# A Type 2 Diabetes Subtype Responsive to ACCORD Intensive Glycemia Treatment

Diabetes Care 2021;44:1410–1418 | https://doi.org/10.2337/dc20-2700

Arshiya Mariam,[1] Galen Miller-Atkins,[1] Kevin M. Pantalone,[2] Robert S. Zimmerman,[2] John Barnard,[1] Michael W. Kattan,[1] Hetal Shah,[3] Howard L. McLeod,[4] Alessandro Doria,[3] Michael J. Wagner,[5] John B. Buse,[6] Alison A. Motsinger-Reif,[7] and Daniel M. Rotroff[1]


## OBJECTIVE

**Current type 2 diabetes (T2D) management contraindicates intensive glycemia treatment in patients with high cardiovascular disease (CVD) risk and is partially motivated by evidence of harms in the Action to Control Cardiovascular Risk in Diabetes (ACCORD) trial. Heterogeneity in response to intensive glycemia treatment has been observed, suggesting potential benefit for some individuals.**

## RESEARCH DESIGN AND METHODS

**ACCORD was a randomized controlled trial that investigated whether intensively treating glycemia in individuals with T2D would reduce CVD outcomes. Using a novel approach to cluster $HbA_{1c}$ trajectories, we identified groups in the intensive glycemia arm with modified CVD risk. Genome-wide analysis and polygenic score (PS) were developed to predict group membership. Mendelian randomization was performed to infer causality.**

## RESULTS

**We identified four clinical groupings in the intensive glycemia arm, and clinical group 4 (C4) displayed fewer CVD (hazard ratio [HR] 0.34; $P = 2.01 \times 10^{-3}$) and microvascular outcomes (HR 0.86; $P = 0.015$) than those receiving standard treatment. A single-nucleotide polymorphism, rs220721, in *MAS1* reached suggestive significance in C4 ($P = 4.34 \times 10^{-7}$). PS predicted C4 with high accuracy (area under the receiver operating characteristic curve 0.98), and this predicted C4 displayed reduced CVD risk with intensive versus standard glycemia treatment (HR 0.53; $P = 4.02 \times 10^{-6}$), but not reduced risk of microvascular outcomes ($P < 0.05$). Mendelian randomization indicated causality between PS, on-trial $HbA_{1c}$, and reduction in CVD outcomes ($P < 0.05$).**

## CONCLUSIONS

**We found evidence of a T2D clinical group in ACCORD that benefited from intensive glycemia treatment, and membership in this group could be predicted using genetic variants. This study generates new hypotheses with implications for precision medicine in T2D and represents an important development in this landmark clinical trial warranting further investigation.**


The Action to Control Cardiovascular Risk in Diabetes (ACCORD) was a landmark trial to examine the effect of intensive glycemia treatment targeting glycated hemoglobin $A_{1c}$ ($HbA_{1c}$) <6% versus more modest therapy targeting $HbA_{1c}$ 7.0–7.9%. The study was conducted in patients with type 2 diabetes (T2D) at high cardiovascular risk, with a primary end point of time to first occurrence of major adverse cardiovascular events (MACE), specifically nonfatal myocardial infarction (MI), nonfatal stroke, or

[1]Department of Quantitative Health Sciences, Lerner Research Institute, Cleveland Clinic, Cleveland, OH
[2]Endocrinology and Metabolism Institute, Cleveland Clinic, Cleveland, OH
[3]Joslin Diabetes Center and Harvard Medical School, Boston, MA
[4]Taneja College of Pharmacy, University of South Florida, Tampa, FL
[5]Center for Pharmacogenomics and Individualized Therapy, University of North Carolina at Chapel Hill, Chapel Hill, NC
[6]Division of Endocrinology, School of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, NC
[7]Biostatistics and Computational Biology Branch, National Institute of Environmental Health Sciences, Research Triangle Park, NC

Corresponding author: Daniel M. Rotroff, rotrofd@ccf.org

Received 2 November 2020 and accepted 23 February 2021

This article contains supplementary material online at https://doi.org/10.2337/figshare.14109806.

A.M. and G.M.-A. contributed equally to this work.

cardiovascular death. Notably, the intensive glycemia arm of the trial was terminated prematurely because of an increase in cardiovascular and overall mortality (1). The findings from ACCORD have had important implications regarding guidelines for glycemic management (2–4). Although there was a significant increase in mortality in the intensive glycemia arm, heterogeneity was observed (5–8). In addition, potential benefits in some measures of neuropathy, eye complications, and microalbuminuria were observed (9). Epidemiological analysis of ACCORD demonstrated that individuals at greatest risk of mortality and MACE were those intensively treated who did not reach the target $HbA_{1c}$ (10).

We previously developed a risk score based on two genetic variants that predicted White individuals with modified risk of cardiovascular mortality and displayed a significant interaction with glycemia control (5). Here, we expanded our approach to a racially diverse population, characterizing a glycemia-responsive group in the intensive glycemia arm of ACCORD using a novel application of dynamic time warping to measure the similarity of patient $HbA_{1c}$ trajectories. We then clustered patients into clinical groups based on $HbA_{1c}$ trajectory and identified a group of patients intensively treated in ACCORD with significantly lower risk of mortality, MACE, and multiple other cardiovascular disease (CVD) outcomes compared with other intensively treated patients. We also demonstrated causal inference regarding this relationship using Mendelian randomization. Furthermore, we performed a genome-wide association study (GWAS) to identify genetic variants associated with membership in a low-risk clinical group and used machine learning to construct a polygenic score (PS) to predict patients likely to benefit from ACCORD-like intensive intervention. Importantly, this study generates new hypotheses that patients predicted to be in this low-risk clinical group have significantly lower risk of CVD outcomes compared with the same predicted group receiving standard glycemia treatment.

## RESEARCH DESIGN AND METHODS

### Study Participants

The ACCORD trial (NCT00000620, ClinicalTrials.gov) had a $2 \times 2$ factorial design and has been previously described (1). Briefly, 10,245 individuals with T2D and either a history of CVD or two or more risk factors for CVD were randomly assigned to receive standard glycemia treatment ($HbA_{1c}$ 7–7.9%; $n = 5{,}119$) or intensive glycemic treatment ($HbA_{1c} < 6\%$; $n = 5{,}126$). Patients were further randomized to determine whether intensively treating blood pressure or dyslipidemia was effective at reducing MACE. Participants were scheduled to have their $HbA_{1c}$ tested at each 4-month visit and had to have $\geq$3 $HbA_{1c}$ measurements recorded to be included in the analysis. Missing values were mean imputed with the $HbA_{1c}$ immediately preceding and following. Trial outcomes included first occurrence of MACE (i.e., nonfatal heart attack, nonfatal stroke, cardiovascular death), all-cause mortality, CVD mortality, congestive heart failure, nonfatal MI, nonfatal stroke, total stroke, expanded macrovascular events, and coronary heart disease. Cohort details are provided in Supplementary Table 1. In addition, microvascular outcomes including neuropathies, nephropathies, and retinopathies were also investigated. The GWAS and PS development were performed using the 8,054 individuals who consented to genetic studies. Phenotypic data are available for download at National Institutes of Health BioLINCC (HLB01041317a).

