View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Carolina Digital Repository

Characterizing the weight-glycemia phenotypes of type 1 diabetes in youth and young adulthood

Anna R Kahkoska ,[1] Crystal T Nguyen,[2] Xiaotong Jiang,[2] Linda A Adair,[1] Shivani Agarwal,[3] Allison E Aiello,[4] Kyle S Burger,[1] John B Buse,[5] Dana Dabelea,[6,7] Lawrence M Dolan,[8] Giuseppina Imperatore,[9] Jean Marie Lawrence,[10] Santica Marcovina,[11] Catherine Pihoker,[12] Beth A Reboussin,[13] Katherine A Sauder,[6,7] Michael R Kosorok,[2,14] Elizabeth J Mayer-Davis[1,5]

To cite: Kahkoska AR, Nguyen CT, Jiang X, *et al*. Characterizing the weight-glycemia phenotypes of type 1 diabetes in youth and young adulthood. *BMJ Open Diab Res Care* 2020;**8**:e000886. doi:10.1136/bmjdrc-2019-000886

► Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/bmjdrc-2019-000886).

MRK and EJM-D contributed equally.

Received 5 September 2019
Revised 27 December 2019
Accepted 4 January 2020

© Author(s) (or their employer(s)) 2020. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

For numbered affiliations see end of article.

Correspondence to
Dr Anna R Kahkoska;
anna_kahkoska@med.unc.edu

ABSTRACT

Introduction Individuals with type 1 diabetes (T1D) present with diverse body weight status and degrees of glycemic control, which may warrant different treatment approaches. We sought to identify subgroups sharing phenotypes based on *both* weight and glycemia and compare characteristics across subgroups.

Research design and methods Participants with T1D in the SEARCH study cohort (n=1817, 6.0–30.4 years) were seen at a follow-up visit >5 years after diagnosis. Hierarchical agglomerative clustering was used to group participants based on five measures summarizing the joint distribution of body mass index z-score (BMIz) and hemoglobin A1c (HbA1c) which were estimated by reinforcement learning tree predictions from 28 covariates. Interpretation of cluster weight status and glycemic control was based on mean BMIz and HbA1c, respectively.

Results The sample was 49.5% female and 55.5% non-Hispanic white (NHW); mean±SD age=17.6±4.5 years, T1D duration=7.8±1.9 years, BMIz=0.61±0.94, and HbA1c=76±21 mmol/mol (9.1±1.9)%. Six weight-glycemia clusters were identified, including four normal weight, one overweight, and one subgroup with obesity. No cluster had a mean HbA1c <58 mmol/mol (7.5%). Cluster 1 (34.0%) was normal weight with the lowest HbA1c and comprised 85% NHW participants with the highest socioeconomic position, insulin pump use, dietary quality, and physical activity. Subgroups with very poor glycemic control (ie, ≥108 mmol/mol (≥12.0%); cluster 4, 4.4%, and cluster 5, 7.5%) and obesity (cluster 6, 15.4%) had a lower proportion of NHW youth, lower socioeconomic position, and reported decreased pump use and poorer health behaviors (overall p<0.01). The overweight subgroup with very poor glycemic control (cluster 5) showed the highest lipids and blood pressure (p<0.01).

Conclusions There are distinct subgroups of youth and young adults with T1D that share weight-glycemia phenotypes. Subgroups may benefit from tailored interventions addressing differences in clinical care, health behaviors, and underlying health inequity.

INTRODUCTION

As the prevalence of obesity increases worldwide, recent data have shown that

Significance of this study

What is already known about this subject?

- Obesity is an increasing issue among youth and young adults with type 1 diabetes.
- However, there is considerable heterogeneity in both weight status and glycemic control across the population, which challenges the use of universal or 'one-size-fits-all' recommendations to optimize weight alongside glycemic control.

What are the new findings?

- We used data-driven methods and focused on cohort of youth and young adults with etiologically defined type 1 diabetes to characterize heterogeneity in the clinical phenotypes based on both weight status and glycemia.
- We found six distinct subgroups of youth and young adults with type 1 diabetes sharing weight-glycemia clinical phenotypes based on measures of body mass index z-score (BMIz) and hemoglobin A1c (HbA1c).
- We found differences in sociodemographic characteristics, clinical factors, and behavioral correlates across subgroups, suggesting that the subgroups with the poorest glycemic control and highest weight are more likely to be non-white, have lower measures of socioeconomic status, and report poor health behaviors.

How might these results change the focus of research or clinical practice?

- We show that the weight-glycemia phenotypes of diabetes present with varying combinations of BMIz and HbA1c with differences in other individual characteristics, including sociodemographic, clinical, and behavioral features. These subgroups of individuals may benefit from similar therapeutic strategies and can be targeted more efficiently as groups for clinical recommendations and interventions to co-optimize the outcomes of weight and glycemia.
- The results implicate a precision medicine approach to address the complicated physiologic relationships between weight and glycemic control across phenotypic subgroups.

the prevalence of overweight and obesity in youth and young adults with type 1 diabetes (T1D) is even higher than in the general population.[1 2] Excess adiposity increases the risk of cardiovascular disease later in life which is already elevated up to 10-fold in persons with T1D. Therefore, there are early efforts to integrate weight management with the complexities of routine T1D care including both weight loss and prevention of overweight and obesity.[3]

However, the rising numbers of youth and young adults with T1D who are overweight or obese have also contributed to the heterogeneity in the T1D patient population. Given that appropriate treatment algorithms may vary markedly across the broad spectrum of body weight and glycemia,[4] the T1D patient population is a good candidate for precision medicine, which matches interventions to different subgroups of patients expected to show a positive response.[5 6] Epidemiological evidence suggests population-level associations between body mass index z-score (BMIz) and hemoglobin A1c (HbA1c)[2 7 8]; however, surprisingly little is known about how weight status and glycemic control are codistributed across the population and interact to form more nuanced clinical phenotypes of T1D. The weight-glycemia phenotype may confer information about goals for treatment and effectiveness of specific therapeutic strategies for optimizing outcomes simultaneously, especially given that weight gain may be an unintended consequence of intensive insulin therapy in some individuals.[9]

Previous work used data-driven approaches to stratify adults with T1D and type 2 diabetes (T2D) into subgroups based on six 'raw' clinical and physiologic features.[10] The resulting subgroups showed differences in progression of T2D and risk for complications.[10] However, few studies have characterized heterogeneity in weight and glycemia *within* the etiologic diagnosis of T1D despite the rising prevalence of overweight and obesity. Therefore, our objective was to use data from a large, diverse cohort of youth and young adults with T1D to identify and characterize subgroups sharing clinical phenotypes of T1D based on weight status, measured by BMIz, and glycemic control, measured by HbA1c.

METHODS

Study population

The SEARCH for Diabetes in Youth study began in 2000 with an overarching objective to describe the incidence and prevalence of youth-onset diabetes in the USA by age, sex, and race/ethnicity. Youth and young adults with diabetes diagnosed <20 years of age ('youth') were identified from a population-based incidence registry network at five US sites (South Carolina; Cincinnati, Ohio and surrounding counties; Colorado with southwestern Native American sites; Seattle, Washington and surrounding counties; and Kaiser Permanente Health Plan, Southern California).[11] A subset of participants with newly diagnosed diabetes between 2002 and 2006 and in 2008 were recruited for a follow-up 'cohort' visit between 2012-2015 if they had attended a baseline visit and had greater than 5 years of diabetes duration. The subset of youth who were included in the SEARCH cohort visit were not significantly different from all other SEARCH youth diagnosed between the years 2002 and 2008 in terms of average diabetes-onset age, demographics, or clinical measures.[12]

Inclusion criteria for this report consisted of incident cases of T1D between 2002 and 2006 and 2008 who attended the SEARCH cohort visit (n=2004). Diabetes type for these analyses was based on an etiological classification using diabetes autoantibodies and estimated insulin sensitivity score (euglycemic clamp-validated equation including waist circumference, HbA1c and triglyceride levels) from the baseline visit.[13] Participants who reported inconsistent or implausible levels of severe hypoglycemia (n=4) and who were missing BMIz or HbA1c measures at the cohort visit (n=183) were excluded.

Research visits

Trained personnel administered questionnaires; measured height, weight, and blood pressure; and obtained fasting blood samples. BMI was defined as weight (kilograms) divided by height (meters^2) and converted to a z-score based on US growth reference data.[14] To facilitate study across youth and young adults, BMIz for individuals >20 years was estimated assuming an age of 20 years (the maximum age represented in the growth reference); this approach has been operationalized in previous SEARCH studies[12 15] and elsewhere.[16] A blood draw occurred after an 8-hour overnight fast, and medications, including short-acting insulin, were withheld the morning of the visit.

