



# HHS Public Access

Author manuscript

*Pediatr Diabetes*. Author manuscript; available in PMC 2021 February 01.

Published in final edited form as:

*Pediatr Diabetes*. 2021 February ; 22(1): 22–30. doi:10.1111/pedi.12981.

## Clinical profile at diagnosis with youth-onset type 1 and type 2 diabetes in two pediatric diabetes registries: SEARCH (United States) and YDR (India)

Christine W Hockett<sup>1</sup>, Pradeep A Praveen<sup>3</sup>, Toan C. Ong<sup>2</sup>, Amutha Anandakumar<sup>4</sup>, Scott P Isom<sup>5</sup>, Elizabeth T Jensen<sup>6</sup>, Ralph B D'Agostino Jr<sup>5</sup>, Richard F Hamman<sup>1</sup>, Elizabeth J Mayer-Davis<sup>7</sup>, Jean M. Lawrence<sup>8</sup>, Catherine Pihoker<sup>9</sup>, Michael G Kahn<sup>2</sup>, Viswanathan Mohan<sup>4</sup>, Nikhil Tandon<sup>3</sup>, Dana Dabelea<sup>1</sup>

<sup>1</sup>Lifecourse Epidemiology of Adiposity and Diabetes Center, Colorado School of Public Health, University of Colorado Denver, Aurora, CO

<sup>2</sup>Department of Pediatrics, University of Colorado, Aurora, CO

<sup>3</sup>Department of Endocrinology & Metabolism, All India Institute of Medical Sciences, New Delhi, India

<sup>4</sup>Dr. Mohan's Diabetes Specialties Centre and Madras Diabetes Research Foundation, Chennai, India

<sup>5</sup>Department of Biostatistics and Bioinformatics, Wake Forest School of Medicine, Winston-Salem, NC

<sup>6</sup>Department of Epidemiology, Wake Forest School of Medicine, Winston-Salem, NC

<sup>7</sup>Department of Nutrition and Medicine, University of North Carolina, Chapel Hill, NC

<sup>8</sup>Department of Research & Evaluation, Kaiser Permanente Southern California, Pasadena, CA

<sup>9</sup>Department of Pediatrics, University of Washington, Seattle, WA

### Abstract

**Background:** Over the last decades, diabetes in youth has increased in both India and the U.S., along with the burden of long-term complications and healthcare costs. However, there are limited standardized population-based data in contemporary youth cohorts for comparison of clinical and demographic characteristics of diabetes for both type 1 (T1D) and type 2 (T2D).

**Methods:** In partnership, we harmonized demographic and clinical data from the SEARCH for Diabetes in Youth (SEARCH) registry in the U.S. and the Registry of People with Diabetes with Youth Age at Onset (YDR) in India to the structure and terminology of the Observational Medical

---

**Correspondence:** Christine W Hockett, PhD, University of Colorado Anschutz Medical Campus, 12474 E. 19<sup>th</sup> Avenue – Mail Stop F426, Aurora, CO 80045, Tel: 605-691-1274, Fax: 303-724-7724, christine.hockett@ucdenver.edu.

**Author Contributions:** N.T, D.D, V.M, R.F.H and E.J.M.-D conceptualized the study and oversaw the data harmonization. C.W.H., P.P., T.C.O., S.P.I., M.G.K. and A.A. harmonized and transformed data into the common data model (OMOP). C.W.H., T.C.O., and S.P.I analysed the data. C.W.H. prepared the first draft of the manuscript. All authors reviewed and edited the manuscript and contributed to discussion. All authors have read and approved the final manuscript.

**Disclosures:** None of the authors have any potential conflicts of interest relevant to the manuscript.

Outcomes Partnership Common Data Model. Data were from youth with T1D and T2D, aged <20 years and newly diagnosed between 2006 and 2010. We compared key characteristics across registries using chi-squared tests and t-tests.

**Results:** In total, there were 9,650 youth with T1D and 2,406 youth with T2D from 2006 to 2012. SEARCH youth were diagnosed at younger ages than YDR youth for T1D and T2D (10.0 vs. 10.5 y,  $p<0.001$  and 14.7 vs. 16.1 y,  $p<0.001$ , respectively). For T2D, SEARCH had a higher proportion of females and significantly lower proportion of youth of high socioeconomic status compared to YDR. For T1D and T2D, SEARCH youth had higher BMI, lower blood pressure, and lower A1c compared to YDR youth.

**Conclusions:** These data offer insights into the demographic and clinical characteristics of diabetes in youth across the two counties. Further research is needed to better understand why these differences exist.

### Keywords

data harmonization; type 1 diabetes; early onset type 2 diabetes; India; United States; registry

---

## INTRODUCTION

Over the last several decades, both type 1 (T1D) and type 2 diabetes (T2D) in youth have been increasing both in India and in the United States (U.S.), in all age groups and both sexes.<sup>1</sup> The epidemiology of diabetes in youth has mainly been focused on T1D. However, with the increase of childhood obesity, unhealthy eating and sedentary lifestyles, pediatric T2D has become increasingly prevalent.<sup>2–5</sup> Many youth with diabetes are at risk for poor glycemic control<sup>6</sup>, diabetic eye and kidney disease, and future cardiovascular disease; this is especially true for youth with T2D, and for racial and ethnic minority youth with T1D in the U.S.<sup>7</sup>

There are limited standardized population-based data in contemporary youth cohorts that provide the ability to compare socio-demographic and clinical characteristics of diabetes in youth between countries. Previous international childhood diabetes registries, EURODIAB and DIAMOND, provided critical international information from over 60 countries regarding T1D among youth aged 0–14 years using standardized protocols to allow for comparison of incidence rates and diabetes-related complications.<sup>8</sup> However, limited international data exist for comparisons of clinical profiles and risk factors for complications for both T1D and T2D in contemporary cohorts. With the steady increase of both T1D and T2D in youth, such information is critical for an improved understanding of potentially modifiable risk factors, and workforce planning.