### Patient Subgrouping

$HbA_{1c}$ trajectories from individuals randomized to intensive glycemia treatment ($n = 4{,}946$) and truncated to each trial outcome were compared using a modified dynamic time-warping approach (etwDTW) that weighted the cost of earlier $HbA_{1c}$ differences more than later differences. Dynamic time warping compares two trajectories and calculates a score for how similar they are to each other. To ensure that $HbA_{1c}$ impacts after a CVD outcome did not bias the results, $HbA_{1c}$ trajectories were truncated to ensure that only pre-CVD event $HbA_{1c}$ values were used. For this reason, nine versions of truncated $HbA_{1c}$s were developed, one for each of the nine CVD outcomes of interest. Through the pairwise comparison of patient-truncated $HbA_{1c}$ trajectories, dissimilarity matrices were generated for each outcome. Unsupervised hierarchical clustering with the Ward method was used to cluster $HbA_{1c}$ trajectories based on etwDTW dissimilarity for each outcome (11,12). Because clustering was performed on the truncated $HbA_{1c}$ trajectories for each outcome, a consensus cluster was developed that integrated clustering for each outcome into a single dendrogram. To do this, the cophenetic distances of the dendrogram of each outcome were summed to generate a cophenetic dissimilarity matrix (13). Hierarchical clustering with the Ward method was used again to cluster $HbA_{1c}$ trajectories across all outcomes based on cophenetic dissimilarity. Individuals who clustered together were considered to be part of the same clinical group; the optimal number of clinical groups ($k$) was determined by comparing the risk of outcome for each dendrogram split using Cox proportional hazards, and the number of clusters preceding a nonsignificant split was selected ($P > 0.05$) (Fig. 1). For each clinical group, composite trajectories were created by averaging the individual $HbA_{1c}$ values across each time point. Additional information regarding patient clustering and etwDTW can be found in the Supplementary Material.

Baseline demographic and medication differences were evaluated using logistic regression and were adjusted for multiple comparisons using a false discovery rate (FDR) approach ($q < 0.01$) (14). CVD and microvascular risks were calculated using Cox proportional hazards, adjusting for sex, race, age, years with diabetes, blood pressure treatment arm, history of CVD, microvascular disorders at baseline, and baseline $HbA_{1c}$. Meta-analyzing correlated values, such as CVD outcomes, as if they are independent can be problematic. Meta-analyses summarizing CVD outcomes, treated as random effects, were performed using the metafor R package (15,16), a multivariate restricted maximum-likelihood model that accounts for the correlation structure of the CVD outcomes. The joint multivariate meta-analysis closely followed the methodology described by Berkey et al. (15), which also takes into account the covariance-variance of the outcomes. The covariance-variance matrix for hazard ratios (HRs) was calculated using the bootstrapping method described by Riley et al. (17).

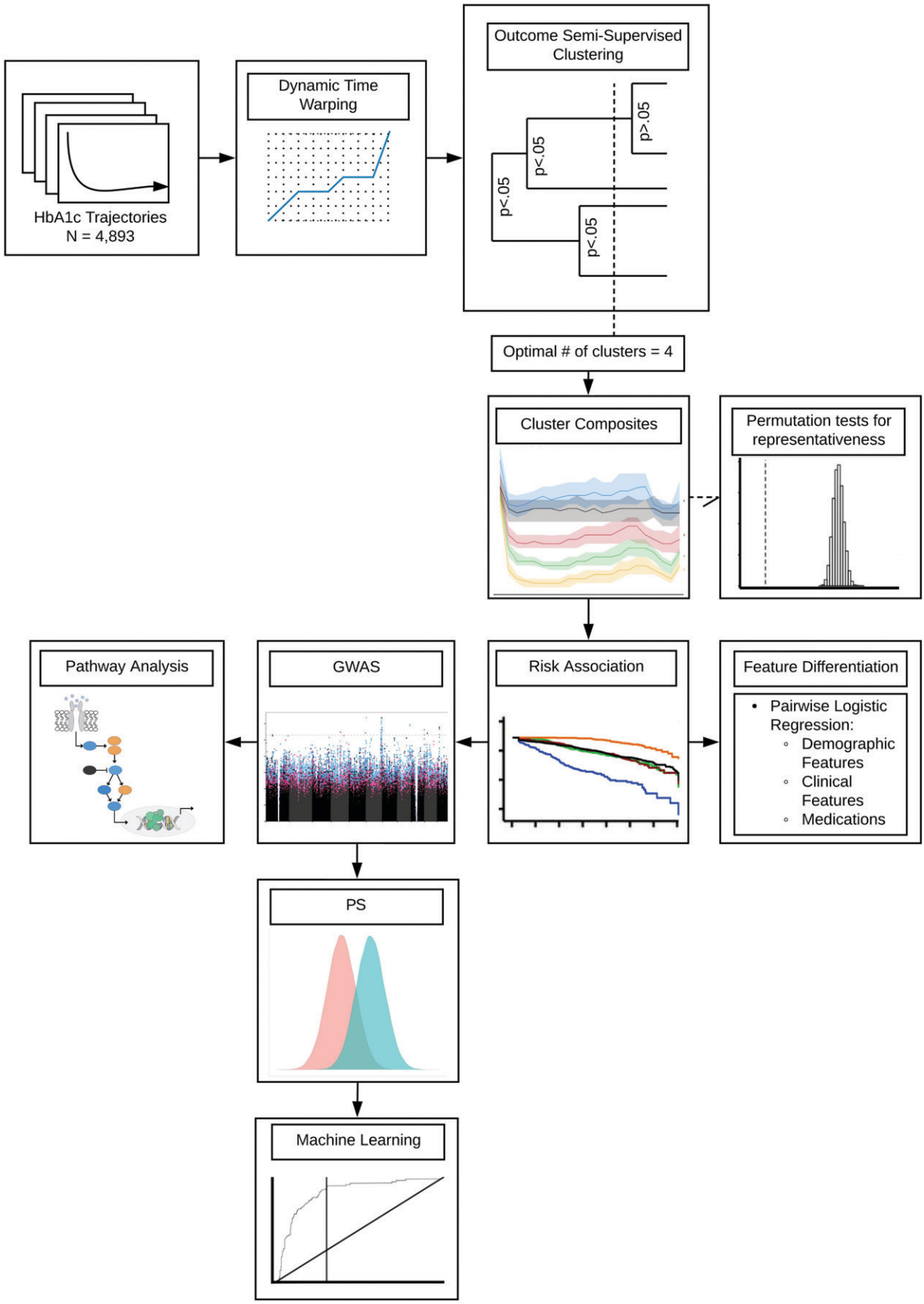


**Figure 1**—Analysis workflow. A workflow describing the process of identifying clinical groups using dynamic time warping, performing the genomic studies, and developing the risk model. PS, polygenic score.

