intercept, time and SNP effects on HbA1c.
- $\varepsilon_{ij(1)} \sim N(0, \sigma_{(1)}^2)$  is the error term, with *specified* variance  $\sigma_{(1)}^2$ .

##### **Observed SBP ( $l=2$ )**

$$Y_{i(2)}(t_{ij}) = \beta_{0(2)} + b_{i0(2)} + (\beta_{1(2)} + b_{i1(2)})t_{ij} + \beta_{g(2)}SNP5_i + \beta_{h(2)}Sex_i + \varepsilon_{ij(2)} \\ = Y_{i(2)}^*(t_{ij}) + \varepsilon_{ij(2)}$$

- $b_{i(2)} = (b_{i0(2)}, b_{i1(2)}) \sim N_2(0, D_{(2)})$  is the vector of random effects for SBP and  $D_{(2)}$  is the *specified*  $2 \times 2$  positive definite covariance matrix.
- $\beta_{(2)} = (\beta_{0(2)}, \beta_{1(2)}, \beta_{g(2)}, \beta_{h(2)})$  are the *specified* fixed parameter values for the intercept, time, SNP and sex effects on SBP.
- $\varepsilon_{ij(2)} \sim N(0, \sigma_{(2)}^2)$  denotes the error term, with *specified* variance  $\sigma_{(2)}^2$ .

##### **Simulated latent risk factor ( $U$ )**

$$U_i(t_{ij}) = \beta_{0(U)} + b_{i0(U)} + (\beta_{1(U)} + b_{i1(U)})t_{ij} + \beta_{g(U)}SNP3_i + \varepsilon_{ij(U)} \\ = U_i^*(t_{ij}) + \varepsilon_{ij(U)}$$

- $b_{i(U)} = (b_{i0(U)}, b_{i1(U)}) \sim N_2(0, D_{(U)})$  is the vector of random effects for  $U$  and  $D_{(U)}$  is the *specified*  $2 \times 2$  positive definite covariance matrix.
- $\beta_{(U)} = (\beta_{0(U)}, \beta_{1(U)}, \beta_{g(U)})$  are *specified* fixed parameter values for the intercept, the time and the SNP effects on  $U$ .
- $\varepsilon_{ij(U)} \sim N(0, \sigma_{(U)}^2)$  is the error term, with *specified* variance  $\sigma_{(U)}^2$ .

**(ii) PH time-to-event models**

*Simulated Time-to-DR (k=1)*

$$\lambda_{i(1)}(t) = \lambda_{0(1)}(t) \times \exp\{\eta_{i(1)} + \alpha_{U(1)}U_i^*(t)\}$$

With:

- $\lambda_{0(1)}(t) = \varphi_{(1)}\xi_{(1)}(\xi_{(1)}t)^{\varphi_{(1)}-1}$ , where  $\varphi_{(1)}$  and  $\xi_{(1)}$  are the *specified* shape and scale parameters of the *specified* Weibull model.
- $\eta_{i(1)} = \gamma_{g(1)}SNP2_i + \alpha_{1(1)}Y_{i(1)}^*(t) + \gamma_{v(1)}T1D\_duration_{i(1)}$  with  $\gamma_{g(1)}$ ,  $\alpha_{1(1)}$  and  $\gamma_{v(1)}$  the *specified* fixed effects of SNP2,  $Y_{i(1)}^*(t)$  and T1D duration on DR.
- $\alpha_{U(1)}$  is the *specified* effect of the shared longitudinal risk factor  $U$  used to induce unexplained dependency between the simulated time-to-T1DC traits.

*Simulated Time-to-DN (k=2)*

$$\lambda_{i(2)}(t) = \lambda_{0(2)}(t) \times \exp\{\eta_{i(2)} + \alpha_{U(2)}U_i^*(t)\}$$

With:

- $\lambda_{0(2)}(t) = \varphi_{(2)}\xi_{(2)}(\xi_{(2)}t)^{\varphi_{(2)}-1}$ , where  $\varphi_{(2)}$  and  $\xi_{(2)}$  the *specified* shape and scale parameters of the *specified* Weibull model.
- $\eta_{i(2)} = \gamma_{g(2)}SNP4_i + \gamma'_{g(2)}SNP5_i + \sum_{l=1}^2 \alpha_{l(1)}Y_{i(l)}^*(t) + \gamma_{v(2)}T1D\_duration_{i(2)}$  with  $\gamma_{g(2)}$ ,  $\gamma'_{g(2)}$ ,  $\alpha_{1(1)}$ ,  $\alpha_{1(2)}$  and  $\gamma_{v(2)}$  are the *specified* fixed effects of SNP2,  $Y_{i(l)}^*(t)$  and T1D duration on DR.
- $\alpha_{U(2)}$  is the *specified* effect of the shared longitudinal risk factor  $U$  used to induce unexplained dependency between the simulated time-to-T1DC traits.

**Procedure for data generation under the causal genetic scenario specified in Figures 2 and 3**

Given the specified parameter values for each longitudinal sub-model:  $\Omega_{(1)} = (\beta'_{(1)}, D_{(1)}, \sigma_{(1)}^2)$ ,  $\Omega_{(2)} = (\beta'_{(2)}, D_{(2)}, \sigma_{(2)}^2)$  and  $\Omega_{(U)} = (\beta'_{(U)}, D_{(U)}, \sigma_{(U)}^2)$ ; and each time-to-event sub-model:  $\Gamma_k = (\varphi_{(k)}, \xi_{(k)}, \gamma_{g(k)}, \alpha_{1(k)}, \alpha_{2(k)}, \gamma_{v(k)}, \alpha_{u(k)})$ , as well as specified minor allele frequencies vector  $p = (p_1, \dots, p_m, \dots, p_M)$  for the  $M$  causal SNPs, the proposed simulation procedure proceeds in three steps. In **Step 1**, we first simulate the longitudinal values for  $U$  for the  $N$  DCCT individuals. In **Step 2**, we simulate the  $R$  replicates of  $M$  genotype data based on observed/simulated longitudinal traits for DCCT individuals. In **Step 3**, we simulate  $R$  replicates of non-independent time-to-event traits for all  $N$  DCCT individuals using SNPs and the trajectories for *observed* (HbA1c, SBP) and *simulated* ( $U$ ) longitudinal risk factors from *Steps 1 & 2* data.

➤ **Step 1: Simulation of  $U$**

For each individual  $i$ , we generate  $U_i \sim N_{n_i}(X_{i(U)}\beta_{(U)}, Z_{i(U)}D_{(U)}Z'_{i(U)} + \sigma_{(U)}^2I_{n_i})$  where  $X_{i(U)} = (1, t_i)$  and  $Z_{i(U)} = (1, t_i)$  are the *specified* design matrices for the fixed and random effects and  $\Omega_{(U)} = (\beta'_{(U)}, D_{(U)}, \sigma_{(U)}^2)$  are the *specified* parameter values.