Laboratory measures

Blood samples were obtained under conditions of metabolic stability, defined as no episodes of diabetic ketoacidosis in the preceding month and the absence of fever and acute infections. They were processed locally and shipped within 24 hours to the central laboratory (Northwest Lipid Metabolism and Diabetes Research Laboratories, Seattle, WA). HbA1c was measured by a dedicated ion exchange high-performance liquid chromatography instrument (TOSOH Bioscience, San Francisco, CA).

Other measures

Demographic measures included sex and self-reported race and ethnicity, categorized as non-Hispanic white, non-Hispanic black, Hispanic, Asian, Native American, Pacific Islander, and other. Highest education by either parent was classified as less than high school degree, high school graduate, some college through associate degree, and bachelor's degree or more. Annual household income was classified as >$75 000, $50 000–75 000, $25 000–49 999 and <$25 000. Healthcare access was

measured by health insurance type, classified as none, private, Medicaid, or other.

Insulin regimen was classified as pumps, long-acting with short/rapid-acting insulin injections with ≥3 injections per day, and any other form of multiple or singular daily injections. Self-reported frequency of self-monitoring of blood glucose (SMBG) was categorized as <1, 1–3, and >3 times per day. Diabetes care provider was classified as pediatric endocrinologist, adult endocrinologist, and all other healthcare providers (pediatrician, family practice doctor, nurse practitioner, and so on).

Depressive symptoms were measured using the Center for Epidemiologic Studies Depression Scale (CESD).[17] Quality of life was measured using the age-appropriate Center for Pediatric Quality of Life Inventory (PedsQL).[18] The CESD and PedsQL were modeled as continuous variables. Physical activity and screen time were assessed using questionnaires. High physical activity was classified as vigorous activity 3–7 days weekly. High screen time was classified as 2 or more hours of screen time per day. Data from a self-administered, validated food frequency questionnaire were available for 1643 participants who were aged >10 years.[19] Dietary quality was assessed by adherence to the Dietary Approaches to Stop Hypertension diet using an index score ranging from 0 to 80.[20]

Statistical methods

We used cluster analysis to identify and characterize subgroups sharing clinical phenotypes of T1D based on weight status and glycemic control. Considerations of the clustering approach are described in detail in the online supplementary appendix. As opposed to unsupervised cluster analysis, where there is no outcome measure or data labels, we wished to perform a semisupervised cluster analysis guided by the two outcomes of interest. A challenge in identifying supervised clusters is that noise in a given outcome may obscure true subgroups of clinical interest.[21] Therefore, rather than cluster individuals based on the observed values of BMIz and HbA1c at the cohort visit, we employed a novel, semisupervised clustering technique to group individuals in SEARCH by five measures of the *joint distribution* of BMIz and HbA1c at the cohort visit: the means and variances of BMIz and HbA1c and their covariance. The five values summarizing the joint distribution of BMIz and HbA1c were predicted for each individual using reinforcement learning trees (RLT), a type of tree-based machine learning technique,[22] and 28 other characterizing variables that were available for each patient (*X*-variables). The 28 *X*-variables were chosen to capture a breadth of individual characteristics available at the cohort visit including sociodemographic, clinical, anthropometric, laboratory, psychosocial and behavioral measures (see online supplementary table S1). Any given *X*-variable was missing from at most 12% of individuals and imputed by a multiple imputation method, missForest.[23] The resulting RLT-estimated outcomes represent smoothed outcome measures, denoised by the *X*-variables, which maintain the individual-level signal with reduced noise or measurement error.[22] Of note, the 28 *X*-variables were only used to predict measures of the joint distribution of BMIz and HbA1c for each individual and were not used directly in the cluster analysis.

The five clustering variables (RLT-predicted means and variances of BMIz and HbA1c and their covariance) were standardized and a hierarchical clustering algorithm with Ward's D2 method and Euclidean distance was applied. The number of clusters was chosen using the NbClust package in R[24] and restricted to considering between four and nine clusters. The smallest cluster was restricted to greater than 50 people for adequate statistical power (>85%) to detect small to medium effects in overall cluster comparisons.[25] Clustering stability was assessed by sequentially omitting individual clusters, one at a time, and evaluating the agreement of the remaining clusters using the adjusted Rand index.[26] For more information on imputation methods, RLT parameters, clustering methods, stability assessments, and additional analyses, see online supplementary appendix.

Clusters were ordered by increasing weight status and then by increasing mean HbA1c. Clusters were named based on mean BMIz and HbA1c using traditional clinical cut-point for ease of interpretation. Cluster weight status was classified as underweight (mean BMIz <−1.64), normal weight (mean BMIz −1.64 to <1.04), overweight (mean BMIz 1.04 to <1.64), and obesity (mean BMIz ≥1.64), corresponding to <5th, 5th to <85th, 85th to <95th, and ≥95th percentiles for age and sex, respectively.[27] Cluster glycemic control was defined as good (mean HbA1c <58 mmol/mol (<7.5%)), moderate (mean HbA1c 58 to <75 mmol/mol (7.5% to <9.0%)), poor (mean HbA1c 75 to <108 mmol/mol (9.0% to <12.0%)), and very poor (mean HbA1c ≥108 mmol/mol (≥12.0%)).[28] Unfortunately, this method of cluster nomenclature cannot represent the weight status and glycemic control of each individual within cluster and instead was selected to facilitate cluster-level phenotypic interpretation and comparisons thereof.

The cross-sectional correlates of each cluster from the follow-up visit were summarized using descriptive statistics. The chosen approach and statistical criteria were part of an a priori specified analysis plan. Data are presented as mean±SD. Overall tests of difference as well as pairwise comparisons were carried out using analysis of variance, t-tests, and χ^2 tests, or Fisher's exact tests, where appropriate. We accounted for multiple comparisons in (1) overall tests of difference, and (2) post hoc pairwise comparisons between individual clusters. Overall tests were corrected via a Bonferroni adjustment using the total number of tests as the denominator. For pairwise comparisons, we report q values, which control for the positive false discovery rate[29] (see online supplementary appendix). P values and q values were evaluated at the 0.05 significance level.

Additional analyses explored clusters based on the observed or 'raw' measures of BMIz and HbA1c for

comparison to the clusters based on the joint distribution of BMIz and HbA1c, denoted Y-clusters. These analyses are described in detail in the online supplementary appendix, see *Additional Analyses*. All data analyses were performed using the statistical analysis software package R V.3.4.1 and SAS V.9.4 (SAS Institute).

RESULTS

The study included 1817 individuals with T1D, with a mean age of 17.6 (6.0–30.4 years) and a mean T1D duration of 7.8 years (table 1). Six weight-glycemia phenotypic clusters were identified based on measures of the joint distribution of BMIz and HbA1c (figure 1).

Based on mean measures, cluster 1 (n=618, 34.0%) was normal weight with moderate glycemic control (mean BMIz 0.59±0.59, mean HbA1c 61±12 mmol/mol (7.7%±1.1%)); although the mean HbA1c was classified as moderate, this subgroup had the highest proportion of youth (38%) with HbA1c <58 mmol/mol (7.5%). Cluster 2 (n=195, 10.7%) was also classified as normal weight with moderate glycemic control but showed a slightly lower mean BMIz and higher mean HbA1c than cluster 1 (mean BMIz −0.68±0.66, mean HbA1c 68±10 mmol/mol (8.4%±0.9%)). Cluster 3 (n=509, 28.0%) was normal weight with poor glycemic control (mean BMIz 0.56±0.62, mean HbA1c 86±12 mmol/mol (10.0%±1.1%)). Cluster 4 (n=79, 4.4%) was normal weight with very poor glycemic control (mean BMIz −1.05±0.83, mean HbA1c 113±15 mmol/mol (12.5%±1.4%)). Cluster 5 (n=137, 7.5%) was overweight with very poor glycemic control (mean BMIz 1.29±0.69, mean HbA1c 109±15 mmol/mol (12.1%±1.5%)). Cluster 6 (n=279, 15.4%) was those with obesity and moderate glycemic control (mean BMIz 1.74±0.42, mean HbA1c 70±11 mmol/mol (8.6%±1.0%)). Figure 2 depicts the density distribution of BMIz and HbA1c within each weight-glycemia cluster.