### GWAS

Genomic DNA was extracted using the FlexiGene DNA Kit (Qiagen, Valencia, CA) (18) and was genotyped with two separate platforms: 6,085 ACCORD participant samples were genotyped on Illumina HumanOmniExpressExome-8 (version 1.0) chips (set 1), and 8,174 participant samples (6,085 from set 1 plus 2,089 samples) were genotyped on Affymetrix Axiom Biobank1 chips (set 2). GRCh37/hg19 was used as the reference genome build, and single-nucleotide polymorphisms (SNPs) with genotyping rates >97% were merged using PLINK (version 1.9). Prephasing and imputation were performed using SHAPEIT2 (version 2.r778) and IMPUTE2 (version 2.3.0), respectively, using set 2 and the 1000 Genomes Phase 1 integrated haplotypes reference panel (19,20). Variants with an info metric <0.5 and minor allele frequency <3% were excluded. SNPs genotyped across all individuals were annotated as GENO; SNPs genotyped in some individuals and imputed in others were analyzed separately and meta-analyzed (because of the two different genotyping platforms) and annotated as META. SNPs imputed across all participants were annotated as IMPU. Genotypic data are available at the National Center for Biotechnology Information database for genotypes and phenotypes (phs001411.v1.p1). The final genotyped data set consisted of 8,054 unique individuals and 1,238,999 variants. Additional details regarding genetic data processing can be found in Marvel et al. (21) and Rotroff et al. (22).

A GWAS was conducted to determine SNPs associated with membership in the intensively treated clinical group with lower risk of CVD outcomes (C4; $n = 1{,}367$) versus all other participants (C1–3; $n = 3{,}579$) using logistic regression. Potentially confounding covariates were selected using a backward selection approach (Supplementary Table 2). Principal component analysis was performed to address population stratification using EIGENSTRAT (version 4.2). For each variant, a logistic regression model was constructed, with clinical group as the outcome and variant as the predictor. Principal components 1–3, sex, years with diabetes, BMI, and use of sulfonylurea, biguanide, thiazolidinedione, or any type of insulin at baseline were included as covariates (Supplementary Table 3). SNPs with $P < 5 \times 10^{-8}$ were considered to have genome-wide significance and $P < 5 \times 10^{-6}$ were considered to be of suggestive significance.

### PS and Predictive Model Development

Several approaches to constructing a PS were evaluated. Sets of SNPs from the GWAS were chosen using a clumping procedure based on 28 unique combinations of hyperparameters (Supplementary Table 4). A PS was derived based on each set of clumped variants and each *P* value threshold (CT-PS). In addition, a penalized logistic regression framework was used to derive a stacked clumping and thresholding PS (SCT-PS) using the R package bigsnpr (23,24). This approach generated multiple PS for each individual, and penalized regression was then used to derive an optimal combination of each PS, from which the weighted allele counts were summed to create the final SCT-PS to predict clinical group membership.

Next, we developed models using genetic and baseline clinical factors to predict patients in the group that benefited from intensive glycemia treatment (C4). Patients randomized to intensive treatment were partitioned into a training set (66%; $n = 2{,}270$) and a test set (33%; $n = 1{,}169$). Cross-validation was performed on the training data; least absolute shrinkage and selection operator was used to select baseline clinical features and their interactions with the PS, and these were included in a logistic regression model (25). Each model was developed using clinical features only, PS only, or clinical features and PS. Models developed on the training data were then tested on the 33% of data withheld as the test set and evaluated based on the area under the receiver operating characteristic curve in the test set. The SCT-PS without clinical variables performed comparable or better than other approaches and was selected because of parsimony and clinical utility. Details of the other models are provided in the Supplementary Material. Both the SNPs and corresponding weights for the CT-PS and SNPs as well as the code to run the SCT-PS can be found at www.github.com/rotroff-lab/accord-C4-ps.

### Mendelian Randomization

Mendelian randomization was performed using the Mendelian randomization R package (version 0.4.2) (26) on standard and intensively treated cohorts, separately. Mendelian randomization attempts to identify causal relationships between an exposure variable and an outcome by using genetic data as instrumental variables to randomize the population. Here, in each treatment arm, each CVD outcome was tested using the SCT-PS as the instrumental variable and median $HbA_{1c}$ before an event as the exposure variable. Significant causal relationships were assessed using the inverse-variance weighted (IVW) method. The IVW estimate is the ratio of the gene-outcome association to the gene-exposure association, standardized by the SE of the gene-outcome association. The IVW estimates for every CVD outcome in each treatment arm were then meta-analyzed to identify overall causal associations.

## RESULTS

### Clustering

Based on the criteria described above, four distinct clinical groups (C1–C4) were identified (Fig. 2*A* and Supplementary Figs. 1 and 2). Characterization of each clinical group regarding outcome incidence, sex, race, and other factors is provided in Supplementary Table 1. The mean $HbA_{1c}$ was lower at 4 months for C1, C2, and C4 at 6.7%, 7.0%, and 6.2%, respectively, compared to 7.6% across the standard glycemia arm (Fig. 2*A*). C3 had highest mean $HbA_{1c}$ of 7.7% at 4 months.

Clinical groups displayed different incidence rates for adverse outcomes (Fig. 2). C4 displayed the lowest risk across multiple outcomes, MACE (HR 0.27; 95% CI 0.19–0.36; $P = 1.57 \times 10^{-16}$), and total mortality (HR 0.33; 0.23–0.48; $P = 2.74 \times 10^{-9}$). C3 displayed the greatest risk for MACE (HR 1.60; 1.31–1.95; $P = 4.58 \times 10^{-6}$) and total mortality (HR 2.52; 2.00–3.17; $P = 3.44 \times 10^{-15}$). Although groups were separated by $HbA_{1c}$ reduction, the etwDTW approach identified groups with differential risks better than stratification by $HbA_{1c}$ quartile (Supplementary Figs. 7 and 8). All intensive clinical groups had significantly increased risk of hypoglycemic events compared with the standard arm, including C4, which had a lower