➤ **Step 2: Simulation of the genetic data**

*Causal SNPs with indirect effects*

We simulate genotypes for SNP1, SNP3 and SNP5 (with MAF vector  $p_m$ ,  $m \in \{1, 3, 5\}$ ) for each individual  $i$  ( $1 \leq i \leq N$ ) and each replicate  $r$  ( $1 \leq r \leq R$ ) under Hardy-Weinberg and linkage equilibrium assumptions. Particularly, for each SNP $m$ , we first draw genotype  $g_i$  for each individual  $i$  from the genotype vector  $g = (0, 1, 2)$  under a multinomial distribution with conditional genotype probabilities  $(\pi_{0i}, \pi_{1i}, \pi_{2i})$ , where each  $\pi_{gi}$  is calculated given the following:

$$\begin{aligned} \pi_{gi} &= P(SNP_i = g_i | \psi_i) = \frac{P(\psi_i, SNP_i = g_i)}{P(\psi_i)} \\ &= \frac{P(\psi_i | SNP_i = g_i) P(SNP_i = g_i)}{P(\psi_i)} \\ &= \frac{P(\psi_i | SNP_i = g_i) P(SNP_i = g_i)}{\sum_{g=0}^2 P(\psi_i | SNP_i = g) P(SNP_i = g)} \end{aligned}$$

With:

- $\psi_i = \begin{cases} (Y_{i(1)} | SNP_{1i} = g_i, t_i, \Omega_{(1)}) & \text{for SNP1} \\ (Y_{i(2)} | SNP_{2i} = g_i, t_i, \Omega_{(2)}) & \text{for SNP2} \\ (U_i | SNP_{3i} = g_i, t_i, \Omega_{(U)}) & \text{for SNP3} \end{cases}$
- $P(\psi_i(t_i) | SNP_i = g_i)$  denotes the probability density function of the multivariate distribution for  $\psi_i$ , with  $\psi_i \sim N_{n_i}(X_i \beta, Z_i D Z_i' + \sigma^2 I_{n_i})$  under the *specified* underlying linear mixed model. Here,  $X_i$  and  $Z_i$  denote the design matrices of the corresponding fixed and random effects.  $D$  is the covariance matrix of the random effects and  $\sigma^2$ , the variance of the error term of the corresponding longitudinal trait  $\psi_i \in \{Y_{i(l)}, U_i\}$ .
- *Unconditional* genotype probability  $P(SNP_i = g)$  is specified by allele frequency of SNP $m$  ( $p_m$ ) under Hardy-Weinberg equilibrium, with  $P(SNP_i = 0) = (1 - p_m)^2$ ,  $P(SNP_i = 1) = 2p_m(1 - p_m)$  and  $P(SNP_i = 2) = p_m^2$ .

*Causal SNPs with direct effects*

For SNP2 ( $p_2$ ) and SNP4 ( $p_4$ ) with *direct* effects on time-to-event trait  $k$ , we simulate genotypes for each individual  $i$  by drawing  $R$  genotypes from the multinomial distribution with genotype probabilities  $(\pi_0, \pi_1, \pi_2)$ , where under the Hardy-Weinberg Equilibrium assumption  $\pi_0 = (1 - p_m)^2$ ,  $\pi_1 = 2p_m(1 - p_m)$  and  $\pi_2 = p_m^2$ . As opposed to the simulation of the SNPs with indirect effects, the simulation of SNPs with direct effects depend on population probabilities that do not vary among individuals.

➤ **Step 3: Simulation of the non-independent time-to-event traits**

The  $K=2$  *non-independent* time-to-T1DC traits  $(T_{i(k)}, \delta_{i(k)})$ , with  $k=1, 2$ ) are generated using the specified hazards functions  $\lambda_{i(k)}(t)$  that share  $U_i^*(t)$  as an unmeasured longitudinal risk factor used to induce unexplained dependency. Specifically, for each individual  $i$  of each data replicate  $r$  from Stage 1:

1. We sample a censoring time  $C_i$  from a Uniform distribution on  $[0, t_{n_i}]$ , where  $t_{n_i}$  denotes the time where the last measurement is collected for individual  $i$ .
2. For each type of event  $k$  with specified parameters  $\Gamma_{(k)}$ :

- a. We compute the *uncensored* event times  $T_{i(k)}^*$  by calculating the inverse of the cumulative of the hazard function using the Brent univariate root-finding method<sup>1,2</sup>.
  - b. We generate the *right-censored* event time  $T_{i(k)}$  and event *indicator*  $\delta_{i(k)}$ , as  $T_{i(k)} = \min(T_{i(k)}^*, C_i)$  and set  $\delta_{i(k)}$  to 1 if  $T_{i(k)}$  occurs before or at  $C_i$  or 0 otherwise ( $\delta_{i(k)} = I(T_{i(k)}^* \leq C_i)$ ).
3. Repeat 1 and 2  $N$  times.
  4. Repeat 1 to 3  $R$  times.

### DCCT datasets used for the simulation study and specified parameters

We use  $N=667$  individuals from the DCCT Conventional treatment group with longitudinal HbA1c ( $Y_{i(1)}$ ) and SBP ( $Y_{i(2)}$ ) values as well as baseline covariates (sex and T1D duration) measured for each individual  $i$ . We exclude all longitudinal values prior to the 6 months visit due to a mild non-linear time trend observed for HbA1c (screening effect) at the beginning of the DCCT study. For a few individuals, we replace missing measures for either HbA1c or SBP by their predicted values from the linear mixed model ignoring the SNP variable. We use mean centered HbA1c and SBP measures.

Specified parameter values  $\Omega_{(1)}$ ,  $\Omega_{(2)}$ ,  $\Omega_{(U)}$  for the longitudinal models and  $\Gamma_{(1)}$ ,  $\Gamma_{(2)}$  for the time-to-event traits models, as well as for the MAFs of the  $M$  SNPs ( $p$ ) for the causal scenario of **Figure 2** are summarized in **Table 2** and in **Table S1**. Particularly, we specify SNP effects in  $\Omega_{(1)}$ ,  $\Omega_{(2)}$ ,  $\Gamma_{(1)}$ ,  $\Gamma_{(2)}$  and MAFs according to the DCCT Genetics Study and/or T1DC literature. For the non-SNP parameter values in  $\Omega_{(1)}$ ,  $\Omega_{(2)}$ , we use the estimates of the linear mixed models fitted without the SNP in DCCT individuals. We specify non-SNP parameter values in  $\Omega_{(U)} = (\beta'_{(U)}, D_{(U)}, \sigma^2_{(U)})$  and  $\Gamma_{(2)}$ . Finally, we specify the shape and scale parameters in  $\Gamma_{(1)}$  and  $\Gamma_{(2)}$  to generate ~54% DR and ~25% DN events.

### Procedure to simulate SNPs under the *global* null genetic scenario

For the genetic scenario under the global null hypothesis, where none of the SNPs is associated with any traits, we simulate  $M$  SNPs with the exact same MAF vector as for the causal SNPs from **Figures 2**. These SNPs are simulated independently of the longitudinal and time-to-event traits with the same procedure as previously described to generate the SNPs with direct effects.