Cluster 1 (normal weight with moderate glycemic control) was the largest cluster, comprising 34.0% of the sample. Based on mean BMIz and HbA1c measures closest to clinical targets, this group was selected as the referent group for individual comparisons. Tables 1 and 2 depict the sociodemographic characteristics and the diabetes care, psychosocial, and behavioral factors according to the six weight-glycemia clusters. Participants in cluster 1 were 46% female, 88% non-Hispanic white and were characterized by the highest measures of socioeconomic position, including 74% having parents with a bachelor's degree or more and 85% with private health insurance. This group also had the highest prevalence of insulin pump use and frequency of SMBG, the lowest level of depressive symptoms, the highest quality of life, the highest dietary quality, and the highest levels of physical activity (overall p<0.001).

One cluster emerged with poor glycemic control (cluster 3, normal weight with poor glycemic control) and two with mean HbA1c >108 mmol/mol (12.0%) (cluster 4, normal weight with very poor glycemic control; cluster 5, overweight with very poor glycemic control). The latter two were the smallest subgroups, comprising approximately 4% and 8% of the sample, respectively. Compared with cluster 1, these clusters included a significantly higher proportion of non-white individuals (58% and 50%), with the highest proportion of non-Hispanic black individuals in cluster 4 (28%) and highest prevalence of Hispanic individuals in cluster 5 (23%) (q<0.001). Clusters 4 and 5 also had lower measures of socioeconomic position and significantly lower insulin pump use and less frequent SMBG (table 1; all q<0.001). Approximately 38% of individuals in cluster 4% and 41% in cluster 5 experienced an episode of diabetic ketoacidosis in the past 6 months compared with 10% in cluster 1. Finally, clusters 4 and 5 were characterized by higher depressive symptoms, lower quality of life, poorer dietary quality, and greater a proportion of high screen time (all q<0.001).

Two clusters were classified as overweight and having obesity (cluster 5, overweight with very poor glycemic control; cluster 6, obesity with moderate glycemic control). Compared with cluster 1, both subgroups contained a higher proportion of females (66% in cluster 5 and 55% in cluster 6) and non-white youth. Cluster 6 was also characterized by moderately lower measures of socioeconomic position compared with cluster 1 (all q<0.001).

Additional post hoc pairwise comparisons were made between cluster 5 and cluster 6, the two overweight/obesity subgroups (online supplementary tables S2–S4). Compared with cluster 6 (obesity with moderate glycemic control), cluster 5 (overweight with poor glycemic control) comprised more female (q=0.028) and non-white participants (q<0.001). Individuals in cluster 5 were older at the follow-up visit (q<0.001) and had lower socioeconomic position (q<0.001 for parental education, income, and insurance type) with no significant differences in diabetes duration (p=0.15). These participants were also less likely to use an insulin pump or report frequent SMBG (q<0.001). There was a higher prevalence of high screen time in cluster 5 (p=0.001) with no significant differences in physical activity (q=0.34).

Table 3 depicts other clinical measures across the weight-glycemia clusters. Compared to cluster 1, clusters 3, 4, 5, and 6 showed significantly higher blood lipid levels. Cluster 5 showed higher total cholesterol, low-density lipoprotein cholesterol, and triglycerides compared with both the referent cluster 1 and cluster 6 (obesity). This group also exhibited higher mean systolic and diastolic blood pressure levels than cluster 1 and a higher mean diastolic blood pressure compared with cluster 6 (all q<0.001).

DISCUSSION

In a large, diverse cohort of youth and young adults with T1D, we found evidence of subgroups that share distinct weight-glycemia phenotypes including varying combinations across BMIz and glycemic control parameters. None of the clusters that were identified had a mean HbA1c <58 mmol/mol to be classified as good glycemic control, underscoring that youth and young adults with

Table 1 Sociodemographic characteristics according to weight-glycemia phenotype clusters 1–6

Characteristics, mean (SD) or n (%)	All n=1817	Weight-glycemia cluster						P value†
		Cluster 1 n=618 (34.0%)	Cluster 2 n=195 (10.7%)	Cluster 3 n=509 (28.0%)	Cluster 4 n=79 (4.4%)	Cluster 5 n=137 (7.5%)	Cluster 6 n=279 (15.4%)	
Weight-glycemia								
BMIz	0.61 (0.94)	0.59 (0.59)	−0.68 (0.65)**	0.56 (0.62)	−1.05 (0.83)**	1.29 (0.69)**	1.74 (0.42)**	<0.0001
HbA1c (mmol/mol)	76 (21)	61 (12)	68 (10)	86 (12)	113 (15)	109 (15)	70 (11)	<0.0001
HbA1c (%)	9.1 (1.9)	7.7 (1.1)	8.4 (0.9)**	10.0 (1.1)**	12.5 (1.4)**	12.1 (1.5)**	8.6 (1.0)**	<0.0001
Weight status‡								<0.0001
Underweight	36 (2.0)	0 (0.0)	17 (8.7)	1 (0.2)	18 (22.8)	0 (0.0)	0 (0.0)	
Normal weight	1152 (63.4)	467 (75.6)	177 (90.8)**	390 (76.6)	61 (77.2)**	47 (34.3)**	10 (3.6)**	
Overweight	390 (21.5)	138 (22.3)	1 (0.5)**	105 (20.6)	0 (0.0)**	42 (30.7)**	104 (37.3)**	
Obesity	239 (13.2)	13 (2.1)	0 (0.0)**	13 (2.6)	0 (0.0)**	48 (35.0)**	165 (59.1)**	
Glycemic control§								<0.0001
Good	306 (16.8)	237 (38.4)	31 (15.9)**	2 (0.4)**	0 (0.0)**	0 (0.0)**	36 (12.9)**	
Moderate	656 (36.1)	312 (50.5)	112 (57.4)**	86 (16.9)**	0 (0.0)**	2 (1.5)**	155 (51.6)**	
Poor	704 (38.8)	69 (11.2)	52 (26.7)**	389 (76.4)**	28 (35.4)**	67 (48.9)**	99 (35.5)**	
Very poor	151 (8.3)	0 (0.0)	0 (0.0)	32 (6.3)**	51 (64.6)**	68 (49.6)**	0 (0.0)**	
Demographic characteristics								
Age at follow-up (years)	17.6 (4.5)	17.6 (5.0)	16.8 (4.6)*	17.3 (4.1)	19.0 (4.0)*	18.8 (3.6)**	17.7 (4.3)	0.008
Age at diagnosis (years)	9.8 (4.1)	9.8 (4.5)	9.3 (4.0)	9.5 (3.8)	11.1 (3.9)*	10.7 (3.3)*	9.7 (3.9)	0.014
Diabetes duration (months)	93.3 (22.8)	92.2 (23.0)	88.8 (23.1)*	94.1 (22.6)	95.2 (22.4)	97.3 (22.1)*	95.1 (22.9)*	0.050
Female	898 (49.4)	282 (45.6)	62 (31.8)**	286 (56.2)**	25 (31.7)*	90 (65.7)**	153 (54.8)*	<0.0001
Race/ethnicity¶								<0.0001
Non-Hispanic white	1380 (76.0)	542 (87.7)	167 (85.6)	351 (69.0)**	46 (58.2)**	69 (50.4)**	205 (73.5)**	
Non-Hispanic black	173 (9.5)	17 (2.8)	7 (3.6)	69 (13.6)**	22 (27.9)**	32 (23.4)**	26 (9.3)**	
Hispanic	222 (12.2)	45 (7.3)	19 (9.7)	75 (14.7)**	11 (13.9)**	31 (22.6)**	41 (14.7)**	
Asian Pacific Islander	28 (1.5)	12 (1.9)**	2 (1.0)**	9 (1.8)**	0 (0.0)**	2 (1.5)**	3 (1.1)**	
Native American	9 (0.5)	1 (0.2)**	0 (0.0)**	2 (0.4)**	0 (0.0)**	3 (2.2)**	3 (1.1)**	
Other	4 (0.2)	0 (0.0)**	0 (0.0)**	3 (0.6)**	0 (0.0)**	0 (0.0)**	1 (0.4)**	
Socioeconomic position								
Parental education								<0.0001
Bachelor's degree or more	956 (52.6)	459 (74.3)	143 (73.3)	182 (35.8)**	19 (24.0)**	33 (24.1)**	120 (43.0)**	
Some college through associate degree	578 (31.8)	121 (19.6)	40 (20.5)	217 (42.6)**	32 (40.5)**	61 (44.5)**	107 (38.4)**	
High school degree	214 (11.8)	32 (5.2)	9 (4.6)	86 (16.9)**	20 (25.3)**	28 (20.4)**	39 (14.0)**	