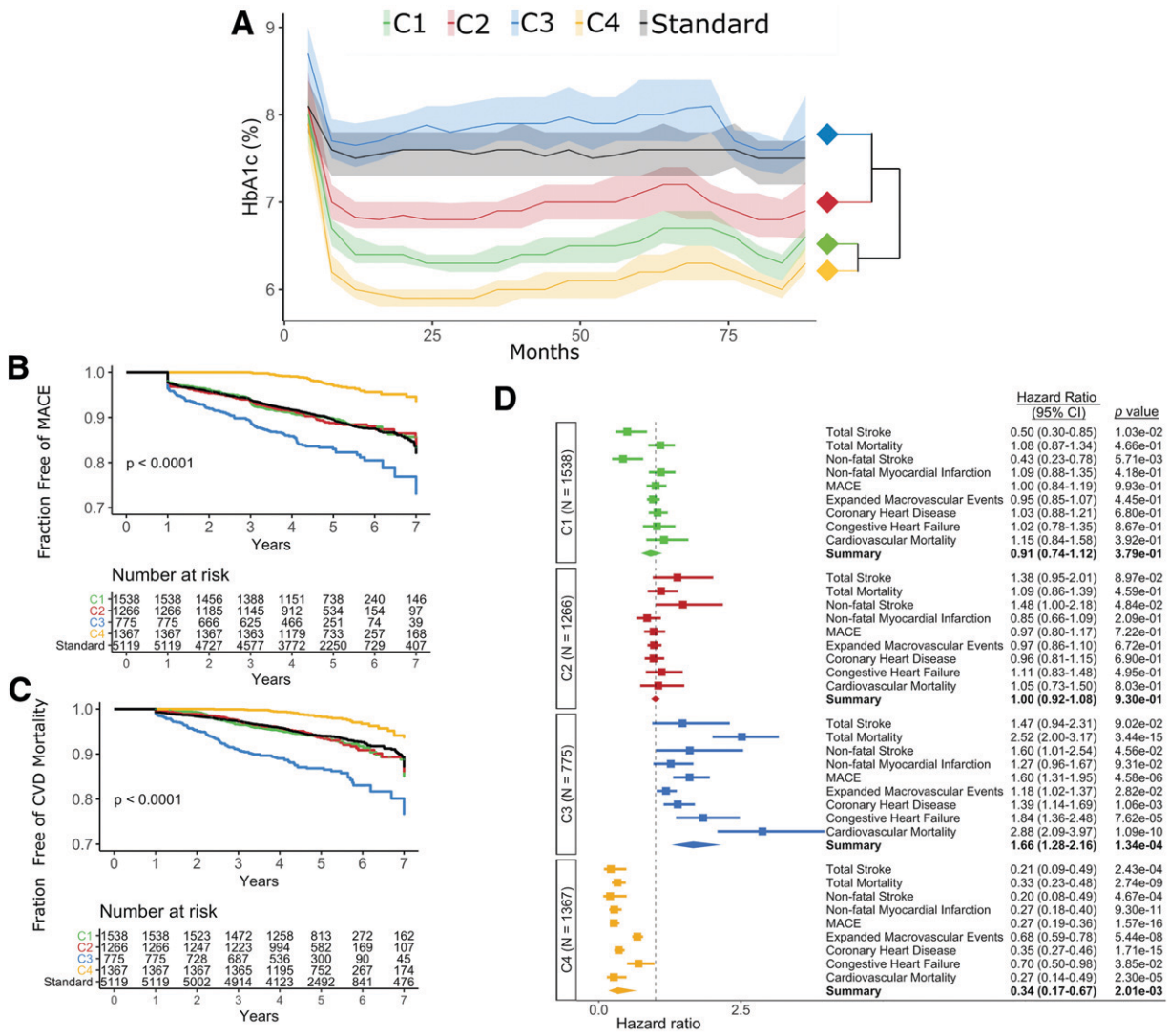


**Figure 2**—Identification of clinical subgroups. (*A*) The composite trajectories from each of the four clinical groups based on $HbA_{1c}$ trajectories compared with the composite trajectory from the standard arm (black). The interval surrounding the composite $HbA_{1c}$ trajectory represents two median absolute deviations of the underlying trajectories. (*B*) Kaplan-Meier curves of each clinical group and standard treatment group for developing MACE, including nonfatal heart attack, nonfatal stroke, or cardiovascular death. (*C*) Kaplan-Meier curves of each clinical group and standard treatment group for developing CVD mortality. (*D*) Forest plot of HRs for each CVD outcome separated by clinical group relative to standard glycemia treatment. Summary HR is the meta-analysis of all outcomes in the cluster after accounting for covariance between outcomes (15).

risk of CVD outcomes (Supplementary Fig. 6). Of the intensive clinical groups, C4 had the lowest risk of severe hypoglycemic events (HR 1.89; 95% CI 1.52–2.35; $P = 1.3 \times 10^{-8}$), whereas C3 had the greatest risk (HR 4.07; 3.26–5.08; $P = 2.2 \times 10^{-35}$). Meta-analysis across outcomes indicated that C4 had the overall lowest risk (HR 0.34; 0.17–0.67; $P = 2.01 \times 10^{-3}$) and C3 had the greatest risk of CVD outcomes (HR 1.66; 1.28–2.16; $P = 1.34 \times 10^{-4}$) (Fig. 2*D*).

C4 also demonstrated a significantly lower risk of microvascular events compared with standard glycemia treatment, including retinopathies, nephropathies, and neuropathies (HR 0.86; 95% CI 0.77–0.97; $P = 0.002$). All individual microvascular outcomes were significantly reduced in C4 compared with standard treatment, with HRs ranging from 0.82 to 0.89 ($P < 0.05$). Similar to CVD risk, C3 demonstrated the highest microvascular risk compared with standard treatment (HR 1.30; 1.16–1.47; $P = 1.2 \times 10^{-5}$].

Clinical and demographic differences were observed between clinical groups (Supplementary Tables 5–8). C4 had greater proportions of men and White individuals than the two high-risk clinical groups, C3 and C2. C4 individuals also had fewer years since their T2D diagnosis ($q < 0.01$) (Supplementary Table 7). After adjustment for years with T2D, increased alcohol intake, fewer eye diseases, and depression at baseline were observed in C4 ($q < 0.01$). After adjusting for years with T2D, compared with those in C1, individuals in C4 were also less likely to use biguanides (60.86% vs. 65.80%; $q = 7.56 \times 10^{-3}$), sulfonylureas (50.69% vs. 58.84%; $q = 2.65 \times 10^{-5}$), and thiazolidinediones (18.14% vs. 24.64%; $q = 6.86 \times 10^{-4}$) (Supplementary Table 6). C4 also had a significantly lower proportion of individuals using insulin at baseline than C2

and C3 ($q < 0.001$). In addition to having significantly lower on-trial $HbA_{1c}$, the individuals in C4 also had lower on-trial systolic blood pressure (SBP), diastolic blood pressure, LDL, and total cholesterol (FDR $P < 0.05$) compared with other clinical groups. There were no significant differences between C4 and other clinical groups for HDL, triglycerides, or very-low-density lipoprotein (FDR $P > 0.05$). Importantly, stratification by $HbA_{1c}$ explained a greater proportion of variation for CVD and microvascular risks compared with LDL and SBP, suggesting that it was the primary risk-reducing factor in C4 (Supplementary Fig. 9).

### GWAS

The GWAS did not identify any SNPs associated with membership in the reduced-risk intensive clinical group, C4, versus other clinical groups based on the genome-wide significance threshold ($P < 5 \times 10^{-8}$) (Fig. 3*A* and *B*). However, 45 SNPs reached suggestive significance ($P < 5.0 \times 10^{-6}$), and the most significant genotyped SNPs were located within the Mas1 proto-oncogene (*MAS1*), neural EGFL-like 1 (*NELL1*), and supervillin (*SVIL*). After adjusting for covariates (refSupplementary Table 3), rs220721, located in *MAS1*, was the SNP most significantly associated with C4 membership ($P = 4.34 \times 10^{-7}$) (Supplementary Table 10). The C allele of rs220721 conferred a 1.38-fold increase in likelihood of C4 membership (Fig. 3*A* and *B* and Supplementary Table 10).