## **Supplementary Information SI-2. Analysis of the DCCT Genetics Study data**

### **The Diabetes Control and Complications Trial (DCCT) data**

The DCCT data was a randomized-controlled trial that demonstrated intensive insulin therapy could prevent and/or delay progression of long-term T1DC.<sup>3</sup> A total of 1,441 T1D patients were recruited between 1983 and 1993 in two cohorts: a primary prevention cohort (726 patients with a short duration of diabetes (1-5 years), exhibiting normal albuminuria (Albumin Excretion Rate (AER) <40 mg/day) and without evidence of retinopathy); a secondary intervention cohort (715 patients with diabetes duration of 1-15 years, with AER<200 mg/day and non-proliferative retinopathy). Members of each cohort were randomly assigned to receive either conventional or intensive treatment and were followed over scheduled visits for the development of complications. Due to significant outcome differences between intensive and conventional treatment groups, the DCCT was stopped prematurely in 1993 (administrative censoring). At the DCCT closeout visit, patients had a mean follow-up time of 6.5 years (range 3 to 9), 99% of the patients completed the study and more than 95% of all scheduled examinations were completed.<sup>3</sup>

Genome-wide genotyping in DCCT subjects was performed subsequently using Illumina 1M and HumanCoreExome Bead Arrays (Illumina, San Diego, CA, USA) and standard quality controls procedures were applied to individuals and genetic markers<sup>4,5</sup>. In this study, we use genotyped individuals with the HumanCoreExome Bead Array with ungenotyped autosomal SNPs imputed using 1000 Genomes<sup>6</sup> data phase 3 (v5) and minimac3<sup>7</sup> (v.1.0.13), as previously described<sup>5</sup>. Largely because of differences in HbA1c between treatment groups since the goal of intensive therapy was to reduce HbA1c in non-diabetic range, our study focuses on  $N=667$  unrelated individuals of European descent ancestry from the conventional treatment group. In the conventional treatment group, longitudinal measurements for HbA1c and SBP were both collected at up to 39 quarterly visits during DCCT follow-up while DR and DN events were collected at annual and semi-annual visits respectively.

We use time to mild DR and time to persistent microalbuminuria, for DR and DN outcomes respectively, as previously defined in the motivating GWAS of HbA1c<sup>4</sup> (see **Table SI-2b**).

**Table SI-2b. Time-to-event traits analyzed in the DCCT Genetics Study**

Time-to-event outcome	Name	Outcome definition
DR	Time to mild DR	Time from DCCT baseline to mild non-proliferative diabetic retinopathy (EDTRS step 4, patient level 35/<35)
DN	Time to persistent microalbuminuria	Time from DCCT baseline to the first of two consecutive visits with Albumin Excretion Rate >30 mg/day

Out of the 667 DCCT individuals, we analyze  $N=516$  subjects with genotype data, without mild retinopathy at DCCT baseline (or prior) or without DN event at DCCT baseline. By the time of the DCCT close-out visit, 297 (57.6%) experienced a DR event, 61 (11.8%) a DN event, including 47 subjects (9.1%) that experienced both events.

## Selection of the 307 candidate SNPs from the literature

We identified 322 independent SNPs ( $r^2 < 0.8$ ,  $MAF \geq 5\%$  in European ancestry individuals) associated with HbA1c, SBP, DR and/or DN outcomes as reported in the literature<sup>4,9-14</sup>. For HbA1c, DR or DN we selected SNPs reported at the suggestive significance level of  $P^* = 10^{-6}$  by GWAS in T1D individuals or SNP associations reported by GWAS conducted in T2D individuals or general populations and confirmed in T1D individuals of European ancestry at the nominal significance level<sup>4,9-14</sup>. For the SBP SNP list, we selected surrogate SNPs reported in the largest meta-analysis conducted in the general population of European ancestry at the conventional genome-wide significance level  $P^* = 5 \times 10^{-8}$  because large-scale GWAS results of SBP in individuals with diabetes were lacking at the time of our analyses. At total, 307 biallelic SNPs with imputation quality score  $R^2 \geq 0.50$  and  $MAF \geq 5\%$ , and after pruning on linkage disequilibrium ( $r^2 < 0.8$ ) using LDlink<sup>8</sup> in 1000 Genomes phase 3 European-ancestry population were analyzed in DCCT (See **Table S8** for a full list of the SNPs).

## Analysis of the DCCT data

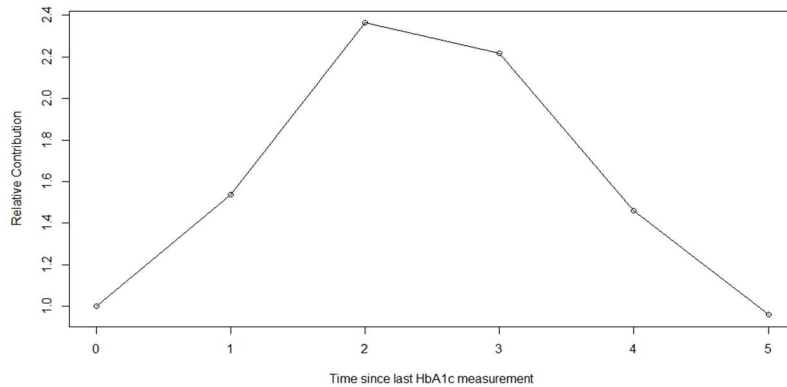
We fit JM, CM-obs for each SNP, one at a time, including baseline covariates (age at diagnosis, T1D duration, cohort, gender, and year of entry in the DCCT study). For the covariate year of entry in DCCT, we group the patients into four consecutive strata with homogeneous number of individuals (*i.e.* 1983-1984, 1985-1986, 1987, using 1988-1989 as the reference category). In Stage 1 of JM, bivariate longitudinal mixed-effects models for HbA1c and SBP are fitted using all available measures at quarterly visits from DCCT baseline to the close-out visit; HbA1c and SBP trajectories are fitted for each individual. In Stage 2 of JM, trajectory values interpolated to the start time of each risk interval are then used as time-varying covariates in the Cox PH frailty model. Time-to-event sub-models are fitted using annual records of DR and DN events. Because the DR and DN outcomes were assessed with different frequency of visits in DCCT (semi-annually for DR and annually for DN), we assign each DR event to the one-year interval visit that included the observed time-to-event. We use  $B=500$  bootstraps of DCCT individuals to compute empirical variance-covariance matrices for parameters estimated by the joint model and CM-obs.

## Association structures for HbA1c

Given established cumulative effects of HbA1c on T1DC traits<sup>15,16</sup>, we compare joint model results obtained with contemporaneous HbA1c value with joint model results obtained using time-weighted cumulative and updated cumulative mean HbA1c effects on T1DC<sup>15,16</sup>. Under each cumulative association profiles, the time-to-event sub-models are fitted by substituting the fitted trajectory of the HbA1c by a summary function of the prior fitted values from DCCT baseline up to the beginning of each risk interval (**Table SI-2**). While the updated mean association structure assumes an equal weighting for all fitted HbA1c measures at prior visits from baseline, the time-weighted cumulative HbA1c effect association structure assumes different weights for each visit. Here, we use a time-weighted formulation that considers all HbA1c values from DCCT baseline up to 5 years prior to the start of each risk interval following previous DCCT data analysis<sup>15</sup>. We extracted weights from this study<sup>15</sup> (**Figure SI-2**) and recalibrated such that the sum of the weights is equal to 1 for each risk interval (**Table SI-2**).