Continued

Table 1 Continued

Characteristics, mean (SD) or n (%)	All n=1817	Weight-glycemia cluster						
		Cluster 1 n=618 (34.0%)	Cluster 2 n=195 (10.7%)	Cluster 3 n=509 (28.0%)	Cluster 4 n=79 (4.4%)	Cluster 5 n=137 (7.5%)	Cluster 6 n=279 (15.4%)	P value†
Less than high school degree	69 (3.8)	6 (1.0)	3 (1.5)	24 (4.7)**	8 (10.1)**	15 (11.0)**	13 (4.7)**	
Household Income								<0.0001
>$75 000	682 (37.5)	320 (51.8)	98 (50.2)	138 (27.1)**	7 (7.6)**	22 (16.1)**	98 (25.1)**	
$50 000–$75 000	271 (14.9)	99 (16.0)	37 (19.0)	57 (11.2)**	6 (7.6)**	22 (16.1)**	50 (17.9)**	
$25 000–$49 000	297 (16.4)	71 (11.5)	26 (13.3)	110 (21.6)**	21 (26.6)**	20 (14.6)**	49 (17.6)**	
<$25 000	277 (15.2)	4 (7.4)	16 (8.2)	103 (20.2)**	25 (31.7)**	47 (34.3)**	40 (14.3)**	
Don't know/refused to answer	290 (16.0)	82 (13.3)	18 (9.2)	101 (19.8)**	21 (26.6)**	26 (19.0)**	42 (15.1)**	
Health insurance type								<0.0001
Private	1309 (72.0)	527 (85.3)	152 (78.0)*	326 (64.1)**	38 (48.1)**	64 (46.7)**	202 (72.4)**	
Medicaid/Medicare	360 (19.8)	53 (8.6)	28 (14.4)*	140 (27.5)**	33 (41.8)**	47 (34.3)**	59 (21.2)**	
None	65 (3.6)	10 (1.6)	3 (1.5)*	24 (4.7)**	5 (6.3)**	16 (11.7)**	7 (2.5)**	
Other	83 (4.6)	28 (4.5)	12 (6.2)*	19 (3.7)**	3 (3.8)**	10 (7.3)**	11 (3.9)**	

All measures are from the cohort visit, unless stated otherwise.
*Significant pairwise comparison to cluster 1, where q<0.05. **Significant pairwise comparison to cluster 1, where q<0.001. Controlled for the positive false discovery rate.
†Bonferroni-corrected p value for overall test of difference, based on use of analysis of variance (ANOVA), χ^2 or Fisher's exact test as appropriate.
‡Weight status defined based on BMIz. Underweight was defined as cluster mean BMIz <−1.64 corresponding to the 5th percentile for age and sex. Normal weight was defined as cluster mean BMIz ≥−1.64 and <1.04, corresponding to ≥ the 5th and <85th percentiles for age and sex. Overweight was defined as cluster mean BMIz ≥1.04 and <1.64, corresponding to ≥85th percentile and <95th percentile for age and sex. Obesity was defined as cluster mean BMIz ≥1.64 corresponding to ≥95th percentile for age and sex.
§Glycemic control was based on HbA1c and defined as good (mean HbA1c <58 mmol/mol (<7.5%)), moderate (mean HbA1c 58 to <75 mmol/mol (7.5% to <9.0%)), poor (mean HbA1c 75 to <108 mmol/mol (9.0% to <12.0%)), and very poor (mean HbA1c ≥108 mmol/mol (≥12.0%)).
¶Self-reported race and ethnicity were collected using 2000 US Census questions.
BMIz, body mass index z-score; HbA1c, hemoglobin A1c.

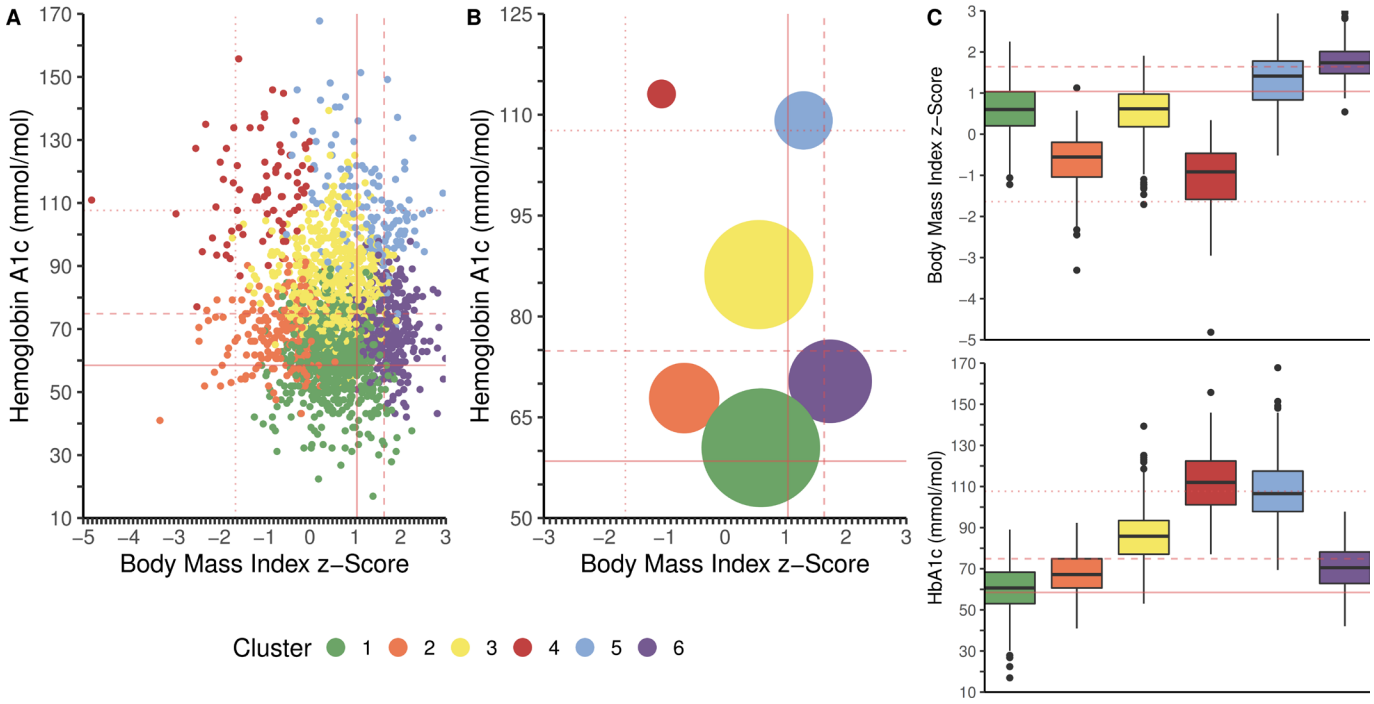

Figure 1 Weight-glycemia phenotypic clusters from the SEARCH for Diabetes in Youth study. Participants were clustered based on the joint distribution of body mass index z-score (BMIz) and hemoglobin A1c (HbA1c) at the 5+ years' cohort visit of the SEARCH study. (A) Scatter plot by BMIz and HbA1c; each point represents an individual. (B) Bubble plot by BMIz and HbA1c; size of circle represents number of individuals within the cluster. (C) Box and whisker plot for BMIz and HbA1c. On the x-axis, the dotted line denotes the BMIz cut-off for underweight (BMIz <–1.64, corresponding to <5th percentile for age and sex), the solid line denotes BMIz cut-off for overweight (BMIz ≥1.04, corresponding to ≥85th percentile for age and sex), and the dashed lined denotes the BMIz cut-off for obesity (BMIz ≥1.64, corresponding to ≥95th percentile for age and sex). On the y-axis, the solid line denotes HbA1c cut-off for moderate glycemic control (HbA1c ≥7.5% (58 mmol/mol)), the dashed line denotes the HbA1c cut-off for poor glycemic control (HbA1c ≥75 mmol/mol (9.0%)), and the dotted line denotes the HbA1c cut-off for very poor glycemic control (HbA1c ≥108 mmol/mol (12.0%)). Clusters include: cluster 1 (n=618, 34.0%): normal weight with moderate glycemic control (mean BMIz 0.59±0.59, mean HbA1c 61±12 mmol/mol (7.7%±1.1%)); cluster 2 (n=195, 10.7%): normal weight with moderate glycemic control (mean BMIz –0.68±0.66, mean HbA1c 68±10 mmol/mol (8.4%±0.9%)); cluster 3 (n=509, 28.0%): normal weight with poor glycemic control (mean BMIz 0.56±0.62, mean HbA1c 86±12 mmol/mol (10.0%±1.1%)); cluster 4 (n=79, 4.4%): normal weight with poor glycemic control (mean BMIz –1.05±0.83, mean HbA1c 113±15 mmol/mol (12.5%±1.4%)); cluster 5 (n=137, 7.5%): overweight with poor glycemic control (mean BMIz 1.29±0.69, mean HbA1c 109±15 mmol/mol (12.1%±1.5%)); cluster 6 (n=279, 15.4%): obesity with moderate glycemic control (mean BMIz 1.74±0.42, mean HbA1c 70±11 mmol/mol (8.6%±1.0%)).