Therefore, the harmonization of the SEARCH for Diabetes in Youth (SEARCH) registry in the U.S. with the Registry of People with Diabetes with Youth Age at Onset (YDR) in India is a critical step towards building capacity to understand global burden and trends in youth-onset diabetes in contemporary cohorts from countries with populations with differing demographic, socio-economic and behavioral characteristics. Here we describe the methods used for the harmonization of the SEARCH and YDR registries and compare the basic

demographic and clinical profiles around the time of diagnosis with T1D and T2D in U.S. and Indian youth.

## METHODS

This work was a collaborative partnership between SEARCH for Diabetes in Youth (U.S.) and the Registry of People with Diabetes with Youth Age at Onset (India).

### Registries and Data Definitions

**SEARCH for Diabetes in Youth**—SEARCH is a multiethnic, population-based registry with five sites across the U.S. ascertaining physician-diagnosed non-gestational incident diabetes cases among individuals aged 19 years or younger. Detailed information about SEARCH is published elsewhere.<sup>1,3,7</sup> Each site conducts active surveillance under the Health Insurance Portability and Accountability Act (HIPAA) waivers of consent using networks of endocrinologists, healthcare providers, hospitals and community health centers, and clinical and administrative data systems along with electronic medical records. Cases are confirmed as valid after review of medical records or by the referring physician. All registered participants are asked to complete an Initial Participant Survey (IPS) (average response rate for incident cases 2006–2012, 82%). Participants diagnosed in 2006, 2008 and 2012 were invited to participate in an in-person baseline research visit (IPV) (average response rates between 46% and 65%), where data were obtained on socio-demographic characteristics, height, weight, medications, glucose control and other risk factors for diabetes-related complications, including laboratory measurements. Blood samples (A1c) were taken and analysed at a central laboratory. For the purposes of this manuscript, the IPV baseline research visit is referred to as the baseline visit.

**Registry of People with Diabetes with Youth Age at Onset (YDR)**—The YDR registry is an observational multicenter clinic-based registry enlisting all cases of physician-diagnosed diabetes, diagnosed at the age of 25 years or younger, who were registered at a designated reporting center on or after January 1, 2000, residing within assigned geographical areas. More detailed information about YDR is published elsewhere.<sup>9</sup> Individuals are classified into various diabetes categories based on the assessment of the principal investigator at the reporting center using symptom-based clinical criteria agreed upon by the registry expert group prior to initiation of data collection in 2006. YDR data collection is coordinated by the Indian Council of Medical Research through regional collaborating centers and their interacting reporting centers. All individuals have a proforma (registration and clinical extract) completed by the participant and physician to obtain information on socio-demographics, clinical profile, anthropometrics and laboratory measurements of the individual. Data from the period 2000–2006 were collected retrospectively in a structured format from medical records; while data from 2006–2012 were collected prospectively and completed by both the participant and physician at the time of registration, which is referred to here as the baseline visit. There are eight regional collaborating centers across India who provide cases to YDR. For this project, data from three of the eight collaborating centers (one in Chennai (Madras Diabetes Research Foundation) and two in New Delhi (All India Institute of Medical Sciences (AIIMS) and the

Univeristy College of Medical Sciences, Delhi) were used. For the purposes of this manuscript, the baseline registration visit is referred to as the baseline visit.

### Data Harmonization

The data harmonization process included a series of operations to extract, transform, and load source data to be syntactically and semantically harmonized to the structure and terminology of the target Common Data Model (CDM).<sup>10</sup> Our method to harmonize data included two major steps: schema mapping and concept mapping. Schema mapping aligns data elements (tables and fields) in the source data to the target CDM. Concept mapping semantically maps the locally encoded values contained in each of these data elements to a standard value set. For this project, we selected the Observational Medical Outcomes Partnership (OMOP)<sup>11</sup> Common Data Model (v5) as the target CDM. OMOP is an internationally supported CDM for clinical data harmonization with rich support for standard terminologies and terminology mappings.<sup>12,13</sup> Using such a CDM allows each registry to maintain data securely without requiring data transfer to a common analytic system, thus ensuring confidentiality.

Even though most of the data elements and terminologies in the source data model exist in the OMOP CDM, there were diabetes-specific or local data elements and terminologies which needed to be harmonized. For example, diabetes duration, or insulin regimen, are not standard OMOP terms, but are common diabetes-specific data elements. For these variables, we extended the OMOP data model by creating custom data fields or concepts. The limitation of this approach is that the custom data fields or concepts are only available to this project, affecting the generality of the queries. However, in the future, efforts can be made to introduce these project-specific concepts to the community of OMOP users, if desired. The deliverables resulting from this process included the concept mapping file, list of custom concepts and registry-specific Structure Query Language (SQL) scripts which use both the mapping file and custom concepts to perform the data transformation. Using these materials, each participating registry then harmonized their datasets to the OMOP CDM, resulting in a harmonized OMOP dataset within each registries data systems.

### Data Collection and Variable Definitions

**Demographic Characteristics**—Age at diagnosis (in years) was calculated using participant's date of birth and date of diagnosis, which were obtained from self-report and medical record abstractions, respectively. Age at baseline visit was calculated using the participant's date of birth and date of the baseline visit. Diabetes duration (in months) is the time from the date of diabetes diagnosis to the date of the baseline visit. Youth's sex and race (White, Asian Indian, Other Asian, Black, American Indian or Alaska Native, Native Hawaiian or Other Pacific Islander, Multiple Race, Unknown) were self-reported. Race was not collected on YDR participants, therefore all YDR participants were categorized as Asian Indian.

Due to the various methods used to define socioeconomic status (SES) within and between countries, there was no variable that matched one-to-one between the two registries. Therefore, in this manuscript SES was defined by self-reported insurance type and

household income in SEARCH and by type of hospital in YDR. For SEARCH, having private health insurance, regardless of income, was equivalent to high SES and having Other, None, or Government insurance was equivalent to low SES. If the participant was missing data on health insurance status but had a household income less than \$50,000/yr, the participant was categorized as low SES. For YDR, having been registered into YDR by a private hospital was assumed to be high SES and having been registered into YDR by a public hospital was assumed to be low SES.