### Predictive Model for Intensive Glycemia–Responsive T2D Subtype

Predictive models based on genetic factors, with or without baseline clinical factors, were developed to predict patients likely to be in clinical group C4 and therefore have a reduced risk of MACE, mortality, and other CVD outcomes. Receiver operating characteristic curves for all evaluated models are presented in Fig. 3*C*. Reported results are based on the 33% of the intensive arm withheld from model training ($n = 1{,}169$).

The SCT-PS models outperformed other approaches, and no appreciable advantage was observed with inclusion of baseline clinical features (area under the receiver operating characteristic curve 0.98 vs. 0.99) (Fig. 3*C*). The SCT-PS model had sensitivity, specificity, and balanced accuracy of 95%, 93%, and 94%, respectively, when applied to the withheld test set (Supplementary Table 13).

Next, we evaluated whether the C4 subtype benefited from intensive glycemia treatment rather than displaying reduced risk regardless of treatment strategy. Applying the SCT-PS to identify predicted C4 patients in the intensive arm (in the withheld test set) and in the standard arm, we demonstrated that predicted C4 patients who received intensive glycemia treatment had significantly reduced risk compared with predicted C4 patients receiving standard treatment for MACE (HR 0.45; 95% CI 0.28–0.72; $P = 9.56 \times 10^{-4}$), coronary

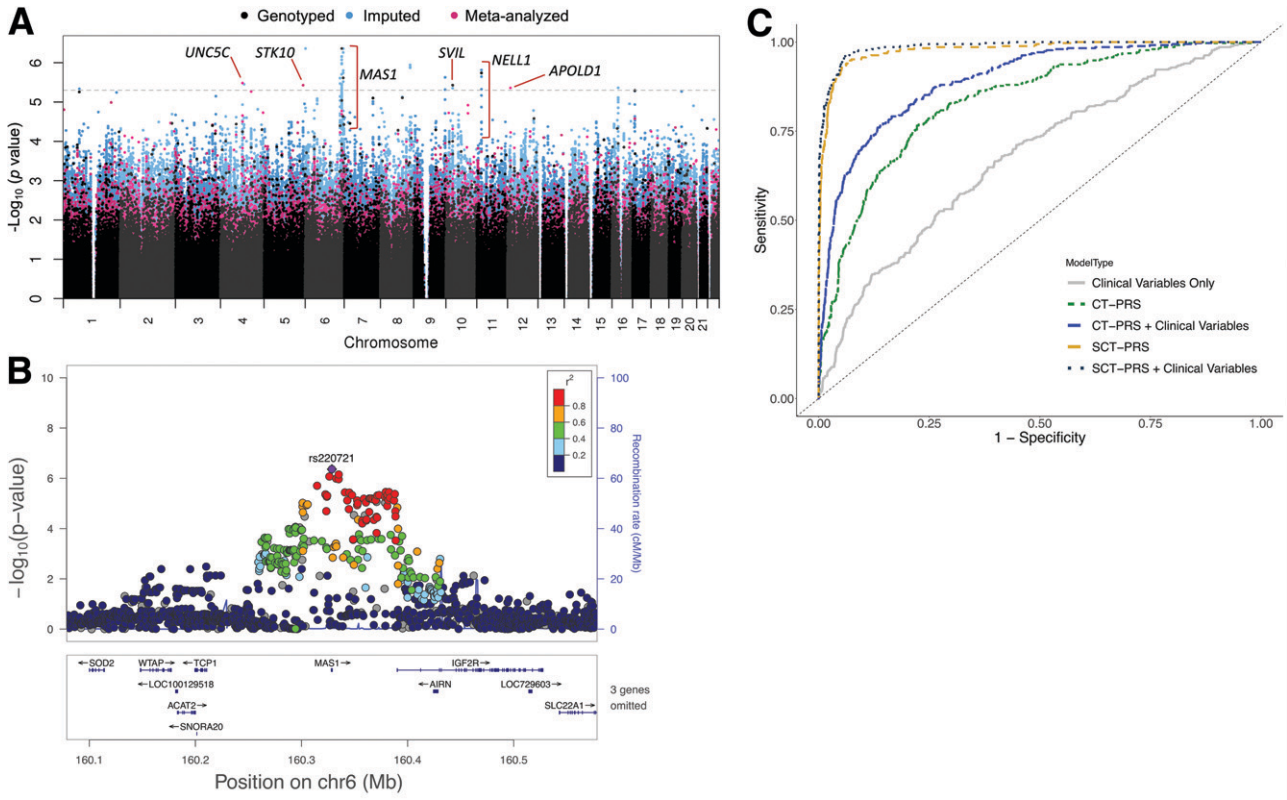


**Figure 3**—GWA analysis of C4. (*A*) Manhattan plot for SNP associations with membership in C4 compared with all other groups. Dashed lines represent thresholds for suggestive significance ($P < 5 \times 10^{-6}$). (*B*) LocusZoom plot of SNPs located in *MAS1*. (*C*) Receiver operating characteristic curve for a logistic regression model containing baseline clinical features only, CT-PS (CT-PRS) only, CT-PS and baseline clinical features, SCT-PS (SCT-PRS) only, and SCT-PS and baseline clinical features. The model combining SCT-PS with baseline clinical features outperformed the other models, with an area under the curve (AUC) of 0.99. However, the SCT-PS only model performed nearly as well (AUC 0.98) and was selected as the best model based on parsimony. All model results can be found in Supplementary Table 13

heart disease (HR 0.52; 0.34–0.79; $P = 0.002$), total mortality (HR 0.53; 0.28–0.99; $P = 0.047$), and nonfatal MI (HR 0.47; 0.26–0.83; $P = 0.01$). Moreover, the risk across all CVD outcomes was significantly reduced in predicted C4 individuals receiving intensive treatment (HR 0.53; 0.40–0.69; $P = 4.02 \times 10^{-6}$) (Fig. 4*A*). Sensitivity analysis and propensity score matching demonstrated that the SCT-PS reduction in risk was stable across sample-size imbalances and potential demographic differences between cohorts (Supplementary Figs. 15 and 16). A comparison with the non-C4 predicted patients can be seen in Supplementary Fig. 17. The SCT-PS outperformed in White patients (HR 0.41; 0.26–0.65; $P = 1.0 \times 10^{-4}$), but significant reductions in risk were also observed in non-White patients (HR 0.71; 0.52–0.97; $P = 0.03$) (Supplementary Figs. 18 and 19). Comparison with our previous 2-SNP PS for cardiovascular mortality in White patients and the SCT-PS with and without clinical variables is shown in Supplementary Figs. 21 and 22, indicating that the SCT-PS outperformed for outcomes other than cardiovascular mortality, particularly when all races were included. Interestingly, the predicted C4 group in the test set failed to reach a statistically significant reduction in microvascular outcomes compared with the predicted C4 group receiving standard treatment ($P < 0.05$) (Supplementary Fig. 20).