T1D are not meeting the targets put forward by the American Diabetes Association and International Society for Pediatric and Adolescent Diabetes.[30] Based on mean measures, four clusters were largely normal weight, with the remaining two clusters classified as overweight and having obesity, although there were individuals across all weight status categories who were captured in each cluster based on measures of the joint distribution. Examination of the latter two subgroups reveals that while overweight and poor glycemic control can co-occur in young people with diabetes (ie, the weight-glycemia phenotype of cluster 5, comprising 8% of the sample), obesity is not always associated and does not necessarily account for those with poor or very poor glycemic control (ie,

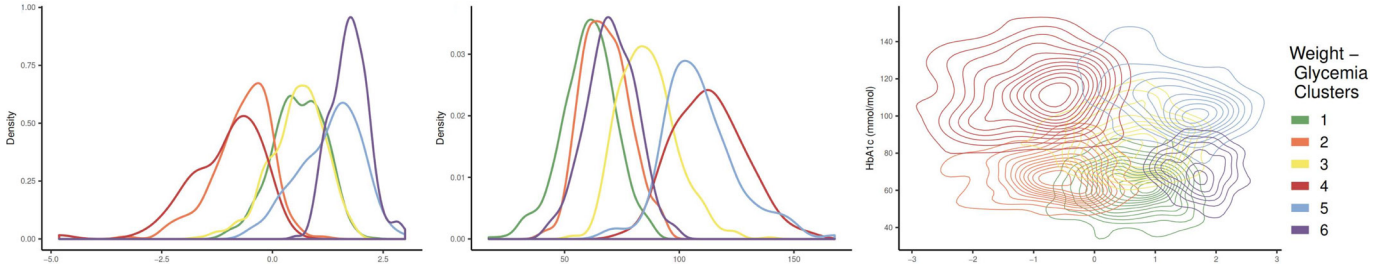

Figure 2 Density distribution plots of body mass index z-score (BMIz) and hemoglobin A1c (HbA1c) by weight-glycemia phenotypic clusters. From left to right: density distribution of BMIz by weight-glycemia cluster, density distribution of HbA1c by weight-glycemia cluster, density distribution plot of BMIz and HbA1c by weight-glycemia cluster.

Table 2 Diabetes care, psychosocial, and behavioral factors according to weight-glycemia clusters 1–6

Characteristics, mean (SD) or n (%)	All n=1817	Weight-glycemia cluster						P value†
		Cluster 1 n=618 (34.0%)	Cluster 2 n=195 (10.7%)	Cluster 3 n=509 (28.0%)	Cluster 4 n=79 (4.4%)	Cluster 5 n=137 (7.5%)	Cluster 6 n=279 (15.4%)	
Diabetes care factors								
Insulin regimen								<0.0001
Pump	1036 (57.0)	446 (72.2)	125 (64.1)*	246 (48.3)**	22 (27.9)**	29 (21.2)**	168 (60.2)**	
Long+short/rapid insulin, 3+ times/day	341 (18.8)	84 (13.6)	38 (19.5)*	100 (19.7)**	25 (31.7)**	44 (32.1)**	50 (17.9)**	
Long+other combination‡	440 (24.2)	88 (14.2)	32 (16.4)*	163 (32.0)**	32 (40.5)**	64 (46.7)**	61 (21.9)**	
Insulin dose (daily units/kg)	0.86 (0.38)	0.80 (0.40)	0.78 (0.31)*	0.90 (0.34)**	1.00 (0.42)**	1.01 (0.48)**	0.84 (0.32)*	<0.0001
Blood glucose monitoring frequency								<0.0001
<1/day	107 (5.9)	15 (2.4)	8 (4.1)*	39 (7.7)**	9 (11.4)**	25 (18.3)**	11 (3.9)**	
2–4×/day	501 (27.6)	90 (14.6)	46 (23.6)*	197 (38.7)**	40 (50.6)**	62 (45.3)**	66 (23.7)**	
>4×/day	1209 (66.5)	513 (83.0)	141 (72.3)*	273 (53.6)**	30 (38.0)**	50 (36.5)**	202 (72.4)**	
Acute complications (past 6 months)§								
1+ severe hypoglycemic episode	130 (7.2)	56 (9.1)	10 (5.1)	28 (5.5)*	4 (5.1)	6 (4.4)	26 (9.3)	0.273
1+ diabetic ketoacidosis episode	332 (18.3)	61 (9.9)	25 (12.8)	124 (24.4)**	30 (38.0)**	56 (40.9)**	36 (12.9)	<0.0001
Diabetes care provider								0.507
Pediatric endocrinologist	1007 (55.4)	332 (53.7)	108 (55.4)	303 (58.5)*	40 (50.6)	67 (48.9)	157 (56.5)	
Adult endocrinologist	344 (18.9)	136 (22.0)	31 (15.9)	82 (16.1)*	15 (19.0)	26 (19.0)	54 (19.4)	
All other¶	466 (25.6)	150 (24.3)	56 (28.7)	124 (24.4)*	24 (30.4)	44 (31.1)	68 (24.4)	
Psychosocial factors								
Depressive symptoms (CESD score)††	10.1 (8.7)	7.8 (7.4)	8.3 (7.8)	12.1 (8.8)**	12.7 (11.0)**	15.3 (10.2)**	9.8 (8.4)**	<0.0001
Quality of life (PedsQL score)‡‡	82.3 (13.3)	85.3 (12.1)	85.4 (11.0)	79.9 (13.3)**	77.7 (16.8)**	75.9 (14.2)**	82.1 (13.1)**	<0.0001
Lifestyle behavioral factors								
Adherence to DASH diet§§	46.5 (9.4)	48.7 (9.1)	46.7 (8.6)*	45.3 (9.5)**	44.0 (9.0)**	44.9 (8.8)**	45.2 (9.8)**	<0.0001
Total energy intake (kcal)	1699 (788)	1694 (760)	1623 (688)	1746 (860)	1960 (1153)*	1791 (764)*	1559 (623)*	0.010
Total energy from fat (%)	37.5 (6.2)	36.8 (6.0)	37.3 (6.6)	38.0 (6.3)*	37.5 (5.9)	38.0 (6.1)	38.0 (6.1)	0.188
Total energy from carbohydrate (%)	48.1 (8.2)	49.1 (7.8)	48.3 (8.5)	47.6 (8.6)*	48.2 (7.8)	47.0 (8.3)*	47.1 (7.9)**	0.060
Total energy from protein (%)	16.1 (2.6)	16.1 (2.5)	16.3 (2.5)	16.0 (2.7)	15.7 (2.4)	16.4 (2.8)	16.5 (2.5)	0.300
Physically active¶¶	1063 (58.5)	429 (69.4)	118 (60.5)*	264 (51.9)**	36 (45.6)**	69 (50.4)**	147 (52.7)**	<0.0001
High screen time¶¶	924 (50.9)	234 (37.9)	78 (40.0)	309 (60.7)**	57 (72.2)**	96 (70.1)**	150 (53.8)**	<0.0001

Continued

Table 2 Continued

Characteristics, mean (SD) or n (%)	All n=1817	Weight-glycemia cluster						P value†
		Cluster 1 n=618 (34.0%)	Cluster 2 n=195 (10.7%)	Cluster 3 n=509 (28.0%)	Cluster 4 n=79 (4.4%)	Cluster 5 n=137 (7.5%)	Cluster 6 n=279 (15.4%)	

All measures are from the cohort visit.
Missing data: n=4 for 1+ diabetic ketoacidosis episodes. n=174 for DASH score, total energy, total energy from carbohydrate, total energy from protein, and total energy from fat. Missingness not different across clusters (p>0.05).
*Significant pairwise comparison to cluster 1, where q<0.05.
†Bonferroni-corrected p value for overall test of difference, based on use of analysis of variance (ANOVA), χ^2 or Fisher's exact test as appropriate.
‡Includes 2+ times/day or any insulin combination (excluding long), 3+ times/day or any insulin(s) taken 1×/day, or any insulin combination (excluding long) 2+/day.
§Self-reported, in the past 6 months.
¶Includes family practice doctor, general practice doctor, internist, nurse diabetes educator, nurse practitioner/physician's assistant, dietician/nutritionist, other, unknown, and none.
**Significant pairwise comparison to Ccluster 1, where q<0.001. Controlled for the positive false discovery rate.
††Center for Epidemiologic Studies Depression Scale, total score.
‡‡Pediatric s Quality of Life Inventory, total score.
§§Dietary Approaches to Stop Hypertension diet diet, total score.
¶¶Physically active defined as exercise 3–7 days per week. High screen time defined as 2+ hours of screen- time per day.