**Clinical Characteristics**—Diabetes type (T1D and T2D) and date of diagnosis were obtained from medical records. For SEARCH participants, height, weight, and blood pressure were measured by a trained and certified staff member using a standard protocol on children aged 3 years or older at the baseline visit.<sup>14</sup> Hemoglobin A1c (A1c) was measured from a fasting blood draw taken at the baseline visit. Blood samples were obtained only if there was no episode of diabetic ketoacidosis within the prior month. Specimens were processed at the site and shipped within 24 h to the Northwest Lipid Metabolism and Diabetes Research Laboratories in Seattle, Washington, which serves as the study's central laboratory. For YDR, height, weight, and blood pressure were obtained clinically at the respective reporting centers, using standardized protocols. A1c was obtained from the most recent clinical encounter prior to the baseline visit. Blood samples were analyzed for A1c levels locally at the respective reporting center.

Height, weight, and blood pressure measurements were obtained from participants who completed a baseline visit for both registries. Body mass index (BMI) was calculated as  $\text{kg}/\text{m}^2$  and BMI z-scores were calculated using the World Health Organization Child Growth Standards reference data.<sup>15</sup> BMI categories (underweight, normal, overweight, obese) were determined using the BMI z-scores, age and the weight status cutoffs defined by the WHO in 2007.<sup>16</sup> Blood pressure percentiles were calculated using blood pressure, height and the percentile data provided by the 4<sup>th</sup> National Heart Lung, and Blood Institute (NHLBI) report on the Diagnosis, Evaluation, and Treatment of High Blood Pressure in Children and Adolescents.<sup>17</sup> Hypertension was defined as SBP or DBP percentiles greater than 95<sup>th</sup> percentile for participants 3–18 years of age; SBP  $\geq 140$  mmHg or DBP  $\geq 90$  mmHg for participants older than 18 years of age. Hemoglobin A1c concentrations were grouped into three categories (less than 7.5% (9.4 mmol/L), 7.5% to 9.0% (9.4 to 11.8 mmol/L), and greater than 9.0% (11.8 mmol/L)).

**Statistical Methods**—This report includes incident cases of diabetes that were ascertained by each registry within 30-months after the end of the year of diagnosis. We created two sample populations to use in analyses - one for demographic characteristics, which included incident cases diagnosed between 2006 and 2012 among youth aged  $<20$  years at diagnosis with T1D or T2D from both registries, and another for the clinical characteristics, which included all YDR incident cases and a subset of SEARCH participants who completed a baseline visit. Compared to those in the first sample population (all incident cases between 2006–2012), the subset of SEARCH participants who completed a baseline visit were not significantly different in terms of sex, age diabetes type.<sup>7</sup>

After creating and executing OMOP specific queries at each site, aggregate results were received. Demographic and clinical characteristics across registries were compared using chi-squared tests and t-tests.

## RESULTS

### Data Harmonization

We reviewed a total of 684 variables to be harmonized between the SEARCH and YDR registries. Seventy-four (10.8%) variables were identified as ‘variables of interest’; of these variables, 41 (55.4%) variables were able to be mapped to standard OMOP terminologies and 33 (44.6%) variables were mapped to custom terminologies. Figure 1 shows the distribution of the number of variables that were reviewed and mapped in each registry. Supplementary Tables 1 and 2 show the schema and concept mappings for this project.

### Demographic and Clinical Characteristics

There were 9,650 youth, aged 0–19 years, with newly diagnosed T1D (n=7,546 for SEARCH; n=2,104 for YDR) and 2,406 youth with newly diagnosed T2D (n=2,179 for SEARCH; n=227 for YDR) from 2006 to 2012. For SEARCH, the average duration of diabetes at the baseline visit was 10.3 months for T1D and 13.3 months for T2D. For YDR, the average duration of diabetes at the baseline visit was 6.2 months for T1D and 7.0 months for T2D.

Demographic characteristics for SEARCH and YDR youth by diabetes type are presented in Table 1. Youth in the SEARCH registry were diagnosed at younger ages than youth in the YDR registry for both T1D and T2D (10.0 vs. 10.5 y,  $p<0.001$  and 14.7 vs. 16.1 y,  $p<0.001$ , respectively). The proportion of females with T2D was higher in SEARCH than in YDR (62.8 vs. 48.9%,  $p<0.001$ , respectively). For T1D, 53.6% of SEARCH youth were of high SES compared to 60.8% of YDR youth ( $p<0.001$ ). For T2D, 23.6% of SEARCH youth were of high SES compared to 88.5% of YDR youth ( $p<0.001$ ).

Clinical characteristics for SEARCH and YDR youth by diabetes type are presented in Table 2. There were 1,899 youth with T1D and 384 youth with T2D who completed a SEARCH baseline research visit. There were 2,104 youth with T1D and 227 youth with T2D who completed a YDR baseline visit. A total of 85.4% of SEARCH youth with T2D were overweight or obese compared to only 58.2% of YDR youth ( $p<0.001$ ). SEARCH youth had significantly lower blood pressure than YDR youth for both types of diabetes SEARCH youth had a lower prevalence of hypertension compared to YDR youth for both T1D and T2D (3.0 vs 14.2%,  $p<0.001$ ; 15.7 vs 23.5%,  $p<0.001$ , respectively). SEARCH youth had significantly lower A1c concentrations at baseline compared to YDR youth for both T1D and T2D (7.8 vs 11.0%,  $p<0.001$ ; 7.2 versus 9.9%,  $p<0.001$ , respectively). Forty-two percent of SEARCH T1D youth, 7.2% of YDR T1D youth, 67.7% of SEARCH T2D youth, and 18.1% of YDR T2D youth had optimal glycemic control, as defined by the American Diabetes Association (A1c  $<7.5\%$ ).<sup>18</sup>

## DISCUSSION

We found important and significant differences in many of the demographic and clinical characteristics for both T1D and T2D between the youth in the SEARCH and YDR registries. Consistent with previous clinic-based studies from various parts of India<sup>19,20</sup>, we observed that T2D made up 10% of all youth diabetes cases, while in the SEARCH registry T2D made up 22% of the diabetes cases. The age at diagnosis was older for YDR youth compared to SEARCH, especially for youth with T2D (16.1 vs. 14.7 y,  $p < 0.001$ , respectively). We also found that SEARCH had a higher proportion of females with T2D compared to YDR. Further research is needed to determine whether the observed differences between the two countries in overall proportion, specifically the age and sex distribution of T2D cases, reflect differences in distribution of risk factors, e.g., greater obesity rates among U.S. youth, differences in pathophysiologic processes leading to T2D in these populations, or both.