**Table SI-2. Three alternative association structures for HbA1c effects on T1DC traits in DCCT**

Parametrization	Time-dependent association structure in the two-stage approach
Contemporaneous (current value)	$f_{1(k)}(\widehat{Y}_{i(1)}^*(t_{ij})) = \widehat{Y}_{i(1)}^*(t_{ij})$
Updated cumulative mean	$f_{1(k)}(\widehat{Y}_{i(1)}^*(t_{ij})) = \frac{1}{j} \sum_{k=1}^j \widehat{Y}_{i(1)}^*(t_{ij})$
Time-weighted cumulative	$f_{1(k)}(\widehat{Y}_{i(1)}^*(t_{ij})) = \begin{cases} \sum_{k=1}^j w_k \widehat{Y}_{i(1)}^*(t_{i(j-k)}) & \text{if } t_{ij} < 5 \\ \sum_{k=1}^K w_k \widehat{Y}_{i(1)}^*(t_{i(j-k)}) & \text{if } t_{ij} \geq 5 \end{cases}$ <p>Here, we use <math>K=5</math> to account for the fitted HbA1c values <i>up</i> to 5 years prior to the current time point <math>t_{ij}</math>, with weights <math>w_k</math> based on Lind et al<sup>15</sup> (see <b>Figure SI-2</b>).</p>



**Figure SI-2. Relative contribution of past HbA1c measures to the risk of mild retinopathy at current time point, based on Lind et al<sup>15</sup>. The X axis represents the time (in year) since the current HbA1c measurement.**



### **Supplementary Information SI-3. Notes on a multi-trait SNP association test for SNP effects estimated under the proposed joint model framework**

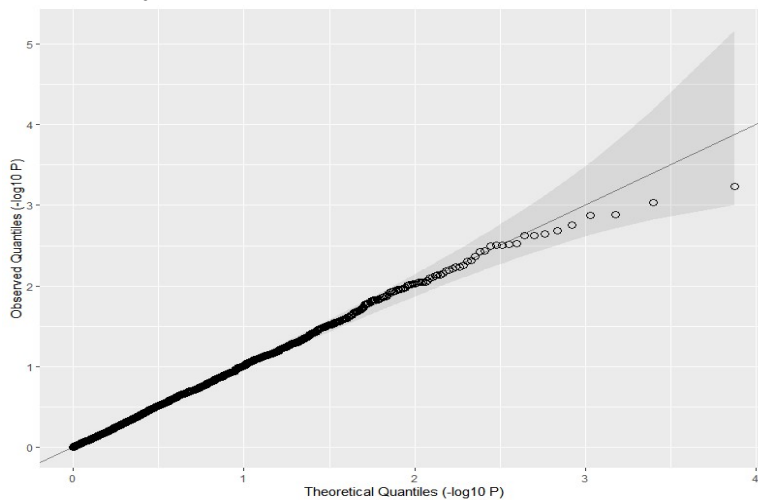
Although our primary aim is to develop inference methods to distinguish among direct and/or indirect SNP associations with each time-to-event trait, the multi-trait aspect of the joint model also lends itself to multi-trait SNP association testing for SNP discovery. In this section, we present a multi-trait SNP association test derived from the parameters of the joint model and report test performance in the DCCT-based simulated data and results in application to DCCT study.

#### **Multi-trait SNP association test**

The proposed multi-trait SNP association test, assesses if at *least* one of the SNP has a non-null effect on any of the  $L+K$  traits (*ie* at least one  $\beta_{g(l)}$  or  $\gamma_{g(k)} \neq 0$ ,  $1 \leq l \leq L$  and  $k \leq l \leq K$ ) against the *global null* hypothesis that all the SNPs have a null effect on the traits (*ie* all  $L+K$  traits  $\beta_{g(l)} = \gamma_{g(k)} = 0$ , with  $1 \leq l \leq L$  and  $k \leq l \leq K$ ). This test is a generalized Wald statistic constructed from the SNP effects vector  $\varphi_g = (\beta_g, \gamma_g)$  and its bootstrap covariance matrix  $\Sigma_g$ ,  $W_{\varphi_g} = \widehat{\varphi}_g^T \Sigma_g^{-1} \widehat{\varphi}_g$  which is assumed to asymptotically follows a  $\chi^2$  distribution with  $L+K$  *df*.

#### **Performances of the multi-trait SNP association test for SNP discovery**

We evaluate the statistical performance (power, type I error) of the multi-trait SNP association test from the joint models (JM-cmp, JM-mis), in comparison to the results from marginal linear mixed models and Cox PH models (referred as MM). Under the *global null* hypothesis, the multi-trait SNP association test does not show any marked departure from its expected distribution ( $\chi^2$  with four degrees of freedom), as shown in the following Quantile-Quantile plot (**Figure SI-3a**). Furthermore, type I error under the *global null* genetic scenario appears well controlled, respectively: 0.048 at  $P^* = 0.05$  and 0.0097 at  $P^* = 0.01$ .

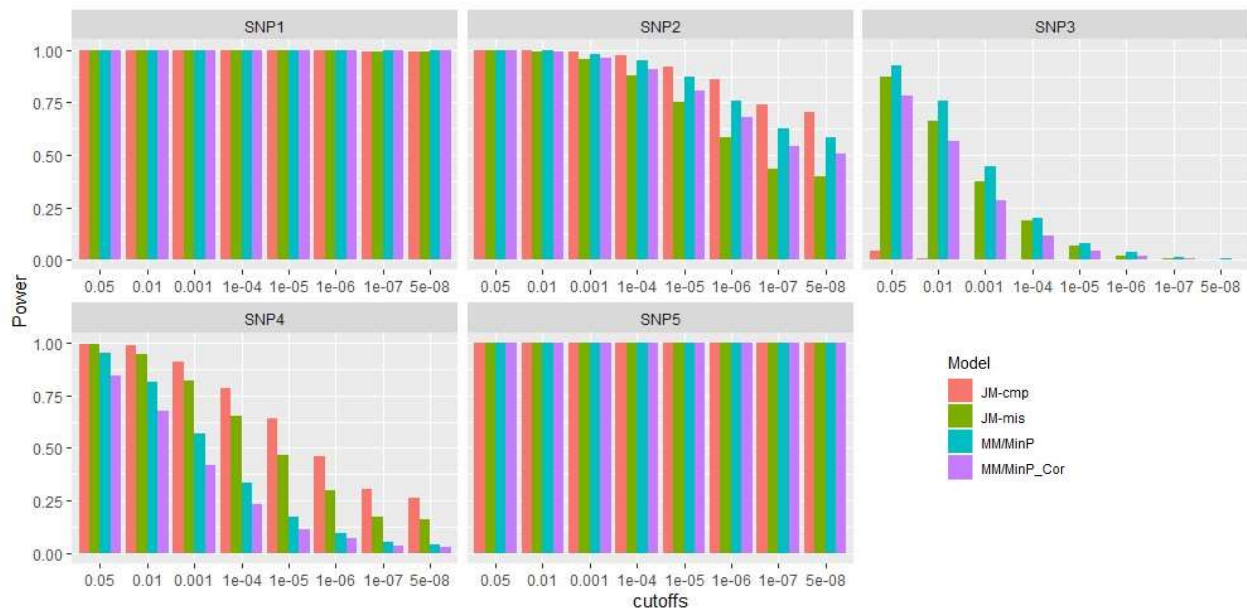


**Figure SI-3a. Quantile-Quantile plot of the multi-trait SNP association test p values (JM-cmp), assessed under the *global null* genetic scenario. P-values for all SNPs, were pooled, yielding to  $R=5000$  replicates.**