Table 3 Clinical characteristics according to weight-glycemia phenotype clusters 1–6

Characteristics, mean (SD) or n (%)	All n=1817	Weight-glycemia cluster						P value†
		Cluster 1 n=618 (34.0%)	Cluster 2 n=195 (10.7%)	Cluster 3 n=509 (28.0%)	Cluster 4 n=79 (4.4%)	Cluster 5 n=137 (7.5%)	Cluster 6 n=279 (15.4%)	
Lipids								
Total cholesterol (mg/dL)	169.6 (34.7)	156.2 (25.7)	154.5 (25.7)	177.0 (31.0)**	185.2 (37.6)**	207.9 (51.9)**	172.4 (29.9)**	<0.0001
HDL cholesterol (mg/dL)	55.2 (13.7)	56.5 (13.2)	57.2 (14.2)	57.0 (13.8)*	53.5 (12.6)*	50.6 (13.3)**	50.5 (12.3)**	<0.0001
LDL cholesterol (mg/dL)	96.1 (27.9)	86.6 (21.5)	83.5 (21.8)*	100.6 (26.1)**	104.9 (28.8)**	120.3 (38.4)**	103.1 (27.0)**	<0.0001
VLDL cholesterol (mg/dL)	18.2 (12.5)	13.1 (5.0)	13.8 (5.8)	19.5 (10.5)**	27.0 (16.5)**	36.5 (25.4)**	18.7 (9.4)**	<0.0001
Triglycerides (mg/dL)	92.3 (70.8)	65.6 (25.0)	68.8 (29.1)	97.7 (52.4)**	141.7 (123.4)**	195.0 (152.3)**	93.5 (47.1)**	<0.0001
Blood pressure								
Systolic blood pressure (mm Hg)	106.0 (10.9)	104.9 (10.6)	102.7 (11.2)*	105.3 (10.1)	104.4 (11.4)	111.1 (9.8)**	110.2 (11.6)**	<0.0001
Diastolic blood pressure (mm Hg)	68.5 (8.9)	66.7 (8.5)	66.0 (8.9)	55.6 (23.5)**	69.2 (9.5)*	73.1 (8.2)**	71.1 (32.9)**	<0.0001

All measures are from the cohort visit.
*Significant pairwise comparison to cluster 1, where q<0.05. **Significant pairwise comparison to cluster 1, where q<0.001. Controlled for the positive false discovery rate.
†Bonferroni-corrected p value for overall test of difference, based on use of analysis of variance (ANOVA), χ^2 or Fisher's exact test as appropriate.
HDL, high-density lipoprotein; LDL, low-density lipoprotein; VLDL, very-low-density lipoprotein.

the weight-glycemia phenotype of cluster 6, comprising 15% of the sample). This is consistent with previous work demonstrating a U-shaped association between HbA1c and BMI standard deviation score (SDS), with the highest HbA1c levels among children and adolescents with T1D who were classified as underweight and obese.[8]

Clinical recommendations for individuals with HbA1c levels above target may differ based on weight status, especially given the complicated physiologic relationships between weight and glycemia.[31] For example, approaches for under or normal weight individuals with elevated HbA1c may be centered on insulin intensification, while approaches for overweight individuals could balance the glycemic benefits of insulin intensification with the potential for weight gain via concurrent behavioral, pharmacological, or surgical interventions.[32] Although the American Diabetes Association Standards of Care currently state that the risks and benefits of agents adjunctive to insulin therapy in T1D require further evaluation, early studies have shown improvements in glycemic control as well as body weight associated with metformin, GLP1-receptor agonists, and SGLT-2 inhibitors.[32 33] Subgroups with adequate glycemic control and overweight or obesity may be specifically targeted for more intensive behavioral modifications such as dietary interventions to address obesity with special attention towards avoiding hypoglycemia associated with decreased caloric or carbohydrate intake.[31] Other aspects of clinical care may be tailored according to subgroup as well. For example, while all youth should receive timely and regular screening for complications of diabetes, particular subgroups may warrant earlier interventions, including the introduction of renoprotective regimens or statin medications. Ultimately, larger and intentionally designed trials will be required to move from understanding observational phenotypes to devising their therapeutic approaches. In particular, as the treatment landscape of T1D expands to include novel non-insulin adjuvants[32] as well as hybrid closed loop and fully automated insulin delivery systems,[34] capturing treatment response across subgroups both in terms of changes in body weight and glycemic control will provide critical data to inform how phenotypes may be modified to better predict heterogeneity in response to specific treatment approaches.

Given the high risk for long-term complications, we focus our discussion on clusters 4 and 5, the subgroups with the poorest glycemic control (HbA1c >108 mmol/mol), as well as cluster 6, the subgroup with obesity and moderate glycemic control. Together, they account for approximately 27% of the study population.

The results corroborate previous reports that glycemic control differs by race and ethnicity among youth and young adults with T1D[4 28] and is also associated with lower measures of socioeconomic position including parental education, income, and health insurance type. The results also substantiate other studies showing that lower household income and parental education level associate with overweight/obesity status in T1D[35] and are consistent with reports that youth with T1D who are of Hispanic ethnicity and females are at the highest risk of overweight or obesity.[36] This finding is particularly concerning given recent data suggesting that the incidence of T1D is increasing most rapidly among Hispanic youth.[37]

Weight-glycemia clusters also showed significant differences in several aspects of clinical care, psychosocial outcomes, and health behaviors that were measured concurrently with BMIz and HbA1c. In our study, the best mean glycemic control was associated with higher use of insulin pump therapy[38] and increased frequency of blood glucose monitoring.[39]

The psychosocial correlates of clusters were consistent with previous studies, including a positive relationship between mean HbA1c levels and mean depressive symptoms and a negative association between mean HbA1c levels and mean perceived quality of life measures.[40] Differences in potentially modifiable behavioral factors underscore that physically active lifestyle and decreased sedentary time are associated with lower BMI and percentage of body fat among people with T1D.[35] Unfortunately, overall dietary quality measures were low, even among youth and young adults with the most favorable weight-glycemia phenotype, consistent with previous findings.[19]

The significant differences in clinical parameters across weight-glycemia clusters implicate potential disparity in long-term cardiovascular disease risk across these subgroups.[9] The combination of very high HbA1c and increased insulin needs of cluster 5, the overweight subgroup with very poor glycemic control, is suggestive of insulin resistance.[9 36] This group also exhibited the worst cardiovascular disease risk profile including elevated lipid and blood pressure levels. Laboratory measures were significantly elevated compared with cluster 6, despite the higher mean BMIz of cluster 6. More work is needed to understand how adiposity and hyperglycemia jointly contribute to cardiovascular disease risk profiles.

One of the most striking results is the pattern with which all other demographic, socioeconomic, clinical care, psychosocial, and behavioral factors track across the clusters derived from measures of the joint distribution of weight and glycemia. It is relevant to note that non-white race/ethnicity, lower socioeconomic position and healthcare access, and poorer psychosocial well-being have been shown to be associated with each other and with glycemic control elsewhere in SEARCH studies.[38 40] A maximally effective precision medicine approach to co-optimize weight and glycemia will concern itself with underlying biology as well as characteristics of individuals and resource constraints that may influence outcomes over time. Although the cross-sectional cluster analysis is not designed for causal conclusions, future research is needed to develop the specific interventional strategies to impact weight and glycemia outcomes that considers the close relationships among these economic, social,

and cultural factors. In the meantime, efforts should also be made to identify youth soon after diagnosis who may have sociodemographic risk with subsequent efforts to ensure access and affordability of necessary diabetes medications and supplies.