Socioeconomic status is a complex, multidimensional concept that can be defined in many different ways (i.e. income, education, or social class). The majority of T1D cases in both registries were categorized as high SES. In developed countries, such as Europe and the U.S., research has shown that higher SES (i.e. higher income and higher education) at both an individual and neighborhood level is associated with a higher prevalence of T1D.<sup>21–23</sup> However, there is a lack of data on the relationship between SES and T1D in developing countries such as India, specifically in youth, and our results suggest a similar pattern. For T2D cases, the majority of cases in YDR were classified as high SES, while in SEARCH the majority of T2D cases were classified as low SES. In the U.S. and other developed countries, low SES is a known risk factor for the development of obesity and T2D<sup>24–26</sup>. This relationship has been explained by obesity-related factors, such as physical inactivity and the limited access to and consumption of healthy foods. However, in developing countries, the burden of obesity and T2D is higher among those with a high SES.<sup>27–29</sup> Our results are consistent with these reports. The impact of SES on the prevalence of T2D in India is clearly shown in the Chennai Urban Population Study.<sup>27</sup> In that study, the higher SES group had two-fold higher prevalence of T2D compared to the lower SES group.<sup>27</sup> The reasons for the higher prevalence of T2D among higher SES has also been attributed to the consumption of an unhealthy diet and lack of physical activity. This is also supported in a systematic review by Allen et al. that found in low-income and lower-middle income countries, such as India, higher SES groups tended to have higher levels of physical inactivity and consume more fats, salt, and processed foods than low SES groups.<sup>30</sup>

BMI z-scores in SEARCH youth with T1D and T2D were significantly higher compared to YDR youth. A recent article by Hsu et al. discusses the need for different BMI cut points to identify at-risk Asian Americans for T2D screening<sup>31</sup>. This is supported by previous studies that showed the association of BMI with T2D risk is shifted to lower BMI values for Asians<sup>32</sup>, and at similar BMI levels, diabetes prevalence is higher among Asian compared to Whites<sup>33</sup>. Our results are consistent with these prior studies. However, a recent study of Asian adults found that glucose intolerance and T2D remained elevated after adjustment for body composition, suggesting that it may not just be obesity that leads to a higher prevalence of diabetes in Asians.<sup>34</sup> On the other hand, our results may just reflect differences in the

underlying BMI distribution among the general population of youth in the two countries. More research is needed to better understand these differences.

We showed that on average SEARCH youth had significantly lower blood pressure than YDR youth for both T1D and T2D. Prior SEARCH data support this finding by showing that the prevalence of elevated blood pressure is higher in minority youth with T1D and T2D, specifically Asian Pacific Islanders, when compared to Non-Hispanic Whites.<sup>14</sup> We also observed that SEARCH youth had a lower prevalence of classified hypertension compared to YDR for T1D and T2D. This lower prevalence of elevated blood pressure and hypertension in the SEARCH population may be due to differences in measurement. SEARCH blood pressure was obtained using a research protocol, whereas YDR blood pressure data were from clinical measurements. It is possible that differences in treatment exist between the two countries, where the U.S. may be more aggressive in providing treatments for hypertension or pre-hypertension in youth. However, prior SEARCH data suggest that very few US youth with diabetes are treated for hypertension.<sup>14</sup> Unfortunately, in this study, we were not able to explore differences in treatment of hypertension between the two both registries. Alternatively, similar to the BMI results, these results may also result from potential differences in the underlying distribution of blood pressure levels among the general population of youth in the two countries. More research is needed to better understand these differences.

Poor glycemic control, as assessed by A1c levels, is associated with the development and progression of microvascular complications of diabetes in adults and children with T1D and T2D<sup>35,36</sup>. Previous studies, both in the US and internationally, have shown that Asian ethnic groups have poorer glycemic control than non-Hispanic whites<sup>6,37</sup>. This is consistent with our findings; however, the differences are larger in our study. This may be due, in part, to the methodological differences between the two studies. In SEARCH, A1c concentrations were measured at the baseline research visit through a venous sample and analyzed in a central standardized laboratory. In YDR, A1c concentrations were measured clinically at the last visit prior to the YDR baseline visit, without the use of a central standardized laboratory. Due to the variability seen in the various A1c assays<sup>38</sup>, these differences should be interpreted with caution. Additionally, in both registries, the participants' A1c concentrations reflect glycemic control prior to the baseline visit and thus may reflect differences in clinical care patterns between the two countries. While SEARCH is an observational research registry, with little role in clinical decision making, YDR is a clinical registry with a primary goal of optimizing clinical care for registered participants. It will be of interest to compare A1c concentrations between the two registries at follow-up, when treatment patterns are likely to have been optimized for YDR participants. Additional information on A1c levels and associated treatment regimens are presented in another article by Anandakumar et al. within this special edition.