### Mendelian Randomization

Mendelian randomization was used to test whether causality could be inferred between the reduction in glycemia and CVD outcomes through an interaction with the SCT-PS and was stratified by individuals randomized to standard and intensive treatments (Supplementary Fig. 14). For individuals receiving standard treatment, the IVW estimate failed to reach statistical significance for any of the individual CVD outcomes ($P > 0.05$), and the meta-analyzed estimate across CVD outcomes also failed to reach statistical significance ($P = 0.08$). However, in individuals randomized to intensive treatment, there was a statistically significant relationship observed between glycemia lowering and reduction in $HbA_{1c}$ through the interaction of the SCT-PS, where MACE, total mortality, CVD mortality, nonfatal MI, and nonfatal stroke reached statistical significance ($P < 0.005$). In the same individuals, the meta-analyzed estimate across all CVD outcomes was also statistically significant ($P = 6.71 \times 10^{-23}$), indicating that there is a causal relationship between glycemia reduction and CVD outcomes through interaction of SCT-PS and treatment (Supplementary Fig. 14).

## CONCLUSIONS

The failure of ACCORD to show benefit from intensive glycemia treatment has significantly affected clinical management of patients with advanced T2D and risk of CVD. However, individual heterogeneity within a complex disease classification may complicate findings from research studies and clinical trials (27–29). Here, we propose that a genetic subtype with T2D exists that is responsive to intensive glycemia treatment, despite the lack of efficacy observed across the overall cohort.

It has been recognized that patients treated intensively in ACCORD who attained a lower $HbA_{1c}$ had reduced risk of death compared to those treated intensively with little $HbA_{1c}$ reduction, albeit this risk reduction was insignificantly lower than with standard treatment (30). However, to our knowledge, this is the first study showing a genetic subgroup (C4) in ACCORD with statistically significant reductions in risk across all CVD outcomes (HR 0.34; 95% CI 0.17–0.67; $P = 2.01 \times 10^{-3}$) (Fig. 2). Significant reductions were observed for each

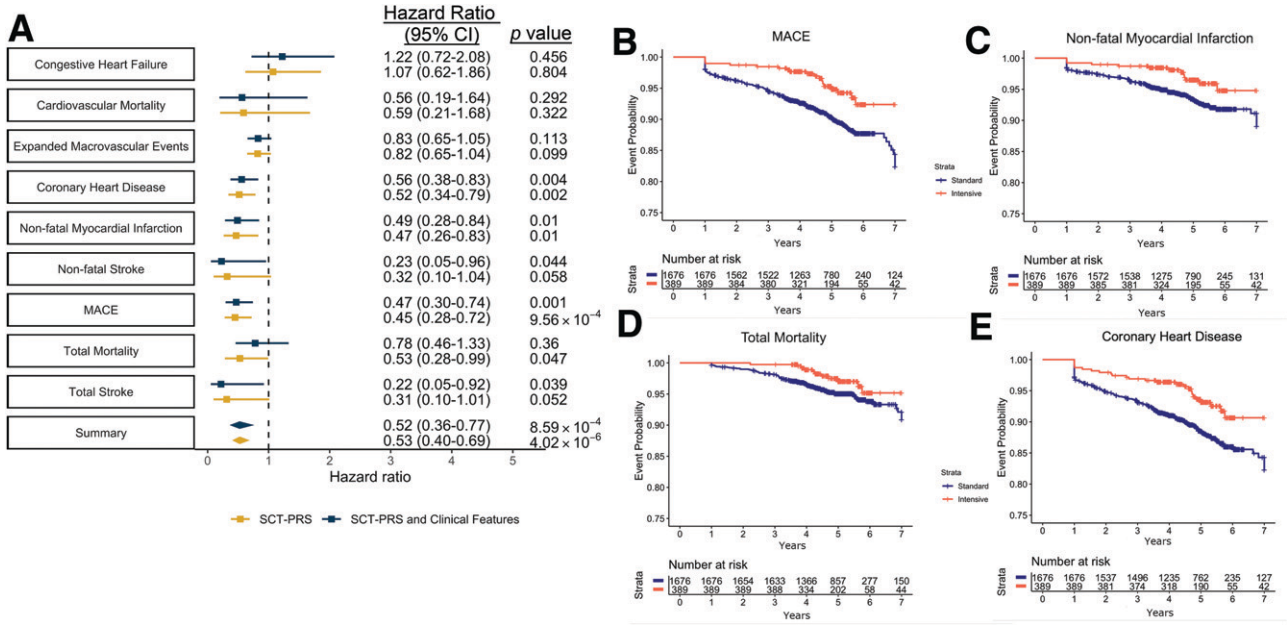


**Figure 4**—Comparison between predicted C4 patients receiving standard vs. intensive glycemia treatment. (*A*) Forest plot of HRs of individuals predicted to be in C4 who received intensive glycemia treatment (from withheld test set) compared with those predicted to be in C4 who received standard treatment. Results from the SCT-PS (SCT-PRS) model with and without baseline clinical factors are shown. Kaplan-Meier curves are shown comparing, in those predicted to be in C4 who received intensive glycemia treatment compared with those predicted to be C4 who received standard treatment, incidence of MACE (*B*), nonfatal MI (*C*), total mortality (*D*), and coronary heart disease (*E*).

individual CVD outcome, with the most significant reductions for total and nonfatal stroke, with HRs of 0.21 ($P = 2.43 \times 10^{-4}$) and 0.021 ($P = 4.67 \times 10^{-4}$), respectively. Cardiac events–nonfatal MI and coronary heart disease also displayed large reductions in risk, with HRs of 0.27 ($P < 9.30 \times 10^{-11}$) and 0.35 ($P = 1.71 \times 10^{-15}$), respectively. Interestingly, congestive heart failure showed the weakest reduction in risk, with an HR of 0.70 (0.50–0.98; $P = 0.04$) (Fig. 2*D*). We also observed a significant reduction across all microvascular outcomes (HR 0.86; 0.77–0.97; $P = 0.002$). However, when the SCT-PS for C4 was applied to the withheld test set of patients, the predicted C4 groups showed a benefit in CVD outcomes (HR 0.53; 0.40–0.69; $P = 4.02 \times 10^{-6}$), but no statistically significant benefit in microvascular outcomes was seen ($P > 0.05$) (Supplementary Fig. 20). It may be that C4 does not benefit from reduced microvascular outcomes with intensive glycemia treatment, or it may be that the withheld test set was underpowered to detect the relatively smaller effect sizes observed for microvascular outcomes (HR 0.86; Supplementary Fig. 5) compared with the effect sizes for CVD outcomes (HR 0.34; Fig. 2*D*).