The study has several weaknesses. Despite the collective use of gap statistics and supporting graphs, selection of the number of clusters is subjective. Additional external validation studies are required to understand the generalizability of major weight-glycemia phenotypes across other observational cohort studies of T1D. In addition, BMIz was used as a proxy for weight status, although this measure may not reflect adiposity,[41 42] and the large age range necessitated imputation of BMIz for participants over 20 years of age using z-score data for age 20 years. To assess for differential error of BMIz on participant age (ie, youth vs young adults), we stratified the sample by age at follow-up visit (<21 years, n=1399, ≥21 years, n=418) and independently evaluated clusters in each sample (online supplementary table S5, figure S1). Despite significant differences in sample sizes, we found largely consistent clustering results in both strata, suggesting that the measure of BMIz did not bias the nature of the clusters across different age ranges. Future work could use validated predictive equations to estimate body fat percentage[43] and derive 'adiposity-glycemia' clusters using the predicted variable. The current study is cross-sectional and cannot elucidate temporal associations with the weight-glycemia phenotypes nor the longitudinal clinical outcomes; future studies should explore whether subgroups develop different rates or patterns in the emergence of T1D complications and cardiovascular disease risk factors to inform clinical utility of this weight-glycemia phenotype. Along these lines, the prognostic value of T1D phenotypes may be enhanced by incorporation of other clinical variables representing cardiovascular disease risk factors to inform risk stratification with regard to cardiovascular disease. The study also has several strengths. One is inherent in the analytic design; this approach to characterize a phenotype based on two outcomes allows real-life phenotypes to emerge rather than forcing a fit based on a priori clinical cut-points for weight and glycemic control. In additional analyses, the six weight-glycemia clusters were compared with strata of the same sample defined by clinical cut-points for overweight/obesity and poor glycemic control (see online supplementary tables S6 and S7). The strata corroborated main descriptive results (ie, differences in sociodemographic characteristics across subgroups with differing levels of glycemic control), providing face validity to the weight-glycemia clusters. However, the use of a priori cut-points was found to be less well suited to identify subgroups sharing clinically significant yet more nuanced weight-glycemia phenotypes who may otherwise distinguish themselves in a clustering approach, such as the subgroups with very poor glycemic control. For example, clinical cut-points collapsed all individuals in clusters 3, 4, and 5 in the same strata of glycemic control, despite the notable differences in glycemia (refer to these subgroups in relation to the dashed line denoting poor glycemic control at HbA1c 75 of mmol/mol (9.0%) in figure 1B.) Given the relatively recent emergence of overweight and obesity in T1D, the data-driven approach is particularly well suited to explore underlying subgroups within the SEARCH cohort, from which the population-specific utility of existing clinical cut-points can be better evaluated. A further strength of the study was the novel semisupervised statistical methods used to identify a phenotype based on two clinical outcomes and their relationship to each other, using all patient information to adjust for potential measurement error and within-person heterogeneity. In additional analyses, the Y-clusters showed multiple nodes of density and larger within-cluster distribution of BMIz and HbA1c (online supplementary table S8, supplementary figure S2), suggesting higher within-cluster variability due to noise in the raw outcomes that obscures underlying clustering structure in the data. The advantage of clusters driven by predicted measures of the *joint distribution* is that this method uses *X*-variables to denoise the raw outcome measures, thereby maximizing data available in the cohort study to understand the underlying variance in weight and glycemia, and their relationship as a clinical phenotype. Finally, to our knowledge, this is the first study to investigate the spectrum of the weight-glycemia phenotypes of T1D and their broad epidemiologic correlates using the large, diverse SEARCH cohort. The study complements previous efforts to address heterogeneity in adult diabetes[10] with a focus on T1D in a younger age range to inform earlier interventions.

In conclusion, we show that the heterogeneous population of youth and young adults with T1D comprised identifiable subgroups with shared weight-glycemia clinical phenotypes based on measures of the joint distribution of BMIz and HbA1c. Importantly, overweight and obesity present with varying degrees of glycemic control in this population, implicating different therapeutic and clinical strategies to concurrently address weight and glycemia across subgroups. To this end, a precision medicine framework may facilitate a systems-based approach to address health inequity and deliver targeted strategies needed to optimize obesity and dysglycemia, particularly when both are poorly controlled.

Author affiliations
[1]Department of Nutrition, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA
[2]Department of Biostatistics, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA
[3]Center for Diabetes Translational Research, Albert Einstein College of Medicine, Bronx, New York, USA
[4]Department of Epidemiology, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA
[5]Department of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA
[6]Department of Epidemiology, Colorado School of Public Health, Aurora, Colorado, USA

[7]Department of Pediatrics, School of Medicine, University of Colorado, Aurora, Colorado, USA
[8]Division of Endocrinology, Cincinnati Children's Hospital Medical Center, Cincinnati, Ohio, USA
[9]Division of Diabetes Translation, Centers of Disease Control and Prevention, Atlanta, Georgia
[10]Department of Research and Evaluation, Kaiser Permanente Southern California, Pasadena, Southern California, USA
[11]Northwest Lipid Metabolism and Diabetes Research Laboratories, Department of Medicine, University of Washington, Seattle, Washington, USA
[12]Department of Biostatistical Sciences, Wake Forest School of Medicine, Winston-Salem, North Carolina, USA
[13]Department of Pediatrics, University of Washington, Seattle, Washington, USA
[14]Department of Statistics and Operations Research, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA

Acknowledgements The SEARCH for Diabetes in Youth study is indebted to the many youth and their families, and their healthcare providers, whose participation made this study possible.

Contributors ARK, CTN, XJ, MRK, and EJMD designed the analyses. ARK and CTN conducted the analyses. ARK and EJMD drafted the initial manuscript. LAA, SA, AEA, KSB, JBB, DD, LMD, GI, JML, SM, CP, BAR, and KAS reviewed all analyses and provided critical review of the manuscript. GI—a representative of one of the cosponsors—reviewed all analyses and provided critical review of the manuscript. ARK and EJMD had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. MRK and EJMD are cosenior authors.

Funding Grant support (SEARCH 4): The SEARCH for Diabetes in Youth cohort study (1UC4DK108173) is funded by the National Institutes of Health, National Institute of Diabetes and Digestive and Kidney Diseases and supported by the Centers for Disease Control and Prevention. Grant support (SEARCH 3): The SEARCH for Diabetes in Youth is funded by the Centers for Disease Control and Prevention (PA numbers 00097, DP-05-069, and DP-10-001) and supported by the National Institute of Diabetes and Digestive and Kidney Diseases. Site contract numbers: Kaiser Permanente Southern California (U18DP006133, U48/CCU919219, U01 DP000246, and U18DP002714), University of Colorado Denver (U18DP006139, U48/CCU819241-3, U01 DP000247, and U18DP000247-06A1), Cincinnati's Children's Hospital Medical Center (U18DP006134, U48/CCU519239, U01 DP000248, and 1U18DP002709), University of North Carolina at Chapel Hill (U18DP006138, U48/CCU419249, U01 DP000254, and U18DP002708), Seattle Children's Hospital (U18DP006136, U58/CCU019235-4, U01 DP000244, and U18DP002710-01), Wake Forest University School of Medicine (U18DP006131, U48/CCU919219, U01 DP000250, and 200-2010-35171). The authors wish to acknowledge the involvement of the Kaiser Permanente Southern California's Clinical Research Center (funded by Kaiser Foundation Health Plan and supported in part by the Southern California Permanente Medical Group); the South Carolina Clinical and Translational Research Institute at the Medical University of South Carolina, NIH/National Center for Advancing Translational Sciences (NCATS) grant numbers UL1 TR000062, UL1Tr001450; Seattle Children's Hospital and the University of Washington, NIH/NCATS grant number UL1 TR00423; University of Colorado Pediatric Clinical and Translational Research Center, NIH/NCATS grant number UL1 TR000154; the Barbara Davis Center at the University of Colorado at Denver (DERC NIH grant number P30 DK57516); the University of Cincinnati, NIH/NCATS grant numbers UL1 TR000077, UL1 TR001425; and the Children with Medical Handicaps program managed by the Ohio Department of Health. This study includes data provided by the Ohio Department of Health, which should not be considered an endorsement of this study or its conclusions. ARK is supported by the National Institute of Diabetes and Digestive and Kidney Disease of the National Institutes of Health under award number F30DK113728. MRK, CTN, XJ, and JBB are supported by funding from NC Tracs (the CTSA at UNC): UL1TR002489.