Our study has limitations and strengths. The benefit of harmonizing data between registries is that only one query is required to be executed against all local OMOP datasets, which returns results in the same data structure and semantics, making results adequate to answer the proposed research questions. However, harmonization is a complex process and the most challenging part is being able to ensure the understanding of the structure and meaning of



each data element from each data source. For some data elements, the harmonization can be easily achieved (i.e. sex); in others, the harmonization is more difficult and requires in-depth discussion and agreement about the importance of the differences in the definition or collection method used for a variable. Other harmonization challenges included basic differences in definitions due to differences in the registry types, health care models and care processes between countries, and differences in data collection and validation methods. For example, for diabetes-related complications SEARCH collected self-reported complications using a survey, while YDR collected physician-diagnosed complications from the medical record; this difference in the methodology of data collection lead to uncomparable data. If harmonization is not done correctly, then inappropriate conclusions may occur. Additionally, to protect patient data privacy, only summary data were included in the results of the queries. Another limitation was the availability of YDR data, as only three of the eight YDR sites across India were used for analysis. Therefore, the data used may not be representative of India as a whole. Lastly, there was a large proportion of missing data for specific characteristics (SES for SEARCH participants and A1c for YDR participants). However, given that SES of SEARCH participants who participate in the baseline research visit is higher than the general population (e.g., higher proportion have completed college and report household income >\$50,000/year), it is likely that those missing SES in SEARCH are of low SES, thus only increasing the differences observed and reported here. Within this special edition, Jensen et al look further into the differences in incident rates of youth diabetes (T1D and T2D) between U.S. and India by sex and age at diagnosis and Praveen et al. explore the differences in the the prevalence of DKA between the two registries.

In conclusion, our data offer insights into the demographic and clinical characteristics of diabetes in youth across the two counties, possibly providing clues regarding underlying etiological drivers and/or differences in patterns of care. Further research and collaboration are needed to advance the insights into potential causes and provide access to improved treatments, in order to improve the quality of life for all youth with diabetes.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgements:

The SEARCH for Diabetes in Youth Study is indebted to the many youth and their families and their health care providers, whose participation made this study possible. YDR acknowledges the patients enrolled and the participation of the reporting centres contributing data to YDR. SEARCH for Diabetes in Youth (SEARCH) registry in the U.S. is funded by the Centers for Disease Control and Prevention (CDC) and the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK). The Registry of People with Diabetes with Youth Age at Onset (YDR) in India is funded by the Indian Council of Medical Research (ICMR).

**Grant support:** Study supported by the National Institutes of Health (R21DK105869-02) and the Indian Council of Medical Research.

**SEARCH 3/4:** 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 & Translational Research Institute, at the Medical University of South Carolina, NIH/National Center for Advancing Translational Sciences (NCATS) grant number UL1 TR000062, UL1 Tr001450; 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 number 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.

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.

Grant Support (SEARCH 3): 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.

Grant Support (SEARCH 4): The Population Based Registry of Diabetes in Youth Study (1U18DP006131, U18DP006133, U18DP006134, U18DP006136, U18DP006138, U18DP006139) is funded by the Centers for Disease Control and Prevention and supported by the National Institutes of Health, National Institute of Diabetes and Digestive and Kidney Diseases.

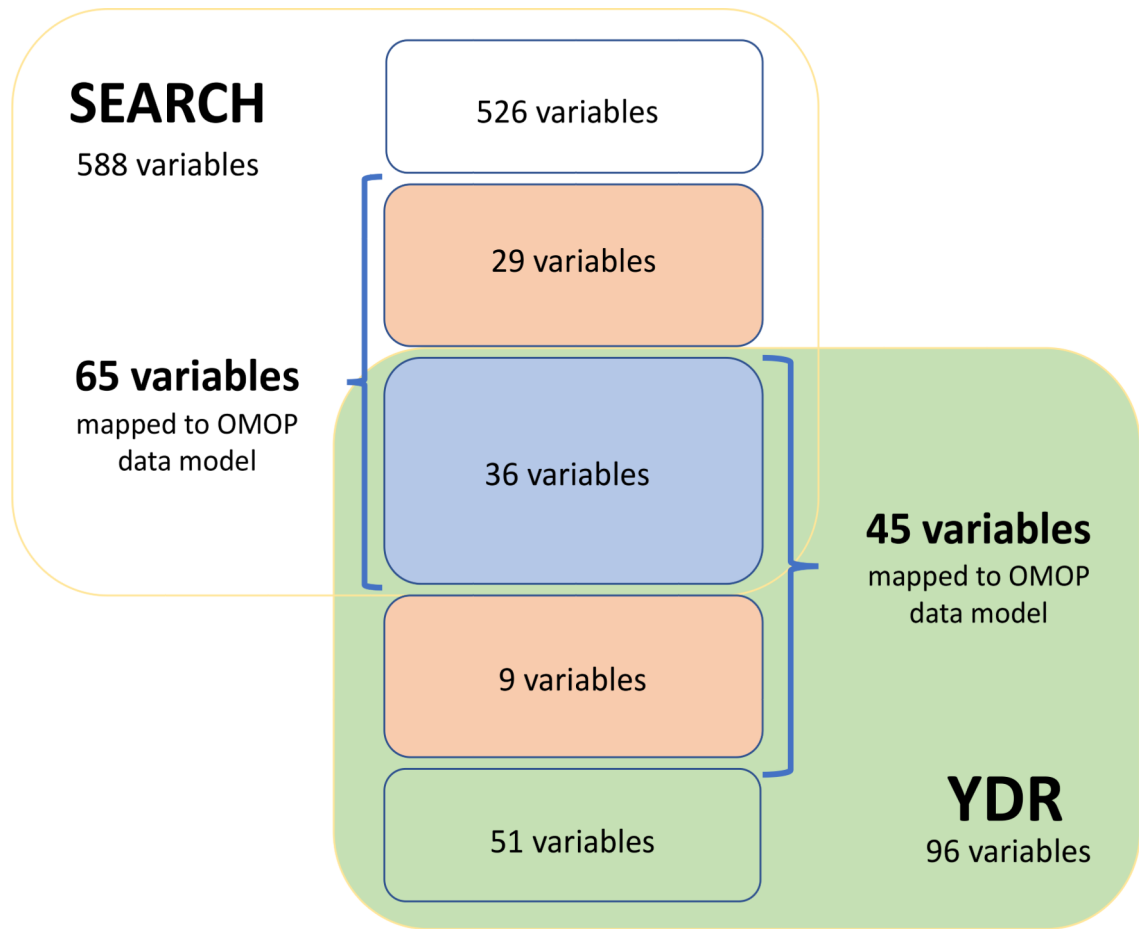
Sites (SEARCH 3/4): Kaiser Permanente Southern California (U18DP006133, U48/CCU919219, U01 DP000246, and U18DP002714), University of Colorado Denver (U18DP006139, U48/CCU819241-3, U01 DP000247, and U18DP000247-06A1), Cincinnati's Children's Hospital Medical Center (U18DP006134, U48/CCU519239, U01 DP000248, and U18DP002709), 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)