Under the *causal* genetic scenario, when the JM is fully specified (JM-cmp), as shown in **Figure SI-3b**, the power of the proposed multi-trait SNP association test appears higher (or similar) than the power based on the minimum *P-value* from MM in all scenarios of SNP associations (MM\_MinP uncorrected and MM\_MinP\_cor corrected for the four traits analyzed separately assuming that the traits are independent, which is unrealistic but represents an extreme case of conservative *P-value*). When the joint model is mis-specified (JM-mis), the multi-trait SNP

association test is still as or more powerful than the MM\_MinP for all causal SNPs except for SNP2 where the power reduction is more pronounced at more stringent significant levels ( $P^* \leq 10^{-5}$ ). Multi-trait association results for SNP3 also show attenuated power compared to the uncorrected MM\_MinP; but its power remain higher than the power of the conservative MM\_MinP\_cor. Except for SNP5 that has both direct and indirect effects on time-to-DN, the other simulated causal SNPs have either a *single* direct or a *single* indirect effect on the time-to-event traits; the latter do not represent scenarios of SNP associations where we would expect large power improvement from the proposed multi-trait SNP association test (largest power improvement would be expected for SNPs with effects on multiple traits).

**Figure SI-3b. Power of the multi-trait SNP association tests ( $P_{mult}$ ) under the joint models (JM-cmp, JM-mis) to detect a SNP association, compared to the minimum SNP  $P$ -value from the marginal models (MM) fitted separately for each trait corrected and uncorrected for the number of traits tested (Min\_P\_Cor and MinP). Power is assessed using  $R=1000$  replicates of  $N=667$  DCCT subjects simulated under the *causal* genetic scenario from Figure 2.**



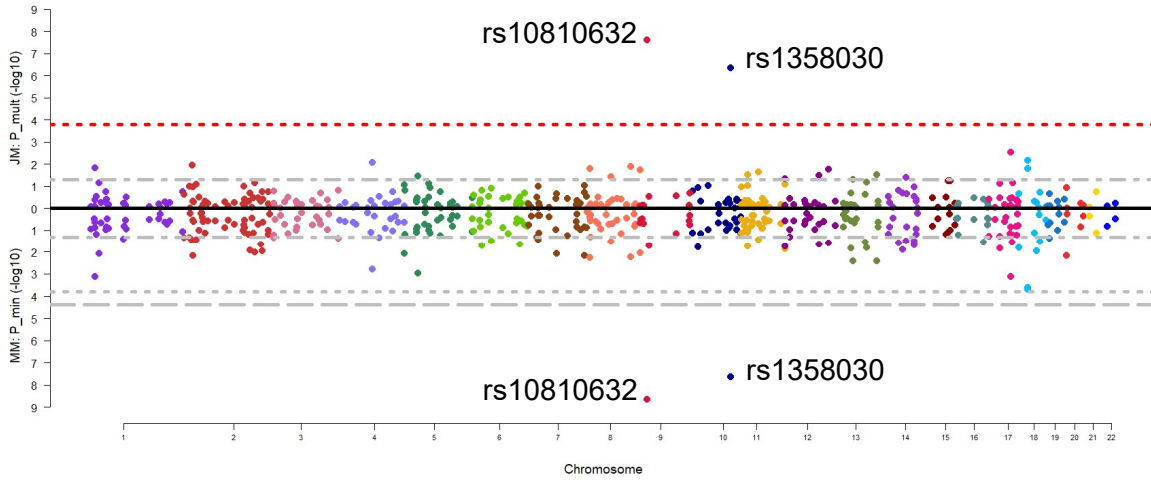
### Application of the multi-trait SNP association test in DCCT

Out of the 307 SNPs analyzed in DCCT individuals with the joint model, we identify two SNPs reaching the Bonferroni corrected significance threshold for the effective number of SNPs tested ( $P^*=1.7 \times 10^{-4}$ , Figure SI-3c A). These SNPs are rs10810632 (*BNC2*,  $P_{mult}=9.1 \times 10^{-8}$ ) and rs1358030 (nearby *SORCS1*,  $P_{mult}=3.4 \times 10^{-7}$ ) and are among the SNPs reported associated with HbA1c and time-to-T1DC traits in the previous GWAS of HbA1c in the Conventional treatment group of DCCT<sup>4</sup>. We obtained similar conclusions with the two other compared association structures for HbA1c on T1DC (Figure SI-3c, panels B and C).

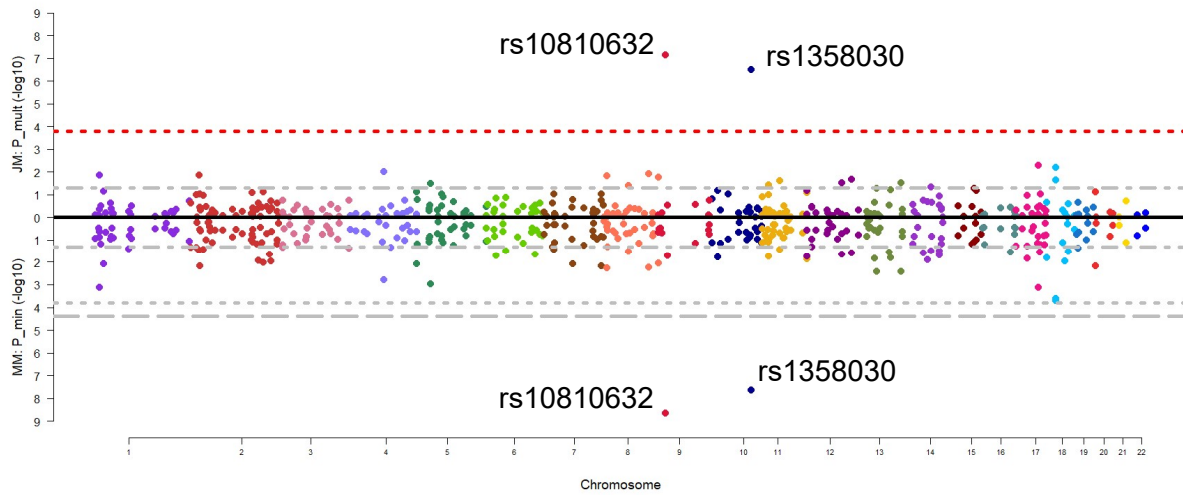
**Figure SI-3c. Mirror-plots of the multi-trait SNP association test  $P$ -values ( $-\log_{10}$ ) from the joint model using alternatively each association structure for HbA1c effects on time-to-T1DC traits.** For each Manhattan plot, the upper panel represents the  $P$ -values ( $-\log_{10}$ ) of the multi-trait SNP association test, and the lower panel represents the results of the minimum  $P$ -value from the separate analysis of each trait, included for comparison. On each upper panel, the horizontal lines represent the significance levels (in red: Bonferroni-corrected thresholds ( $P^* = 1.7 \times 10^{-4}$ ) and in grey the nominal significance levels). On each lower panel, the grey dashed lines represent the nominal significance and Bonferroni-corrected significance levels for the effective number of SNPs tested and the Bonferroni-corrected significance level further corrected for the four traits tested (under the assumption of independent traits).