Disclaimer The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention and the National Institute of Diabetes and Digestive and Kidney Diseases.

Competing interests None declared.

Patient consent for publication Not required.

Ethics approval The study was approved by institutional review boards with jurisdiction, and the parent, adolescent or young adult, or both provided consent or assent for all participants, as appropriate.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data may be obtained from a third party and are not publicly available. The data that support the findings of this study are available from the SEARCH for Diabetes in Youth study, but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Data are however available from the authors upon reasonable request and with permission of the SEARCH for Diabetes in Youth study.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

ORCID iD
Anna R Kahkoska http://orcid.org/0000-0003-2701-101X

REFERENCES

1 Liu LL, Lawrence JM, Davis C, *et al*. Prevalence of overweight and obesity in youth with diabetes in USA: the search for diabetes in youth study. *Pediatr Diabetes* 2010;11:4–11.
2 DuBose SN, Hermann JM, Tamborlane WV, *et al*. Obesity in Youth with Type 1 Diabetes in Germany, Austria, and the United States. *J Pediatr* 2015;167:627–32.
3 Mottalib A, Kasetty M, Mar JY, *et al*. Weight management in patients with type 1 diabetes and obesity. *Curr Diab Rep* 2017;17:92.
4 Klonoff DC. *Precision medicine for managing diabetes*. . SAGE Publications, 2015: 9. 3–7.
5 Kosorok MR, Medicine P. Annual reviews of statistics and its application. In Press 2018.
6 Trusheim MR, Berndt ER, Douglas FL. Stratified medicine: strategic and economic implications of combining drugs and clinical biomarkers. *Nat Rev Drug Discov* 2007;6:287–93.
7 Nansel TR, Lipsky LM, Iannotti RJ. Cross-Sectional and longitudinal relationships of body mass index with glycemic control in children and adolescents with type 1 diabetes mellitus. *Diabetes Res Clin Pract* 2013;100:126–32.
8 Maffeis C, Birkebaek NH, Konstantinova M, *et al*. Prevalence of underweight, overweight, and obesity in children and adolescents with type 1 diabetes: data from the International sweet registry. *Pediatr Diabetes* 2018;19:1211–20.
9 Purnell JQ, Braffett BH, Zinman B, *et al*. Impact of excessive weight gain on cardiovascular outcomes in type 1 diabetes: results from the diabetes control and complications Trial/Epidemiology of diabetes interventions and complications (DCCT/EDIC) study. *Diabetes Care* 2017;40:1756–62.
10 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;6:361–9.
11 Hamman RF, Bell RA, Dabelea D, *et al*. The search for diabetes in youth study: rationale, findings, and future directions. *Diabetes Care* 2014;37:3336–44.
12 Dabelea D, Stafford JM, Mayer-Davis EJ, *et al*. Association of type 1 diabetes vs type 2 diabetes diagnosed during childhood and adolescence with complications during teenage years and young adulthood. *JAMA* 2017;317:825–35.
13 Dabelea D, Pihoker C, Talton JW, *et al*. Etiological approach to characterization of diabetes type. *Diabetes care* 2011;34:1628–33.
14 Kuczmarski RJ, Ogden CL, Grummer-Strawn LM, *et al*. Cdc growth charts: United States. *Adv Data* 2000;314:1–27.
15 Sauder KA, Stafford JM, Mayer-Davis EJ, *et al*. Co-Occurrence of early diabetes-related complications in adolescents and young adults with type 1 diabetes: an observational cohort study. *Lancet Child Adolesc Health* 2019;3:35–43.
16 Esbenshade AJ, Simmons JH, Koyama T, *et al*. Body mass index and blood pressure changes over the course of treatment of pediatric acute lymphoblastic leukemia. *Pediatr Blood Cancer* 2011;56:372–8.
17 Radloff LS. The CES-D scale: a self-report depression scale for research in the general population. *Appl Psychol Meas* 1977;1:385–401.
18 Varni JW, Burwinkle TM, Jacobs JR, *et al*. The PedsQL™ in type 1 and type 2 diabetes. *Diabetes care* 2003;26:631–7.
19 Mayer-Davis EJ, Nichols M, Liese AD, *et al*. Dietary intake among youth with diabetes: the search for diabetes in youth study. *J Am Diet Assoc* 2006;106:689–97.

20 Windhauser MM, Ernst DB, Karanja NM, *et al*. Translating the dietary approaches to stop hypertension diet from research to practice: dietary and behavior change techniques. DASH Collaborative Research Group. *J Am Diet Assoc* 1999;99:S90–5.
21 Bair E. Semi-supervised clustering methods. *Wiley Interdiscip Rev Comput Stat* 2013;5:349–61.
22 Zhu R, Zeng D, Kosorok MR. Reinforcement learning trees. *J Am Stat Assoc* 2015;110:1770–84.
23 Stekhoven DJ, Bühlmann P. MissForest--non-parametric missing value imputation for mixed-type data. *Bioinformatics* 2012;28:112–8.
24 Charrad M, Ghazzali N, Boiteau V, *et al*. NbClust: An *R* Package for Determining the Relevant Number of Clusters in a Data Set. *J Stat Softw* 2014;61:1–36.
25 Faul F, Erdfelder E, Lang A-G, *et al*. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav Res Methods* 2007;39:175–91.
26 Steinley D. Properties of the Hubert-Arabie adjusted Rand index. *Psychol Methods* 2004;9:386–96.
27 Wang Y, Chen H-J. *Use of percentiles and z-scores in anthropometry. Handbook of anthropometry*. Springer, 2012: 29–48.
28 Petitti DB, Klingensmith GJ, Bell RA, *et al*. Glycemic control in youth with diabetes: the search for diabetes in youth study. *J Pediatr* 2009;155:668–72.
29 Storey JD. The positive false discovery rate: a Bayesian interpretation and the Q -value. *Ann Statist* 2003;31:2013–35.
30 DiMeglio LA, Acerini CL, Codner E, *et al*. Glycemic control targets and glucose monitoring for children, adolescents, and young adults with diabetes. *Pediatric diabetes* 2018.
31 Corbin KD, Driscoll KA, Pratley RE, *et al*. Obesity in type 1 diabetes: pathophysiology, clinical impact, and mechanisms. *Endocr Rev* 2018;39:629–63.
32 Harris K, Boland C, Meade L, *et al*. Adjunctive therapy for glucose control in patients with type 1 diabetes. *Diabetes Metab Syndr Obes* 2018;11:159–73.
33 American Diabetes Association. 9. Pharmacologic Approaches to Glycemic Treatment: *Standards of Medical Care in Diabetes—2020*. *Diabetes Care* 2020;43:S98–110.
34 Boughton CK, Hovorka R. Advances in artificial pancreas systems. *Sci Transl Med* 2019;11:eaaw4949.
35 Driscoll KA, Corbin KD, Maahs DM, *et al*. Biopsychosocial aspects of weight management in type 1 diabetes: a review and next steps. *Curr Diab Rep* 2017;17:58.
36 Minges KE, Whittemore R, Weinzimer SA, *et al*. Correlates of overweight and obesity in 5529 adolescents with type 1 diabetes: the T1D exchange clinic registry. *Diabetes Res Clin Pract* 2017;126:68–78.
37 Mayer-Davis EJ, Lawrence JM, Dabelea D, *et al*. Incidence trends of type 1 and type 2 diabetes among Youths, 2002-2012. *N Engl J Med* 2017;376:1419–29.
38 Paris CA, Imperatore G, Klingensmith G, *et al*. Predictors of insulin regimens and impact on outcomes in youth with type 1 diabetes: the search for diabetes in youth study. *J Pediatr* 2009;155:183–9.
39 Levine B-S, Anderson BJ, Butler DA, *et al*. Predictors of glycemic control and short-term adverse outcomes in youth with type 1 diabetes. *J Pediatr* 2001;139:197–203.
40 Lawrence JM, Yi-Frazier JP, Black MH, *et al*. Demographic and clinical correlates of diabetes-related quality of life among youth with type 1 diabetes. *J Pediatr* 2012;161:201–7.
41 Garn SM, Leonard WR, Hawthorne VM. Three limitations of the body mass index. *Am J Clin Nutr* 1986;44:996–7.
42 Kragelund C, Omland T. A farewell to body-mass index? *The Lancet* 2005;366:1589–91.
43 Stevens J, Ou F-S, Cai J, *et al*. Prediction of percent body fat measurements in Americans 8 years and older. *Int J Obes* 2016;40:587–94.