## REFERENCES

1. Hamman RF, Bell RA, Dabelea D, et al. The SEARCH for Diabetes in Youth Study: Rationale, Findings, and Future Directions. *Diabetes Care*. 2014;37(12):3336–3344. doi:10.2337/dc14-0574 [PubMed: 25414389]
2. Patterson C, Guariguata L, Dahlquist G, Soltész G, Ogle G, Silink M. Diabetes in the young – a global view and worldwide estimates of numbers of children with type 1 diabetes. *Diabetes Res Clin Pract*. 2014;103(2):161–175. doi:10.1016/j.diabres.2013.11.005 [PubMed: 24331235]
3. Mayer-Davis EJ, Dabelea D, Lawrence JM. Incidence Trends of Type 1 and Type 2 Diabetes among Youths, 2002–2012. *N Engl J Med*. 2017;377(3):301. doi:10.1056/NEJMc1706291
4. Pettitt DJ, Talton J, Dabelea D, et al. Prevalence of Diabetes in U.S. Youth in 2009: The SEARCH for Diabetes in Youth Study. *Diabetes Care*. 2014;37(2):402–408. doi:10.2337/dc13-1838 [PubMed: 24041677]
5. Pradeepa R, Mohan V. Prevalence of type 2 diabetes and its complications in India and economic costs to the nation. *Eur J Clin Nutr*. 2017;71(7):816–824. doi:10.1038/ejcn.2017.40 [PubMed: 28422124]
6. Pettitt DB, Klingensmith GJ, Bell RA, et al. Glycemic control in youth with diabetes: the SEARCH for diabetes in Youth Study. *J Pediatr*. 2009;155(5):668–672.e1-3. doi:10.1016/j.jpeds.2009.05.025 [PubMed: 19643434]
7. 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(8):825–835. doi:10.1001/jama.2017.0686 [PubMed: 28245334]
8. Soltész G, Patterson CC, Dahlquist G, EURODIAB Study Group. Worldwide childhood type 1 diabetes incidence--what can we learn from epidemiology? *Pediatr Diabetes*. 2007;8 Suppl 6:6–14. doi:10.1111/j.1399-5448.2007.00280.x [PubMed: 17727380]
9. Praveen PA, Madhu SV, Mohan V, et al. Registry of Youth Onset Diabetes in India (YDR). *J Diabetes Sci Technol*. 2016;10(5):1034–1041. doi:10.1177/1932296816645121 [PubMed: 27179010]
10. Taniar D, Chen L. Integrations of Data Warehousing, Data Mining and Database Technologies: Innovative Approaches. IGI Global; 2001 <https://www.igi-global.com/book/integrations-data-warehousing-data-mining/45967>. Accessed March 5, 2019.

11. Definition and DDLs for the OMOP Common Data Model (CDM): OHDSI/CommonDataModel. Observational Health Data Sciences and Informatics; 2018 <https://github.com/OHDSI/CommonDataModel>. Accessed December 6, 2018.
12. Kahn MG, Batson D, Schilling LM. Data model considerations for clinical effectiveness researchers. *Med Care*. 2012;50 Suppl:S60–67. doi:10.1097/MLR.0b013e318259bfff4 [PubMed: 22692260]
13. Hripcsak G, Duke JD, Shah NH, et al. Observational Health Data Sciences and Informatics (OHDSI): Opportunities for Observational Researchers. *Stud Health Technol Inform*. 2015;216:574–578. [PubMed: 26262116]
14. Rodriguez BL, Dabelea D, Liese AD, et al. Prevalence and correlates of elevated blood pressure in youth with diabetes mellitus: the SEARCH for diabetes in youth study. *J Pediatr*. 2010;157(2):245–251.e1. doi:10.1016/j.jpeds.2010.02.021 [PubMed: 20394942]
15. WHO | The WHO Child Growth Standards. WHO <http://www.who.int/childgrowth/en/>. Accessed November 26, 2018.
16. de Onis M Development of a WHO growth reference for school-aged children and adolescents. *Bull World Health Organ*. 2007;85(09):660–667. doi:10.2471/BLT.07.043497 [PubMed: 18026621]
17. National High Blood Pressure Education Program Working Group on High Blood Pressure in Children and Adolescents. The fourth report on the diagnosis, evaluation, and treatment of high blood pressure in children and adolescents. *Pediatrics*. 2004;114(2 Suppl 4th Report):555–576. [PubMed: 15286277]
18. American Diabetes Association. 6. Glycemic Targets: Standards of Medical Care in Diabetes—2019. *Diabetes Care*. 2019;42(Supplement 1):S61–S70. doi:10.2337/dc19-S006 [PubMed: 30559232]
19. Unnikrishnan AG, Bhatia E, Bhatia V, et al. Type 1 diabetes versus type 2 diabetes with onset in persons younger than 20 years of age. *Ann N Y Acad Sci*. 2008;1150:239–244. doi:10.1196/annals.1447.056 [PubMed: 19120303]
20. Bhatia V, Arya V, Dabadghao P, et al. Etiology and Outcome of Childhood and Adolescent Diabetes Mellitus in North India. *J Pediatr Endocrinol Metab*. 2011;17(7):993–1000. doi:10.1515/JPEM.2004.17.7.993
21. Borchers AT, Uibo R, Gershwin ME. The geoepidemiology of type 1 diabetes. *Autoimmun Rev*. 2010;9(5):A355–365. doi:10.1016/j.autrev.2009.12.003 [PubMed: 19969107]
22. Liese AD, Puett RC, Lamichhane AP, et al. Neighborhood level risk factors for type 1 diabetes in youth: the SEARCH case-control study. *Int J Health Geogr*. 2012;11:1. doi:10.1186/1476-072X-11-1 [PubMed: 22230476]
23. Siemiatycki J, Colle E, Campbell S, Dewar R, Aubert D, Belmonte MM. Incidence of IDDM in Montreal by ethnic group and by social class and comparisons with ethnic groups living elsewhere. *Diabetes*. 1988;37(8):1096–1102. [PubMed: 3391344]
24. Connolly V, Unwin N, Sherriff P, Bilous R, Kelly W. Diabetes prevalence and socioeconomic status: a population based study showing increased prevalence of type 2 diabetes mellitus in deprived areas. *J Epidemiol Community Health*. 2000;54(3):173–177. doi:10.1136/jech.54.3.173 [PubMed: 10746110]
25. Robbins JM, Vaccarino V, Zhang H, Kasl SV. Socioeconomic status and diagnosed diabetes incidence. *Diabetes Res Clin Pract*. 2005;68(3):230–236. doi:10.1016/j.diabres.2004.09.007 [PubMed: 15936465]
26. Pulgaron ER, Delamater AM. Obesity and Type 2 Diabetes in Children: Epidemiology and Treatment. *Curr Diab Rep*. 2014;14(8):508. doi:10.1007/s11892-014-0508-y [PubMed: 24919749]
27. Mohan V, Shanthirani S, Deepa R, et al. Intra-urban differences in the prevalence of the metabolic syndrome in southern India -- the Chennai Urban Population Study (CUPS No. 4). *Diabet Med J Br Diabet Assoc*. 2001;18(4):280–287.
28. Mbanya JC, Ngogang J, Salah JN, Minkoulou E, Balkau B. Prevalence of NIDDM and impaired glucose tolerance in a rural and an urban population in Cameroon. *Diabetologia*. 1997;40(7):824–829. doi:10.1007/s001250050755 [PubMed: 9243104]