(see Figure on next page)

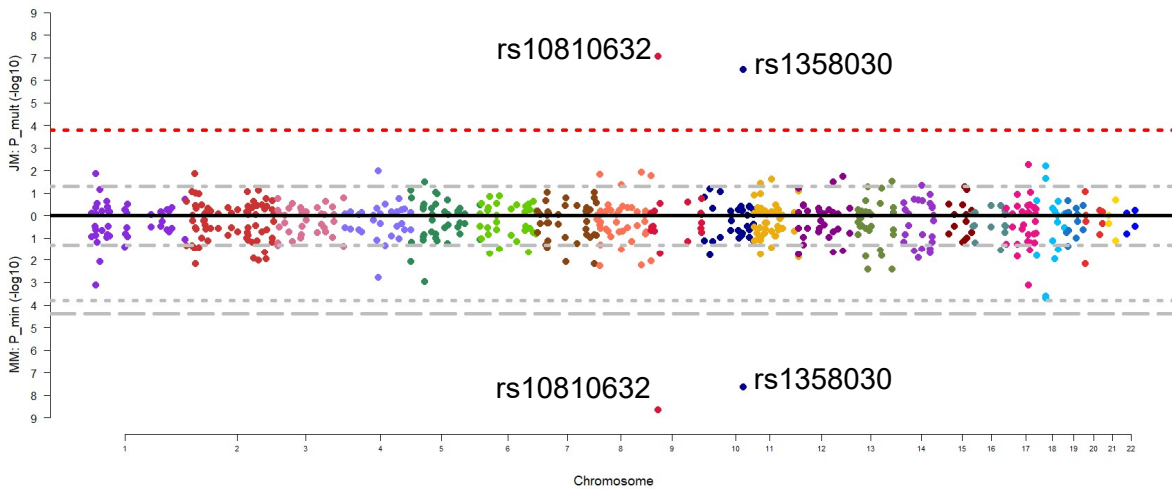
### A. Contemporaneous HbA1c effects



### B. Updated cumulative mean HbA1c effects



### C. Time-weighted cumulative HbA1c effects.



## List of Supplementary Tables that are provided in a separate document

**Table S1.** Other specified parameters in the simulation study.

**Table S2.** Empirical power of single-parameter SNP association tests using separate analysis of each trait and joint models (JM-cmp, JM-mis), assessed using  $R=1000$  replicates of  $N=667$  DCCT subjects simulated under the *causal* genetic scenario from Figure 2

**Table S3.** Estimation accuracy of the parameters of interest for causal SNP1 with indirect effects on both T1DC traits induced via HbA1c longitudinal risk factor, assessed using  $R=1000$  replicates of  $N=667$  DCCT subjects simulated under the *causal* genetic scenario from Figure 2

**Table S4.** Estimation accuracy of the parameters of interest in the time-to-event sub-models for SNP2 and SNP4 having direct associations on DR and on DN, assessed using  $R=1000$  replicates of  $N=667$  DCCT subjects simulated under the *causal* genetic scenario from Figure 2

**Table S5.** Estimation accuracy of the overall effects of SNP3 with indirect effects on T1DC traits induced via  $U$ , assessed using  $R=1000$  replicates of  $N=667$  DCCT subjects simulated under the *causal* genetic scenario from Figure 2

**Table S6.** Estimation accuracy of parameters of interest for SNP5 with a direct effect and an indirect effect on Time-to-DN via SBP, assessed using  $R=1000$  replicates of  $N=667$  DCCT subjects simulated under the *causal* genetic scenario from Figure 2

**Table S7.** Type-I errors for tests of SNP effects on each trait from the joint model (JM-cmp), assessed using  $R=1000$  replicates of  $N=667$  DCCT subjects simulated under the *global null* genetic scenario

**Table S8.** SNP association results from marginal analysis in DCCT subjects for the 307 SNPs from the literature

**Table S9.** Results from the joint model fitted in DCCT subjects with alternative association structures for HbA1c effects on T1DC traits and without genetic variable

**Table S10.** Results from CM-obs fitted in DCCT subjects with alternative association structures for HbA1c effects on T1DC traits and without genetic variable

## List of Supplementary Figures that are provided in a separate document

**Figure S1.** SNP effects estimated by the joint models (JM-cmp, JM-mis, JM-sep) for each SNP on (a.) HbA1c ( $\beta_{g(1)}$ ), (b.) SBP ( $\beta_{g(2)}$ ), (c.) Time-to-DR ( $\gamma_{g(1)}$ ) and (d.) Time-to-DN ( $\gamma_{g(2)}$ ), assessed using  $R=1000$  replicates of  $N=667$  DCCT subjects simulated under the *causal* genetic scenario from Figure 2

**Figure S2.** Quantile-quantile plots of the  $P$ -values from the Wald tests for each SNP effect in longitudinal (upper) and time-to-event components (bottom) of the complete joint model (JM-cmp), assessed using  $R=1000$  replicates of  $N=667$  DCCT subjects simulated under the *global null* genetic scenario

**Figure S3.** Classification rates of each SNP as direct and/or indirect association with each time-to-event trait through each longitudinal trait, assessed using  $R=1000$  replicates of  $N=667$  DCCT subjects simulated under the *global null* genetic scenario

**Figure S4.** Classification rates of each SNP as direct and/or indirect association with each pair of time-to-event/longitudinal traits, assessed using  $R=1000$  replicates of  $N=667$  DCCT subjects simulated under the *causal* genetic scenario from Figure 2

**Figure S5.** Classification of direct and/or indirect SNP association for 307 candidate SNPs tested in  $N=516$  DCCT subjects using the contemporaneous association structure and the two alternative *cumulative* association structures (updated cumulative mean and time-weighted cumulative) for HbA1c effects on T1DC traits

## List of the participants of the DCCT/EDIC Research Group (as of January 1, 2021)

*Study Chairpersons* – D.M. Nathan (chair), B. Zinman (vice-chair); *Past*: O. Crofford; *Deceased*: S. Genuth

*Editor, EDIC Publications* – D.M. Nathan

### Clinical Centers

Case Western Reserve University – *Current*: R. Gubitosi-Klug, L. Mayer, J. Wood, D. Miller, A. Nayate, M. Novak, S. Pendegast, L. Singerman, D. Weiss, H. Zegarra; *Past*: E. Brown, P. Crawford, M. Palmert, P. Pugsley, J. Quin, S. Smith-Brewer; *Deceased*: W. Dahms, S. Genuth, J. McConnell

Weill Cornell Medical College – *Current*: N.S. Gregory, R. Hanna, R. Chan, S. Kiss, A. Orlin, M. Rubin; *Past*: S. Barron, B. Bosco, D. Brillon, S. Chang, A. Dwoskin, M. Heinemann, L. Jovanovic, M.E. Lackaye, T. Lee, B. Levy, V. Reppucci, M. Richardson; *Deceased*: R. Campbell

Henry Ford Health System – *Current*: A. Bhan, J.K. Jones, D. Kruger, P.A. Edwards, H. Remtema; *Past*: E. Angus, A. Galprin, M. McLellan, A. Thomas; *Deceased*: J.D. Carey, F. Whitehouse

International Diabetes Center – *Current*: R. Bergenstal, S. Dunnigan, M. Johnson, A. Carlson, ; *Past*: R. Birk, P. Callahan, G. Castle, R. Cuddihy, M. Franz, D. Freking, L. Gill, J. Gott, K. Gunyou, P. Hollander, D. Kendall, J. Laechelt, S. List, W. Mestrezat, J. Nelson, B. Olson, N. Rude, M. Spencer, L. Thomas; *Deceased*: D. Etzwiler, K. Morgan