We investigated whether genetic variants indicated a different genetic profile for individuals responsive to intensive treatment (i.e., C4). Although no SNPs reached genome-wide significance ($P < 5 \times 10^{-8}$), several SNPs reached the predefined threshold for suggestive significance ($P < 5 \times 10^{-6}$) and could be considered candidates for follow-up studies. The association with Rs220721, located in *MAS1*, and C4 membership reached suggestive significance ($P = 4.34 \times 10^{-7}$). Although substantial literature implicates the MAS1 receptor in CVD, and to some extent T2D, these molecular pathways are complex and not well understood (31). The MAS1 receptor is a constitutively active GPCR expressed in many tissues and interacts with angiotensin-(1–7), a MAS1 agonist, and may play a role in ischemic stroke and CVD (31,32). Angiotensin-(1–7) regulates insulin secretion through MAS-dependent cAMP signaling in pancreatic islet cells and reduced hyperglycemia in a rat model of T2D (33,34). Pharmacological antagonism and $Mas^{(-/-)}$ mice displayed significant reductions in insulin secretion, suggesting a potential role for MAS1 in glycemic response (33). SNPs in *MAS1* are in linkage disequilibrium with IGF-2 receptor (*IGF2R*) (Fig. 3*B*), and genetic variation in *IGF2R* has been previously associated with coronary heart disease and shown to affect circulating levels of IGF2R, which has been associated with T2D (35–37). Additional studies to determine the role of genetic variation in *MAS1* or *IGF2R* in relation to intensive glycemia response are needed.

These findings generate new hypotheses regarding the existence of a genetic subtype within T2D (C4) that is responsive to intensive glycemia treatment. Through developing an SCT-PS with 94% balanced accuracy in the withheld test set, we demonstrated that predicted C4 patients who received intensive treatment had a reduction in CVD outcomes versus the predicted C4 patients who received standard treatment (HR 0.53; 95% CI 0.40–0.69; $P = 4.02 \times 10^{-6}$). This overall reduction in CVD outcomes in the predicted C4 group was largely driven by reductions in nonfatal MI and coronary heart disease (Fig. 4*A*). It also seems that $HbA_{1c}$ stratification was the largest contributor to the reduction in CVD outcomes, compared with relatively smaller contributions from LDL and SBP control (Supplementary Fig. 9). Furthermore, Mendelian randomization demonstrated evidence of a causal relationship between the SCT-PS, median on-trial $HbA_{1c}$, and CVD outcomes for patients receiving intensive treatment ($P < 0.001$) but not standard treatment ($P > 0.05$) (Supplementary Fig. 14). Together, this indicates there was indeed benefit to receiving intensive glycemia treatment for some patients (Fig. 4*A*). These results are further supported by sensitivity analysis and propensity score matching, which demonstrated the robust nature of the model (Supplemental Material). Lastly, the SCT-PS seemed to outperform a previous cardiovascular mortality PS to predict, in a racially diverse population, a group that benefits from intensive glycemia treatment (Supplementary Fig. 22). With additional validation, it may be possible to accurately identify patients in this subtype likely to benefit from intensive glycemia treatment.

These results have not been replicated in an independent cohort, which is made difficult by the limited number of trials with available genetic data that have evaluated intensive glycemia treatment in patients with high CVD risk and longer duration of disease. Therefore, it is unknown at this point if these results would be generalizable beyond ACCORD. Since ACCORD, medications such as GLP-1 agonists and SGLT2 inhibitors have become available, and these are also beneficial for CVD risk, which may have different impacts on outcomes for these subtypes. Given safety concerns, it is unlikely that a trial like ACCORD would be repeated. However, we did infer a causal relationship between the SCT-PS, $HbA_{1c}$ response with intensive treatment, and reduction in CVD risk. Other studies have proposed distinct etiological subtypes of T2D (38–40). However, it is not clear if the subtypes here represent distinct T2D etiologies or just subtypes relevant for treatment response. Although additional work is needed to validate these findings, they indicate important developments in this landmark clinical trial with implications for precision medicine and suggest that additional review by advisory groups is warranted.

**Acknowledgments.** The authors thank all of the ACCORD/ACCORDion investigators for their role in conducting and organizing this important study.

**Funding.** A.M., and G.M.-A., and D.M.R. were supported in part by a Clinical and Translational Science Collaborative of Cleveland grant from the National Center for Advancing Translational Sciences component of the National Institutes of Health (KL2TR002547). A.A.M.-R. was supported by intramural funds from the National Institute of Environmental Health Sciences. The ACCORD genome-wide association analysis was supported by National Heart, Lung, and Blood Institute (NHLBI), National Institutes of Health (NIH), grants HL110400 (to A.D.) and HL110380 (to J.B.B., M.J.W., D.M.R., and A.A.M.-R.) and National Institute of Diabetes and Digestive and Kidney Diseases, NIH, grant DK36836 (Advanced Genomics and Genetics Core of the Diabetes Research Center at the Joslin Diabetes Center).

**Duality of Interest.** R.S.Z., M.W.K., and D.M.R. have received research funding from Novo Nordisk. K.M.P. has received research support from Bayer, Merck, and Novo Nordisk; speaker honoraria from AstraZeneca, Corcept Therapeutics, Merck, and Novo Nordisk; and consulting honoraria from AstraZeneca, Bayer, Corcept Therapeutics, Merck, Novo Nordisk, and Sanofi. No other potential conflicts of interest relevant to this article were reported.

**Author Contributions.** A.M. and G.M.-A. performed data analysis and wrote the manuscript. K.M.P., R.S.Z., J.B., M.W.K., H.L.M., A.D., M.J.W., J.B.B., and A.A.M.-R. contributed to the analysis plan, data interpretation, and writing and editing of the manuscript. H.S. contributed to data analysis and quality control. D.M.R. was the principal investigator, performed research, wrote the manuscript, and coordinated the study. All authors critically reviewed the report and approved the final version. D.M.R. is the guarantor of this work and, as such, had full access to all the data in the study and takes responsibility for the integrity and the accuracy of the data analysis.