29. abu Sayeed M, Ali L, Hussain MZ, Rumi MA, Banu A, Azad Khan AK. Effect of socioeconomic risk factors on the difference in prevalence of diabetes between rural and urban populations in Bangladesh. *Diabetes Care*. 1997;20(4):551–555. [PubMed: 9096979]
30. Allen L, Williams J, Townsend N, et al. Socioeconomic status and non-communicable disease behavioural risk factors in low-income and lower-middle-income countries: a systematic review. *Lancet Glob Health*. 2017;5(3):e277–e289. doi:10.1016/S2214-109X(17)30058-X [PubMed: 28193397]
31. Hsu WC, Araneta MRG, Kanaya AM, Chiang JL, Fujimoto W. BMI Cut Points to Identify At-Risk Asian Americans for Type 2 Diabetes Screening. *Diabetes Care*. 2015;38(1):150–158. doi:10.2337/dc14-2391 [PubMed: 25538311]
32. Chan JCN, Malik V, Jia W, et al. Diabetes in Asia: epidemiology, risk factors, and pathophysiology. *JAMA*. 2009;301(20):2129–2140. doi:10.1001/jama.2009.726 [PubMed: 19470990]
33. King GL, McNeely MJ, Thorpe LE, et al. Understanding and Addressing Unique Needs of Diabetes in Asian Americans, Native Hawaiians, and Pacific Islanders. *Diabetes Care*. 2012;35(5):1181–1188. doi:10.2337/dc12-0210 [PubMed: 22517939]
34. Flowers E, Lin F, Kandula NR, et al. Body Composition and Diabetes Risk in South Asians: Findings From the MASALA and MESA Studies. *Diabetes Care*. 2019;42(11):2019–2026. doi:10.2337/dc18-1510
35. Diabetes Control and Complications Trial Research Group, Nathan DM, Genuth S, et al. The effect of intensive treatment of diabetes on the development and progression of long-term complications in insulin-dependent diabetes mellitus. *N Engl J Med*. 1993;329(14):977–986. doi:10.1056/NEJM199309303291401 [PubMed: 8366922]
36. Intensive blood-glucose control with sulphonylureas or insulin compared with conventional treatment and risk of complications in patients with type 2 diabetes (UKPDS 33). UK Prospective Diabetes Study (UKPDS) Group. *Lancet Lond Engl*. 1998;352(9131):837–853.
37. Herman WH, Dungan KM, Wolffenbittel BHR, et al. Racial and Ethnic Differences in Mean Plasma Glucose, Hemoglobin A1c, and 1,5-Anhydroglucitol in Over 2000 Patients with Type 2 Diabetes. *J Clin Endocrinol Metab*. 2009;94(5):1689–1694. doi:10.1210/jc.2008-1940 [PubMed: 19276235]
38. Chan CL, McFann K, Newnes L, Nadeau KJ, Zeitler PS, Kelsey M. Hemoglobin A1c Assay Variations and Implications for Diabetes Screening in Obese Youth. *Pediatr Diabetes*. 2014;15(8):557–563. doi:10.1111/peidi.12132 [PubMed: 24636682]



**Figure 1.**

A diagram showing the number of variables identified as possible matches between the SEARCH (yellow square) and YDR (green square) registries and how many were mapped to the OMOP data model (within the blue brackets) and of those mapped to the OMOP data model, how many were mapped to the standard OMOP terminologies (blue box) and how many were mapped to custom terminologies (orange boxes).