Joslin Diabetes Center – *Current*: L.P. Aiello, E. Golden, P. Arrigg, R. Beaser, L. Bestourous, J. Cavallerano, R. Cavicchi, O. Ganda, O. Hamdy, T. Murtha, D. Schlossman, S. Shah, G. Sharuk, P. Silva, P. Silver, M. Stockman, J. Sun, E. Weimann; *Past*: V. Asuquo, A. Jacobson, R. Kirby, L. Rand, J. Rosenzweig, H. Wolpert

Massachusetts General Hospital – *Current*: D.M. Nathan, M.E. Larkin, M. Cayford, A. deManbey, L. Gurry, J. Heier, A. Joseph, F. Leandre, K. Martin, C. Shah, C. Stevens, N. Thangthaeng; *Past*: E. Anderson, H. Bode, S. Brink, M. Christofi, C. Cornish, D. Cros, S. Crowell, L. Delahanty, K. Folino, S. Fritz, C. Gauthier-Kelly, J. Godine, C. Haggan, K. Hansen, P. Lou, J. Lynch, C. McKittrick, D. Moore, D. Norman, M. Ong, E. Ryan, C. Taylor, D. Zimble

Mayo Clinic – *Current*: A. Vella, A. Zipse, A. Barkmeier; *Past*: B. French, M. Haymond, J. Mortenson, J. Pach, R. Rizza, L. Schmidt, W.F. Schwenk, F.J. Service, R. Woodwick, G. Ziegler; *Deceased*: R. Colligan, A. Lucas, B. Zimmerman

Medical University of South Carolina – *Current*: H. Karanchi, L. Spillers, J. Fernandes, K. Hermayer, S. Kwon, K. Lee, M. Lopes-Virella, T. Lyons, M. Nutaitis; *Past*: A. Blevins, M. Bracey, S. Caulder, J. Colwell, S. Elsing, A. Farr, D. Lee, P. Lindsey, L. Luttrell, R. Mayfield, J. Parker, N. Patel, C. Pittman, J. Selby, J. Soule, M. Szpiech, T. Thompson, D. Wood, S. Yacoub-Wasef

Northwestern University – *Current*: A. Wallia, M. Hartmuller, S. Ajroud-Driss, P. Astelford, A. Degillio, M. Gill, L. Jampol, C. Johnson, L. Kaminski, N. Leloude, A. Lyon, R. Mirza, D. Ryan, E. Simjanoski, Z. Strugula; *Past*: D. Adelman, S. Colson, M. Molitch, B. Schaefer

University of California, San Diego – *Current*: S. Mudaliar, G. Lorenzi, O. Kolterman, M. Goldbaum; *Past*: T. Clark, M. Giotta, I. Grant, K. Jones, R. Lyon, M. Prince, R. Reed, M. Swenson; *Deceased*: G. Friedenber

University of Iowa – *Current*: W.I. Sivitz, B. Vittetoe, J. Kramer; *Past*: M. Bayless, C. Fountain, R. Hoffman, J. MacIndoe, N. Olson, H. Schrott, L. Snetselaar, T. Weingeist, R. Zeitler

University of Maryland – *Current*: R. Miller, S. Johnsonbaugh; *Past*: M. Carney, D. Counts, T. Donner, J. Gordon, M. Hebdon, R. Hemady, B. Jones, A. Kowarski, R. Liss, S. Mendley, D. Ostrowski, M. Patronas, P. Salemi, S. Steidl

University of Michigan – *Current*: W.H. Herman, R. Pop-Busui, C.L. Martin, P. Lee, J. W. Albers, E.L. Feldman; *Past*: N. Burkhart, D.A. Greene, T. Sandford, M.J. Stevens; *Deceased*: J. Floyd

University of Minnesota – *Current*: J. Bantle, M. Rhodes, D. Koozekanani, S. Montezuma, J. Terry; *Past*: N. Flaherty, F. Goetz, C. Kwong, L. McKenzie, M. Mech, J. Olson, B. Rogness, T. Strand, R. Warhol, N. Wimmergren

University of Missouri – *Current*: D. Goldstein, D. Hainsworth, S. Hitt; *Deceased*: J. Giangiacomo

University of New Mexico – *Current*: D.S. Schade, J.L. Canady, R.B. Avery, M.R. Burge, J.E. Chapin, A. Das, L.H. Ketaj; *Past*: D. Hornbeck, C. Johannes, J. Rich, M.L. Schluter

University of Pennsylvania – *Current*: M. Schutta, P.A. Bourne, A. Brucker; *Past*: S. Braunstein, B.J. Maschak-Carey, S. Schwartz; *Deceased*: L. Baker

University of Pittsburgh – *Current*: T. Orchard, L. Cimino, D. Rubinstein; *Past*: D. Becker, B. Doft, D. Finegold, K. Kelly, L. Lobes, N. Silvers, T. Songer, D. Steinberg, L. Steranchak, J. Wesche; *Deceased*: A. Drash

University of South Florida – *Current*: J.I. Malone, A. Morrison, M.L. Bernal, P.R. Pavan; *Past*: L. Babbione, T.J. DeClue, N. Grove, D. McMillan, H. Solc, E.A. Tanaka, J. Vaccaro-Kish

University of Tennessee – *Current*: S. Dagogo-Jack, C. Wigley, S. Huddleston, A. Patel; *Past*: M. Bryer-Ash, E. Chaum, A. Iannacone, H. Lambeth, D. Meyer, S. Moser, M.B. Murphy, H. Ricks, S. Schussler, S. Yoser; *Deceased*: A. Kitabchi

University of Texas – *Current*: P. Raskin, S. Strowig, YG. He, E. Mendelson, RL. Ufret-Vincenty; *Past*: M. Basco; *Deceased*: S. Cercone

University of Toronto – *Current*: B.A. Perkins, B. Zinman, A. Barnie, N. Bakshi, M. Brent, R. Devenyi, K. Koushan, M. Mandelcorn, F. Perdikaris, L. Tuason; *Past*: D. Daneman, R. Ehrlich, S. Ferguson, A. Gordon, K. Perlman, S. Rogers

University of Washington – *Current*: I. Hirsch, R. Fahlstrom, L. Van Ottingham, I.H. de Boer, L. Olmos de Koo; *Past*: S. Catton, J. Ginsberg, J. Kinyoun, J. Palmer

University of Western Ontario – *Current*: C. McDonald, M. Driscoll, J. Bylsma, T. Sheidow; *Past*: W. Brown, C. Canny, P. Colby, S. Debrabandere, J. Dupre, J. Harth, I. Hramiak, M. Jenner, J. Mahon, D. Nicolle, N.W. Rodger, T. Smith



Vanderbilt University – *Current*: M. May, J. Lipps Hagan, T. Adkins, A. Agarwal, C. Lovell;  
*Past*: S. Feman, R. Lorenz, R. Ramker; *Deceased*: L. Survant

Washington University, St. Louis – *Current*: N.H. White, L. Levandoski; *Deceased*: I. Boniuk, J. Santiago

Yale University – *Current*: W. Tamborlane, P. Gatcomb, K. Stoessel; *Past*: J. Ahern, K. Fong, P. Ossorio, P. Ramos