## References


1. Gerstein HC, Miller ME, Byington RP, et al.; Action to Control Cardiovascular Risk in Diabetes Study Group. Effects of intensive glucose lowering in type 2 diabetes. N Engl J Med 2008; 358:2545–2559
2. Garber AJ, Abrahamson MJ, Barzilay JI, et al. Consensus statement by the American Association of Clinical Endocrinologists and American College of Endocrinology on the comprehensive type 2 diabetes management algorithm–2019 executive summary [published correction appears in Endocr Pract. 2019;25(2):204]. Endocr Pract 2019;25:69–100
3. American Diabetes Association. 6. Glycemic targets: standards of medical care in diabetes—2019. Diabetes Care 2019;42(Suppl. 1):S61–S70
4. Genuth S, Ismail-Beigi F. Clinical implications of the ACCORD trial. J Clin Endocrinol Metab 2012;97:41–48
5. Shah HS, Gao H, Morieri ML, et al. Genetic predictors of cardiovascular mortality during intensive glycemic control in type 2 diabetes: findings from the ACCORD clinical trial. Diabetes Care 2016;39:1915–1924
6. Basu S, Raghavan S, Wexler DJ, Berkowitz SA. Characteristics associated with decreased or increased mortality risk from glycemic therapy among patients with type 2 diabetes and high cardiovascular risk: machine learning analysis of the ACCORD trial. Diabetes Care 2018;41:604–612
7. Hempe JM, Liu S, Myers L, McCarter RJ, Buse JB, Fonseca V. The hemoglobin glycation index identifies subpopulations with harms or benefits from intensive treatment in the ACCORD trial. Diabetes Care 2015;38:1067–1074
8. Riddle MC, Ambrosius WT, Brillon DJ, et al.; Action to Control Cardiovascular Risk in Diabetes investigators. Epidemiologic relationships between A1C and all-cause mortality during a median 3.4-year follow-up of glycemic treatment in the ACCORD trial. Diabetes Care 2010;33:983–990
9. Ismail-Beigi F, Craven T, Banerji MA, et al.; ACCORD trial group. Effect of intensive treatment of hyperglycaemia on microvascular outcomes in type 2 diabetes: an analysis of the ACCORD randomised trial. Lancet 2010;376:419–430
10. Riddle MC, Karl DM. Individualizing targets and tactics for high-risk patients with type 2 diabetes: practical lessons from ACCORD and other cardiovascular trials. Diabetes Care 2012; 35:2100–2107
11. Petitjean F, Forestier G, Webb GI, Nicholson AE, Chen Y, Keogh E. Faster and more accurate classification of time series by exploiting a novel dynamic time warping averaging algorithm. Knowl Inf Syst 2016;47:1–26
12. Rani S, Sikka G. Recent techniques of clustering of time series data: a survey. Int J Comput Appl 2012;52:1–9
13. Mirzaei A, Rahmati M, Ahmadi M. A new method for hierarchical clustering combination. Intell Data Anal 2008;12:549–571
14. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc B 1995; 57:289–300
15. Berkey CS, Hoaglin DC, Antczak-Bouckoms A, Mosteller F, Colditz GA. Meta-analysis of multiple outcomes by regression with random effects. Stat Med 1998;17:2537–2550
16. Viechtbauer W. Conducting meta-analyses in R with the metafor package. J Stat Softw 2010; 36:1–48
17. Riley RD, Price MJ, Jackson D, et al. Multivariate meta-analysis using individual participant data. Res Synth Methods 2015;6:157–174
18. Mychaleckyj JC, Farber EA, Chmielewski J, et al. Buffy coat specimens remain viable as a DNA source for highly multiplexed genome-wide genetic tests after long term storage. J Transl Med 2011;9:91
19. Howie B, Marchini J, Stephens M. Genotype imputation with thousands of genomes. G3 (Bethesda) 2011;1:457–470
20. Delaneau O, Zagury J-F, Marchini J. Improved whole-chromosome phasing for disease and population genetic studies. Nat Methods 2013;10: 5–6
21. Marvel SW, Rotroff DM, Wagner MJ, et al.; ACCORD/ACCORDion investigators. Common and rare genetic markers of lipid variation in subjects with type 2 diabetes from the ACCORD clinical trial. PeerJ 2017;5:e3187
22. Rotroff DM, Yee SW, Zhou K, et al.; MetGen investigators; ACCORD/ACCORDion investigators. Genetic variants in CPA6 and PRPF31 are associated with variation in response to metformin in individuals with type 2 diabetes. Diabetes 2018; 67:1428–1440
23. Privé F, Aschard H, Blum MGB. Efficient implementation of penalized regression for genetic risk prediction. Genetics 2019;212:65–74
24. Privé F, Aschard H, Ziyatdinov A, Blum MGB. Efficient analysis of large-scale genome-wide data with two R packages: bigstatsr and bigsnpr. Bioinformatics 2018;34:2781–2787
25. Tibshirani R. Regression shrinkage and selection via the lasso. J R Stat Soc B 1996;58:267–288
26. Yavorska OO, Burgess S. Mendelian Randomization: an R package for performing Mendelian randomization analyses using summarized data. Int J Epidemiol 2017;46:1734–1739
27. Fisher R, Pusztai L, Swanton C. Cancer heterogeneity: implications for targeted therapeutics. Br J Cancer 2013;108:479–485
28. Ryan J, Fransquet P, Wrigglesworth J, Lacaze P. Phenotypic heterogeneity in dementia: a challenge for epidemiology and biomarker studies. Front Public Health 2018;6:181
29. Manchia M, Cullis J, Turecki G, Rouleau GA, Uher R, Alda M. The impact of phenotypic and genetic heterogeneity on results of genome wide association studies of complex diseases. PLoS One 2013;8:e76295
30. Riddle MC. Effects of intensive glucose lowering in the management of patients with type 2 diabetes mellitus in the Action to Control Cardiovascular Risk in Diabetes (ACCORD) trial. Circulation 2010;122:844–846
31. Karnik SS, Singh KD, Tirupula K, Unal H. Significance of angiotensin 1-7 coupling with MAS1 receptor and other GPCRs to the renin-angiotensin system: IUPHAR Review 22. Br J Pharmacol 2017;174:737–753
32. Jiang T, Gao L, Lu J, Zhang Y-D. ACE2-Ang-(1-7)-Mas axis in brain: a potential target for prevention and treatment of ischemic stroke. Curr Neuropharmacol 2013;11:209–217
33. Sahr A, Wolke C, Maczewsky J, et al. The angiotensin-(1-7)/Mas axis improves pancreatic β-cell function in vitro and in vivo. Endocrinology 2016;157:4677–4690
34. Santos SHS, Giani JF, Burghi V, et al. Oral administration of angiotensin-(1-7) ameliorates type 2 diabetes in rats. J Mol Med (Berl) 2014;92: 255–265
35. Chanprasertyothin S, Jongjaroenprasert W, Ongphiphadhanakul B. The association of soluble IGF2R and IGF2R gene polymorphism with type 2 diabetes. J Diabetes Res 2015;2015:216383
36. van der Harst P, Verweij N. Identification of 64 novel genetic loci provides an expanded view on the genetic architecture of coronary artery disease. Circ Res 2018;122:433–443.
37. MacArthur J, Bowler E, Cerezo M, et al. The new NHGRI-EBI Catalog of published genome-wide association studies (GWAS Catalog). Nucleic Acids Res 2017;45(D1):D896–D901
38. Ahlqvist E, Storm P, Käräjämäki A, et al. Novel subgroups of adult-onset diabetes and their association with outcomes: a data-driven cluster analysis of six variables. Lancet Diabetes Endocrinol 2018;6:361–369
39. Li L, Cheng W-Y, Glicksberg BS, et al. Identification of type 2 diabetes subgroups through topological analysis of patient similarity. Sci Transl Med 2015;7:311ra174
40. Udler MS, Kim J, von Grotthuss M, et al.; Christopher D. Anderson on behalf of METASTROKE and the ISGC. Type 2 diabetes genetic loci informed by multi-trait associations point to disease mechanisms and subtypes: a soft clustering analysis. PLoS Med 2018;15:e1002654