**Table 1.** Demographic characteristics for SEARCH (N=9,725) and YDR (N=2,331) youth by diabetes type, incident years 2006–2012

Variables	Type 1 diabetes		Type 2 diabetes		p-value
	SEARCH 2006–2012 n=7,546	YDR 2006–2012 n=2,104	SEARCH 2006–2012 n=4,179	YDR 2006–2012 n=227	
Age at diagnosis (yrs), mean (SD)	10.0 (4.5)	10.5 (4.9)	14.7 (2.8)	16.1 (2.8)	<0.001
Age at diagnosis (yrs), n (%)					<0.001
0–4	1203 (15.9)	330 (15.7)	5 (0.2)	0	
5–9	2436 (32.3)	646 (30.7)	105 (4.8)	10 (4.4)	
10–14	2733 (36.2)	688 (32.7)	986 (45.3)	58 (25.6)	
15–19	1174 (15.6)	440 (20.9)	1083 (49.7)	159 (70.0)	
Gender, n (%)					<0.001
Female	3526 (46.7)	990 (47.1)	1369 (62.8)	111 (48.9)	
Male	4020 (53.3)	1114 (52.9)	810 (37.2)	116 (51.1)	
Race, n (%)					
White	5496 (72.8)		489 (22.4)		
Asian Indian	26 (0.3)	2104 (100.0)	4 (0.2)	227 (100.0)	
Black or African American	808 (10.7)		790 (36.3)		
Other Asian	92 (1.2)		83 (3.8)		
American Indian or Alaska Native	54 (0.7)		167 (7.7)		
Native Hawaiian or Other Pacific Islander	11 (0.1)		19 (0.9)		
Multiple Race	343 (4.5)		62 (2.8)		
Unknown	716 (9.5)		565 (25.9)		
Socioeconomic status, n (%)					<0.001
High	4047 (53.6)	1279 (60.8)	515 (23.6)	201 (88.5)	
Low	1770 (23.5)	825 (39.2)	802 (36.8)	26 (11.5)	
Missing	1729 (22.9)	0 (0.0)	862 (39.6)	0 (0.0)	

**Table 2.** Clinical characteristics for SEARCH (N=2,283) and YDR (N=2,331) youth by diabetes type, incident years 2006, 08, 12 for SEARCH and 2006–2012 for YDR

Variables	Type 1 diabetes			Type 2 diabetes		
	SEARCH <sup>^</sup> 2006–2012 n=1,899	YDR 2006–2012 n=2,104	p-value	SEARCH <sup>^</sup> 2006–2012 n=384	YDR 2006–2012 n=227	p-value
Age at diagnosis (yrs), mean (SD)	10.1 (4.3)	10.5 (4.9)	<0.001	14.4 (2.7)	16.1 (2.8)	<0.001
Age at baseline visit (yrs), n (%)			<0.001			<0.001
<b>0–4</b>	198 (10.4)	296 (14.1)		0	0 (0.0)	
<b>5–9</b>	548 (28.9)	612 (29.1)		7 (1.8)	8 (3.5)	
<b>10–14</b>	786 (41.4)	698 (33.2)		157 (40.9)	48 (21.1)	
<b>15–19</b>	354 (18.6)	443 (21.1)		199 (51.8)	151 (66.5)	
<b>20+</b>	13 (0.7)	55 (2.6)		21 (5.5)	20 (8.8)	
Duration of diabetes (mos), mean (SD)	10.3 (7.4)	6.2 (9.7)	<0.001	13.3 (8.2)	7.0 (10.5)	<0.001
Height (cm), mean (SD)	146.6 (23.0)	144.5 (24.3)	<0.001	166.0 (11.4)	156.9 (18.0)	<0.001
Weight (kg), mean (SD)	45.8 (21.6)	32.9 (16.6)	<0.001	98.4 (27.1)	68.6 (21.5)	<0.001
BMI z-score (WHO), mean (SD)	0.69 (1.23)	-0.54 (1.65)	<0.001	3.17 (1.15)	1.65 (2.10)	<0.001
BMI Categories, n (%)			<0.001			<0.001
<b>Underweight</b>	11 (0.6)	319 (15.2)		0 (0.0)	9 (4.0)	
<b>Normal</b>	1144 (60.2)	1280 (60.8)		10 (2.6)	46 (20.3)	
<b>Overweight</b>	400 (21.1)	159 (7.6)		25 (6.5)	49 (21.6)	
<b>Obese</b>	231 (12.2)	90 (4.3)		303 (78.9)	83 (36.6)	
<b>Missing</b>	113 (6.0)	256 (12.2)		46 (12.0)	40 (17.6)	
Blood pressure, mean (SD)			<0.001			<0.001
<b>SBP</b>	99.7 (11.8)	105.5 (15.0)		115.9 (12.6)	119.2 (14.5)	
<b>DBP</b>	63.3 (9.6)	69.0 (10.0)		72.0 (10.1)	78.3 (9.3)	
Hypertension <sup>*</sup> , n (%)	53 (3.0)	171 (14.2)	<0.001	59 (15.7)	46 (23.5)	<0.001
A1c, mean (SD)	7.8 (1.7)	11.0 (2.9)	<0.001	7.2 (2.1)	9.9 (2.8)	<0.001
A1c Categories, n (%)			<0.001			<0.001
<b>Less than 7.5</b>	797 (42.0)	151 (7.2)		260 (67.7)	41 (18.1)	
<b>7.5–9.0</b>	620 (32.6)	214 (10.2)		50 (13.0)	29 (12.8)	

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Variables	Type 1 diabetes		Type 2 diabetes		p-value	p-value
	SEARCH <sup>^</sup>	YDR	SEARCH <sup>^</sup>	YDR		
	2006–2012 n=1,899	2006–2012 n=2,104	2006–2012 n=384	2006–2012 n=227		
Greater than 9.0	355 (18.7)	959 (45.6)	64 (16.7)	98 (43.2)		
Missing	127 (6.7)	780 (37.1)	10 (2.6)	59 (26.0)		

<sup>^</sup> a subset of SEARCH participants diagnosed in 2006, 2008, and 2012 who completed a baseline visit (a SEARCH baseline research visit)

\* SEARCH T1D (n=1795), T2D (n=377); YDR T1D (n=1204), T2D (n=196)