Albert Einstein – *Past*: J. Brown-Friday, J. Crandall, H. Engel, S. Engel, H. Martinez, M. Phillips, M. Reid, H. Shamoon, J. Sheindlin

### **Clinical Coordinating Center**

Case Western Reserve University – *Current*: R. Gubitosi-Klug, L. Mayer, C. Beck, K. Farrell, P. Gaston; *Past*: S. Genuth, M. Palmert, J. Quin, R. Trail; *Deceased*: W. Dahms

### **Data Coordinating Center**

George Washington University, The Biostatistics Center – J. Lachin, I. Bebu, B. Braffett, J. Backlund, L. Diminick, L. El ghormli, X. Gao, D. Kenny, K. Klumpp, M. Lin, V. Trapani; *Past*: K. Anderson, K. Chan, P. Cleary, A. Determan, L. Dews, W. Hsu, P. McGee, H. Pan, B. Petty, D. Rosenberg, B. Rutledge, W. Sun, S. Villavicencio, N. Younes; *Deceased*: C. Williams

### **National Institute of Diabetes and Digestive and Kidney Disease**

National Institute of Diabetes and Digestive and Kidney Disease Program Office – E. Leschek; *Past*: C. Cowie, C. Siebert

### **EDIC Core Central Units**

Central Biochemistry Laboratory (University of Minnesota) – M. Steffes, A. Karger, J. Seegmiller, V. Arends; *Past*: J. Bucksa, B. Chavers, A. Killeen, M. Nowicki, A. Saenger

Central ECG Reading Unit (Wake Forest School of Medicine) – Y. Pokharel, M. Barr, C. Campbell, S. Hensley, J. Hu, L. Keasler, Y. Li, T. Taylor, Z.M. Zhang; *Past*: R. Prineas, E.Z. Soliman

Central Ophthalmologic Reading Unit (University of Wisconsin) – B. Blodi, R. Danis, D. Lawrence, H. Wabers; *Past*: M. Burger, M. Davis, J. Dingleline, V. Gama, S. Gangaputra, L. Hubbard, S. Neill, R. Sussman

Central Neuropsychological Reading Unit (NYU Winthrop Hospital, University of Pittsburgh) – A. Jacobson, C. Ryan, D. Saporito; *Past*: B. Burzuk, E. Cupelli, M. Geckle, D. Sandstrom, F. Thoma, T. Williams, T. Woodfill

## References

1. Crowther, M.J., and Lambert, P.C. (2013). Simulating biologically plausible complex survival data. *Stat. Med.* *32*, 4118–4134.
2. G., D., and Brent, R.P. (1974). Algorithms for minimization without derivatives. *IEEE Trans. Automat. Contr.* *28*, 865.
3. The Diabetes Control and Complications Trial Research Group (1993). The effect of intensive treatment of diabetes on the development and progression of long-term complications in insulin-dependent diabetes mellitus. *N. Engl. J. Med.* *329*, 977–986.
4. Paterson, A.D., Waggott, D., Borch-Johnsen, K., Hosseini, S.M., Shen, E., Sylvestre, M.P., Wong, I., Bharaj, B., Cleary, P.A., Lachin, J.M., et al. (2010). A genome-wide association study identifies a novel major locus for glycemic control in type 1 diabetes, as measured by both A1C and glucose. *Diabetes* *59*, 539–549.
5. Roshandel, D., Gubitosi-Klug, R., Bull, S.B., Canty, A.J., Pezzolesi, M.G., King, G.L., Keenan, H.A., Snell-Bergeon, J.K., Maahs, D.M., Klein, R., et al. (2018). Meta-genome-wide association studies identify a locus on chromosome 1 and multiple variants in the MHC region for serum C-peptide in type 1 diabetes. *Diabetologia* *61*, 1098–1111.
6. 1000 Genomes Project Consortium, Auton, A., Brooks, L.D., Durbin, R.M., Garrison, E.P., Kang, H.M., Korbel, J.O., Marchini, J.L., McCarthy, S., McVean, G.A., et al. (2015). A global reference for human genetic variation. *Nature* *526*, 68–74.
7. Das, S., Forer, L., Schönherr, S., Sidore, C., Locke, A.E., Kwong, A., Vrieze, S.I., Chew, E.Y., Levy, S., McGue, M., et al. (2016). Next-generation genotype imputation service and methods. *Nat. Genet.* *48*, 1284–1287.
8. Machiela, M.J., and Chanock, S.J. (2015). LDlink: A web-based application for exploring population-specific haplotype structure and linking correlated alleles of possible functional variants. *Bioinformatics* *31*, 3555–3557.
9. Wheeler, E., Leong, A., Liu, C.-T.T., Hivert, M.-F.F., Strawbridge, R.J., Podmore, C., Li, M., Yao, J., Sim, X., Hong, J., et al. (2017). Impact of common genetic determinants of Hemoglobin A1c on type 2 diabetes risk and diagnosis in ancestrally diverse populations: A transethnic genome-wide meta-analysis. *PLoS Med.* *14*, e1002383.
10. Evangelou, E., Warren, H.R., Mosen-Ansorena, D., Mifsud, B., Pazoki, R., Gao, H., Ntritsos, G., Dimou, N., Cabrera, C.P., Karaman, I., et al. (2018). Genetic analysis of over 1 million people identifies 535 new loci associated with blood pressure traits. *Nat. Genet.* *50*, 1412–1425.
11. Hosseini, S.M., Borch-Johnsen, K., Sun, L., Canty, A.J., Bull, S.B., Klein, B.E.K., Klein, R., and Paterson, A.D. (2015). The association of previously reported polymorphisms for microvascular complications in a meta-analysis of diabetic retinopathy. *Hum. Genet.* *134*, 247–257.
12. Pollack, S., Igo, R.P., Jensen, R.A., Christiansen, M., Li, X., Cheng, C.-Y., Ng, M.C.Y., Smith, A. V, Rossin, E.J., Segrè, A. V, et al. (2019). Multiethnic genome-wide association study of diabetic retinopathy using liability threshold modeling of duration of diabetes and glycemic control. *Diabetes* *68*, 441–456.
13. Grassi, M.A., Tikhomirov, A., Ramalingam, S., Below, J.E., Cox, N.J., and Nicolae, D.L. (2011). Genome-wide meta-analysis for severe diabetic retinopathy. *Hum. Mol. Genet.* *20*, 2472–2481.
14. Sandholm, N., Salem, R.M., McKnight, A.J., Brennan, E.P., Forsblom, C., Isakova, T., McKay,

G.J., Williams, W.W., Sadlier, D.M., Mäkinen, V.-P.P., et al. (2012). New susceptibility loci associated with kidney disease in type 1 diabetes. *PLoS Genet.* 8, e1002921.

15. Lind, M., Odén, A., Fahlén, M., and Eliasson, B. (2010). The shape of the metabolic memory of HbA1c: Re-analysing the DCCT with respect to time-dependent effects. *Diabetologia* 53, 1093–1098.

16. Lind, M., Odén, A., Fahlén, M., and Eliasson, B. (1995). The relationship of glycemic exposure (HbA(1c)) to the risk of development and progression of retinopathy in the diabetes control and complications trial. *Diabetes* 53, 1093–1098.